

УЭЙН ВИНСТОН



# БИЗНЕС-МОДЕЛИРОВАНИЕ И АНАЛИЗ ДАННЫХ

РЕШЕНИЕ АКТУАЛЬНЫХ ЗАДАЧ С ПОМОЩЬЮ

# Microsoft Excel

- КАК ВЫЧИСЛИТЬ ВНУТРЕННЮЮ СТАВКУ ДОХОДНОСТИ
- КАКАЯ ЦЕНА ОБЕСПЕЧИТ МАКСИМАЛЬНУЮ ПРИБЫЛЬ
- ЧЕРЕЗ СКОЛЬКО ЛЕТ ОКУПИТСЯ ПРОЕКТ
- КАКОЙ ПРОЕКТ ВЫБРАТЬ ДЛЯ РЕАЛИЗАЦИИ
- ГЕОГРАФИЧЕСКИЕ ДАННЫЕ И ФОРМУЛЫ EXCEL

6-Е ИЗДАНИЕ

БЕСТСЕЛЛЕР AMAZON





# **Microsoft Excel 2019 Data Analysis and Business Modeling**

**Sixth Edition**

**Wayne L. Winston**



УЭЙН ВИНСТОН

# БИЗНЕС-МОДЕЛИРОВАНИЕ И АНАЛИЗ ДАННЫХ

РЕШЕНИЕ АКТУАЛЬНЫХ ЗАДАЧ С ПОМОЩЬЮ

# Microsoft Excel

6-Е ИЗДАНИЕ



Санкт-Петербург • Москва • Екатеринбург • Воронеж  
Нижний Новгород • Ростов-на-Дону • Самара • Минск

2021

ББК 32.973.23-018.2  
УДК 004.67  
В49

### **Винстон Уэйн**

**В49** Бизнес-моделирование и анализ данных. Решение актуальных задач с помощью Microsoft Excel. 6-е издание. — СПб.: Питер, 2021. — 944 с.: ил. — (Серия «IT для бизнеса»).

ISBN 978-5-4461-1446-7

Уэйн Винстон научит вас быстро анализировать данные, принимать решения, подводить итоги, составлять отчеты, обрабатывать данные и строить аналитические модели в Microsoft Excel 2019 и Office 365. В новом шестом издании вас ждут более 800 бизнес-задач, основанных на реальных ситуациях, а также обсуждение новых инструментов и функций. Где бы вы ни работали — в крупной корпорации, небольшой компании, государственной или некоммерческой структуре, — это поможет вам увеличить прибыль, снизить издержки или эффективно управлять производством.

Прочитав эту книгу, вы сможете спрогнозировать результаты выборов, научитесь определять точки безубыточности, рассчитывать вероятность выигрыша в кости или победы любимой команды в турнире.

Хотите обогнать конкурентов? Решайте в Excel реальные задачи!

**16+** (В соответствии с Федеральным законом от 29 декабря 2010 г. № 436-ФЗ.)

ББК 32.973.23-018.2  
УДК 004.67

Права на издание получены по соглашению с Pearson Education Inc. Все права защищены. Никакая часть данной книги не может быть воспроизведена в какой бы то ни было форме без письменного разрешения владельцев авторских прав.

Информация, содержащаяся в данной книге, получена из источников, рассматриваемых издательством как надежные. Тем не менее, имея в виду возможные человеческие или технические ошибки, издательство не может гарантировать абсолютную точность и полноту приводимых сведений и не несет ответственности за возможные ошибки, связанные с использованием книги. Издательство не несет ответственности за доступность материалов, ссылки на которые вы можете найти в этой книге. На момент подготовки книги к изданию все ссылки на интернет-ресурсы были действующими.

ISBN 978-1509305889 англ.  
ISBN 978-5-4461-1446-7

© 2019 by Pearson Education, Inc.  
© Перевод на русский язык ООО Издательство «Питер», 2021  
© Издание на русском языке, оформление ООО Издательство «Питер», 2021  
© Серия «IT для бизнеса», 2021

# Оглавление

<b>Оба вторе .....</b>	<b>16</b>
<b>Введение .....</b>	<b>17</b>
Новое в этом издании .....	18
Что необходимо знать для работы с книгой .....	19
Как работать с книгой.....	19
Сопроводительный сайт .....	20
Благодарности .....	20
Поддержка и обратная связь.....	21
От издательства.....	21
<b>Глава 1   Основы моделирования в электронных таблицах.....</b>	<b>22</b>
Ответы на вопросы .....	22
Задания .....	30
<b>Глава 2   Имена диапазонов .....</b>	<b>31</b>
Как создать именованный диапазон? .....	31
Ответы на вопросы .....	36
Задания .....	41
<b>Глава 3   Функции поиска .....</b>	<b>43</b>
Синтаксис функций поиска .....	43
Ответы на вопросы .....	44
Задания .....	48
<b>Глава 4   Функция ИНДЕКС .....</b>	<b>51</b>
Синтаксис функции ИНДЕКС .....	51
Ответы на вопросы .....	51
Задания .....	53
<b>Глава 5   Функция ПОИСКПОЗ .....</b>	<b>54</b>
Ответы на вопросы .....	56
Задания .....	60

<b>Глава 6</b>	<b>Текстовые функции и инструмент Мгновенное заполнение.....</b>	<b>61</b>
	Синтаксис текстовых функций .....	62
	Ответы на вопросы .....	66
	Задания .....	77
<b>Глава 7</b>	<b>Даты и функции работы с датами .....</b>	<b>80</b>
	Ответы на вопросы .....	81
	Задания .....	86
<b>Глава 8</b>	<b>Оценка инвестиций по чистой приведенной стоимости .....</b>	<b>89</b>
	Ответы на вопросы .....	90
	Задания .....	94
<b>Глава 9</b>	<b>Внутренняя ставка доходности .....</b>	<b>96</b>
	Ответы на вопросы .....	97
	Задания .....	102
<b>Глава 10</b>	<b>Еще несколько финансовых функций Excel .....</b>	<b>104</b>
	Ответы на вопросы .....	104
	Задания .....	114
<b>Глава 11</b>	<b>Циклические ссылки .....</b>	<b>119</b>
	Ответы на вопросы .....	119
	Задания .....	122
<b>Глава 12</b>	<b>Функции ЕСЛИ, ЕСЛИОШИБКА, ЕСЛИМН, ВЫБОР и ПЕРЕКЛЮЧ... ..</b>	<b>125</b>
	Ответы на вопросы .....	126
	Задания .....	145
<b>Глава 13</b>	<b>Время и функции времени .....</b>	<b>151</b>
	Ответы на вопросы .....	151
	Задания .....	157
<b>Глава 14</b>	<b>Команда Специальная вставка .....</b>	<b>158</b>
	Ответы на вопросы .....	158
	Задания .....	164
<b>Глава 15</b>	<b>Трехмерные формулы и гиперссылки .....</b>	<b>165</b>
	Ответы на вопросы .....	165
	Задания .....	169

<b>Глава 16</b>	<b>Инструменты проверки зависимостей и надстройка Inquire .....</b>	<b>170</b>
	Ответы на вопросы .....	173
	Задания .....	181
<b>Глава 17</b>	<b>Анализ чувствительности с помощью таблиц данных .....</b>	<b>182</b>
	Ответы на вопросы .....	183
	Задания .....	191
<b>Глава 18</b>	<b>Инструмент Подбор параметра.....</b>	<b>195</b>
	Ответы на вопросы .....	196
	Задания .....	198
<b>Глава 19</b>	<b>Анализ чувствительности с помощью Диспетчера сценариев...</b>	<b>201</b>
	Ответ на вопрос.....	201
	Задания .....	206
<b>Глава 20</b>	<b>Функции СЧЁТЕСЛИ, СЧЁТЕСЛИМН, СЧЁТ, СЧЁТЗ и СЧИТАТЬПУСТОТЫ .....</b>	<b>207</b>
	Ответы на вопросы .....	209
	Задания .....	213
<b>Глава 21</b>	<b>Функции СУММЕСЛИ, СРЗНАЧЕСЛИ, СУММЕСЛИМН, СРЗНАЧЕСЛИМН, МАКСЕСЛИ и МИНЕСЛИ .....</b>	<b>215</b>
	Ответы на вопросы .....	217
	Задания .....	221
<b>Глава 22</b>	<b>Функция СМЕЩ .....</b>	<b>223</b>
	Ответы на вопросы .....	224
	Задания .....	235
<b>Глава 23</b>	<b>Функция ДВССЫЛ .....</b>	<b>238</b>
	Ответы на вопросы .....	239
	Задания .....	248
<b>Глава 24</b>	<b>Условное форматирование .....</b>	<b>249</b>
	Ответы на вопросы .....	251
	Задания .....	276
<b>Глава 25</b>	<b>Сортировка в Excel.....</b>	<b>281</b>
	Ответы на вопросы .....	281
	Задания .....	289

<b>Глава 26</b>	<b>Таблицы .....</b>	<b>291</b>
	Ответы на вопросы .....	291
	Задания .....	305
<b>Глава 27</b>	<b>Счетчики, полосы прокрутки, переключатели, флажки, группы и поля со списками .....</b>	<b>307</b>
	Ответы на вопросы .....	308
	Задания .....	318
<b>Глава 28</b>	<b>Революция в аналитике .....</b>	<b>320</b>
	Ответы на вопросы .....	320
<b>Глава 29</b>	<b>Введение в оптимизацию с надстройкой Поиск решения .....</b>	<b>326</b>
	Задания .....	330
<b>Глава 30</b>	<b>Поиск решения при определении оптимального ассортимента продукции.....</b>	<b>331</b>
	Ответы на вопросы .....	331
	Задания .....	341
<b>Глава 31</b>	<b>Поиск решения при планировании расписания работы сотрудников .....</b>	<b>344</b>
	Ответ на вопрос.....	344
	Задания .....	347
<b>Глава 32</b>	<b>Поиск решения для задач транспортировки и распределения ....</b>	<b>349</b>
	Ответ на вопрос.....	349
	Задания .....	352
<b>Глава 33</b>	<b>Поиск решения для бюджетирования капиталовложений.....</b>	<b>355</b>
	Ответ на вопрос.....	355
	Задания .....	360
<b>Глава 34</b>	<b>Поиск решения при финансовом планировании .....</b>	<b>363</b>
	Ответы на вопросы .....	364
	Задания .....	368
<b>Глава 35</b>	<b>Поиск решения при оценке спортивных команд .....</b>	<b>370</b>
	Ответ на вопрос.....	371
	Задания .....	374

<b>Глава 36</b>	<b>Расположение складов по методу ОПГ с несколькими начальными точками и согласно эволюционному поиску решения .....</b>	<b>376</b>
	Метод ОПГ с несколькими начальными точками и эволюционный поиск решения.....	376
	Ответы на вопросы .....	381
	Задания .....	385
<b>Глава 37</b>	<b>Штрафы и эволюционный поиск решения .....</b>	<b>386</b>
	Ответы на вопросы .....	386
	Задания .....	390
<b>Глава 38</b>	<b>Задача коммивояжера.....</b>	<b>393</b>
	Ответы на вопросы .....	393
	Задания .....	396
<b>Глава 39</b>	<b>Импорт данных из текстового файла или документа.....</b>	<b>398</b>
	Ответ на вопрос.....	398
	Задания .....	403
<b>Глава 40</b>	<b>Инструмент Получить и преобразовать данные .....</b>	<b>404</b>
	Ответы на вопросы .....	405
	Задания .....	410
<b>Глава 41</b>	<b>Типы данных «Акции» и «География».....</b>	<b>411</b>
	Ответы на вопросы .....	411
	Задания .....	415
<b>Глава 42</b>	<b>Проверка достоверности данных .....</b>	<b>416</b>
	Ответы на вопросы .....	416
	Задания .....	422
<b>Глава 43</b>	<b>Обобщение данных на гистограммах и диаграммах Парето .....</b>	<b>425</b>
	Ответы на вопросы .....	425
	Задания .....	437
<b>Глава 44</b>	<b>Обобщение данных с помощью описательной статистики .....</b>	<b>439</b>
	Ответы на вопросы .....	440
	Задания .....	457
<b>Глава 45</b>	<b>Сводные таблицы и срезы для описания данных.....</b>	<b>460</b>
	Ответы на вопросы .....	461
	Задания .....	505



<b>Глава 46</b>	<b>Модель данных .....</b>	<b>509</b>
	Ответы на вопросы .....	509
	Задания .....	516
<b>Глава 47</b>	<b>Power Pivot .....</b>	<b>517</b>
	Ответы на вопросы .....	518
	Задания .....	531
<b>Глава 48</b>	<b>2D-картограммы и 3D-карты Power Map .....</b>	<b>532</b>
	Ответы на вопросы .....	532
	Задания .....	543
<b>Глава 49</b>	<b>Спарклайны.....</b>	<b>544</b>
	Ответы на вопросы .....	544
	Задания .....	548
<b>Глава 50</b>	<b>Обработка данных с помощью статистических функций для баз данных .....</b>	<b>549</b>
	Ответы на вопросы .....	551
	Задания .....	556
<b>Глава 51</b>	<b>Фильтрация данных и удаление дубликатов .....</b>	<b>558</b>
	Ответы на вопросы .....	560
	Задания .....	572
<b>Глава 52</b>	<b>Консолидация данных.....</b>	<b>573</b>
	Ответ на вопрос.....	573
	Задания .....	577
<b>Глава 53</b>	<b>Создание промежуточных итогов .....</b>	<b>578</b>
	Ответы на вопросы .....	578
	Задания .....	582
<b>Глава 54</b>	<b>Приемы работы с диаграммами .....</b>	<b>583</b>
	Ответы на вопросы .....	584
	Задания .....	622
<b>Глава 55</b>	<b>Оценка линейных зависимостей .....</b>	<b>625</b>
	Ответы на вопросы .....	627
	Задания .....	632

<b>Глава 56</b>	<b>Моделирование экспоненциального роста.....</b>	<b>634</b>
	Ответ на вопрос.....	635
	Задания .....	637
<b>Глава 57</b>	<b>Степенная кривая .....</b>	<b>638</b>
	Ответ на вопрос.....	640
	Задания .....	643
<b>Глава 58</b>	<b>Представление зависимостей с помощью корреляции .....</b>	<b>645</b>
	Ответ на вопрос.....	647
	Задания .....	650
<b>Глава 59</b>	<b>Введение во множественную регрессию.....</b>	<b>652</b>
	Ответы на вопросы .....	652
<b>Глава 60</b>	<b>Включение качественных факторов во множественную регрессию.....</b>	<b>659</b>
	Ответы на вопросы .....	659
<b>Глава 61</b>	<b>Моделирование нелинейных характеристик и взаимосвязей ...</b>	<b>670</b>
	Ответы на вопросы .....	670
	Задания к главам 59–61 .....	674
<b>Глава 62</b>	<b>Однофакторный дисперсионный анализ.....</b>	<b>678</b>
	Ответы на вопросы .....	679
	Задания .....	683
<b>Глава 63</b>	<b>Рандомизированные блоки и двухфакторный дисперсионный анализ .....</b>	<b>684</b>
	Ответы на вопросы .....	685
	Задания .....	693
<b>Глава 64</b>	<b>Скользящие средние для временных рядов.....</b>	<b>694</b>
	Ответ на вопрос.....	694
	Задание .....	696
<b>Глава 65</b>	<b>Метод Винтерса и Лист прогноза.....</b>	<b>697</b>
	Характеристики временных рядов.....	697
	Определение параметров .....	698
	Определение начальных параметров для метода Винтерса.....	698
	Вычисление сглаживающих постоянных.....	699

	Инструмент Excel Лист прогноза .....	701
	Задания .....	704
<b>Глава 66</b>	<b>Метод прогнозирования «по отношению к скользящему среднему» .....</b>	<b>705</b>
	Ответы на вопросы .....	705
	Задание .....	708
<b>Глава 67</b>	<b>Прогноз для особых случаев .....</b>	<b>709</b>
	Ответы на вопросы .....	709
	Задания .....	717
<b>Глава 68</b>	<b>Введение в теорию вероятности .....</b>	<b>718</b>
	Ответы на вопросы .....	718
	Задания .....	725
<b>Глава 69</b>	<b>Введение в случайные величины.....</b>	<b>728</b>
	Ответы на вопросы .....	728
	Задания .....	732
<b>Глава 70</b>	<b>Биномиальные, гипергеометрические и отрицательные биномиальные случайные величины .....</b>	<b>733</b>
	Ответы на вопросы .....	734
	Задания .....	740
<b>Глава 71</b>	<b>Пуассоновская и экспоненциальная случайные величины .....</b>	<b>742</b>
	Ответы на вопросы .....	742
	Задания .....	745
<b>Глава 72</b>	<b>Нормальная случайная величина и Z-оценка .....</b>	<b>747</b>
	Ответы на вопросы .....	747
	Задания .....	755
<b>Глава 73</b>	<b>Распределение Вейбулла и бета-распределение: моделирование надежности механизмов и продолжительности работы .....</b>	<b>758</b>
	Ответы на вопросы .....	758
	Задания .....	763
<b>Глава 74</b>	<b>Создание вероятностных высказываний на основе прогнозов .</b>	<b>764</b>
	Ответы на вопросы .....	765
	Задания .....	767

<b>Глава 75</b>	<b>Логарифмически нормальная случайная величина в моделировании курса акций.....</b>	<b>768</b>
	Ответы на вопросы .....	768
	Задания .....	772
<b>Глава 76</b>	<b>Импорт в Excel истории торгов (загрузка биржевых данных)...</b>	<b>773</b>
	Ответы на вопросы .....	773
	Задания .....	776
<b>Глава 77</b>	<b>Введение в моделирование по методу Монте-Карло.....</b>	<b>777</b>
	Ответы на вопросы .....	777
	Задания .....	786
<b>Глава 78</b>	<b>Вычисление оптимальной цены предложения .....</b>	<b>788</b>
	Ответы на вопросы .....	788
	Задания .....	792
<b>Глава 79</b>	<b>Моделирование цен на акции и распределения средств между активами.....</b>	<b>794</b>
	Ответы на вопросы .....	795
	Задания .....	803
<b>Глава 80</b>	<b>Развлечения и игры: моделирование вероятностей для азартных игр и спортивных соревнований .....</b>	<b>804</b>
	Ответы на вопросы .....	804
	Задания .....	811
<b>Глава 81</b>	<b>Анализ данных с помощью повторной выборки.....</b>	<b>813</b>
	Ответ на вопрос.....	813
	Задания .....	816
<b>Глава 82</b>	<b>Ценообразование опционов.....</b>	<b>818</b>
	Ответы на вопросы .....	818
	Задания .....	830
<b>Глава 83</b>	<b>Определение потребительской ценности .....</b>	<b>833</b>
	Ответы на вопросы .....	833
	Задания .....	837
<b>Глава 84</b>	<b>Оптимальный размер заказа в модели управления запасами...</b>	<b>839</b>
	Ответы на вопросы .....	839
	Задания .....	843

<b>Глава 85</b>	<b>Построение моделей управления запасами для неопределенного спроса .....</b>	<b>844</b>
	Ответы на вопросы .....	845
	Задания .....	850
<b>Глава 86</b>	<b>Теория массового обслуживания (теория очередей) .....</b>	<b>851</b>
	Ответы на вопросы .....	851
	Задания .....	856
<b>Глава 87</b>	<b>Оценка кривой спроса .....</b>	<b>858</b>
	Ответы на вопросы .....	858
	Задания .....	862
<b>Глава 88</b>	<b>Ценообразование продуктов с сопутствующими товарами .....</b>	<b>863</b>
	Ответ на вопрос .....	863
	Задания .....	866
<b>Глава 89</b>	<b>Ценообразование продуктов с помощью субъективно определяемого спроса .....</b>	<b>868</b>
	Ответы на вопросы .....	868
	Задания .....	871
<b>Глава 90</b>	<b>Нелинейное ценообразование .....</b>	<b>874</b>
	Ответы на вопросы .....	874
	Задания .....	882
<b>Глава 91</b>	<b>Формулы массива и функции, возвращающие массив .....</b>	<b>883</b>
	Ответы на вопросы .....	884
	Задания .....	903
<b>Глава 92</b>	<b>Запись макросов .....</b>	<b>907</b>
	Ответы на вопросы .....	907
	Задания .....	918
<b>Глава 93</b>	<b>Продвинутый анализ чувствительности .....</b>	<b>919</b>
	Ответ на вопрос .....	919
	Задания .....	921

*Вивиан, Джен и Грегу.*

*Вы замечательные, и я вас очень люблю!*

# Об авторе



**Уэйн Л. Винстон** — почетный профессор Школы бизнеса Kelley School of Business при Университете Индианы. Он также преподавал в Университете Хьюстона и Уэйк-Форест. Он получил свыше 40 наград для преподавателей, обучил работе в Excel и моделированию в среде Excel множество бизнес-аналитиков из компаний, входящих в Fortune 500, бухгалтерских фирм, армии и флота США. Двукратный чемпион игры *Jeopardy!* и соавтор системы мониторинга игроков, применяемой Марком Кубаном и командой НБА «Даллас Мэверикс».



# Введение

Работаете ли вы в крупной корпорации, в небольшой компании, в правительственной или некоммерческой структуре — если вы читаете эту книгу, то, скорее всего, вы используете Microsoft для выполнения своих повседневных задач. Возможно, вы занимаетесь тем, что подводите итоги, составляете отчеты и анализируете данные. Или же вы строите аналитические модели, призванные помочь работодателю увеличить прибыль, снизить издержки или эффективно управлять операциями.

Начиная с 1999 г. я обучил продуктивному использованию программы Microsoft Excel тысячи аналитиков в таких организациях, как Abbott Labs, консалтинговая компания Booz Allen Hamilton, Bristol-Myers Squibb, Broadcom, Cisco Systems, Deloitte Consulting, Drugstore.com, eBay, Eli Lilly, Ford, General Electric, General Motors, Intel, Microsoft, MGM Hotels Morgan Stanley, NCR, Owens Corning, Pfizer, Proctor & Gamble, PWC, Schlumberger, Tellabs, 3M, Армия США, Министерство обороны США, Военно-морской флот США и Verizon. Студенты часто говорят мне, что инструменты и методы, которые они освоили на занятиях, сэкономили им массу рабочего времени благодаря тому, что важные бизнес-задачи теперь анализируются быстрее и рациональнее.

Методы решения многих бизнес-задач, которые описаны в этой книге, я использую и сам, когда консультирую компании. Например, мы с Excel помогли менеджерам баскетбольных команд НБА «Даллас Мэверикс» и «Нью-Йорк Никербокерс» оценить судейство, игроков и позиции игроков на поле. Последние 20 лет я также веду занятия по бизнес-моделированию и анализу данных в Excel для студентов, обучающихся по программам MBA в Школе бизнеса Kelley School of Business при Университете Индианы, в Бауэровском колледже бизнеса Bauer College of Business при Университете Хьюстона и в Уэйк-Форест. (Мой преподавательский опыт подтвержден более чем 45 наградами для преподавателей, включая шесть наград Школы для преподавателей программ MBA.) Хочу заметить, что 95% студентов MBA из Университета Индианы выбрали мой курс моделирования в электронных таблицах, который даже не входит в обязательную программу.

В этой книге я попытался изложить этот популярный курс так, чтобы его мог пройти каждый. Эта книга научит вас более эффективной работе в Excel, и вот почему:

- Материалы протестированы на тысячах аналитиков из компаний, входящих в Fortune 500, и правительственных организаций.
- Книга написана в разговорном стиле. Я считаю, что такой стиль позволяет перенести дух успешной работы в аудитории на печатные страницы.
- Для обучения я использую задания, упрощающие освоение концепций. В заданиях разбираются реальные ситуации, часто встречающиеся на практике. Многие задания основаны на вопросах, которые задавали мне работники компаний, входящих в Fortune 500.

Я даю все инструкции, необходимые для изучения моих методов работы в Excel. Вы будете читать мои пояснения и отслеживать процесс решения по примерам листов Excel. Кроме того, я разместил файлы шаблонов для заданий на сопроводительном сайте книги ( [MicrosoftPressStore.com/Excel2019DataAnalysis/downloads](https://MicrosoftPressStore.com/Excel2019DataAnalysis/downloads)). Вы можете использовать эти шаблоны для работы прямо в Excel и самостоятельно выполнить все задания.

Как правило, главы невелики по размеру и посвящены чему-то одному. Я старался сделать так, чтобы каждую главу можно было освоить за два часа. Вопросы в начале каждой главы дадут вам общее представление о тех проблемах, которые вы будете в состоянии разрешить после освоения темы данной главы.

Помимо формул Excel вы безболезненно освоите некоторые важные разделы математики. Например, вы познакомитесь со статистическими методами, прогнозированием, моделями оптимизации, моделированием по методу Монте-Карло, построением моделей управления запасами и теорией очередей. Также вы получите представление о таких новшествах в деловом мышлении, как реальные опционы, потребительская ценность и математические модели ценообразования.

В конце каждой главы я поместил список практических задач (всего их более 800), с которыми вы можете работать самостоятельно. Многие задачи основаны на реальных ситуациях, с которыми имеют дело бизнес-аналитики компаний из списка Fortune 500. Эти задачи помогут до конца понять изложенный в главе материал. Ответы к задачам находятся на сопроводительном сайте книги.

И, самое главное, обучение прежде всего должно доставлять удовольствие. Прочитав эту книгу, вы узнаете, как прогнозировать президентские выборы в США, как определить точки безубыточности для футбольных матчей, как вычислить вероятность выигрыша в кости и вероятность победы определенной команды в турнире Национальной ассоциации студенческого спорта (NCAA). Благодаря подобным интересным примерам вы получите множество интересных и важных сведений о том, как решать бизнес-задачи с использованием Excel.

---

## ПРИМЕЧАНИЕ

Для работы с этой книгой необходимо иметь программу Microsoft Excel 2019 или Office 365. Однако для большей части книги будет вполне достаточно Excel 2013 или 2016. Предыдущие издания этой книги можно использовать с программами Excel 2003, Excel 2007 или Excel 2010.

---

## Новое в этом издании

В это издание книги внесены следующие изменения:

- Добавлена новая глава (глава 40) об инструменте «Получить и преобразовать данные» (Get&Transform).
- Добавлена новая глава (глава 41) о новых типах данных — «География» и «Акции».
- Обсуждение шести новых функций, входящих в Office 365, — ЕСЛИМН, МАКСЕСЛИ, МИНЕСЛИ, СМЕЩ, ОБЪЕДИНИТЬ и СЦЕП.

- Обсуждение (в главе 6) функции **ТЕКСТ**.
- Обсуждение (в главе 48) картограмм и 3D-карт Power Map.
- Обсуждение (в главе 65) инструмента «Лист прогноза».
- Обсуждение (в главе 12) функции **ВЫБОР**.
- Обсуждение (в главе 76) загрузки данных по различным акциям одновременно.
- Обсуждение (в главе 93) продвинутого анализа чувствительности.

## Что необходимо знать для работы с книгой

Для выполнения описанных в книге заданий вам не нужно знать Excel досконально. В принципе, достаточно уметь выполнять два ключевых действия.

- **Вводить формулы.** Формулы всегда начинаются со знака равенства (=). Следует также знать знаки основных математических операторов. Например, звездочка (\*) используется для умножения, слеш (/) — для деления, а знак вставки (^) — для возведения в степень.
- **Работать со ссылками на ячейки.** Важно помнить, что при копировании формулы, содержащей ссылку на ячейку, оформленную как **\$A\$4** (для создания абсолютной ссылки используется знак доллара), там, куда вы скопируете формулу, формула будет продолжать ссылаться на ячейку **A4**. При копировании формулы с такой ссылкой, как **\$A4** (смешанная ссылка), номер столбца останется прежним, а номер строки изменится. И, наконец, при копировании формулы со ссылкой вида **A4** (относительная ссылка) изменится и номер строки, и номер столбца ячейки, ссылку на которую содержит формула.

Эти понятия подробно описаны в главе 1.

## Как работать с книгой

Для выполнения заданий, представленных в книге, вы можете выбрать один из двух подходов.

- Можно открыть файл шаблона, соответствующий изучаемому заданию, и шаг за шагом выполнять задания по мере прочтения книги. Вы удивитесь, насколько легким окажется процесс обучения и как много всего вы сможете узнать и запомнить. Именно такой подход я использую при обучении студентов.
- Вместо того чтобы работать с шаблоном, можно читать пояснения в книге к окончательному варианту каждого файла с примером.

## Сопроводительный сайт

У этой книги есть сопроводительный сайт, на котором предоставлен доступ ко всем используемым в заданиях файлам (на сайт выложены как окончательные варианты книг Excel, так и исходные шаблоны, с которыми можно работать самостоятельно).

Книги Excel и шаблоны лежат в папках, названных поглавно. К заданиям, которые помещены в конце каждой главы, есть файлы примеров и ответы. Каждому файлу с ответами присвоено имя, по которому его можно легко идентифицировать. Например, файл с ответом на задание 2 к главе 10 называется `s10_2.xlsx`.

Для работы с заданиями необходимо скопировать файлы примеров на свой компьютер. Эти файлы и другие данные вы можете скачать по адресу:

[MicrosoftPressStore.com/Excel2019DataAnalysis/downloads](http://MicrosoftPressStore.com/Excel2019DataAnalysis/downloads)

Для загрузки файлов откройте страницу в браузере и следуйте инструкции.

## Благодарности

Я бесконечно благодарен Дженнифер Скуг (Jennifer Skoog) и Норму Тонине (Norm Tonina), которые поверили в меня и первыми пригласили вести занятия по Excel для Microsoft Finance. В частности, Дженнифер помогла мне составить план и методику занятий, на основе которых написана книга. Кейт Ланге (Keith Lange) из компании Eli Lilly, Пэт Китинг (Pat Keating) и Дуг Хопп (Doug Hoppe) из корпорации Cisco Systems, а также Дэннис Фуллер (Dennis Fuller) из армии США помогли мне прояснить мое понимание того, как следует преподавать анализ и моделирование данных.

Редакторы проекта Рик и Шарлотта Куген (Rick and Charlotte Kughen) — превосходно справились с копированием и редактурой рукописи. Выпускающий редактор Лоретта Йейтс (Loretta Yates) буквально нянчилась с проектом до самого его успешного завершения, а еще она сразу же отвечает на письма! Также я благодарен своим слушателям из организаций, где преподавал, и студентам из Школы бизнеса Kelley School of Business при Университете Индианы и из Бауэровского колледжа бизнеса Bauer College of Business при Университете Хьюстона. Они научили меня таким вещам об Excel, которые я сам не знал.

Алекс Блантон (Alex Blanton), ранее работавший в Microsoft Press, горячо приветствовал этот проект с самого начала и разделял мою точку зрения на создание текста в непринужденном стиле, ориентированного на бизнес-аналитиков.

И наконец, моя любимая талантливая жена Вивиан (Vivian) и мои замечательные дети, Дженнифер (Jennifer) и Грегори (Gregory), мирились с тем, что я проводил долгие часы в выходные дни за клавиатурой.

## Поддержка и обратная связь

Ниже вы найдете информацию об опечатках, клиентской поддержке и о том, куда направлять ваши отзывы и предложения.

## Контакты

Давайте продолжим наше общение. Twitter: <http://twitter.com/MicrosoftPress>.

## Найденные опечатки и поддержка

Мы приложили все усилия к тому, чтобы информация в книге и на сопроводительном сайте была точной и правильной.

Сведения о найденных опечатках и соответствующих исправлениях печатаются на этой странице (на английском языке):

[MicrosoftPressStore.com/Excel2019DataAnalysis/errata](http://MicrosoftPressStore.com/Excel2019DataAnalysis/errata)

Если вы найдете какую-нибудь новую ошибку, пожалуйста, сообщите нам о ней через эту же страницу.

Если вам требуется дополнительная помощь, напишите в клиентский отдел «Microsoft Press Book» на электронный адрес [microsoftpresscs@pearson.com](mailto:microsoftpresscs@pearson.com).

Пожалуйста, обратите внимание, что по указанным выше адресам не предлагается поддержка программных и аппаратных продуктов Microsoft, поэтому за таковой вам следует обращаться сюда: <http://support.microsoft.com>.

## От издательства

Ваши замечания, предложения, вопросы отправляйте по адресу [comp@piter.com](mailto:comp@piter.com) (издательство «Питер», компьютерная редакция).

Мы будем рады узнать ваше мнение!

На веб-сайте издательства [www.piter.com](http://www.piter.com) вы найдете подробную информацию о наших книгах. Когда мы только начинали перевод книги, перед нами встала сложная задача: подготовить книгу таким образом, чтобы ей могли пользоваться владельцы как английской, так и русской версии Excel. У большинства пользователей в России и других странах СНГ установлена русская версия Excel, однако многие интернациональные компании используют только английские версии Excel, поэтому все команды, формулы, функции и элементы интерфейса даются на двух языках.

В книге сохранены авторские скриншоты интерфейса, так как большая их часть показывает процесс решения авторских задач. Перевод элементов интерфейса дается в тексте, в скобках сохранены оригинальные названия, английские названия функций приводятся в начале глав, далее по тексту идет только русское название.

Обратите внимание: в английской версии Excel десятичные дроби отделяются от целого числа точкой, в русской версии — запятой. Числа с точкой не воспринимаются в русскоязычных версиях Excel, как числовой формат. В файлах для самостоятельной работы используются числа с точкой.

Задачи для самостоятельной работы можно скачать по адресу: [MicrosoftPressStore.com/Excel2019DataAnalysis/downloads](http://MicrosoftPressStore.com/Excel2019DataAnalysis/downloads).

## ГЛАВА 1

# Основы моделирования в электронных таблицах

### Обсуждаемые вопросы

- Как эффективно определить понедельные выплаты всех моих сотрудников?
- Как эффективно определить, сколько пекарня должна каждому из своих поставщиков?
- Как оценить, какое количество посетителей будет иметь новый фитнес-клуб через 10 лет?
- Как правильно задавать порядок выполнения операций в Excel?
- Как определить влияние на прибыль изменения в цене и себестоимости единицы товара для кофейни по соседству?

Мой друг Деннис Фуллер сказал как-то: «Электронная таблица для бизнес-аналитика — это то же, что холст для художника». Никто из нас не может написать шедевр, подобный «Звездной ночи» Ван Гюга, но, зная Excel, мы можем начать с чистого листа электронной таблицы и создать свой собственный шедевр, моделирующий виртуально любую ситуацию. Для многих барьером к овладению Excel является понимание того, как в Excel работают формулы. В этой главе мы разберем несколько различных простых моделей таблиц, которые помогут вам перейти к созданию сложных моделей в Excel.

## Ответы на вопросы

### ❓ Как эффективно определить понедельные выплаты всех моих сотрудников?

В файле *Wagestemp.xlsx* (в папке *Templates*), показанном на рис. 1.1, заданы количество отработанных часов и почасовая оплата нескольких сотрудников. Необходимо определить выплату каждому сотруднику, а также общую сумму отработанных часов и среднюю почасовую оплату.

Для вычисления недельной зарплаты сотрудника Лука необходимо перемножить значения в ячейках C4 и D4. Для этого в ячейку E4 вы помещаете формулу  $=C4*D4$ .

	B	C	D	E	F
1					
2					
3	<b>Employee</b>	<b>Hours</b>	<b>Wage Per Hour</b>	<b>Weekly Salary</b>	
4	Luka Abrus	49	\$ 10.00	\$ 490.00	=C4*D4
5	Terry Adams	36	\$ 13.00	\$ 468.00	=C5*D5
6	David Ahs	43	\$ 14.00	\$ 602.00	=C6*D6
7	Kim Akers	35	\$ 10.00	\$ 350.00	=C7*D7
8	Ties Arts	38	\$ 9.00	\$ 342.00	=C8*D8
9	Kamil Amerih	38	\$ 14.00	\$ 532.00	=C9*D9
10	Amy Alberts	42	\$ 11.00	\$ 462.00	=C10*D10
11	Matt Berg	39	\$ 9.00	\$ 351.00	=C11*D11
12	<b>Totals</b>	320			
13		=SUM(C4:C11)			
14			<b>Average Salary</b>	\$ 449.63	
15				=AVERAGE(E4:E11)	

**Рис. 1.1.** Расчет недельной платежной ведомости

Для вычисления зарплаты Тергу можно было бы перейти в ячейку E5 и ввести формулу `=C5*D5`, но можно воспользоваться командой Excel **Копировать (Copy)**, что позволит легко вычислить недельную зарплату каждого сотрудника, будь их хоть миллион. (Excel 2007 и более поздние версии поддерживают 1 048 576 строк!) Просто перейдите в ячейку E4 и нажмите **Ctrl+C** для копирования формулы. Затем выделите диапазон E5:E11, или примените комбинацию **Ctrl+V**, или нажмите **Enter**, чтобы применить эту формулу к диапазону E5:E11. Кроме того, вы можете скопировать формулу из E4 в диапазон E5:E11, наведя курсор на маленький квадратик в нижнем правом углу ячейки E4, и, после того как курсор изменит свой вид на тонкое черное перекрестье, нажать левую кнопку мыши и перетащить нужную формулу в диапазон E5:E11. В каждой ячейке, куда скопирована формула, Excel перемножит два значения слева от этой ячейки в столбце E. Отметим, что мы будем часто использовать функцию Excel **Ф.ТЕКСТ (FORMULATEXT)** (появившуюся в Excel 2013), чтобы текст формул был виден в электронной таблице. Например, если ввести в ячейке F4 формулу `=FORMULATEXT(E4)`, то в ней отобразится формула из E4.

Кроме того, вы можете вычислить в ячейке C12 (вы можете это увидеть в файле *Wagesfinal.xlsx*) общую сумму отработанных часов за неделю с помощью формулы `=SUM(C4:C11)`. В ячейке E14 я вычислил среднюю зарплату работника по формуле `=CP3HACH(E4:E11)`.

### Как эффективно определить, сколько пекарня должна каждому из своих поставщиков?

В файле *Bakery1temp.xlsx* задана цена за фунт, которую пекарня платит каждому из шести поставщиков сахара, масла и муки. Нам нужно узнать сумму, которая выплачивается каждому поставщику сахара, муки и масла. Также нам нужно определить суммарную выплату.



Как показано на рис. 1.2, в ячейке E23 мы вычисляем сумму, выплаченную за сахар поставщику Supplier 1, умножая стоимость фунта сахара, запрошенную Supplier 1, на количество сахара, закупленного у этого поставщика. Для этого мы используем формулу  $=E5*E14$ .

Чтобы вычислить сумму, выплаченную каждому поставщику за каждый продукт, мы можем воспользоваться любым из приведенных далее способов:

- Выделите ячейку E23 и, нажав Ctrl+C, выберите диапазон E23:G28, а затем нажмите Ctrl+V.
- Выделите ячейку E23 и, нажав Ctrl+C, выберите диапазон E23:G28, а затем нажмите Enter.
- Выделите ячейку E23 и, после того как вид курсора изменится на перекрестье, перетащите формулу в F23:G23. Затем перетащите диапазон E23:G23 в E24:G28.

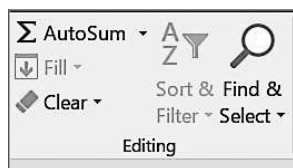
	C	D	E	F	G	H	I	J
2								
3		Prices						
4			Sugar	Butter	Flour			
5		Supplier 1	\$ 0.32	\$ 1.57	\$ 0.11			
6		Supplier 2	\$ 0.35	\$ 1.54	\$ 0.10			
7		Supplier 3	\$ 0.25	\$ 1.54	\$ 0.21			
8		Supplier 4	\$ 0.29	\$ 1.24	\$ 0.10			
9		Supplier 5	\$ 0.35	\$ 1.30	\$ 0.18			
10		Supplier 6	\$ 0.27	\$ 1.42	\$ 0.15			
11								
12		Quantity						
13			Sugar	Butter	Flour			
14		Supplier 1	364	391	220			
15		Supplier 2	387	245	314			
16		Supplier 3	290	211	200			
17		Supplier 4	340	265	330			
18		Supplier 5	261	345	246			
19		Supplier 6	365	232	390			
20								
21		Cost						
22			Sugar	Butter	Flour	Total		
23	=E5*E14	Supplier 1	\$ 116.48	\$ 613.87	\$ 24.20	\$754.55	=SUM(E23:G23)	
24		Supplier 2	\$ 135.45	\$ 377.30	\$ 31.40	\$544.15	=SUM(E24:G24)	
25		Supplier 3	\$ 72.50	\$ 324.94	\$ 42.00	\$439.44	=SUM(E25:G25)	
26		Supplier 4	\$ 98.60	\$ 328.60	\$ 33.00	\$460.20	=SUM(E26:G26)	
27		Supplier 5	\$ 91.35	\$ 448.50	\$ 44.28	\$584.13	=SUM(E27:G27)	
28		Supplier 6	\$ 98.55	\$ 329.44	\$ 58.50	\$486.49	=SUM(E28:G28)	
29		Total	\$ 612.93	\$ 2,422.65	\$ 233.38			
30			=SUM(E23:E28)	=SUM(F23:F28)	=SUM(G23:G28)			

**Рис. 1.2.** Вычисление выплат пекарни: цены у поставщиков различаются

Чтобы найти сумму, выплаченную каждому из поставщиков, введите в H23 формулу  $=\text{СУММ}(E23:G23)$  и скопируйте эту формулу в H24:H28. Введя формулу  $=\text{СУММ}(E23:E28)$  в ячейку E29 и скопировав эту формулу в F29:H29, вы получите общую сумму, выплаченную за каждый продукт.

Более быстрый способ задать эти суммы — это выбрать диапазон H23:H28 и (с нажатой клавишей Ctrl) выбрать диапазон E30:G30. Затем просто нажмите кнопку

Сумма (AutoSum, значок  $\Sigma$ ) в группе Редактирование (Editing) на вкладке Главная (Home), как показано на рис. 1.3.



**Рис. 1.3.** Использование кнопки AutoSum для суммирования группы чисел

Команда **Сумма** выбирает (не всегда верно, так что будьте внимательны!) диапазон ячеек, которые вы хотите просуммировать, чтобы заполнить выделенные ячейки. **Сумма** может сэкономить вам пять секунд времени!

В файле Bakery2temp.xlsx, изображенном на рис. 1.4, мы исходим из предположения, что все поставщики запрашивают одинаковую цену на каждый из продуктов. Снова вычислите сумму, выплаченную каждому из поставщиков за каждый продукт, и общую сумму, которую пекарня выплатит каждому поставщику.

	C	D	E	F	G	H
11			Price			
12			\$0.40	\$1.20	\$0.12	
13		Quantity	Sugar	Butter	Flour	
14		Supplier 1	364	391	220	
15		Supplier 2	387	245	314	
16		Supplier 3	290	211	200	
17		Supplier 4	340	265	330	
18		Supplier 5	261	345	246	
19		Supplier 6	365	232	390	
20						
21		Cost				
22			Sugar	Butter	Flour	Total
23	=E\$12*E14	Supplier 1	\$ 145.60	\$ 469.20	\$ 26.40	\$ 641.20
24	=E\$12*E15	Supplier 2	\$ 154.80	\$ 294.00	\$ 37.68	\$ 486.48
25	=E\$12*E16	Supplier 3	\$ 116.00	\$ 253.20	\$ 24.00	\$ 393.20
26	=E\$12*E17	Supplier 4	\$ 136.00	\$ 318.00	\$ 39.60	\$ 493.60
27	=E\$12*E18	Supplier 5	\$ 104.40	\$ 414.00	\$ 29.52	\$ 547.92
28	=E\$12*E19	Supplier 6	\$ 146.00	\$ 278.40	\$ 46.80	\$ 471.20
29		Total	\$ 802.80	\$ 2,026.80	\$ 204.00	

**Рис. 1.4.** Вычисление выплат пекарни: цены у поставщиков одинаковые

Действуя, как предлагалось ранее, неискушенный читатель может ввести в ячейку E23 формулу =E12\*E14 и скопировать эту формулу в диапазон E23:G28. К сожалению, в строках от 24 до 28 ссылки на строки 12, и 14 будут изменяться. При копировании этой формулы мы хотели бы, чтобы ссылка на строку 14 менялась, а на 12 — нет, потому что нам нужно брать цену каждого продукта из строки 12.

Чтобы этого добиться, мы помещаем перед 12 знак \$. Это называется абсолютной адресацией или блокированием строки. Когда в формуле перед номером строки стоит знак доллара и эта формула копируется, номер строки останется неизменным. Поэтому мы вводим в ячейку E23 формулу  $=\$12*E14$ .

Простой способ добавить знак доллара — использовать клавишу F4. Если выделите часть формулы и будете несколько раз нажимать F4, Excel циклически добавит сначала знак доллара к строке и к столбцу, затем только к строке, затем только к столбцу, а затем совсем уберет знак доллара.

### ❓ Как оценить, какое количество посетителей будет иметь новый фитнес-клуб через 10 лет?

Ответ вы найдете в файле Chapter1customer.xlsx, приведенном на рис. 1.5. Теперь нам требуется создать модель с чистого листа. Модели в электронных таблицах содержат исходные данные или предположения, которые мы используем, чтобы вычислить требуемые выходные значения. В базовой исходной модели оценки посетителей нам необходимы три исходных значения:

- Число посетителей, посещающих клуб, на начало года 1 (Year 1).
- Коэффициент оттока клиентов: доля посетителей на начало года (не считая новых посетителей), которые перестают посещать клуб, каждый год.
- Число новых посетителей, приходящих каждый год.

	B	C	D	E	F	G	H	I
1								
2	Start	100						
3	Newperyear	20						
4	Churnrate	0.15						
5								
6								
7	Year	Start customers	New Customers	Quits	End Customers	Start Formulas	Quit Formulas	End Formulas
8	1	100	20	15	105	=C2	=\$C\$4*C8	=C8+D8-E8
9	2	105	20	15.75	109.25	=F8	=\$C\$4*C9	=C9+D9-E9
10	3	109.25	20	16.39	112.8625	=F9	=\$C\$4*C10	=C10+D10-E10
11	4	112.8625	20	16.93	115.933125	=F10	=\$C\$4*C11	=C11+D11-E11
12	5	115.933125	20	17.39	118.5431563	=F11	=\$C\$4*C12	=C12+D12-E12
13	6	118.543156	20	17.78	120.7616828	=F12	=\$C\$4*C13	=C13+D13-E13
14	7	120.761683	20	18.11	122.6474304	=F13	=\$C\$4*C14	=C14+D14-E14
15	8	122.64743	20	18.4	124.2503158	=F14	=\$C\$4*C15	=C15+D15-E15
16	9	124.250316	20	18.64	125.6127685	=F15	=\$C\$4*C16	=C16+D16-E16
17	10	125.612768	20	18.84	126.7708532	=F16	=\$C\$4*C17	=C17+D17-E17

**Рис. 1.5.** Прогнозирование числа посетителей с помощью построения модели с входными и выходными значениями

Мы ввели значения этих исходных данных в ячейки C2:C4. Важно следить, чтобы исходные данные в электронной таблице были отделены от выходных значений

и никогда не программировались формулами Excel. Отделение в электронной таблице исходных данных от выходных значений позволяет легко определять, как изменение в исходных данных влияет на выходные значения.

В строках 8–17 рассчитывается количество посетителей на конец года путем сложения количества новых посетителей с количеством посетителей в начале года за вычетом посетителей, которые перестают посещать клуб. В ячейках C2:C4 содержатся входные значения этой таблицы. Ключевыми отношениями нашей модели оценки числа посетителей являются следующие:

- (Посетители на конец года  $t$  (End Year  $t$  Customers)) = (Посетители на начало года  $t$  (Start Year  $t$  Customers)) + (Новые посетители года  $t$  (New Year  $t$  Customers)) – (Переставшие ходить (Year Customer quits)).
- (Посетители на начало года 1 (Start Year 1 Customers)) = значение в ячейке C2.

Другим ключом к решению этой задачи является понимание того, что нам нужно отслеживать в течение каждого года:

- число посетителей на начало года;
- новые посетители, приходящие каждый год;
- посетители, переставшие ходить;
- число посетителей на конец года.

В ячейке C8 мы вычисляем начальное число посетителей в первый год (Year 1) по формуле =C2. Затем в столбце D мы повторно вводим число новых посетителей для каждого года, копируя значение из D8 в D9:D17 с помощью формулы =\$C\$3 или C\$3.

Обратите внимание, что цифре 3 должен предшествовать знак доллара; в противном случае при копировании формулы из ячейки D8 обращение к 3 изменится, что приведет к неверным результатам. Перед символом C знак доллара можно как вставлять, так и не вставлять, так как мы не копируем эту формулу в другой столбец.

Число посетителей, переставших посещать клуб каждый год, — это число посетителей в начале года, умноженное на коэффициент оттока клиентов. Поэтому в столбце E мы вычисляем количество посетителей, отсеивающихся каждый год, копируя из E8 в E9:E18 формулу =\$C\$4\*C8 или C\$4\*C8. Отметим, что здесь перед цифрой 8 мы не используем знак доллара, потому что при копировании мы хотим, чтобы 8 изменялась на 9, 10 и т. д.

Количество посетителей на конец каждого года вычисляется как сумма посетителей на начало этого года плюс новые клиенты, из которой вычитается количество покинувших клуб клиентов. Скопировав из F8 в F9:F18 формулу =C8+D8-F8, мы вычисляем число посетителей на конец каждого года.

Для годов 2–10 число клиентов на начало года равно числу клиентов в конце прошлого года, поэтому копируем из C9 в C10:C17 формулу =F8. Мы получим (не пере-

живайте по поводу дробных чисел), что через 10 лет в нашем фитнес-клубе будет около 127 посетителей.

Проницательный читатель может возразить, что на самом деле мы не знаем коэффициента оттока клиентов и количества новых посетителей каждый год. Это действительно так. Мы должны выполнить анализ чувствительности, чтобы выяснить, как изменения числа новых клиентов и коэффициента ежегодного оттока клиентов меняют конечное число посетителей 10-го года. В главе 17 мы изучим, как использовать таблицы данных для выполнения такого анализа чувствительности.

### ❓ Как правильно задавать порядок выполнения операций в Excel?

Сложные формулы в Excel часто содержат множество сложных математических операций, таких как возведение в степень, умножение и деление. При вычислении формул Excel следует правилам выполнения математических операций:

- Сначала выполняются операции в скобках.
- Затем выполняются все возведения в степень, следуя слева направо.
- Затем выполняются все умножения и деления, следуя слева направо.
- После этого выполняются все сложения и вычитания, следуя слева направо.

Например, Excel будет вычислять формулу  $=3+6*(5+4)/3-7$  в следующем порядке.

- $3+6 \times 9/3 - 7$  (скобки сняты).
- $3 + 54/3 - 7$  (умножение).
- $3 + 18 - 7$  (деление).
- $21 - 7$  (суммирование).
- 14 (вычитание).

Или, например, предположим, нам нужно извлечь квадратный корень из значения ежегодного процентного роста продаж наших продуктов (см. файл PEMDAStemp.xlsx и рис. 1.6).

	D	E	F	G	H	I
			Square Root of			
3	This Year	Next Year	Improvement	Wrong	Right Formula	Wrong Formula
4	100	150	0.707106781	5	$=((E4-D4)/D4)^{0.5}$	$=(E4-D4)/D4^{0.5}$
5	200	2000	3	127.279	$=((E5-D5)/D5)^{0.5}$	$=(E5-D5)/D5^{0.5}$
6	80	400	2	35.7771	$=((E6-D6)/D6)^{0.5}$	$=(E6-D6)/D6^{0.5}$

**Рис. 1.6.** Пример применения порядка выполнения операций

В F4 мы ввели правильную формулу,  $=((E4-D4)/D4)^{0.5}$ , и скопировали ее в F5:F6. Эта формула заставляет Excel вычислить процентную долю (выраженную в виде дроби) роста продаж (.5) и затем извлечь квадратный корень. Окончательное значение .707 (квадратный корень из .5) правильно. Обратите внимание: знак ^ (находится на клавиатуре на цифре 6) — это знак возведения в степень.

В G4 мы ввели неправильную формулу,  $= (E4 - D4) / D4^{.5}$ .

По этой формуле сначала вычисляется  $E4 - D4 = 50$ , а затем извлекается квадратный корень из D4 (10). После этого конечный (неверный) результат такой:  $50 / 10 = 5$ .

### ? Как определить влияние на прибыль изменения в цене и себестоимости единицы товара для кофейни по соседству?

Ключом к пониманию того, как изменение цены влияет на прибыль, является правильная оценка кривой спроса. Кривая спроса показывает, как изменения в цене меняют спрос на продукт. Давайте предположим, что дневной спрос на кофе в кофейне составляет  $100 - 15 \cdot \text{Цена в долларах}$ . (См. главы 87–89, где подробнее рассматривается оценка кривой спроса.) В файле *Coffee.xlsx* (рис. 1.7) показано, как дневная прибыль зависит от того, как меняются себестоимость и цена чашки кофе.

	E	F	G	H	I	J	K	L
1								
2	Demand=100-15*price			Profit=(Price-Unit Cost)*Demand				
3								
4	Demandslope	15						
5	Demandintercept	100						
6								
7								
8		Price						
9	Cost	\$2.00	\$2.50	\$3.00	\$3.50	\$4.00	\$4.50	\$5.00
10	\$0.50	\$105.00	\$125.00	\$137.50	\$142.50	\$140.00	\$130.00	\$112.50
11	\$1.00	\$70.00	\$93.75	\$110.00	\$118.75	\$120.00	\$113.75	\$100.00
12	\$1.50	\$35.00	\$62.50	\$82.50	\$95.00	\$100.00	\$97.50	\$87.50
13	\$2.00	\$0.00	\$31.25	\$55.00	\$71.25	\$80.00	\$81.25	\$75.00
14								
15								
16		=(\$F\$5-\$F\$4*\$F\$9)*(\$F\$9-\$E10)						

**Рис. 1.7.** Зависимость спроса от цены и себестоимости

Мы исходим из предположения, что себестоимость чашки кофе варьируется между \$0,50 и \$2,00, а цена чашки кофе варьируется между \$2,00 и \$5,00. Чтобы определить прибыль для каждой комбинации цена/себестоимость, нам потребуется ввести в F10 формулу  $= ($F$5 - $F$4 * $F$9) * ($F$9 - $E10)$  и скопировать эту формулу из F10 в F10:L13.

- Ссылки на F5 и F4 — абсолютные, так как нам нужно, чтобы ни строка, ни столбец не менялись при копировании формулы.
- Ссылка на цену (ячейка F9) требует использования знака доллара (или блокирования строки), так как нам всегда требуется, чтобы цена бралась из строки 9.
- Ссылка на себестоимость (E10) требует, чтобы перед буквой столбца стоял знак доллара, так как нам всегда требуется, чтобы себестоимость бралась из столбца E.

Мы получим, например, что если себестоимость равна \$1,50, а мы продаем чашку за \$4,00, наша прибыль составит \$100:  $(100 - 4 \cdot 15) \cdot (4 - 1,5) = \$100$ .

Заметьте, что для каждого значения себестоимости цена, при которой прибыль максимальна, выделяется желтым фоном. В главе 24 вы узнаете, как использовать условное форматирование для создания такого стильного выделения.

## Задания

1. Проводя занятия, я дал пять домашних заданий, каждое оценив в 25 баллов, и три теста, оценив каждый в 100 баллов. Я вычисляю итоговую оценку студента, присваивая весовой коэффициент в 75% для тестов и 25% для домашних заданий. Создайте таблицу для вычисления итоговой оценки студента, которая позволит вам изменять весовой коэффициент, присвоенный тестам.
2. Индекс массы тела человека (ИМТ, BMI) вычисляется так:  $BMI = 703 \cdot \text{Weight} / \text{Height}^2$ . Создайте таблицу вычисления ИМТ человека.
3. Последовательность Фибоначчи определяется следующим образом:  $F_0 = 0$ ,  $F_x = 1$ , а для  $n$  больше 1 —  $F_{N+x} = F_N + F_{n-1}$ . Создайте таблицу для вычисления последовательности Фибоначчи. Покажите, что для больших значений  $N$  отношение последовательных чисел Фибоначчи стремится к золотому сечению (1,62).
4. Знаменитый эффект бабочки утверждает, что если бабочка взмахнет крыльями на Таити, это маленькое событие может вызвать ураган в Техасе. Предположим, что погода во время  $t$  всегда имеет значение между 0 и 1 и управляется выражением  $x_{t+1} = 4 \times x_t \times (1 - x_t)$ . Для  $x_t = 0,3$  и  $x_t = 0,3000001$  вычислите  $x_1, x_2, \dots, x_{50}$ . Как ваши вычисления иллюстрируют эффект бабочки?
5. Зарыбление озера оценивается в настоящее время в 12 230 особей. Каждый год коэффициент рождений у рыб равен 1,2, а коэффициент смертей равен 0,7. Покажите, что если улов каждый год будет составлять 6115 особей, количество рыбы в озере будет оставаться постоянным.
6. Коэффициент Джини обычно используется для измерения расслоения общества по уровню годового дохода. Если доходы  $n$  человек перечислены по возрастанию ( $x_1$  = наименьший доход,  $x_n$  = наибольший доход), то коэффициент Джини вычисляется по формуле:

$$G = \frac{n+1}{n} - \frac{2 \sum_{i=1}^n (n+1-i)x_i}{n \sum_{i=1}^n x_i}.$$

Создайте таблицу, определяющую коэффициент Джини для группы из пяти человек.



## ГЛАВА 2

# Имена диапазонов

### Обсуждаемые вопросы

- Необходимо вычислить общий объем продаж в штатах Аризона, Калифорния, Монтана, Нью-Йорк и Нью-Джерси. Можно ли для вычисления общего объема продаж воспользоваться формулой  $AZ+CA+MT+NY+NJ$  вместо формулы  $\text{СУММ}(A21:A25)$  и получить правильный ответ?
- Для чего нужна формула  $\text{СРЗНАЧ}(A:A)$ ?
- Чем различаются имена с областью действия «книга» и «лист»?
- Мне начинают нравиться имена диапазонов. Я стал использовать имена диапазонов во многих книгах, которые я создаю в офисе. Однако эти имена не появляются в формулах. Как добиться отображения недавно созданных имен диапазонов в ранее созданных формулах?
- Каким образом можно вставить список имен всех диапазонов (и представляемых ими ячеек) в лист?
- Предполагаемый годовой доход вычисляется как кратный прошлогоднему доходу. Может ли формула выглядеть как  $(1+\text{прирост})*\text{предыдущий\_год}$ ?
- Для каждого дня недели даны почасовая оплата и количество отработанных часов. Можно ли вычислить итоговую сумму оплаты за каждый день по формуле  $\text{почасовая\_оплата}*\text{часы}$ ?

Возможно, вам доводилось работать с листами, в которых использовалась, например, формула  $=\text{СУММ}(A5000:A5049)$ . В этом случае вам приходилось догадываться, что находится в ячейках  $A5000:A5049$ . Если в ячейках  $A5000:A5049$  содержатся объемы продаж по всем штатам США, не кажется ли вам формула  $=\text{СУММ}(\text{USSales})$  более понятной? В данной главе рассказывается, как присваивать имена отдельным ячейкам и диапазонам ячеек, а также вставлять имена диапазонов в формулы.

## Как создать именованный диапазон?

Создать именованный диапазон можно тремя способами:

- ввести имя диапазона в поле **Имя (Name)**;
- выбрать на вкладке **Формулы (Formulas)** в группе **Определенные имена (Defined Names)** инструмент **Создать из выделенного (Create from Selection)**;

- выбрать на вкладке **Формулы (Formulas)** в группе **Определенные имена (Defined Names)** инструментов **Присвоить имя (Define Name)** или **Диспетчер имен (Name Manager)**.

## Создание имени диапазона в поле Имя

Поле **Имя (Name)** (рис. 2.1) находится прямо над меткой столбца **A**, слева от поля **Строка формул (Formula bar)**. Для создания имени диапазона выделите ячейку или диапазон ячеек, которым требуется присвоить имя, затем перейдите в поле **Имя (Name)** и введите имя диапазона. Имя диапазона создастся, когда вы нажмете **Enter**. При нажатии в поле **Имя (Name)** на стрелку появятся имена диапазонов, которые есть в текущей книге. Клавиша **F3** открывает диалоговое окно **Вставка имени (Paste Name)** с именами всех диапазонов. Если выбрать в поле **Имя (Name)** имя диапазона, то все ячейки, соответствующие этому диапазону, отметятся автоматически. Так вы можете убедиться, что правильно выбрали ячейку или диапазон ячеек для указанного имени. Имена диапазонов не чувствительны к регистру букв.

Скажем, нам нужно присвоить ячейке **F3** имя **east**, а ячейке **F4** — имя **west** (см. рис. 2.2 и файл **Eastwestempt.xlsx**). Выделите ячейку **F3**, введите **east** в поле **Имя (Name)** и нажмите **Enter**. Выделите ячейку **F4**, введите **west** в поле **Имя (Name)** и нажмите **Enter**. Теперь в какой-нибудь другой ячейке для ссылки на ячейку **F3** можно указывать **=east**, а не **=F3**. То есть вместо любой ссылки **east** в формуле будет автоматически подставлено значение из ячейки **F3**.



**Рис. 2.1.** Можно создать имя диапазона, выбрав диапазон ячеек, который вы хотите назвать, и введя имя в поле **Имя**

	E	F
1		
2		
3	east	5
4	west	10

**Рис. 2.2.** Присвоение ячейкам **F3** и **F4** имен **east** и **west**

Предположим, необходимо присвоить имя **data** прямоугольному диапазону ячеек (например, **A1:B4**). Выделите диапазон ячеек **A1:B4**, введите **data** в поле **Имя (Name)** и нажмите **Enter**. Теперь мы можем вычислить среднее значение содержимого ячеек **A1:B4** с помощью функции **СРЗНАЧ (AVERAGE)** по формуле **=СРЗНАЧ(data)** (см. файл **Data.xlsx** и рис. 2.3).

Иногда требуется присвоить имя диапазону ячеек, состоящему из нескольких несмежных прямоугольных диапазонов. Например, на рис. 2.4 и в файле **Noncontigtemp.xlsx** показан диапазон с именем **noncontig**, состоящий из ячеек **B3:C4**, **E6:G7** и **B10:C10**. Выделите любой из трех прямоугольников (здесь **B3:C4**). С нажатым **Ctrl** выделите оставшиеся два диапазона (**E6:G7** и **B10:C10**). Отпустите **Ctrl**, введите **noncontig** в поле **Имя (Name)** и нажмите **Enter**. Теперь в любой формуле имя

noncontig указывает на содержимое ячеек B3:C4, E6:G7 и B10:C10. Например, введя формулу =СРЗНАЧ(noncontig) в ячейку E11, мы получим значение 4,75 (сумма 12 чисел в заданном диапазоне равна 57 и  $57/12 = 4,75$ ).

	A	B	C	D
1	1	2		
2	3	2		
3	1	1		
4	2	-1		
5			1.375	
6		cell C5	=AVERAGE(data)	
7		contains		

Рис. 2.3. Присвоение диапазону A1:B4 имени data

	B	C	D	E	F	G	H
1							
2							
3	1	2					
4	3	4					
5							
6				6	7	10	
7				8	9	1	
8							
9							
10	2	4					
11				4.75 =AVERAGE(noncontig)			

Рис. 2.4. Присвоение имени несмежному диапазону ячеек

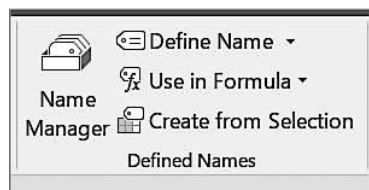
## Создание имен с помощью инструмента Создать из выделенного

На листе Statestemp.xlsx указаны мартовские продажи для каждого из 50 штатов США. На рис. 2.5 вы видите фрагмент этих данных. Требуется присвоить каждой ячейке в диапазоне B6:B55 правильную двухбуквенную аббревиатуру — сокращенное название штата. Сначала выделите диапазон A6:B55 и на вкладке Формулы (Formulas) в группе Определенные имена (Defined Names) выберите Создать из выделенного (Create from Selection), как показано на рис. 2.6, а затем в открывшемся диалоговом окне установите флажок В столбце слева (Left column) (рис. 2.7) и нажмите ОК.

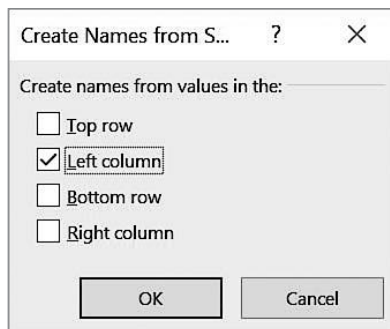
Теперь имена в первом столбце выделенного диапазона связаны с ячейками во втором столбце выделенного диапазона. Таким образом, ячейке B6 присвоено имя диапазона AL, ячейке B7 — AK и т. д. Мы бы устали создавать имена таких диапазонов с помощью поля Имя (Name)! Нажмите на стрелку в поле Имя (Name) и убедитесь, что все имена диапазонов созданы.

	A	B	C	D	E
1					
2					
3					
4					
5	State	March Sales			
6	AL	\$ 915.00			
7	AK	\$ 741.00			
8	AZ	\$ 566.00		with cells	\$ 2,976.00
9	AR	\$ 754.00		with ranges	2976
10	CA	\$ 687.00			
11	CO	\$ 757.00			
12	CT	\$ 786.00			

**Рис. 2.5.** После присвоения имен ячейкам, содержащим продажи и сокращенные названия штатов, при ссылке на ячейку вместо буквы столбца и номера строки можно использовать соответствующее сокращенное название



**Рис. 2.6.** Выберите Создать из выделенного



**Рис. 2.7.** Установите флажок в столбце слева

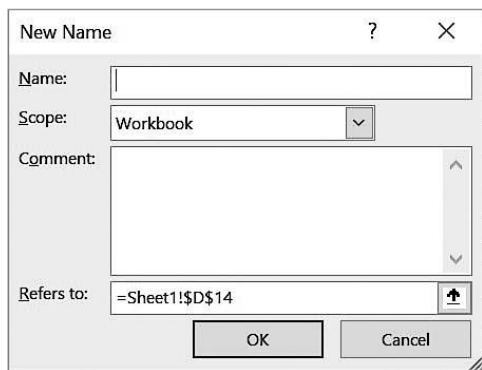
## Создание имен диапазонов с помощью инструмента Присвоить имя

Откроем диалоговое окно Создание имени (New Name), показанное на рис. 2.8. Для этого на вкладке **Формулы** (Formulas) в группе **Определенные имена** (Defined Names) нажмите **Присвоить имя** (Define Name) в меню, показанном на рис. 2.6.

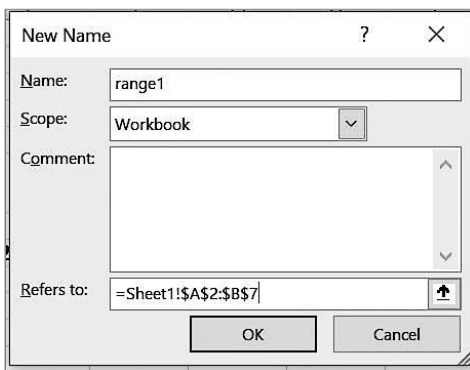
Предположим, нам нужно присвоить имя **range1** (в именах диапазонов регистр букв не учитывается) диапазону ячеек **A2:B7**. Введите **range1** в поле **Имя** (Name) и выберите диапазон или введите **=A2:B7** в поле **Диапазон** (Refers To). Диалоговое окно **Создание имени** (New Name) должно выглядеть так, как на рис. 2.9. Нажмите **ОК**.

Нажав на стрелку раскрывающегося списка в поле **Область** (Scope), вы можете выбрать строку **Книга** (Workbook) или любой лист в книге, указав тем самым область

действия имени. Более подробно этот вопрос обсуждается далее в этой главе, а пока выберите область действия по умолчанию — Книга (Workbook). Кроме того, к любым именам диапазонов можно добавить комментарии.



**Рис. 2.8.** Диалоговое окно Создание имени до задания каких-либо имен



**Рис. 2.9.** Диалоговое окно Создание имени при создании имени диапазона

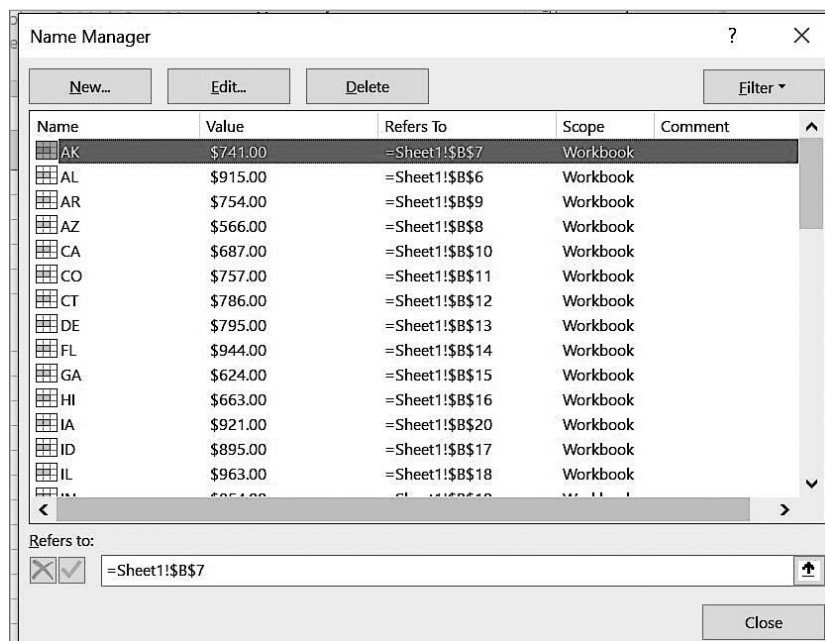
## Диспетчер имен

Если теперь вы нажмете на стрелку в поле **Имя (Name)**, то увидите в раскрывающемся списке имя **range1** (и все остальные созданные ранее диапазоны). В Microsoft Excel 2019 есть простой способ изменения или удаления имен диапазонов. На вкладке **Формулы (Formulas)** выберите группу **Определенные имена (Defined Names)** и откройте **Диспетчер имен (Name Manager)**, показанный на рис. 2.6. Появится список имен всех диапазонов. Например, так выглядит диалоговое окно **Диспетчер имен (Name Manager)** для файла *States.xlsx* (рис. 2.10).

Для изменения имени диапазона дважды щелкните кнопкой мыши на имени этого диапазона или выделите его и нажмите кнопку **Изменить (Edit)**. После этого можно изменить не только имя диапазона, но и его область действия, а также поменять ячейки в диапазоне.

Чтобы удалить некоторое подмножество имен диапазонов, выделите имена диапазонов: если имена диапазонов идут последовательно, выделите первое имя в группе имен, которую требуется удалить, затем, удерживая **Shift**, выделите последнее имя в группе. Если требуемые имена не следуют друг за другом, можно выделить любое из имен, которое необходимо удалить, а далее, удерживая **Ctrl**, выделить остальные требуемые имена диапазонов. Затем нажмите кнопку **Удалить (Delete)**.

Теперь рассмотрим несколько конкретных примеров использования имен диапазонов.



**Рис. 2.10.** Диалоговое окно Диспетчер имен для файла States.xlsx

## Ответы на вопросы

**?** Необходимо вычислить общий объем продаж в штатах Аризона, Калифорния, Монтана, Нью-Йорк и Нью-Джерси. Можно ли для вычисления общего объема продаж воспользоваться формулой **AZ+CA+MT+NY+NJ** вместо формулы **СУММ(A21:A25)** и получить правильный ответ?

Вернемся к файлу **States.xlsx**, в котором двухбуквенные аббревиатуры были присвоены как имена диапазонов соответствующим объемам продаж. Для вычисления общего объема продаж в Алабаме, на Аляске, в Аризоне и в Арканзасе можно воспользоваться формулой **=СУММ(B6:B9)**. Однако если указать ячейки **B6, B7, B8** и **B9**, формула будет выглядеть так: **=AL+AK+AZ+AR**. Последняя запись, конечно, гораздо нагляднее.

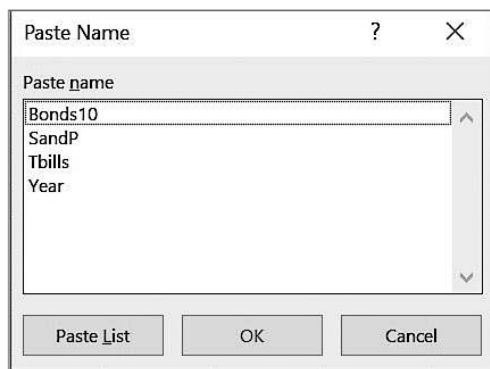
В качестве другого примера использования имен диапазонов рассмотрим файл **Historicalinvesttemp.xlsx** (рис. 2.11), в котором содержится годовая процентная доходность акций, казначейских векселей и облигаций. (На этом рисунке часть строк скрыта; данные заканчиваются в строке 89.)

Выделите диапазон ячеек **B1:D89** и на вкладке **Формулы (Formulas)** в группе **Определенные имена (Defined Names)** выберите **Создать из выделенного (Create from Selection)**. В этом примере имена диапазона указаны в строке выше (Top row). Диа-

пазон B2:B89 получает имя **Акции (Stocks)**, диапазон C2:C89 — имя **Векселя (Tbills)** и диапазон D2:D89 — имя **Облигации (Bonds10)**. Затем введите в ячейке =СРЗНАЧ(, но перед тем как вводить диапазон, можно нажать F3, и откроется диалоговое окно **Вставка имени (Paste Name)**, показанное на рис. 2.12.

	A	B	C	D
1	<b>Year</b>	<b>SandP</b>	<b>Tbills</b>	<b>Bonds10</b>
2	1928	43.81%	3.08%	0.84%
3	1929	-8.30%	3.16%	4.20%
4	1930	-25.12%	4.55%	4.54%
5	1931	-43.84%	2.31%	-2.56%
6	1932	-8.64%	1.07%	8.79%
7	1933	49.98%	0.96%	1.86%
83	2009	25.94%	0.14%	-11.12%
84	2010	14.82%	0.13%	8.46%
85	2011	2.10%	0.03%	16.04%
86	2012	15.89%	0.05%	2.97%
87	2013	32.15%	0.07%	-9.10%
88	2014	13.52%	0.05%	10.75%
89	2015	1.38%	0.21%	1.28%
90	2016	11.77%	0.51%	0.69%
91	2017	21.64%	1.39%	2.80%

**Рис. 2.11.** Ретроспективные данные по инвестициям



**Рис. 2.12.** Добавление имени диапазона в формулу в диалоговом окне Вставка имени

Теперь в окне **Вставка имени (Paste Name)** выберите из списка имя **Акции** и нажмите **OK**. После ввода закрывающей скобки в формуле =СРЗНАЧ(Акции) автоматически будет рассчитано среднее значение доходности акций ( 11,41 %). Прелесть этого подхода состоит в том, что даже не помня точно, где находятся данные, можно работать с данными о доходности акций в любом месте книги!

Было бы упущением с моей стороны не упомянуть о такой интересной возможности Microsoft Excel 2019, как автозавершение формул (AutoComplete). После ввода =СРЗНАЧ(В автоматически появится список диапазонов и функций, имена которых начинаются с **В**. Для завершения ввода имени диапазона дважды щелкните на имени **Векселя**, и затем вам останется только ввести закрывающую скобку.

### ❓ Для чего нужна формула СРЗНАЧ(А:А)?

При использовании в формуле имени столбца (в формате **А:А**, **С:С** и т. д.) весь столбец обрабатывается в Excel как именованный диапазон. Например, по формуле =СРЗНАЧ(А:А) вычисляется среднее значение всех чисел в столбце **А**. Использование имени диапазона для целого столбца очень эффективно при частом вводе новых данных в столбец. Например, если столбец **А** содержит данные о ежемесячных продажах продукта, то новые данные добавляются каждый месяц, и по такой формуле вычисляется актуальное среднее значение ежемесячных продаж. Однако

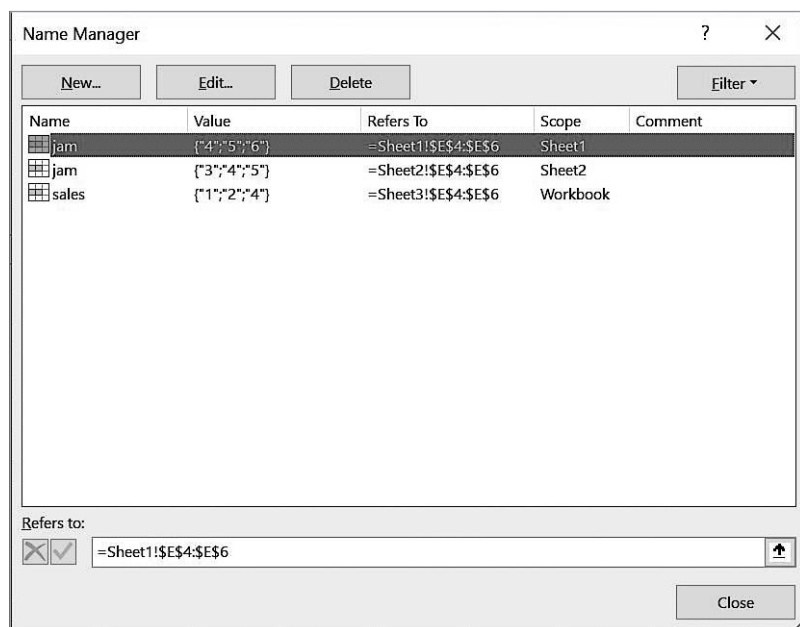
будьте осторожны: если ввести формулу `=СРЗНАЧ(А:А)` в столбец А, то появится сообщение о циклической ссылке, так как значение в ячейке, содержащей формулу расчета среднего, будет зависеть от ячейки, содержащей среднее значение. Способ разрешения циклических ссылок приведен в главе 11. Аналогично по формуле `=СРЗНАЧ(1:1)` рассчитывается среднее значение всех чисел в строке 1.

### ? Чем различаются имена с областью действия «книга» и «лист»?

Понять различие вам поможет файл `Sheetnames.xlsx`. При создании имен с помощью поля **Имя (Name)** областью действия имен по умолчанию становится **Книга**. Предположим, что с помощью поля **Имя (Name)** имя `sales` присвоено диапазону ячеек `E4:E6` на листе **Лист3**, и эти ячейки содержат числа 1, 2 и 4 соответственно.

Если мы введем формулу `=СУММ(sales)` на любом листе, то получим 7, так как областью действия созданных в поле **Имя (Name)** имен является книга. То есть если в любом месте книги указано имя `sales` (областью действия которого является вся книга), то оно указывает на ячейки `E4:E6` на листе **Лист3**.

Теперь введите числа 4, 5, 6 в ячейки `E4:E6` на листе **Лист1** и 3, 4, 5 в ячейки `E4:E6` на листе **Лист2**. Откройте окно **Диспетчер имен (Name Manager)**, присвойте имя `jam` ячейкам `E4:E6` на **Лист1** и определите область действия этого имени как **Лист1**. Перейдите на **Лист2**, откройте окно **Диспетчер имен**, присвойте имя `jam` ячейкам `E4:E6` и определите область действия этого имени как **Лист2**. Диалоговое окно **Диспетчер имен** теперь будет выглядеть как на рис. 2.13.



**Рис. 2.13.** Имена на уровне листа и на уровне книги в диалоговом окне **Диспетчер имен**



Что произойдет, если ввести формулу `=СУММ(jam)` на каждом из трех листов? На листе **Лист1** будут просуммированы значения ячеек **E4:E6**. Так как в этих ячейках содержатся числа 4, 5 и 6, получится 15. На листе **Лист2** также будут просуммированы значения ячеек **E4:E6**, что в сумме даст  $3 + 4 + 5 = 12$ . Однако на листе **Лист3** вычисление по формуле `=СУММ(jam)` вызовет ошибку `#ИМЯ?`, поскольку на этом листе отсутствует диапазон с именем `jam`. Если где-либо на листе **Лист3** ввести формулу `=СУММ(Лист2!jam)`, Excel распознает имя на уровне листа, которое представляет диапазон ячеек **E4:E6** листа **Лист2**, и в результате получится  $3 + 4 + 5 = 12$ . То есть если вы указываете перед именем диапазона нужное имя листа с восклицательным знаком (!), Excel обращается к диапазону на другом листе.

**?** Мне начинают нравиться имена диапазонов. Я стал определять имена диапазонов во многих книгах, которые я создаю в офисе. Однако эти имена не появляются в формулах. Как добиться отображения недавно созданных имен диапазонов в ранее созданных формулах?

Рассмотрим файл `Applynames.xlsx` и рис. 2.14.

	E	F
1		
2		
3	price	\$5.00
4	demand	8500
5	unitcost	\$4.00
6	fixed cost	\$3,000.00
7	profit	\$5,500.00

**Рис. 2.14.** Новые имена диапазонов в старых формулах

В ячейке **F3** на листе **Лист1** указана цена продукта, а в ячейке **F4** — потребность в продукте  $=10000-300*\text{F3}$ . Ячейки **F5** и **F6** содержат себестоимость единицы продукции и постоянные затраты соответственно. Прибыль вычисляется в ячейке **F7** по формуле  $=\text{F4}*(\text{F3}-\text{F5})-\text{F6}$ . В диапазон **E3:E7** введены новые имена. Выделите диапазон **E3:F7**, затем присвойте ячейкам следующие имена: ячейке **F3** — цена, ячейке **F4** — потребность, ячейке **F5** — себестоимость, ячейке **F6** — затраты и ячейке **F7** — прибыль. Для этого на вкладке **Формулы (Formulas)** используйте **Создать из выделенного (Create from Selection)** и флажок в столбце слева (**Left column**). Теперь имена созданных диапазонов необходимо отобразить в формулах ячеек **F4** и **F7**. Для этого выделите диапазон, для которого они создаются (в данном случае **F4:F7**). Затем на вкладке **Формулы (Formulas)** в группе **Определенные имена (Defined Names)** нажмите стрелку **Присвоить имя (Define Name)** и выберите инструмент **Применить имена (Apply Names)**. Выделите в окне имена, которые требуется применить, и нажмите **OK**. Обратите внимание, что в ячейке **F4** теперь находится формула  $=10000-300*\text{цена}$ , а в ячейке **F7** формула  $=\text{потребность}*(\text{цена}-\text{себестоимость})-\text{затраты}$ , что и требовалось.

Если вам нужно применить имена диапазонов ко всему листу, выделите его кнопкой **Выделить все (Select All)** на пересечении заголовков столбцов и строк.

❓ **Каким образом можно вставить список имен всех диапазонов (и представляемых ими ячеек) в лист?**

Нажмите F3. Откроется окно **Вставка имени (Paste Name)**. Теперь нажмите кнопку **Все имена (Paste List)** — см. рис. 2.12. На листе, начиная с текущей ячейки, появится список имен диапазонов и соответствующих им ячеек.

❓ **Предполагаемый годовой доход вычисляется как кратный прошлогоднему доходу. Может ли формула выглядеть как  $(1+\text{прирост})^*\text{предыдущий\_год}$ ?**

Решение этой проблемы ищите в файле `Last year.xlsx`. Как показано на рис. 2.15, требуется вычислить доходы за 2014–2021 гг с приростом 10% в год, начиная с базового уровня \$300 млн в 2014 г.

В поле **Имя (Name)** присвойте ячейке B3 имя **прирост**. А теперь самое интересное! Переместите курсор в ячейку B7 и на вкладке **Формулы (Formulas)** в группе **Определенные имена (Defined Names)** выберите **Присвоить имя (Define Name)** — откроется диалоговое окно **Редактировать имя (Edit Name)**. Введите в него данные, как показано на рис. 2.16; на рисунке видно, как, при положении курсора в ячейке B7, мы можем создать имя **предыдущий\_год**, которое всегда будет ссылаться на ячейку ровно на одну строку выше текущей.

	A	B	C
1			
2			
3	growth	0.1	
4			
5		revenue	
6	2011	300	
7	2012	330	=lastyear*(1+growth)
8	2013	363	=lastyear*(1+growth)
9	2014	399.3	=lastyear*(1+growth)
10	2015	439.23	=lastyear*(1+growth)
11	2016	483.153	=lastyear*(1+growth)
12	2017	531.4683	=lastyear*(1+growth)
13	2018	584.6151	=lastyear*(1+growth)

**Рис. 2.15.** Создание имени диапазона для предыдущего года

**Рис. 2.16.** Создание имени диапазона для подсчета продаж предыдущего года

Поскольку мы сейчас на ячейке B7, Excel интерпретирует имя диапазона как указывающее на ячейку, находящуюся над текущей ячейкой. Конечно, это не работает, если в ссылке на ячейку B6 останется знак доллара. Теперь если мы введем в ячейку B7 формулу **=предыдущий\_год\*(1+прирост)** и скопируем ее в диапазон

B8:B13, каждая ячейка будет содержать нужную нам формулу, по которой содержимое ячейки, находящейся над активной, будет умножаться на 1,1.

❓ Для каждого дня недели даны почасовая оплата и количество отработанных часов. Можно ли вычислить итоговую сумму оплаты за каждый день по формуле `почасовая_оплата*часы`?

Как показано на рис. 2.17 (см. файл `Namedrows.xlsx`), строка 12 содержит данные о почасовой оплате по дням недели, а строка 13 — количество отработанных часов за каждый день.

	E	F	G	H	I	J	K	L
11		Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
12	wage	\$ 5.00	\$ 6.00	\$ 7.00	\$ 8.00	\$ 9.00	\$ 15.00	\$ 15.00
13	hours	55	65	75	65	77	88	36
14	payroll	\$275.00	\$390.00	\$525.00	\$520.00	\$693.00	\$1,320.00	\$540.00

**Рис. 2.17.** Использование имен диапазонов для ссылки на различные строки

Выберите строку 12 (щелкнув на 12) и в поле **Имя (Name)** введите имя `почасовая_оплата`. Выберите строку 13 и введите в поле **Имя (Name)** имя `часы`. Если теперь в ячейку F14 ввести формулу `=почасовая_оплата*часы` и скопировать эту формулу в диапазон G14:L14, то в каждом столбце появится произведение двух множителей — почасовой оплаты и отработанных часов.

## Замечания

- В Excel нельзя использовать буквы `г` и `с` в качестве имен диапазонов.
- Если вы используете **Создать из выделенного (Create from Selection)**, пробелы в созданном имени автоматически будут заменены на подчеркивание (`_`). Например, имя `Product 1` будет создано как `Product_1`.
- Имена диапазонов не могут начинаться с цифр или выглядеть как ссылка на ячейку. Например, нельзя использовать имена `3Q` и `A4`. Поскольку Excel поддерживает более 16 000 столбцов, такие имена, как `cat1`, являются недопустимыми, поскольку существует ячейка `CAT1`. Если вы попытаетесь присвоить ячейке имя `CAT1`, появится сообщение о том, что вы ввели недопустимое имя. Возможно, самый наилучший вариант — назвать ячейку `cat1_`.
- Единственные символы, которые можно использовать в именах диапазонов, это точка (`.`) и подчеркивание (`_`).

## Задания

1. Файл `Stock.xlsx` содержит данные о ежемесячной доходности акций General Motors и Microsoft. Присвойте имена диапазонам, содержащим ежемесячную

доходность для каждой акции, и вычислите среднемесячную доходность каждой акции.

2. Создайте новую пустую книгу и присвойте имя **Red** диапазону, содержащему ячейки **A1:B3** и **A6:B8**.
3. Если задать широту и долготу любых двух городов, в файле **Citydistances.xlsx** будет вычислено расстояние между этими двумя городами. Определите имена диапазонов для широты и долготы каждого города и проверьте, чтобы эти имена отображались в формуле для расчета расстояния.
4. Файл **Sharedata.xlsx** содержит количество акций (**shares**) для каждого вида акций и цену одной акции (**price**). Вычислите стоимость акций каждого вида по формуле **=shares\*price**.
5. Создайте имя диапазона для расчета среднего значения продаж за последние пять лет. Предполагается, что годовые продажи перечислены в единственном столбце.

## ГЛАВА 3

# Функции поиска

### Обсуждаемые вопросы

- Как создать формулу для вычисления налоговых ставок на основе дохода?
- Как посмотреть цену продукта по идентификатору продукта?
- Допустим, что цена продукта изменяется со временем. Известна дата продажи продукта. Как создать формулу для вычисления цены продукта?

## Синтаксис функций поиска

Функция поиска в Excel позволяет просматривать значения в диапазонах на листах книги. В Microsoft Excel 2016 доступен как вертикальный (с помощью функции ВПР), так и горизонтальный просмотр (с помощью функции ГПР). При вертикальном просмотре операция поиска начинается с первого столбца диапазона. При горизонтальном просмотре поиск начинается с первой строки диапазона. Я сосредоточусь на функции ВПР, поскольку в большинстве формул используется вертикальный просмотр.

### Синтаксис функции ВПР

Ниже приведен синтаксис функции ВПР (VLOOKUP). В квадратных скобках ( [ ] ) указаны необязательные аргументы.

ВПР(искомое\_значение;таблица;номер\_столбца;[интервальный\_просмотр])

- **искомое\_значение** (lookup\_value) — значение для поиска в первом столбце таблицы;
- **таблица** (table\_range) — диапазон таблицы. В него входит первый столбец, в котором выполняется поиск искомого значения, и любые другие столбцы, в которых требуется просмотреть результаты расчетов по формулам;
- **номер\_столбца** (column\_index) — номер столбца в таблице, из которого функция поиска возвращает значение;
- **интервальный\_просмотр** (range\_lookup) является необязательным аргументом. Он устанавливает точное или приблизительное совпадение. Если значение **интервальный\_просмотр** равно ИСТИНА или опущено, первый столбец диапазона таблицы должен быть отсортирован по возрастанию. Если значение **интерваль-**

ный\_просмотр равно ИСТИНА или опущено и в первом столбце таблицы найдено точное совпадение, Excel основывает поиск на табличной строке, в которой найдено точное совпадение. Если значение интервальный\_просмотр равно ИСТИНА или опущено и нет точного совпадения, то поиск основывается на наибольшем значении в первом столбце, не превышающем искомое значение. Если значение интервальный\_просмотр равно ЛОЖЬ и в первом столбце таблицы найдено точное совпадение, поиск основывается на табличной строке, в которой найдено точное совпадение. Если точного совпадения нет, Excel выдает сообщение об ошибке #Н/Д (недоступно). В главе 12 вы узнаете, как использовать функцию ЕСЛИОШИБКА (IFERROR), чтобы Excel не выдавал подобной ошибки (#Н/Д). Обратите внимание, что значение 1 для аргумента интервальный\_просмотр эквивалентно значению ИСТИНА, а значение 0 эквивалентно значению ЛОЖЬ.

## Синтаксис функции ГПР

Функция ГПР (HLOOKUP) ищет значение в первой строке (а не в первом столбце) таблицы. Для функции ГПР используйте синтаксис функции ВПР, только поменяйте номер\_столбца на номер\_строки.

Теперь рассмотрим несколько интересных примеров с использованием функций поиска.

## Ответы на вопросы

### ❓ Как создать формулу для вычисления налоговых ставок на основе дохода?

Вот как работает функция ВПР с первым столбцом таблицы, состоящим из чисел, отсортированных по возрастанию. Допустим, налоговая ставка зависит от дохода, как показано в таблице.

Уровень доходов	Налоговая ставка
\$0–\$9999	15%
\$10 000–\$29 999	30%
\$30 000–\$99 999	34%
\$100 000 и выше	40%

Как создать формулу для расчета налоговой ставки см. в файле `Lookup.xlsx` и на рис. 3.1.

Сначала в диапазон ячеек D6:E9 мы ввели соответствующие данные (налоговые ставки и точки прерывания). Таблице D6:E9 присвоено имя `lookup`. Рекомендуется всегда присваивать имена ячейкам, используемым в качестве диапазона таблицы.

	C	D	E	F	G
1					
2					
3	<b>Lookup Tables</b>				
4					
5		<b>Income</b>	<b>Tax rate</b>		<b>Lookup=D6:E9</b>
6		0	0.15		
7		10000	0.3		
8		30000	0.34		
9		100000	0.4		
10					
11			<b>TRUE</b>	<b>FALSE</b>	
12		<b>Income</b>	<b>Rate</b>		
13		-1000	#N/A	#N/A	
14		30000	0.34	0.34	
15		29000	0.3	#N/A	
16		98000	0.34	#N/A	
17		104000	0.4	#N/A	

**Рис. 3.1.** Функция поиска для расчета налоговой ставки.  
Числа в первом столбце таблицы отсортированы по возрастанию

Тогда нам не нужно помнить точное местонахождение таблицы, а при копировании любой формулы, включающей функцию поиска, диапазон поиска всегда будет правильным. Затем для демонстрации работы функции поиска я ввел в диапазон D13:D17 значения доходов. Налоговая ставка для уровней доходов, указанных в диапазоне D13:D17, была рассчитана путем копирования формулы =ВПР(D13;Lookup;2;ИСТИНА) из E13 в E14:E17.

Проверьте, как работает поиск в ячейках E13:E17. Обратите внимание, что ответ всегда берется из второго столбца таблицы, поскольку в формуле указан номер столбца 2.

- В D13 доход, равный -1000, вызвал ошибку #Н/Д, поскольку такая сумма меньше самого низкого уровня доходов в первом столбце таблицы. Если вам нужно связать налоговую ставку 15% с доходом -\$1000, замените 0 в ячейке D6 числом, не большим -1000.
- В D14 доход в \$30 000 точно совпадает со значением в первом столбце таблицы, поэтому функция возвращает значение ставки 34%.
- В D15 доход в \$29 000 не имеет точного совпадения в первом столбце таблицы. Это означает, что функция поиска остановилась на самом большем числе в первом столбце таблицы, не превышающем \$29 000; в данном случае это \$10 000. Функция возвратила налоговую ставку из второго столбца таблицы, соответствующую \$10 000, то есть 30%.
- В D16 доход в \$98 000 не имеет точного совпадения в первом столбце таблицы. Функция поиска остановилась на самом большем числе в первом столбце таб-

лицы, не превышающем \$98 000. Она возвратила налоговую ставку из второго столбца таблицы, соответствующую \$30 000, то есть 34%.

- В D17 доход в \$104 000 не имеет точного совпадения в первом столбце таблицы. Функция поиска остановилась на самом большом числе в первом столбце таблицы, не превышающем \$104 000, что означает возврат налоговой ставки из второго столбца таблицы, соответствующей \$100 000, то есть 40%.

В F13:F17 я изменил значение аргумента `интервальный_просмотр` с `ИСТИНА` на `ЛОЖЬ`, и скопировал из ячейки F13 в F14:F17 формулу `=ВПР(D13;Lookup;2;ЛОЖЬ)`. Ячейка F14 по-прежнему содержит налоговую ставку 34%, поскольку в первом столбце таблицы имеется точное совпадение — 30000. Во всех других ячейках F13:F17 Excel выдал ошибку `#Н/Д`, поскольку уровни дохода в этих ячейках не имеют точного совпадения в первом столбце таблицы.

### ❓ Как посмотреть цену продукта по идентификатору продукта?

Нередко бывает так, что числа в первом столбце не отсортированы по возрастанию. Например, там могут быть перечислены идентификаторы продуктов или имена сотрудников. Из своей практики я знаю, что тысячи финансовых аналитиков не умеют пользоваться функцией поиска, если в первом столбце числа не указаны в возрастающем порядке. Здесь нужно помнить только одно простое правило: для аргумента `интервальный_просмотр` указывайте значение `ЛОЖЬ`.

Например, файл `Lookup.xlsx` (рис. 3.2) содержит цену для пяти продуктов, указанных по идентификаторам. Каким создать формулу, которая по идентификатору продукта выдавала бы цену продукта?

	A	H	I	J
9				<b>Lookup2=H11:I15</b>
10		<b>Product ID</b>	<b>Price</b>	
11		A134	\$ 3.50	
12		B242	\$ 4.20	
13		X212	\$ 4.80	
14		C413	\$ 5.00	
15		B2211	\$ 5.20	
16				
17		<b>ID</b>	<b>Price</b>	
18		B2211	3.5	
19		B2211	5.2	

**Рис. 3.2.** Поиск цен по идентификаторам продуктов. Если таблица не отсортирована по возрастанию, укажите в формуле значение `ЛОЖЬ` для последнего аргумента функции поиска

Многие ввели бы в ячейку I18 формулу `=ВПР(H18;Lookup2;2)`. Но, если вы опустите четвертый аргумент (`интервальный_просмотр`), предполагается, что его значение



равно **ИСТИНА**. Тогда будет возвращена неправильная цена (3,50), поскольку идентификаторы продуктов в таблице **Lookup2** (ячейки H11:I15) перечислены не в алфавитном порядке. Если ввести в ячейку I18 формулу **=ВПР(H18;Lookup2;2;ЛОЖЬ)**, цена будет возвращена правильная (5,20).

Для поиска зарплаты сотрудника по его фамилии или идентификационному номеру укажите значение **ЛОЖЬ** в формуле аналогичным образом.

Кстати, на рис. 3.2 видно, что столбцы A–G скрыты. Чтобы скрыть нужные столбцы, сначала выделите их. Затем на вкладке **Главная (Home)** в группе **Ячейки (Cells)** раскройте список **Формат (Format)**, в **Видимость (Visibility)** выберите **Скрыть** или **отобразить (Hide & Unhide)** и затем **Скрыть столбцы (Hide Columns)**.

**?** Допустим, что цена продукта изменяется со временем. Известна дата продажи продукта. Как создать формулу для вычисления цены продукта?

Допустим, цена продукта зависит от даты его продажи. Каким образом можно написать для формулы функцию поиска правильной цены продукта? Например, для следующей таблицы?

Дата продажи	Цена
Январь–апрель 2005 г.	\$98
Май–июль 2005 г.	\$105
Август–декабрь 2005 г.	\$112

Напишите формулу для определения правильной цены продукта для любой из дат продажи продукта в 2005 г. Для разнообразия воспользуйтесь функцией **ГПР**. Даты изменения цены приведены в первой строке таблицы (см. файл **Datelookup.xlsx** и рис. 3.3).

	A	B	C	D
1				
2	<b>Date</b>	1/1/2005	5/1/2005	8/1/2005
3	<b>Price</b>	98	105	112
4				
5			<b>Lookup:B2:D3</b>	
6				
7		<b>Date</b>	<b>Price</b>	
8		1/4/2005	98	
9		5/10/2005	105	
10		9/12/2005	112	
11		5/1/2005	105	
12				

**Рис. 3.3.** Функция **ГПР** для определения цены, изменяющейся в зависимости от даты продажи

Я скопировал формулу `=ГПР(B8,lookup,2,ИСТИНА)` из C8 в C9:C11. Формула пытается сравнить даты из столбца B с первой строкой диапазона B2:D3. Для любой даты между 01.01.05 и 30.04.05 функция поиска выбирает 01.01.05 и возвращает значение цены из B3; для любой даты между 01.05.05 и 31.07.05 функция поиска выбирает 1 мая и возвращает значение цены из C3; и для любой даты после 01.08.05 выбирается 1 августа и возвращается значение цены из D3.

## Задания

1. В файле `Hr.xlsx` указаны идентификационные номера, оклады и стаж работников. Создайте формулу для поиска зарплаты работника по идентификационному коду. Создайте еще одну формулу, находящую стаж работника по его идентификационному коду.
2. В файле `Assign.xlsx` работники распределены по четырем группам. Там же указано, насколько данный работник пригоден для работы в каждой группе (по шкале от 0 до 10). Напишите формулу, которая вычисляет, насколько пригоден каждый работник для группы, в которую его назначили.
3. Вы подумываете о рекламе продуктов Microsoft на спортивном канале. Чем больше рекламных объявлений вы покупаете, тем ниже цена одного объявления (см. таблицу ниже).

Количество рекламных объявлений	Цена одного рекламного объявления
1–5	\$12 000
6–10	\$11 000
11–20	\$10 000
21 и более	\$9000

Например, при покупке 8 рекламных объявлений за одно объявление придется выложить \$11 000, а при покупке 14 объявлений — \$10 000. Напишите формулу, рассчитывающую общие затраты для любого количества рекламных объявлений.

4. Вы подумываете о рекламе продуктов Microsoft в популярной музыкальной телепрограмме. За первую группу рекламных объявлений вы платите фиксированную сумму, но если вы покупаете больше рекламных объявлений, цена за каждое объявление снижается (см. таблицу ниже).

Количество рекламных объявлений	Цена одного рекламного объявления
1–5	\$12 000
6–10	\$11 000
11–20	\$10 000
21 и более	\$9000

Например, при покупке 8 рекламных объявлений первые 5 объявлений обойдутся в \$12 000 каждое, а оставшиеся — по \$11 000. При покупке 14 объявлений первые 5 стоят \$12 000, следующие 5 по \$11 000 и оставшиеся 4 по \$10 000. Напишите формулу расчета общих затрат на покупку любого количества объявлений. Подсказка: вам потребуются минимум три столбца в таблице и две функции поиска в формуле.

5. В следующей таблице приведены ежегодные процентные ставки по предоставляемым банком кредитам на 1, 5, 10 и 30 лет.

Срок кредитования, лет	Процентная ставка, %
1	6
5	7
10	9
30	10

Если кредит в банке взят на любой срок от 1 до 30 летне указанный в таблице, процентная ставка рассчитывается путем интерполяции между ставками, указанными в таблице. Например, если кредит взят на 15 лет, процентная ставка определяется по формуле:

$$\frac{1}{4}(9) + \frac{3}{4}(10) = 9,75\%.$$

Напишите формулу расчета процентной ставки по кредиту для любого срока от 1 года до 30 лет.

6. Расстояние между двумя городами США (Аляску и Гавайи не учитываем) можно вычислить приблизительно по следующей формуле:

$$69 \times \sqrt{(\text{lat}_1 - \text{lat}_2)^2 + (\text{long}_1 - \text{long}_2)^2}.$$

В файле **Citydata.xlsx** приведен список городов США с шириной и долготой. Создайте таблицу, вычисляющую расстояние между любыми двумя городами из данного списка.

7. На первом листе книги **Pinevalley.xlsx** указаны зарплаты сотрудников Университета Pine Valley, на втором листе указан их возраст, на третьем — стаж работы. Создайте четвертый лист, содержащий зарплату, возраст и стаж каждого сотрудника.
8. В файле **Lookupmultiplecolumns.xlsx** приведены данные продаж в магазине электроники. Имя продавца будет введено в B17. Создайте формулу Excel, которую можно скопировать из C17 в D17:F17 и которая извлекает данные о том, сколько этот продавец продал радиотоваров, в C17, телевизоров — в D17, принтеров — в E17 и CD — в F17.

9. В файле **Grades.xlsx** приведены экзаменационные оценки студентов. Допустим, соотношение баллов и оценок следующее (см. табл.).

Баллы	Оценка
Ниже 60	F
60–69	D
70–79	C
80–89	B
90 и выше	A

При помощи формулы Excel определите оценку каждого студента по баллам.

10. В файле **Employees.xlsx** вы найдете рейтинг (по шкале от 0 до 10), присвоенный каждым из 35 сотрудников трем заданиям. В файле также указаны задания, выданные каждому сотруднику. С помощью формулы определите рейтинг, который каждый сотрудник присвоил назначенному ему заданию.
11. Допустим, что за один доллар дают 1000 иен, 5 песо или 0,7 евро. Создайте электронную таблицу, в которую можно ввести сумму в долларах США, указать валюту и затем перевести введенную сумму в указанную валюту.
12. В файле **Qb2013.xlsx** содержится статистика по квотербэкам (QB) национальной футбольной лиги за сезон 2013 г. Напишите в ячейках J2 и K2 формулы, возвращающие количество забитых мячей и перехватов мяча конкретного нападающего, когда вы вводите его имя в ячейке I2.
13. Файл **NBAplayers.xlsx** содержит данные о возрасте и заработной плате нескольких игроков НБА. Введите формулы в ячейках J5:K50, которые возвращают возраст и зарплату каждого игрока.
14. В столбце F файла **Hardware.xlsx** содержатся коды хозтоваров, а в столбце G — цены этих товаров. В столбцах M–O перечислены количество и цена, по которой магазин хозтоваров закупил различные товары. Определите общую стоимость закупок магазина хозтоваров.

## ГЛАВА 4

# Функция ИНДЕКС

### Обсуждаемые вопросы

- У меня есть список расстояний между различными городами США. Как должна выглядеть функция, возвращающая расстояние между Сиэтлом и Майами?
- Можно ли создать формулу со ссылкой на весь столбец, содержащий сведения о расстоянии между Сиэтлом и другими городами?

## Синтаксис функции ИНДЕКС

Функция **ИНДЕКС** (**INDEX**) возвращает содержимое ячейки в указанной строке и указанном столбце внутри массива чисел. Наиболее часто для функции **ИНДЕКС** используется следующий синтаксис:

**ИНДЕКС(массив;номер\_строки;номер\_столбца)**

Например, формула **=ИНДЕКС(A1:D12;2;3)** возвращает содержимое ячейки во второй строке третьего столбца массива A1:D12. Это ячейка C2.

## Ответы на вопросы

**?** У меня есть список расстояний между различными городами США. Как должна выглядеть функция, возвращающая расстояние между Сиэтлом и Майами?

В файле **Index.xlsx** (рис. 4.1) содержатся данные о расстоянии между восемью городами США. Диапазону C10:J17 с этими данными присвоено имя **Distances**.

Допустим, вы хотите ввести в ячейку расстояние между Бостоном и Денвером. Поскольку расстояния от Бостона до других городов перечислены в первой строке массива **Distances**, а расстояния до Денвера указаны в четвертом столбце массива, формула выглядит так: **=ИНДЕКС(distances;1;4)**. В результате Бостон и Денвер разделяет 1991 миля. Аналогично расстояние (гораздо более значительное) между Сиэтлом и Майами можно узнать по формуле **=ИНДЕКС(distances;6;8)**. Сиэтл и Майами разделяют 3389 миль.

	A	B	C	D	E	F	G	H	I	J
1										
2										
3										
4		Boston-Denver	1991			T Dist to Seattle	15221			
5		Seattle- Miami	3389							
6										
7										
8										
9			Boston	Chicago	Dallas	Denver	LA	Miami	Phoenix	Seattle
10	1	Boston	0	983	1815		3036	1539	2664	2612
11	2	Chicago	983	0	1205	1050	2112	1390	1729	2052
12	3	Dallas	1815	1205	0	801	1425	1332	1027	2404
13	4	Denver	1991	1050	801	0	1174	2100	836	1373
14	5	LA	3036	2112	1425	1174	0	2757	398	1909
15	6	Miami	1539	1390	1332	2100	2757	0	2359	3389
16	7	Phoenix	2664	1729	1027	836	398	2359	0	1482
17	8	Seattle	2612	2052	2404	1373	1909	3389	1482	0

**Рис. 4.1.** Использование функции ИНДЕКС для определения расстояния между городами

Представьте, что футбольная команда из Сиэтла отправляется в турне, в котором они проводят игры в Финиксе, Лос-Анджелесе, Денвере, Далласе и Чикаго. В конце выездной серии игр команда возвращается в Сиэтл. Легко ли подсчитать, сколько миль они проедут в течение турне? Как показано на рис. 4.2, следует просто перечислить города в порядке посещения ( 8-7-5-4-3-2-8), начиная и заканчивая Сиэтлом, и скопировать формулу =ИНДЕКС(distances;C21;C22) в ячейки с D21 по D26. Формула в ячейке D21 определяет расстояние между Сиэтлом и Финиксом (город номер 7), формула в ячейке D22 — расстояние между Финиксом и Лос-Анджелесом и т. д. Всего команда проедет 7112 миль. И просто забавы ради покажите с помощью функции ИНДЕКС, что команда «Майами Хит» путешествует за сезон больше любой другой команды НБА.

	C	D
19	Road Trip!!	
20	City	Distance
21	8	1482
22	7	398
23	5	1174
24	4	801
25	3	1205
26	2	2052
27	8	
28	Total	7112

**Рис. 4.2.** Расстояния для турне футбольной команды из Сиэтла

❓ Можно ли создать формулу со ссылкой на весь столбец, содержащий сведения о расстояниях между Сиэтлом и другими городами?

Функция ИНДЕКС позволяет сослаться на всю строку или весь столбец какого-либо массива. Если установить значение 0 для номера строки, то функция ИНДЕКС

будет ссылаться на указанный столбец. А если установить 0 для номера столбца, то функция **ИНДЕКС** ссылается на указанную строку. Предположим, необходимо определить суммарное расстояние между всеми перечисленными городами и Сиэтлом. Можно ввести любую из следующих формул:

=СУММ(ИНДЕКС(distances;8;0))

=СУММ(ИНДЕКС(distances;0;8))

В первой формуле суммируются числа в восьмой строке (строка 17) массива **Distances**; во второй формуле суммируются числа в восьмом столбце (столбец J) массива **Distances**. И в том, и в другом случае суммарное расстояние от Сиэтла до других городов и от этих городов до Сиэтла составляет 15 221 миль (см. рис. 4.1).

## Задания

1. С помощью функции **ИНДЕКС** определите расстояние между Лос-Анджелесом и Финиксом, а также расстояние между Денвером и Майами.
2. С помощью функции **ИНДЕКС** вычислите суммарное расстояние от Далласа до других городов, перечисленных на рис. 4.1.
3. Джерри Джонс с командой «Далласские Ковбои» отправляются в путешествие на машине в Чикаго, Денвер, Лос-Анджелес, Финикс и Сиэтл. Сколько миль им предстоит проехать?
4. В файле **Product.xlsx** содержатся данные о продаже шести продуктов по месяцам. С помощью функции **ИНДЕКС** покажите продажи продукта 2 за март. Используйте функцию **ИНДЕКС** для подсчета общего объема продаж за апрель.
5. В файле **Nbadistances.xlsx** показаны расстояния между любыми двумя спортивными аренами НБА. Допустим, вы находитесь в Атланте и вам нужно посетить арены в указанном порядке и затем вернуться в Атланту. Какое расстояние вы проедете?
6. Используйте функцию **ИНДЕКС** для решения задания 10 из главы 3. Вернемся к этому заданию: в файле **Employees.xlsx** вы найдете рейтинг (по шкале от 0 до 10), присвоенный каждым из 35 сотрудников трем заданиям. В файле также указаны задания, выданные каждому сотруднику. С помощью формулы определите рейтинг, который каждый сотрудник присвоил назначенному ему заданию.

## ГЛАВА 5

# Функция ПОИСКПОЗ

### Обсуждаемые вопросы

- Как создать формулу, возвращающую объем продаж какого-либо продукта за определенный месяц, если даны ежемесячные продажи для нескольких продуктов? Например, сколько продукта 2 было продано за июнь?
- Как создать формулу, определяющую бейсболиста с самой высокой зарплатой, если дан список зарплат игроков? А бейсболиста с пятой по величине зарплатой?
- Как создать формулу, определяющую период окупаемости для первоначальных инвестиционных вложений, если известны ежегодные денежные потоки от реализации инвестиционного проекта?

Предположим, что у вас есть лист с 5000 имен в 5000 строках. Необходимо найти имя **John Doe**, про которое известно, что оно встречается в списке только один раз. Хотите узнать формулу, возвращающую номер строки, содержащей искомое имя? Функция **ПОИСКПОЗ (MATCH)** позволяет найти первое вхождение заданной текстовой строки или числа в пределах указанного массива. Используйте функцию **ПОИСКПОЗ** вместо функции поиска, если вам нужно найти положение числа в диапазоне, а не значение в определенной ячейке. Синтаксис функции **ПОИСКПОЗ**:

**ПОИСКПОЗ(искомое\_значение;просматриваемый\_массив;[тип\_сопоставления])**

В приведенном далее пояснении предполагается, что все ячейки массива расположены в одном и том же столбце. Здесь:

- **искомое\_значение (lookup\_value)** — это значение, с которым сопоставляются значения просматриваемого массива;
- **просматриваемый\_массив (lookup\_range)** — это диапазон, который просматривается для сопоставления с искомым значением. Просматриваемый массив должен быть либо строкой либо столбцом;
- **тип\_сопоставления (match\_type) = 1** — просматриваемый диапазон должен состоять из чисел, отсортированных по возрастанию. Функция **ПОИСКПОЗ** возвращает позицию строки в просматриваемом массиве (по отношению к верхней части массива), которая содержит наибольшее значение в массиве, превышающее искомое значение или равное ему;
- **тип\_сопоставления = -1** — просматриваемый диапазон должен состоять из чисел, отсортированных по убыванию. Функция **ПОИСКПОЗ** возвращает позицию строки в просматриваемом массиве (по отношению к верхней части массива),



которая содержит последнее значение в массиве, превышающее искомое значение или равное ему;

- тип\_сопоставления = 0 — возвращает местоположение строки в просматриваемом диапазоне, содержащей первое точное соответствие искомому значению. (В главе 20 рассматривается поиск второго или третьего совпадения.) Если точного соответствия не существует и тип сопоставления равен 0, появляется сообщение об ошибке #Н/Д. Обычно в функции ПОИСКПОЗ указывают тип\_сопоставления = 0, но если этот аргумент не задан, считается, что тип\_сопоставления = 1. Указывайте тип\_сопоставления = 0, если содержимое ячеек просматриваемого массива не отсортировано. Такая ситуация возникает чаще всего.

В файле Matchex.xlsx (рис. 5.1) представлены три примера синтаксиса функции ПОИСКПОЗ.

	A	B	C	D	E	F	G
1							
2							
3							
4		Boston			-5		6
5		Chicago			-4		5
6		Dallas			-3		4
7		Denver			-1		3
8		LA			3		-1
9		Miami			4		-3
10		Phoenix			5		-4
11		Seattle			6		-5
12				last number<=0	4	last number>=-4	7
13	Boston	1					
14	Phoenix	7					
15	Pho*	7					

**Рис. 5.1.** Определение положения значения в диапазоне с помощью функции ПОИСКПОЗ

В ячейке B13 формула =ПОИСКПОЗ("Бостон";B4:B11;0) возвращает значение 1, поскольку значение **Бостон** содержится в первой строке диапазона B4:B11. Текстовые значения должны быть заключены в кавычки ( " "). В ячейке B14 формула =ПОИСКПОЗ("Финикс";B4:B11;0) возвращает значение 7, поскольку ячейка B10 (седьмая ячейка в диапазоне B4:B11) является первой ячейкой в просматриваемом диапазоне, значение которой совпадает с **Финикс**. В ячейке E12 формула =ПОИСКПОЗ(0;E4:E11;1) возвращает значение 4, так как последнее число, не превышающее 0 в диапазоне E4:E11, находится в ячейке E7 (четвертой ячейке в просматриваемом массиве). В ячейке G12 формула =ПОИСКПОЗ(-4;G4:G11;-1) возвращает значение 7, поскольку последнее число в диапазоне G4:G11, превышающее или равное -4, находится в ячейке G10 (седьмой ячейке в просматриваемом массиве).

Функция ПОИСКПОЗ работает и с неточными совпадениями. Например, в ячейке B15 формула =ПОИСКПОЗ("Фин\*";B4:B11;0) возвращает значение 7. Звездочка обрабатывается как подстановочный знак, следовательно, в диапазоне B4:B11 выпол-

няется поиск первой текстовой строки, начинающейся с символов **Фин.** Кстати, этот же прием работает и для функции поиска. Например, в задании по поиску цены в главе 3 формула **=ВПР("х\*";lookup2;2)** возвратит цену продукта X212 (4,80).

Если просматриваемый массив содержится в строке, функция возвращает относительную позицию первого совпадения в массиве, двигаясь слева направо. Как я покажу далее, функция **ПОИСКПОЗ** особенно эффективна в сочетании с другими функциями Excel, такими как **ВПР**, **ИНДЕКС** и **МАКС (MAX)**.

## Ответы на вопросы

**?** Как создать формулу, возвращающую объем продаж какого-либо продукта за определенный месяц, если даны ежемесячные продажи для нескольких продуктов? Например, сколько продукта 2 было продано за июнь?

В файле **Productlookup.xlsx** (рис. 5.2) указаны объемы продаж с января по июнь для четырех фигурок НБА. Как создать формулу которая вычисляла бы продажи конкретного продукта за определенный месяц? Хитрый прием, который нам поможет, — использование одной функции **ПОИСКПОЗ** для поиска строки с указанным продуктом и другой функции **ПОИСКПОЗ** для поиска столбца с заданным месяцем. Затем вы можете применить функцию **ИНДЕКС** для возврата объема продаж продукта за этот месяц.

	A	B	C	D	E	F	G
1							
2							
3		January	February	March	April	May	June
4	LeBron	831	685	550	965	842	804
5	Kobe	719	504	965	816	639	814
6	MJ	916	906	851	912	964	710
7	Curry	844	509	991	851	742	817
8							
9	Product	Month	Row # of product	Column # of month	Product Sales		
10	Kobe	June	2	6	814		

**Рис. 5.2.** Функция **ПОИСКПОЗ** может использоваться в сочетании с функциями **ИНДЕКС** и **ВПР**

В диапазоне **B4:G7** с именем **Sales** содержатся данные об объемах продаж кукол. Продукт, данные о котором требуется найти, указан в ячейке **A10**, а месяц — в ячейке **B10**. В ячейку **C10** введена формула **=ПОИСКПОЗ(A10;A4:A7;0)** для определения номера строки в диапазоне **Sales** с данными по кукле **Kobe**. Далее в ячейке **D10** находится формула **=ПОИСКПОЗ(B10;B3:G3;0)** для определения номера столбца в диапазоне **Sales** с данными по продажам за июнь. Теперь, когда номера строки и столбца, в кото-

рых содержатся требуемые данные по продажам, известны, можно воспользоваться формулой **=ИНДЕКС(Sales;C10;D10)** в ячейке **E10** для получения требуемых данных. Для получения дополнительной информации о функции **ИНДЕКС** см. главу 4.

**?** Как создать формулу, определяющую бейсболиста с самой высокой зарплатой, если дан список зарплат игроков? А бейсболиста с пятой по величине зарплатой?

В файле **Baseball.xlsx** (рис. 5.3) перечислены зарплаты, выплаченные 401 ведущим игрокам бейсбольной лиги в течение сезона 2001 г. Данные по зарплате не отсортированы. Требуется составить формулу, возвращающую имя игрока с самой высокой зарплатой, а также имя игрока с пятой по величине зарплатой.

Для поиска игрока с самой высокой зарплатой выполните следующее:

1. С помощью функции **МАКС** определите значение самой высокой зарплаты.
2. С помощью функции **ПОИСКПОЗ** определите строку, содержащую имя игрока с самой высокой зарплатой.
3. С помощью функции **ВПР** (обрабатывающей строку данных с зарплатой игрока) найдите имя игрока.

Данные о зарплатах игроков содержатся в диапазоне **C12:C412** с именем **Salaries**. Диапазон, указанный в функции **ВПР** (диапазон **A12:C412**), называется **Lookup**.

	A	B	C	D
6		name	Alex Rodriguez	dl-Derek Jeter
7			highest	5th highest
8		player position	345	232
9		amount	22000000	12600000
10				
11		name	salary	
12	1	dl-Mo Vaughn	13166667	
13	2	Tim Salmon	5683013	
14	3	Garret Anderson	4500000	
15	4	Darin Erstad	3450000	
16	5	Troy Percival	3400000	
17	6	Ismael Valdes	2500000	
18	7	Pat Rapp	2000000	
19	8	Glenallen Hill	1500000	
20	9	Troy Glaus	1250000	
21	10	Shigetoshi Hasegawa	1150000	
22	11	Scott Spiezio	1125000	
23	12	Orlando Palmeiro	900000	
24	13	Alan Levine	715000	
25	14	Mike Holtz	705000	
26	15	Jorge Fabregas	500000	
27	16	Benji Gil	350000	

**Рис. 5.3.** Поиск и отображение самого большого значения в списке с помощью функций **МАКС**, **ПОИСКПОЗ** и **ВПР**

В ячейке C9 по формуле `=МАКС(salaries)` найдите самую высокую зарплату (22 млн долларов). Затем в ячейке C8 по формуле `=ПОИСКПОЗ(C9;Salaries;0)` определите номер игрока с самой высокой зарплатой. Так как зарплаты не отсортированы ни по возрастанию, ни по убыванию, используется аргумент `тип_сопоставления = 0`. Самую высокую зарплату имеет игрок с номером 345. Наконец, в ячейке C6 по формуле `=ВПР(C8;Lookup;2)` найдите имя игрока во втором столбце диапазона `Lookup`. Неудивительно, что самым высокооплачиваемым игроком в 2001 г. был Алекс Родригес (Alex Rodriguez).

Для поиска имени игрока с пятой по величине зарплатой необходима функция поиска пятого по величине числа в массиве. Эту работу выполнит функция `НАИБОЛЬШИЙ (LARGE)`. Ее синтаксис: `НАИБОЛЬШИЙ(массив,к)`. В таком виде она возвращает *k*-е по величине число в диапазоне. Таким образом, формула `=НАИБОЛЬШИЙ(salaries;5)` в ячейке D9 дает пятую по величине зарплату (12,6 млн долларов). Действуя по описанному выше алгоритму вы обнаружите, что игроком с пятой по величине зарплатой был Дерек Джетер (Derek Jeter). (Буквы `dl` перед его именем указывают, что в начале сезона Джетер был в списке травмированных игроков.) Функция `НАИМЕНЬШИЙ (SMALL)` в формуле `=НАИМЕНЬШИЙ(salaries;5)` служит для поиска пятой снизу по величине зарплат.

### ❓ Как создать формулу, определяющую период окупаемости для первоначальных инвестиционных вложений, если известны ежегодные денежные потоки от реализации инвестиционного проекта?

В файле `Payback.xlsx` (рис. 5.4) показаны прогнозируемые денежные потоки для инвестиционного проекта в следующие 15 лет. Предположим, что в нулевой период потребовался отток денежных средств в размере 100 млн долларов. За первый год проект вызвал приток денежных средств в размере 14 млн долларов. Ожидается, что денежные потоки будут расти на 10% в год. Сколько лет пройдет, прежде чем инвестиции в проект окупятся?

Количество лет, необходимых для возврата первоначальных инвестиционных затрат на проект, называется *периодом окупаемости*. В высокотехнологичных отраслях период окупаемости часто используется для ранжирования инвестиционных проектов. (В главе 8 показано, что использование периода окупаемости как меры качества инвестиционных проектов является некорректным, поскольку при этом игнорируется изменение стоимости денег во времени.) В данный момент, однако, сосредоточимся на способе определения периода окупаемости для этой упрощенной инвестиционной модели.

Для определения периода окупаемости проекта выполните следующее:

1. В столбце В вычислите денежный поток для каждого года.
2. В столбце С вычислите суммарный денежный поток для каждого года.

Теперь с помощью функции `ПОИСКПОЗ` (укажите `тип_сопоставления` равным 1) определите номер строки первого года с положительным значением суммарного денежного потока. Это вычисление позволит получить период окупаемости.

	A	B	C	D	E
1	Year 1 cash flow	14			Payback Period
2	Growth	0.1			6
3	Initial Invetment	-100			
4	Year	Annual Cash flow	Cumulative cash flow		
5	0	-100	-100		
6	1	14	-86		
7	2	15.4	-70.6		
8	3	16.94	-53.66		
9	4	18.634	-35.026		
10	5	20.4974	-14.5286		
11	6	22.54714	8.01854		
12	7	24.801854	32.820394		
13	8	27.2820394	60.1024334		
14	9	30.01024334	90.11267674		
15	10	33.01126767	123.1239444		
16	11	36.31239444	159.4363389		
17	12	39.94363389	199.3799727		
18	13	43.93799727	243.31797		
19	14	48.331797	291.649767		
20	15	53.1649767	344.8147437		

**Рис. 5.4.** Вычисление периода окупаемости инвестиций с помощью функции ПОИСКПОЗ

Ячейкам диапазона B1:B3 присвоены имена, соответствующие тексту в ячейках A1:A3. Денежный поток в нулевой период, а именно начальные инвестиции со знаком «минус», введены в ячейку B5. Денежный приток первого года указан в ячейке B6. Денежные потоки со второго года по пятнадцатый вычисляются путем копирования формулы =B6\*(1+прирост) из ячейки B7 в ячейки B8:B20.

При расчете суммарного денежного потока для нулевого периода используйте формулу из ячейки B5 в ячейке C5. Для последующих лет суммарный денежный поток можно вычислить по формуле:

$$\text{суммарный\_поток\_за\_год } t = (\text{суммарный\_поток\_за\_год } t - 1) + \text{поток\_за\_год } t.$$

Для реализации этого отношения скопируйте формулу =C5+B6 из ячейки C6 в ячейки C7:C20.

Для расчета периода окупаемости с помощью функции ПОИСКПОЗ (укажите тип сопоставления равным 1) вычислите последнюю строку диапазона C5:C20, содержащую значение меньше 0. Этот расчет всегда дает период окупаемости. Если последней строкой в C5:C20, содержащей отрицательное значение, является шестая строка в диапазоне (год 5), то седьмое значение соответствует суммарному денежному потоку первого года, когда проект начнет приносить прибыль. Поскольку отсчет идет с нулевого года, проект окупится на шестой год. Таким образом, формула =ПОИСКПОЗ(0;C5:C20;1) в ячейке E2 дает срок окупаемости (6 лет). Если какие-либо денежные потоки после нулевого периода имеют отрицательное значение, этот метод работать не будет и выдаст сообщение об ошибке, потому

что в этом случае диапазон суммарных денежных потоков не будет отсортирован по возрастанию. Используя функцию **ЕСЛИОШИБКА (IFERROR)** (представленную в главе 12), вы можете обеспечить выдачу в ячейке **E2** текстового сообщения (на пример «Срок окупаемости не может быть найден»), если не удастся вычислить срок окупаемости.

## Задания

1. Воспользуйтесь данными о расстояниях между городами США в файле **Index.xlsx** и напишите формулу с функцией **ПОИСКПОЗ** для определения (на основе названий городов) расстояния между любыми двумя городами.
2. В файле **Matchtype1.xlsx** перечислены суммы 30 сделок в хронологическом порядке. Напишите формулу для определения первой сделки, после которой общий объем сделок превысил \$10 000.
3. В файле **Matchthemax.xlsx** содержатся идентификационные коды продуктов и показатели продаж в штуках для 265 продуктов. С помощью формулы с функцией **ПОИСКПОЗ** определите идентификационный код продукта с наибольшим показателем продаж.
4. В файле **Buslist.xlsx** содержатся интервалы времени (в минутах) между автобусами на пересечении 45-й улицы и Парк-авеню в Нью-Йорке. Напишите формулу расчета времени ожидания пассажиром автобуса для любого времени прихода пассажира на остановку (отсчет от первого автобуса). Например, если бы вы пришли на остановку, начиная с настоящего момента, через 12,4 минуты, а автобусы прибыли бы, начиная с настоящего момента, через 5 минут и через 21 минуту, вам пришлось бы ждать автобус  $21 - 12,4 = 8,6$  минуты.
5. В файле **Salesdata.xlsx** находятся данные о количестве компьютеров, проданных каждым продавцом. Напишите формулу, возвращающую количество проданного товара указанным продавцом.
6. Предположим, что функция **ВПР** была удалена из Excel. Объясните, как ее можно заменить с помощью функций **ПОИСКПОЗ** и **ИНДЕКС**.
7. В файле **Baseballproblem7.xlsx** приведены статистики по бейсбольным командам высшей лиги. Вы будете вводить имя команды в ячейку **I2** и статистику в ячейку **J2**. Напишите формулу в ячейке **K2**, возвращающую значение этой статистики для выбранной команды.
8. В файле **Footballproblem8.xlsx** приведены показатели по нападающим НФЛ. Допустим, мы ввели имя нападающего в ячейку **G3** и показатель в ячейку **H3**. Напишите формулу, возвращающую в ячейке **I3** значение показателя для этого игрока.

## ГЛАВА 6

# Текстовые функции и инструмент Мгновенное заполнение

### Обсуждаемые вопросы

- У меня есть лист, каждая ячейка которого содержит описание продукта, идентификатор продукта и цену продукта. Как поместить все описания в столбец A, все идентификаторы в столбец B и все цены в столбец C?
- Каждый день я получаю данные о совокупных продажах в США, которые вычисляются в одной ячейке как сумма продаж на востоке, севере и юге. Как мне извлечь продажи на востоке, севере и юге в отдельные ячейки?
- В конце каждого семестра студенты оценивают мою преподавательскую деятельность по шкале от 1 до 7 баллов. Мне известно, сколько студентов выставили тот или иной балл. Как создать гистограмму оценки качества преподавания?
- Числовые данные были загружены из интернета или из базы данных. При попытке использовать их в вычислениях появляется ошибка #ЗНАЧ!. Как решить эту проблему?
- Текстовые функции — вещь хорошая, но есть ли простой способ (без них) извлечь имена и фамилии из данных, создать список адресов электронной почты из списка имен или выполнить другие стандартные операции с текстовыми данными?
- Что такое символы Unicode?
- Чем новая функция ОБЪЕДИНИТЬ лучше старой функции СЦЕПИТЬ (или &)?
- Мне нравятся текстовые функции, а как пользоваться функцией Excel ТЕКСТ?

Когда кто-нибудь посылает вам данные или вы скачиваете их из интернета, они часто имеют не тот формат, что вам нужен. Например, даты и объемы продаж, которые должны находиться в отдельных ячейках, в скачанных данных по продажам могут быть вместе в одной ячейке. Как привести данные в нужный формат? Для этого нужно научиться пользоваться текстовыми функциями Microsoft Excel. В этой главе я расскажу вам, как пользоваться следующими текстовыми функциями Excel:

- |                     |                       |
|---------------------|-----------------------|
| ○ ЛЕВСИМВ (LEFT);   | ○ ЗАМЕНИТЬ (REPLACE); |
| ○ ПРАВСИМВ (RIGHT); | ○ ЗНАЧЕН (VALUE);     |
| ○ ПСТР (MID);       | ○ ПРОПИСН (UPPER);    |



- СЖПРОБЕЛЫ (TRIM);
- ДЛСТР (LEN);
- НАЙТИ (FIND);
- ПОИСК (SEARCH);
- ПОВТОР (REPT);
- СЦЕПИТЬ (CONCATENATE);
- СТРОЧН (LOWER);
- ПРОПНАЧ (PROPER);
- СИМВОЛ (CHAR);
- ПЕЧСИМВ (CLEAN);
- ПОДСТАВИТЬ (SUBSTITUTE);
- ОБЪЕДИНИТЬ (TEXTJOIN).

Также я продемонстрирую вам, как пользоваться новым инструментом **Мгновенное заполнение (Flash Fill)**, позволяющим, как по волшебству, мгновенно придавать вашим данным желаемый вид. Наконец, вы узнаете о доступном в Excel 2019 обширном наборе символов Unicode.

## Синтаксис текстовых функций

Файл Textfunctions.xlsx (рис. 6.1) содержит примеры текстовых функций. Начнем с небольшого обзора каждой функции, а затем рассмотрим их применение для решения конкретных задач. Затем мы научимся комбинировать функции для более сложной обработки данных.

	A	B	C	D
1	Reggie	Miller		
2				
3	Reggie Miller	Left 4	Regg	
4		Right 4	ller	
5		Trim spaces	Reggie Miller	
6		Number of characters		15
7		Number of characters in trimmed result		13
8		5 characters starting at space 2	eggie	
9		Find first space		7
10		Find first r ( case sensitive)		15
11		Find first r (not case sensitive)		1
12		Combining first and Last Name	Reggie Miller	Reggie Miller
13		Replace gg with nn	Rennie Miller	
14	Text 31	Number 31		
15	31		31	
16		Change to lower case	reggie miller	
17		Change to Upper case	REGGIE MILLER	I LOVE EXCEL 2019!
18		Change to Proper case	Reggie Miller	
19		Replace all spaces by *	I*LOVE*EXCEL*2019!	
20		Replace only third space by *	I LOVE EXCEL*2019!	

**Рис. 6.1.** Примеры использования текстовых функций

## Функция ЛЕВСИМВ

Функция **ЛЕВСИМВ(текст;к)** возвращает первые  $k$  символов текстовой строки. Например, формула **=ЛЕВСИМВ(А3;4)** в ячейке С3 возвращает значение **Regg**.



## Функция ПРАВСИМВ

Функция ПРАВСИМВ(текст;к) возвращает последние *k* символов текстовой строки. Например, формула =ПРАВСИМВ(А3;4) в ячейке С4 возвращает значение ller.

## Функция ПСТР

Функция ПСТР(текст;к,м) возвращает *m* символов текстовой строки, начиная с *k*-го символа. Например, формула =ПСТР(А3;2;5) в ячейке С8 возвращает символы 2–6 из ячейки А3; результат равен eggie.

## Функция СЖПРОБЕЛЫ

Функция СЖПРОБЕЛЫ(текст) удаляет все пробелы из текстовой строки, за исключением одиночных пробелов между словами. Например, в ячейке С5 формула =СЖПРОБЕЛЫ(А3) удаляет два из трех пробелов между Reggie и Miller и возвращает результат Reggie Miller. Функция СЖПРОБЕЛЫ также удаляет пробелы в начале и конце содержимого ячейки.

## Функция ДЛСТР

Функция ДЛСТР(текст) возвращает количество символов в текстовой строке (с пробелами). Например, в ячейке С6 формула =ДЛСТР(А3) возвращает значение 15, поскольку в ячейке А3 15 символов. В ячейке С7 формула =ДЛСТР(С5) возвращает значение 13. Так как в ячейке С5 убрано два пробела, в ней на два символа меньше, чем в исходном тексте в ячейке А3.

## Функции НАЙТИ и ПОИСК

Функция НАЙТИ(искомый\_текст;просматриваемый\_текст;к) возвращает позицию первого символа искомого текста в просматриваемом тексте. Параметр *k* указывает начальную позицию поиска. Функция НАЙТИ учитывает регистр. Синтаксис функции ПОИСК аналогичен синтаксису функции НАЙТИ, но не учитывает регистр. Например, формула =НАЙТИ("r";А3;1) в ячейке С10 возвращает значение 15 — позицию первой строчной буквы *r* в текстовой строке Reggie Miller. (Заглавная *R* игнорируется, так как функция НАЙТИ учитывает регистр.) Формула =ПОИСК("r";А3;1) в ячейке С11 возвращает 1, потому что функция ПОИСК сопоставляет символ *r* с символом как нижнего, так и верхнего регистра. Формула =НАЙТИ(" ";А3;1) в ячейке С9 возвращает 7, поскольку первым пробелом в строке Reggie Miller является седьмой символ.

## Функция ПОВТОР

Функция ПОВТОР используется для повторения текстовой строки указанное число раз. Синтаксис функции: ПОВТОР(текст;число\_повторений). Например, ПОВТОР("|",3) дает в результате |||.

## Функции СЦЕПИТЬ и &

Функция СЦЕПИТЬ(текст1;текст2;...;текст30) предназначена для объединения максимум 30 текстовых строк в одну строку. Вместо функции СЦЕПИТЬ можно воспользоваться оператором &. Например, формула =A1&"&B1 в ячейке C12 возвращает строку Reggie Miller. Формула =СЦЕПИТЬ(A1;" ";B1) в ячейке D12 дает тот же результат.

## Функция ОБЪЕДИНИТЬ

Функция ОБЪЕДИНИТЬ соединяет текст из нескольких диапазонов и/или строк и содержит заданный вами символ-разделитель между текстовыми значениями, которые предстоит объединить. Если символом-разделителем является пустая текстовая строка, то эта функция эффективно сцепит диапазоны. Функция ОБЪЕДИНИТЬ имеется только в Office 365. Синтаксис ее следующий:

ОБЪЕДИНИТЬ(разделитель; игнорировать пустые; текст1; [текст2], ...).

Символ-разделитель должен быть указан обязательно, он отделяет выбранные фрагменты текста друг от друга. Параметр «игнорировать пустые» также обязателен, и если он имеет значение ИСТИНА, то будут игнорироваться пустые ячейки. Функция ОБЪЕДИНИТЬ нужна, если требуется сцепить текст, который должен разделяться одним и тем же символом-разделителем. Например, потребуется функция ОБЪЕДИНИТЬ, если вы хотите объединить текст из 10 ячеек так, чтобы фрагменты отделяла запятая. Функция ОБЪЕДИНИТЬ позволяет не включать запятую после упоминания каждой ячейки, как это было бы в формуле с функцией СЦЕПИТЬ или с &. Подробный пример с применением функции ОБЪЕДИНИТЬ дан ниже в этой главе.

## Функция ТЕКСТ

Функция ТЕКСТ позволяет вам изменять форматирование внешнего вида чисел или даты за счет применения специальных кодов. Синтаксис функции ТЕКСТ следующий: ТЕКСТ (форматируемое значение; "желаемый код форматирования"). Ниже в этой главе даны примеры применения функции ТЕКСТ и правильных кодов форматирования.

## Функция ЗАМЕНИТЬ

Функция ЗАМЕНИТЬ(старый\_текст;k;m;новый\_текст) заменяет новым текстом m символов в старом тексте начиная с k-го символа. Например, формула =ЗАМЕНИТЬ(A3;3;2;"nn") в ячейке C13 заменяет третий и четвертый символы ( gg) в ячейке A3 символами nn. В результате получается Rennie Miller.

## Функция ЗНАЧЕН

Функция ЗНАЧЕН(текст) преобразует текстовую строку, представляющую число, в число. Например, формула =ЗНАЧЕН(A15) в ячейке B15 преобразует текстовую

строку 31 из ячейки A15 в числовое значение 31. Значение 31 в ячейке A15 можно распознать как текст, поскольку оно выровнено по левому краю. Аналогично значение 31 в ячейке B15 можно идентифицировать как число по выравниванию по правому краю.

## Функции ПРОПИСН, СТРОЧН и ПРОПНАЧ

Функция ПРОПИСН(текст) меняет в строке текст все буквы на прописные. Например, в ячейке C17 по формуле =ПРОПИСН(C16) все символы преобразуются в прописные и получается REGGIE MILLER. В ячейке C16 по формуле =СТРОЧН(C12) все прописные буквы заменяются строчными и получается reggie miller. Наконец, в ячейке C18 по формуле =ПРОПНАЧ(C17) восстанавливается правильный регистр, и в результате получается Reggie Miller.

## Функция СИМВОЛ

Функция СИМВОЛ(число) возвращает (для числа от 1 до 255) символ ASCII с указанным номером. Например, СИМВОЛ(65) дает в результате A, СИМВОЛ(66) дает B, и так далее. Список символов ASCII вы найдете в файле ASCIIcharacters.xlsx. Часть списка представлена на рис. 6.2.

	D	E	F	G
1				
2	36 \$		58 :	
3	37 %		59 ;	
4	38 &		60 <	
5	39 '		61 =	
6	40 (		62 >	
7	41 )		63 ?	
8	42 *		64 @	
9	43 +		65 A	
10	44 ,		66 B	
11	45 -		67 C	
12	46 .		68 D	
13	47 /		69 E	
14	48 0		70 F	
15	49 1		71 G	
16	50 2		72 H	
17	51 3		73 I	
18	52 4		74 J	
19	53 5		75 K	
20	54 6		76 L	
21	55 7		77 M	
22	56 8		78 N	
23	57 9		79 O	

Рис. 6.2. Часть списка символов ASCII

## Функция ПЕЧСИМВ

Если вы посмотрите в файл `ASCIIcharacters.xlsx`, то заметите, что отдельные символы, такие как символ под номером 10 (перевод строки), являются невидимыми. Функция **ПЕЧСИМВ** удаляет из ячейки некоторые (но не все) невидимые символы ASCII. Например, функция **ПЕЧСИМВ** не удалит **СИМВОЛ(160)**, являющийся неразрывным пробелом. Как удалить такие проблематичные символы, как **СИМВОЛ(160)**, мы рассмотрим далее в этой главе.

## Функция ПОДСТАВИТЬ

Функция **ПОДСТАВИТЬ** заменяет в ячейке указанный текст, если местоположение текста неизвестно. Синтаксис функции **ПОДСТАВИТЬ**: **ПОДСТАВИТЬ(ячейка;старый\_текст;новый\_текст;[номер\_вхождения])**.

Последний аргумент является необязательным. Если он опущен, то каждое вхождение старого текста в ячейке заменяется новым текстом. Если последний аргумент присутствует (например, его значение равно *n*), тогда только *n*-е вхождение старого текста заменится новым текстом. Чтобы вы поняли, как работает функция **ПОДСТАВИТЬ**, допустим, что мы хотим заменить пробелы в ячейке **C17** звездочками. Сначала введите формулу **=ПОДСТАВИТЬ(D17;" ";"\*")** в ячейку **C19**. Каждый пробел будет заменен звездочкой ( *\** ), и в результате получится **I\*LOVE\*EXCEL\*2019!**. При вводе в ячейку **C20** формулы **=ПОДСТАВИТЬ(D17;" ";"\*");3)** произойдет замена на *\** только третьего пробела, и получится **I LOVE EXCEL\*2019!**.

## Ответы на вопросы

Текстовые функции очень эффективны при решении многих актуальных задач, над которыми бьются бывшие студенты, ныне работающие в крупных корпорациях. Часто для решения проблемы нужно всего лишь объединить несколько текстовых функций в одну формулу.

**❓ У меня есть лист, каждая ячейка которого содержит описание продукта, идентификатор продукта и цену продукта. Как поместить все описания в столбец А, все идентификаторы в столбец В и все цены в столбец С?**

В этом примере идентификатор продукта всегда определяется первыми 12 символами, а цена всегда указана в последних 8 символах (с двумя пробелами после каждой цены). В решении (см. файл `Lenora.xlsx` и рис. 6.4) используются функции **ЛЕВСИМВ**, **ПРАВСИМВ**, **ПСТР**, **ЗНАЧЕН**, **СЖПРОБЕЛЫ**, **ДЛСТР** и **СЦЕПИТЬ**.

Начать хорошо всегда с удаления лишних пробелов. Для этого скопируйте формулу **=СЖПРОБЕЛЫ(A4)** из ячейки **B4** в ячейки **B5:B12**. Единственными удаленными пробелами в столбце **A** оказались пробелы после каждой цены. Убедитесь в этом, поставив курсор в ячейку **A4** и нажав **<F2>** для редактирования ячейки.

В конце ячейки вы увидите два пустых символа. Результат применения функции СЖПРОБЕЛЫ показан на рис. 6.3 в столбце В. Мы знаем, что функция СЖПРОБЕЛЫ удалила два дополнительных пробела в конце ячейки, по длине строк после применения формул =ДЛСТР(A4) и =ДЛСТР(B4). Проверьте себя: в ячейке А4 должно быть 52 символа, а в ячейке В4 — 50 символов.

	A	B
1	length of A4	length of B4
2	52	50
3	Untrimmed	Trimmed
4	32592100AFES CONTROLLERPENTIUM/100,(2)1GB H 304.00	32592100AFES CONTROLLERPENTIUM/100,(2)1GB H 304.00
5	32592100JCP9 DESKTOP UNIT 225.00	32592100JCP9 DESKTOP UNIT 225.00
6	325927008990 DESKTOP WINDOWS NT 4.0 SERVER 232.00	325927008990 DESKTOP WINDOWS NT 4.0 SERVER 232.00
7	325926008990 DESKTOP WINDOWS NT 4.0 WKST 232.00	325926008990 DESKTOP WINDOWS NT 4.0 WKST 232.00
8	325921008990 DESKTOP, DOS OS 232.00	325921008990 DESKTOP, DOS OS 232.00
9	325922008990 DESKTOP, WINDOWS DESKTOP OS 232.00	325922008990 DESKTOP, WINDOWS DESKTOP OS 232.00
10	325925008990 DESKTOP, WINDOWS NT OS 232.00	325925008990 DESKTOP, WINDOWS NT OS 232.00
11	325930008990 MINITOWER, NO OS 232.00	325930008990 MINITOWER, NO OS 232.00
12	32593000KEYY MINI TOWER 232.00	32593000KEYY MINI TOWER 232.00

**Рис. 6.3.** Удаление лишних пробелов с помощью функции СЖПРОБЕЛЫ

Для получения идентификатора продукта нам надо извлечь 12 крайних слева символов из столбца В. Скопируйте формулу =ЛЕВСИМВ(B4;12) из ячейки С4 в ячейки С5:С12. Эта формула извлечет 12 первых символов из ячейки В4 и последующих ячеек и сформирует код продукта (рис. 6.4).

	B	C	D	E	F
1	length of B4				
2	50				
3	Trimmed	Product ID	Price	Product Description	Concatenation
4	32592100AFES CONTROLLERPENTIUM/100,(2)1GB H 304.00	32592100AFES	304	CONTROLLERPENTIUM/100,(2)1GB H	32592100AFES CONTROLLERPENTIUM/100,(2)1GB H 304
5	32592100JCP9 DESKTOP UNIT 225.00	32592100JCP9	225	DESKTOP UNIT	32592100JCP9 DESKTOP UNIT 225
6	325927008990 DESKTOP WINDOWS NT 4.0 SERVER 232.00	325927008990	232	DESKTOP WINDOWS NT 4.0 SERVER	325927008990 DESKTOP WINDOWS NT 4.0 SERVER 232
7	325926008990 DESKTOP WINDOWS NT 4.0 WKST 232.00	325926008990	232	DESKTOP WINDOWS NT 4.0 WKST	325926008990 DESKTOP WINDOWS NT 4.0 WKST 232
8	325921008990 DESKTOP, DOS OS 232.00	325921008990	232	DESKTOP, DOS OS	325921008990 DESKTOP, DOS OS 232
9	325922008990 DESKTOP, WINDOWS DESKTOP OS 232.00	325922008990	232	DESKTOP, WINDOWS DESKTOP OS	325922008990 DESKTOP, WINDOWS DESKTOP OS 232
10	325925008990 DESKTOP, WINDOWS NT OS 232.00	325925008990	232	DESKTOP, WINDOWS NT OS	325925008990 DESKTOP, WINDOWS NT OS 232
11	325930008990 MINITOWER, NO OS 232.00	325930008990	232	MINITOWER, NO OS	325930008990 MINITOWER, NO OS 232
12	32593000KEYY MINI TOWER 232.00	32593000KEYY	232	MINI TOWER	32593000KEYY MINI TOWER 232

**Рис. 6.4.** Извлечение кода, цены и описания продукта из текстовой строки с помощью текстовых функций

Для извлечения цен, про которые известно, что они находятся в последних шести цифрах каждой ячейки, нам нужно извлечь шесть крайних справа символов из каждой ячейки. Я скопировал формулу =ЗНАЧЕН(ПРАВСИМВ(B4;6)) из ячейки D4 в ячейки D5:D12. Я использовал функцию ЗНАЧЕН для преобразования извлеченного текста в числовое значение. Без этого преобразования математические операции над ценами в дальнейшем будут невозможны.

Для извлечения описания продукта придется попотеть сильнее. Изучив данные, вы увидите, что если начать извлечение с тринадцатого символа и продолжать до тех пор, пока не останется шесть символов до конца ячейки, можно получить требуемые данные. Эту работу выполнит формула =ПСТР(B4;13;ДЛСТР(B4)-6-12), скопированная из ячейки E4 в ячейки E5:E12. Функция ДЛСТР(B4) возвращает общее

количество символов в сжатой строке. Формула ПСТР начинает обработку с тринадцатого символа и извлекает количество символов, равное общему количеству минус 12 начальных (код продукта) и минус 6 конечных символов (цена продукта). После операции отбрасывания ненужных символов остаются только символы описания продукта.

Теперь предположим, что у вас есть следующее: идентификатор продукта в столбце C, цена продукта в столбце D и описание продукта в столбце E. Можно ли соединить эти значения и восстановить первоначальный текст?

Текстовые строки можно объединить с помощью функции СЦЕПИТЬ. Копирование формулы =СЦЕПИТЬ(C4;E4;D4) из ячейки F4 в ячейки F5:F12 восстановит исходный (без лишних пробелов) текст (см. рис. 6.4).

Объединение начинается с идентификатора продукта в ячейке C4, затем добавляется описание продукта из ячейки E4 и в конце присоединяется цена из ячейки D4. Вы восстановили текст, описывающий каждый компьютер! Объединить строки также можно при помощи знака &. Исходный текст в одной ячейке может восстановить формула =C4&E4&D4. Обратите внимание, что ячейка E4 содержит один пробел до и после описания продукта. Если бы эти пробелы отсутствовали, можно было бы воспользоваться для их вставки формулой =C4&" "&E4&" "&D4. Пробел между кавычками означает вставку пробела.

Если идентификатор продукта не всегда состоит из 12 символов, такой метод извлечения информации не годится. В таком случае можно извлечь код продукта, применив функцию НАЙТИ для обнаружения позиции первого пробела. Далее следует применить функцию ЛЕВСИМВ для извлечения всех символов слева от пробела. Этот подход показан на примере в следующем разделе.

Если цена не всегда состоит из шести символов, извлечь ее будет сложнее. (См. в задании 15 пример извлечения последнего слова текстовой строки.)

**❓ Каждый день я получаю данные о совокупных продажах в США, которые вычисляются в одной ячейке как сумма продаж на востоке, севере и юге. Как мне извлечь продажи на востоке, севере и юге в отдельные ячейки?**

Этот вопрос прислала мне сотрудница финансового отдела компании Microsoft. Каждый день она получала лист, содержащий такие формулы, как =50+200+400, =5+124+1025, и другие им подобные. Ей нужно было извлечь каждое число в ячейку соответствующего столбца. Например, необходимо было первое число (продажи на востоке) извлечь в ячейку столбца C, второе число (продажи на севере) — в столбец D и третье число (продажи на юге) — в столбец E. Задача непростая, так как мы не знаем точную позицию символов, с которых начинаются второе и третье числа в каждой ячейке. В ячейке A3 данные продаж на севере начинается с четвертого символа. В ячейке A4 — с третьего символа.

Данные для этого примера вы найдете в файле Salesstripping.xlsx (рис. 6.5). Позицию данных продаж для различных регионов можно определить следующим образом:

- продажи на востоке представлены символами слева от первого знака «+»;
- продажи на севере представлены символами между первым и вторым знаками «+»;
- продажи на юге представлены символами справа от второго знака «+».

	A	B	C	D	E	F	G
1	<b>Extracting Sales in Three Regions</b>						
2	<b>East+North+South</b>	<b>First +</b>	<b>Second +</b>	<b>East</b>	<b>North</b>	<b>Total Length</b>	<b>South</b>
3	10+300+400	3	7 10	300		10	400
4	4+36.2+800	2	7 4	36.2		10	800
5	3+23+4005	2	5 3	23		9	4005
6	18+1+57.31	3	5 18	1		10	57.31
7							
8	10	300	400				
9	4	36.2	800				
10	3	23	4005				
11	18	1	57.31				
12							
13	<b>Results of Data Text to Columns</b>						

**Рис. 6.5.** Использование функций НАЙТИ, ЛЕВСИМВ, ДЛСТР и ПСТР для извлечения региональных продаж в отдельные ячейки

Задача легко решается комбинированием функций НАЙТИ, ЛЕВСИМВ, ДЛСТР и ПСТР.

- Используйте команды редактирования для замены всех знаков равенства (=) пробелами. Для этого выделите диапазон A3:A6. Затем на вкладке Главная (Home) в группе Редактирование (Editing) кликните на Найти и выделить (Find & Select), а затем на Заменить... (Replace). В поле Найти (Find What) введите знак равенства, а в поле Заменить на (Replace With) введите пробел. Затем нажмите кнопку Заменить все (Replace All). Замена знаков равенства на пробелы преобразует формулу в текст.
- Примените функцию НАЙТИ для определения позиции обоих знаков «+» в каждой ячейке.

Начните с поиска позиции первого знака «+» для каждого фрагмента данных. Для этого скопируйте формулу =НАЙТИ("+";A3;1) из ячейки B3 в ячейки B4:B6. Поиск второго знака «+» начните с символа, следующего за первым знаком «+», скопировав формулу =НАЙТИ("+";A3;B3+1) из ячейки C3 в ячейки C4:C6.

Для извлечения данных продаж на востоке примените функцию ЛЕВСИМВ (извлечь все символы слева от первого знака «+»), скопировав формулу=ЛЕВСИМВ(A3;B3-1) из ячейки D3 в ячейки D4:D6. Для извлечения продаж на севере примените функцию ПСТР и извлеките все символы между двумя знаками «+». Начните с символа, следующего за первым знаком «+», и извлеките количество символов, равное (позиция\_второго\_знака\_плюс) – (позиция\_первого\_знака\_плюс) -1. Если не



учесть -1, в результат попадет второй знак «+». (Проверьте это самостоятельно.) Таким образом, для извлечения данных продаж на севере скопируйте формулу =ПСТР(А3;В3+1;С3-В3-1) из ячейки Е3 в ячейки Е4:Е6.

Для извлечения данных продаж на юге необходимо извлечь все символы справа от второго знака «+», применив функцию ПРАВСИМВ. Число таких символов равно (общее\_число\_символов\_в\_ячейке) – (позиция\_второго\_знака\_плюс). Общее число символов для каждой ячейки вычислите, скопировав формулу =ДЛСТР(А3) из ячейки F3 в ячейки F4:F6. Наконец, данные продаж на юге можно получить, скопировав формулу =ПРАВСИМВ(А3;F3-С3) из ячейки G3 в ячейки G4:G6.

## Извлечение данных с помощью Мастера распределения текста по столбцам

Есть простой способ извлечь данные по продажам на востоке, севере и юге (и другие аналогичные данные) без применения текстовых функций. Выделите ячейки А3:А6 и затем на вкладке Данные (Data) в группе Работа с данными (Data Tools) кликните на Текст по столбцам (Text To Columns). Выберите формат С разделителями (Delimited), нажмите Далее (Next) и заполните данные в диалоговом окне, как показано на рис. 6.6.

	A	B	C	D	E	F	G
1	<b>Extracting Sales in Three Regions</b>						
2	<b>East+North+South</b>	<b>First +</b>	<b>Second +</b>	<b>East</b>	<b>North</b>	<b>Total Length</b>	<b>South</b>
3	10+300+400	3	7 10	300		10	400
4	4+36.2+800	2	7 4	36.2		10	800
5	3+23+4005	2	5 3	23		9	4005
6	18+1+57.31	3	5 18	1		10	57.31

**Рис. 6.6.** Мастер распределения текста по столбцам

Если мы выберем знак «+» в качестве символа-разделителя, то распределим текст каждой ячейки по столбцам в соответствии с вхождениями знака «+». Разделителями также могут служить знак табуляции, точка с запятой, запятая и пробел. Теперь нажмите Далее (Next), выберите левый верхний угол целевого диапазона (в примере я выбрал ячейку А8) и нажмите Готово (Finish). Результат представлен на рис. 6.7.

	A	B	C
8	10	300	400
9	4	36.2	800
10	3	23	4005
11	18	1	57.31
12			
13	<b>Results of Data Text to Columns</b>		

**Рис. 6.7.** Результат работы Мастера распределения текста по столбцам



**?** В конце каждого семестра студенты оценивают мою преподавательскую деятельность по шкале от 1 до 7 баллов. Мне известно, сколько студентов выставили тот или иной балл. Как создать гистограмму оценки качества преподавания?

В файле `Repeatedhisto.xlsx` находятся оценки моей преподавательской деятельности (по шкале от 1 до 7 баллов). Два человека поставили 1 балл, три человека — 2 балла и т. д. График, обобщающий эти данные, можно построить с помощью функции **ПОВТОР**.

Скопируйте формулу `=ПОВТОР("|";C4)` из ячейки D4 в ячейки D5:D10. По этой формуле символы «|» извлекаются в столбец D в количестве, соответствующем значениям в столбце C. На рис. 6.8 ясно виден перевес хороших оценок (6 и 7 баллов) над плохими (1 и 2 балла). Повторение символа «|» позволяет имитировать гистограмму. Другие вопросы создания гистограмм в Excel обсуждаются в главе 43 «Обобщение данных на гистограммах и диаграммах Парето».

	B	C	D
3	Score	Frequency	
4	1	2	
5	2	3	
6	3	6	
7	4	7	
8	5	9	
9	6	33	
10	7	28	

**Рис. 6.8.** Создание графика с помощью функции **ПОВТОР**

**?** Числовые данные были загружены из интернета или из базы данных. При попытке использовать их в вычислениях появляется ошибка **#ЗНАЧ!**. Как решить эту проблему?

В файле `Cleanexample.xlsx` (рис. 6.9) непечатаемые **СИМВОЛ(10)** и **СИМВОЛ(160)** вставлены перед числом 33 в ячейках E5 и H6 соответственно. В ячейках E8 и H8 мы применили функцию **ЗНАЧЕН**, попытавшись преобразовать содержимое ячеек E5 и H6 в числа, но получили ошибку **#ЗНАЧ!**, которая говорит нам о том, что Excel не понимает, что должен обрабатывать содержимое ячеек как числа. В ячейке E11 находится содержимое ячейки E5, очищенное по формуле `=ПЕЧСИМВ(E5)`. В ячейке E12 находится числовое значение 33, то есть после того как **СИМВОЛ(10)** был удален, содержимое ячейки было обработано по формуле `=ЗНАЧЕН(E11)` без ошибок.

В ячейке H10 мы попытались очистить содержимое ячейки H6 с помощью формулы `=ПЕЧСИМВ(H6)`. В ячейке H11 после применения формулы `=ЗНАЧЕН(H10)` получена ошибка, так как функция **ПЕЧСИМВ** не удалила **СИМВОЛ(160)**. Для устранения ошибки необходимо **СИМВОЛ(160)** заменить пробелом с помощью функции

**ПОДСТАВИТЬ.** В ячейку H14 запишите формулу `=НАЙТИ(СИМВОЛ(160);H10)` и убедитесь, что символ `СИМВОЛ(160)` присутствует в ячейке H10: он стоит на первом месте в ячейке. Если применение функции `НАЙТИ` вызовет ошибку, значит, `СИМВОЛ(160)` в ячейке отсутствует. В ячейке H15 все вхождения этого непечатаемого символа заменяются пробелом с помощью формулы `=ПОДСТАВИТЬ(H6;СИМВОЛ(160);" ")`. В ячейке H16 формула `=ЗНАЧЕН(H15)` дает в результате число 33. Таким образом, теперь над содержимым ячейки H15 можно производить математические операции.

	E	F	G	H	I	J
1						
2						
3						
4	char 10					
5	33			char 160		
6				33		
7						
8	#VALUE!			#VALUE!		
9				CLEAN?		
10	cleaned			33		
11	33			#VALUE!		
12	33			CLEAN and VALUE DO NOT WORK		
13				SUBSTITUTE		
14			FIND CHAR 160	1		
15			SUBSTITUTE	33		
16			VALUE	33		
17				160		

**Рис. 6.9.** Удаление непечатаемых символов с помощью функций `ПЕЧСИМВ` и `ПОДСТАВИТЬ`

**?** **Текстовые функции — вещь хорошая, но есть ли простой способ (без них) извлечь имена и фамилии из данных, создать список адресов электронной почты из списка имен или выполнить другие стандартные операции с текстовыми данными?**

Инструмент **Мгновенное заполнение** (Flash Fill; впервые появился в Excel 2013) использует современные технологии распознавания образов, позволяя моделировать множество задач, ранее выполнявшихся с помощью текстовых функций. Файл `Flashfill.xlsx` демонстрирует применение инструмента **Мгновенное заполнение** в следующих примерах:

- извлечение имен и фамилий (на листе `First And Last`);
- создание адресов электронной почты для университета UXYZ University путем добавления `@UXYZ.edu` к фамилии сотрудника (на листе `Email`);
- извлечение сумм в долларах и центах из списка цен (на листе `Dollars And Cents`).

Для использования инструмента **Мгновенное заполнение** выделите столбец рядом с данными. В первой ячейке, соответствующей имеющимся данным, введите пример того, чем требуется заполнить остальные ячейки. Как правило, если нажать **Enter** и затем **Ctrl+E**, инструмент **Мгновенное заполнение** автоматически правильно заполнит оставшиеся ячейки. Однако иногда для получения безошибочного результата приходится ввести несколько примеров.

Как показано на рис. 6.10, мы собираемся извлечь имена и фамилии сотрудников в столбцы **E** и **F**. Напечатайте **Tricia** в ячейке **E6** листа **First and Last** и нажмите **Enter**. Затем нажмите **Ctrl+E**, и ячейки **E7:E13** автоматически заполнятся именами сотрудников. Аналогично, если напечатать **Lopez** в ячейке **F6**, нажать **Enter** и затем **Ctrl+E**, инструмент **Мгновенное заполнение** запишет фамилии сотрудников в ячейки **F7:F13**.

	D	E	F
5	Full	First	Last
6	Tricia Lopez	Tricia	Lopez
7	Will Wong	Will	Wong
8	Jack Spratt	Jack	Spratt
9	Vivian Hibbits	Vivian	Hibbits
10	Jose Gomez	Jose	Gomez
11	April Chou	April	Chou
12	Tanya Walters	Tanya	Walters
13	James Jones	James	Jones

**Рис. 6.10.** Автоматический ввод имен и фамилий

Как показано на рис. 6.11, мы собираемся создать email-адрес для каждого сотрудника, добавив к его фамилии **@XYZ.edu**. Напечатайте **Lopez@XYZ.edu** в ячейке **E6** листа **email**. Далее нажмите **Enter**, а затем **Ctrl+E**, и инструмент **Мгновенное заполнение** автоматически создаст email-адреса в диапазоне **E7:E13**.

Как показано на рис. 6.12, мы собираемся извлечь сумму в долларах и центах из ячеек **D6:D11** столбца с ценами на листе **Dollars and Cents**. Для начала введите **6** в ячейку **E6** и нажмите **Enter**. После того как вы нажмете **Ctrl+E**, верное количество долларов будет вставлено в ячейки **E7:E11**. Аналогично после ввода **56** в ячейку **F6** и нажатия **Enter** и затем **Ctrl+E** инструмент **Мгновенное заполнение** заполнит ячейки **F7:F11** верным количеством центов.

Инструмент **Мгновенное заполнение** может работать с ошибками, особенно на слабо согласованных данных. Кроме того (в отличие от формул, в том числе с текстовыми функциями), результаты работы инструмента **Мгновенное заполнение** не являются динамическими и не обновляются при изменении исходных данных.

Инструмент **Мгновенное заполнение** можно отключить: на вкладке **Файл (File)** выберите **Параметры (Options)**. Затем выберите раздел **Дополнительно (Advanced)**

и в диалоговом окне **Дополнительные параметры для работы с Excel** в разделе **Параметры правки (Editing Options)** снимите флажок **Автоматически выполнять мгновенное заполнение (Automatically Flash Fill)**.

	D	E
5	<b>Name</b>	<b>Email</b>
6	Tricia Lopez	Lopez@UXYZ.edu
7	Will Wong	Wong@UXYZ.edu
8	Jack Spratt	Spratt@UXYZ.edu
9	Vivian Hibbits	Hibbits@UXYZ.edu
10	Jose Gomez	Gomez@UXYZ.edu
11	April Chou	Chou@UXYZ.edu
12	Tanya Walters	Walters@UXYZ.edu
13	James Jones	Jones@UXYZ.edu

**Рис. 6.11.** Создание email-адресов с помощью инструмента **Мгновенное заполнение**

	D	E	F
5	<b>Price</b>	<b>Dollars</b>	<b>Cents</b>
6	\$6.56	6	56
7	\$7.43	7	43
8	\$9.86	9	86
9	\$15.43	15	43
10	\$173.32	173	32
11	\$4.21	4	21

**Рис. 6.12.** Извлечение суммы в долларах и центах с помощью инструмента **Мгновенное заполнение**

## ? Что такое символы Unicode?

Набор символов Unicode содержит около 120 000 символов, включая множество символов, используемых в научной литературе, и символы из других языков.

	F	G	H	I	J
56	945 α				
57	946 β				
58	947 γ				
59	948 δ				
60	949 ε				
61	950 ζ				
62	951 η				
63	952 θ				
64	953 ι				
65	954 κ				
66	955 λ				
67	956 μ	=UNICHAR(F67)			
68	957 ν				
69	958 ξ				
70	959 ο				
71	960 π				
72	961 ρ				
73	962 ς				
74	963 σ				
75	964 τ				
76	965 υ				

**Рис. 6.13.** Примеры использования функций **UNICODE** и **UNICH**

Каждому символу Unicode соответствует числовой код. (Подробное описание Unicode вы можете найти по адресу <http://www.alanwood.net/unicode/>.)

Например, символы греческого алфавита (важные для носителей греческого языка и для ученых!) имеют коды в районе 900 (см. файл `Unicodefinal.xlsx` и рис. 6.13). Найти символ, соответствующий кодовому значению, можно с помощью функции `ЮНИСИМВ(кодовое_значение)` (`UNICHAR(code number)`). Например, формула `=ЮНИСИМВ(F67)` вернет греческий символ  $\mu$ , так как кодовое значение для  $\mu$  — 956. Узнать кодовое значение символа можно, используя функцию `UNICODE(символ)`. Так, формула `=UNICODE(G67)` вернет кодовое значение для  $\mu$  (956).

### ❓ Чем новая функция **ОБЪЕДИНИТЬ** лучше старой функции **СЦЕПИТЬ** (или `&`)?

В файле `Textjoinfinal.xlsx` (рис. 6.14) продемонстрировано использование новой функции **ОБЪЕДИНИТЬ** (`TEXTJOIN`). Нам нужно поместить в одну ячейку в виде скомбинированные данные, разделенные пробелами, из диапазона ячеек `H2:L2`. В ячейке `M2` это проделано с помощью устаревшей функции `Concatenate`, по формуле `=H2&" "&I2&" "&K2&" "&L2`. В ячейке `G2` применена формула `=TEXTJOIN(" ";ИСТИНА;H2:L2)` с тем же результатом. Первый аргумент функции `" "` гарантирует разделение данных из ячеек `H2:L2` пробелами. Второй аргумент, `ИСТИНА`, обеспечивает игнорирование пустых ячеек (в нашем случае это `J2`). В ячейку `G3` мы ввели формулу `=TEXTJOIN(" ";ЛОЖЬ;H3:L3)`. Результат отличается только включением пробела из пустой ячейки `J3` в нашу цепочку. Преимущество функции **ОБЪЕДИНИТЬ** опять проявляется и в отсутствии необходимости повторять символ-разделитель.

	F	G	H	I	J	K	L	M	N	O	P
1											
2	<code>=TEXTJOIN(" ",TRUE,H2:L2)</code>	Taylor Katy John Adele	Taylor	Katy		John	Adele	Taylor Katy John Adele	<code>=H2&amp;" "&amp;I2&amp;" "&amp;K2&amp;" "&amp;L2</code>		
3	<code>=TEXTJOIN(" ",FALSE,H3:L3)</code>	Taylor Katy John Adele	Taylor	Katy		John	Adele				
4											
5											
6	<code>=LEN(G2)</code>		22								
7	<code>=LEN(G3)</code>		23								

**Рис. 6.14.** Пример использования функции **ОБЪЕДИНИТЬ**

### ❓ Мне нравятся текстовые функции, а как пользоваться функцией Excel **ТЕКСТ**?

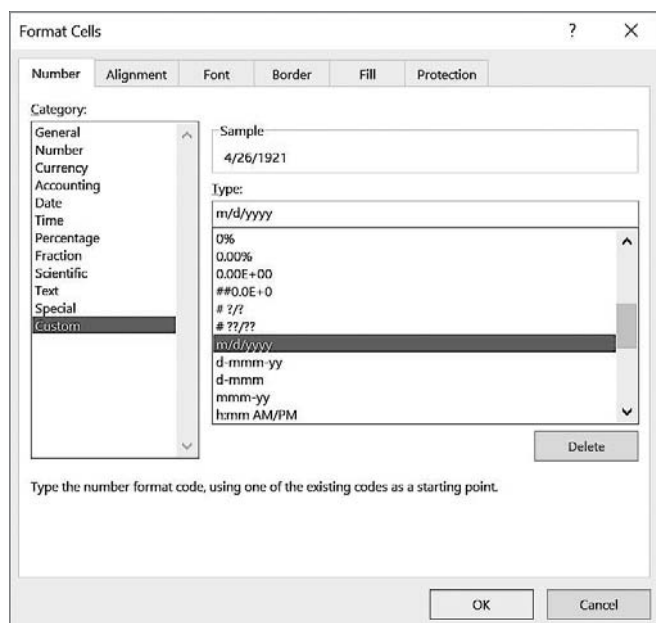
Допустим, вы работаете в отделе кадров, и вам принесли имена и даты рождения сотрудников (рис. 6.15). По каждому сотруднику вы хотите записать фразу типа «Джен родилась 14 октября 1988 г.». В ячейке `F14` мы попробовали сделать это с помощью формулы `D3&" was born on "&E3`. А программа нам отвечает: «Джен родилась 32430-го». Это число дней, минувших с 1 января 1900 г. для даты 14 октября 1988 г. То есть проблема в том, что при конкатенации Excel выбирает неверный формат. Вот когда нас выручит функция **ТЕКСТ**. Если вы знаете правильный код форматирования (а именно `мм/дд/гггг`), вы просто изменяете формулу следу-

ющим образом: D3&" was born on " & ТЕКСТ(Е3,"мм/дд/гггг"), и — вуаля! Формат правильный.

	C	D	E	F	G
1					
2				Wrong	Right
3		Jen	10/14/1988	Jen was born on 32430	Jen was born on 10/14/1988
4		Greg	8/26/1992	Greg was born on 33842	Greg was born on 8/26/1992
5		Wanda	4/26/1921	Wanda was born on 7787	Wanda was born on 4/26/1921
6					
7	Date	Month abbreviation	Month	Weekday Abbreviated	Weekday
8	10/14/1988	Oct	October	Fri	Friday
9	8/26/1992	Aug	August	Wed	Wednesday
10	4/26/1921	Apr	April	Tue	Tuesday

**Рис. 6.15.** Пример использования функции ТЕКСТ

А как найти правильный формат? Просто перейдите в ячейку, где нужен правильный формат (например, E3), и нажмите CTRL+F1, чтобы открыть диалоговое окно Формат ячеек (Format Cells). Выберите вкладку Число (Number) и затем Все форматы (Custom). Вы увидите экран со всякими сложными значками, показанный на рис. 6.16. Просто скопируйте нужный код из поля Тип (Type) (в данном случае, мм/дд/гггг) и вставьте этот код, заключенный в кавычки, как аргумент функции ТЕКСТ.



**Рис. 6.16.** Выбрав Все форматы, можно найти формат любого типа

Как показано в диапазоне ячеек **D8:G10**, с помощью правильного формата вы можете извлечь название месяца или дня недели из даты.

"ммм" возвращает сокращенное название месяца.

"мммм" возвращает полное название месяца.

"ddd" возвращает сокращенное название дня недели.

"dddd" возвращает полное название дня недели.

## Задания

1. В ячейках **B3:B6** файла **Showbiz.xlsx** содержатся фиктивные адреса некоторых знаменитостей. С помощью текстовых функций извлеките имена в один столбец, а адреса — в другой.
2. В файле **IDprice.xlsx** находятся идентификаторы продуктов вместе с ценами. Используя текстовые функции, извлеките идентификаторы и цены в разные столбцы. Затем примените **Распределение текста по столбцам (Convert Text To Columns)** на вкладке **Данные (Data)** для достижения той же цели.
3. В файле **Quarterlygnpdata.xlsx** хранятся квартальные данные по ВВП США (в млрд долларов). Распределите эти данные по трем столбцам следующим образом: в первый столбец извлеките год, во второй — номер квартала и в третий — ВВП.
4. В файле **Textstylesdata.xlsx** хранится информация о фасоне, цвете и размере различных рубашек. Например, первая рубашка имеет фасон 100 (указанный цифрами между двоеточием и дефисом), цвет 65 и размер L. Используя текстовые функции, извлеките в разные столбцы фасон, цвет и размер для каждой рубашки.
5. В файле **Emailproblem.xlsx** даны имена и фамилии новых сотрудников Microsoft. Для создания email-адреса каждому сотруднику возьмите первую букву его имени, фамилию сотрудника и в конце добавьте **@microsoft.com**. Используйте текстовые функции для эффективного создания адресов.
6. В файле **Lineupdata.xlsx** хранится число минут, сыгранных составами из пяти игроков. (Например, первый состав сыграл 10,4 минуты.) С помощью текстовых функций преобразуйте эти данные в форму, пригодную для вычислений; например, преобразуйте **10.4m** в число 10,4.
7. В файле **Reversenames.xlsx** содержатся имена, вторые имена или инициалы и фамилии нескольких людей. Преобразуйте эти имена следующим образом: сначала фамилия, за ней запятая, затем имя и второе имя. Например, преобразуйте **Gregory William Winston** в **Winston, Gregory William**.
8. В файле **Incomefrequency.xlsx** находится распределение стартовых зарплат выпускников MBA колледжа **Faber College**. Обобщите эти данные с помощью частотного графа.

9. Вспомните, что СИМВОЛ(65) дает в результате букву А, СИМВОЛ(66) — букву В и т. д. Воспользовавшись этими знаниями, создайте книгу, заполнив ячейки B1:B26 последовательностью букв А, В, С и так далее до Z.
10. В файле Capitalizefirstletter.xlsx содержатся названия различных песен и фразы, например, The rain in Spain falls mainly in the plain. Убедитесь, что каждое название песни начинается с прописной буквы.
11. В файле Ageofmachine.xlsx хранятся данные в следующем виде:

S/N: 160768, vib roller, 84" smooth drum, canopy Auction: 6/2–4/2005 in Montgomery, Alabama

Это записи о покупках автомобиля. Определите год каждой покупки.

12. При скачивании корпоративных данных с сайта Security and Exchange Commission's EDGAR полученные данные часто выглядят следующим образом:

денежные средства и их эквиваленты: \$31 848, \$31 881.

Как эффективно извлечь денежные средства и их эквиваленты для каждой компании?

13. В файле Lookuptwocolumns.xlsx хранятся данные о модели, годе выпуска и цене для ряда автомобилей. Напишите формулу, позволяющую по введенной модели и году выпуска автомобиля получить его цену.
14. В файле Moviedata.xlsx хранятся названия фильмов, за которыми следует количество DVD-копий, проданных в местном видеомагазине. Извлеките из этих данных названия всех фильмов.
15. В файле Moviedata.xlsx хранятся названия фильмов, за которыми следует количество DVD-копий, проданных в местном видеомагазине. Для каждого фильма извлеките из данных количество проданных копий. Подсказка: возможно, потребуется функция ПОДСТАВИТЬ. Ее синтаксис: ПОДСТАВИТЬ(ячейка;старый\_текст;новый\_текст;[номер\_вхождения]).

Если номер вхождения опущен, то каждое вхождение старого текста в ячейке заменяется новым текстом. Если номер вхождения указан, то только это вхождение старого текста заменяется новым текстом. Например, функция ПОДСТАВИТЬ(A4;1;2) заменит каждую 1 в ячейке A4 на 2, а функция ПОДСТАВИТЬ(A4;1;2;3) заменит только третье вхождение 1 в ячейке A4 на 2.

16. В файле Problem16data.xlsx хранится информация о людях, заполнивших рекламный опросник с пятью вариантами ответов (1 = скорее всего, не куплю, ..., 5 = скорее всего, куплю). Обобщите эти данные графически с помощью символа «\*». Для более наглядного представления на вкладке Главная (Home) откройте в группе Выравнивание (Alignment) раскрывающийся список Ориентация (Orientation) и сделайте текст вертикальным. Затем щелкните правой кнопкой мыши по номеру строки и увеличьте высоту строки. Наконец, в группе Выравнивание (Alignment) выберите Перенести текст (Wrap Text), чтобы сделать текст вашего графика вертикальным.



17. В файле **Problem17data.xlsx** содержатся имена (в виде Mr. John Doe). Используя текстовые функции, извлеките обращение (Mr . или Mrs.) и первое имя каждого человека в отдельные столбцы.
18. В файле **Weirddata.xlsx** находятся три числа, импортированные с веб-сайта. Сложив числа, убедитесь в некорректности результата. Измените данные таким образом, чтобы функция **СУММ** дала правильный результат. Подсказка: воспользуйтесь функцией **НАЙТИ** для поиска невидимого символа!
19. Используя **Мгновенное заполнение**, запишите имена в файле **Flashfilltemplate.xlsx** (в папке **Templates**) строчными буквами.
20. Используя **Мгновенное заполнение**, извлеките число из каждой строки в файле **Movienumbers.xlsx**.
21. Для данных в файле **Movienumbers.xlsx** создайте столбец и с помощью инструмента **Мгновенное заполнение** поместите в него фразу: **The number in the title of this movie is...** и число для каждого названия фильма.
22. В каждой строке файла **Problem22data.xlsx** указано название города, штата и население города. Используя текстовые функции, извлеките название штата в отдельный столбец.
23. В файле **Problem23data.xlsx** строка представляет имя квотербека (QB), его рейтинг и количество результативных передач. Используя текстовые функции, извлеките для каждого игрока количество результативных передач.
24. В каждой строке файла **Problem24data.xlsx** содержится индекс города, его название, штат (везде Texas!), широта и долгота. С помощью формул извлеките из всех строк название города. Можете оттолкнуться от того факта, что индекс всегда состоит из пяти символов. Подсказка: функция **FIND** чувствительна к регистру.
25. В каждой строке файла **Problem25data.xlsx** между дефисом и запятой указано число забитых квотербеком мячей. Примените текстовые функции для извлечения числа забитых мячей для каждого квотербека.
26. В файле **Problem26data.xlsx** содержатся сведения о 2000 крупнейших компаний мира. Если в ячейку **B3** вводится целое число от 1 до 2000, то ячейка **C3** должна вернуть предложение типа «Доход компании, занимающей место № 1 по доходу, равен \$328.21». Примените функцию **ТЕКСТ**, чтобы обеспечить правильный формат суммы в долларах.

## ГЛАВА 7

# Даты и функции работы с датами

### Обсуждаемые вопросы

- Когда я ввожу даты в Excel, то часто получаю число, например 37625, а не дату вида 04.01.2003. Что делать?
- Как с помощью формулы автоматически отобразить сегодняшнюю дату?
- Как определить дату через 50 рабочих дней после другой даты? Как исключить праздничные дни?
- Как определить количество рабочих дней между двумя датами?
- На листе Excel введено 500 дат. Какие формулы нужны, чтобы извлечь из каждой даты месяц, год, день месяца и день недели?
- У меня есть значения года, месяца и дня. Существует ли простой способ восстановления фактической даты?
- У меня бизнес по купле-продаже автомобилей. Для некоторых автомобилей известны даты покупки и продажи. Каким образом можно определить, сколько месяцев эти автомобили пробыли на стоянке?
- Как поместить в лист статичную (неизменяемую) дату?

Для иллюстрации наиболее часто использующихся форматов в Microsoft Excel 2019 предположим, что сегодня 4 января 2004 г. Эту дату можно ввести, например, следующим образом:

- 4/1/2004<sup>1</sup>;
- 4-январь-2004;
- 4 январь 2004;
- 4/1/04;
- 04.01.2004.

Если год указан двумя последними цифрами, и это цифры 30 и больше, то Excel предполагает, что это год XX века; если цифры меньше 30, считается, что они относятся к XXI веку. Например, 1/1/29 будет означать 1 января 2029 г., а 1/1/30 будет обработано как 1 января 1930 г. Год, определяющий обработку даты как даты в XXI веке, увеличивается на единицу каждый год.

---

<sup>1</sup> В русифицированной версии Microsoft Excel 2019 используется формат день-месяц-год. — *Примеч. пер.*

## Ответы на вопросы

- ❓ Когда я ввожу даты в Excel, то часто получаю число, например 37625, а не дату вида 04.01.2003. Что делать?

То, как Excel обрабатывает даты, иногда сбивает с толку новичков. Важно понимать, что в Excel есть два формата отображения дат: день-месяц-год и числовой формат. Дата в числовом формате, например 37625, — это целое положительное число, равное количеству дней между заданной датой и 1 января 1900 г. При подсчете учитывается как текущая дата, так и 1 января 1900 г. Например, в Excel дата 3 января 1900 г. в формате числа отображается как число 3. Таким образом, считается, что промежуток между 1 января 1900 г. и 3 января 1900 г. (включая оба этих дня) составляет три дня.

### ПРИМЕЧАНИЕ

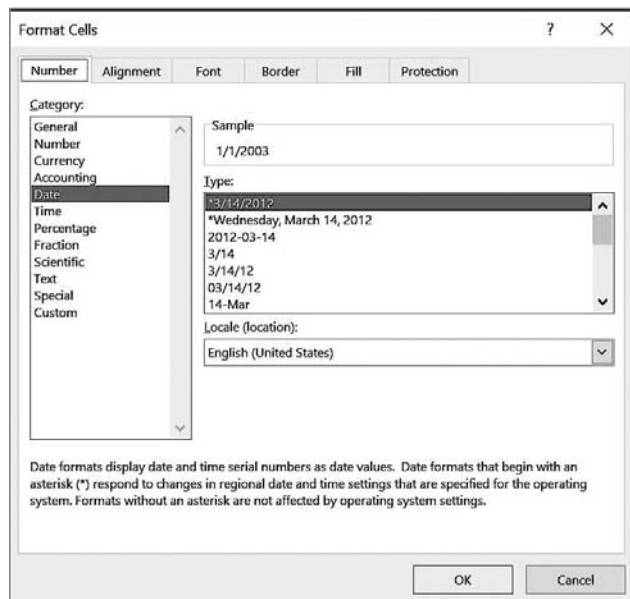
Excel считает, что 1900 г. был високосным, то есть содержал 366 дней. В действительности в нем было 365 дней. Занимательную информацию о происхождении этой ошибки можно прочитать здесь: [www.joelonsoftware.com/items/2006/06/16.html](http://www.joelonsoftware.com/items/2006/06/16.html).

На рис. 7.1 показан лист **serial format** в файле **Dates.xlsx**. Допустим, даты в ячейках D5:D14 даны в числовом формате. Например, значение 37622 в ячейке D5 указывает на дату, приходящуюся на 37 622-й день после 1 января 1900 г. (включая 1 января 1900 г. и сам 37 622-й день). Скопируйте даты в числовом формате в ячейки E5:E14 для отображения в формате день-месяц-год. Выделите диапазон ячеек E5:E14, щелкните правой кнопкой мыши по выделенной области и выберите в контекстном меню **Формат ячеек...** (**Format Cells**). Кстати, диалоговое окно **Формат ячеек** (**Format Cells**) можно открыть в любое время, нажав **Ctrl+1**. Из списка, показанного на рис. 7.2, выберите формат **Дата** (**Date**) и **Тип** (**Type**). Даты в ячейках E5:E14 отобразятся в формате день-месяц-год (рис. 7.1). Если нужно перевести даты в числовой формат, выделите ячейки E5:E14, щелкните правой кнопкой мыши по выделенной области и выберите **Формат ячеек... | Общий** (**Format Cells | General**).

	D	E
4	Dates	Reformatted
5	37622	1/1/2003
6	37623	1/2/2003
7	37624	1/3/2003
8	37625	1/4/2003
9	37626	1/5/2003
10	37627	1/6/2003
11	37628	1/7/2003
12	37629	1/8/2003
13	37630	1/9/2003
14	37631	1/10/2003

**Рис. 7.1.** Преобразование дат из формата десятичного числа в формат день-месяц-год с помощью команды **Формат ячеек**

Простая замена формата даты в ячейке на общий формат приведет к отображению даты в числовом формате. Можно воспользоваться другим способом и применить функцию **ДАТАЗНАЧ** (**DATEVALUE**) с датой, заключенной в кавычки. Например, в файле **Dates.xlsx** на листе **date format** в ячейке **I5** находится формула **=ДАТАЗНАЧ("4/1/2003")**. В результате ее вычисления получается **37625**, что соответствует дате в числовом формате для 4 января 2003 г.



**Рис. 7.2.** Переформатирование даты в формате десятичного числа в формат день-месяц-год в диалоговом окне Формат ячеек

### ❓ Как с помощью формулы автоматически отобразить сегодняшнюю дату?

Пример отображения сегодняшней даты с помощью функции **СЕГОДНЯ** (**TODAY**) приведен в ячейке **C13** на листе **date format** (рис. 7.3). Введите в ячейку формулу **=СЕГОДНЯ()**. Этот скриншот был сделан 15 июня 2016 г.

### ❓ Как определить дату через 50 рабочих дней после другой даты? Как исключить праздничные дни?

Для этого примените функцию **РАБДЕНЬ** (**WORKDAY**). Синтаксис этой функции: **РАБДЕНЬ(нач\_дата;#число\_дней;[праздники])**. Функция **РАБДЕНЬ** отображает дату через столько рабочих дней (рабочий день — это не выходной день) после заданной начальной даты, сколько указано значением аргумента **#число\_дней**. Аргумент **праздники** является необязательным и позволяет исключить из расчета любые даты, указанные в диапазоне ячеек. Таким образом, по формуле **=РАБДЕНЬ(C14;50)** в ячейке **D14** листа **date format** пятидесятым рабочим днем по-

сле 03.01.2003 является 14 марта 2003 г. . Если необходимо учесть праздники День Мартина Лютера Кинга и День независимости, измените формулу на следующую: =РАБДЕНЬ(С14;50;F17:F18) (она приведена в ячейке E14). В этом случае 20.01.2003 не будет учтено, и пятидесятым рабочим днем после 03.01.2003 станет 17 марта 2003 г. Обратите внимание, что вместо ссылки на даты праздников в других ячейках можно ввести даты в формулу , указав каждый праздник в числовом формате, заключенном в фигурные скобки ( { } ). Например, в формуле =РАБДЕНЬ(38500;10;{38600;38680;38711}) для подсчета десятого рабочего дня после даты с номером 38 500 исключены праздники День труда, День благодарения и Рождество в 2005 г.

	B	C	D	E	F	G	H	I
4			Year	Month	Day	Day of Week	Putting date together	Serial Number
5	1/4/2003	1/4/2003	2003	1	4	7	1/4/2003	37625
6	2/1/1901	2/1/1901	1901	2	1	6	2/1/1901	398
7	4-Jan-2003	4-Jan-03	2003	1	4	7	1/4/2003	
8	January 4, 2003	4-Jan-03	2003	1	4	7	1/4/2003	
9	1/4/03	4-Jan-03	2003	1	4	7	1/4/2003	
10		3-Jan-01	1900	1	0	7	1/0/1900	3
11								
12								
			50 workdays from start date	excluding holidays				
13	Today's date	6/11/2018						
14	Start date	1/3/2003	3/14/2003	3/17/2003				
15	Later date	8/4/2003						
16					Holidays			
17	Workdays between (excluding holidays)	150			7/4/2003			
18	Workdays between (no holidays)	152			1/20/2003			

**Рис. 7.3.** Примеры использования функций обработки дат

В Microsoft Excel 2010 появилась функция РАБДЕНЬ.МЕЖД (WORKDAY.INTL), позволяющая выбрать собственное определение рабочего дня. Синтаксис этой функции: РАБДЕНЬ.МЕЖД(нач\_дата;число\_дней;выходные;[праздники]). Для определения выходного дня используется третий аргумент. Коды, приведенные в таблице ниже, определяют как выходные следующие дни:

Код	Выходные дни	Код	Выходные дни
1 или аргумент опущен	суббота, воскресенье	11	только воскресенье
2	воскресенье, понедельник	12	только понедельник
3	понедельник, вторник	13	только вторник
4	вторник, среда	14	только среда
5	среда, четверг	15	только четверг
6	четверг, пятница	16	только пятница
7	пятница, суббота	17	только суббота

Например, в файле `Dates.xlsx` на листе `date format` в ячейке `C27` по формуле `=РАБДЕНЬ.МЕЖД(C23;100;2)` вычислена дата сотого рабочего дня после 14.03.2011 с воскресеньем и понедельником в качестве выходных дней (рис. 7.4). В ячейке `C28` по формуле `=РАБДЕНЬ.МЕЖД(C23;100;11)` вычислена дата сотого рабочего дня после 14.03.2011 для случая, когда выходным является только воскресенье. Кроме того, выходные дни можно определить с помощью строки из семи единиц и нулей; единица означает выходной день, и первая позиция в строке соответствует понедельнику, вторая — вторнику и т. д. Таким образом, в ячейках `D26` и `D27` предыдущие результаты продублированы с помощью формул `= РАБДЕНЬ.МЕЖД(C23;100;"1000001")` и `=РАБДЕНЬ.МЕЖД(C23;100;"0000001")`, соответственно.

	B	C	D
21	<b>International Functions</b>		
22			
23	<b>Start Date</b>	<b>3/14/2011</b>	
24	<b>End Date</b>	<b>8/16/2012</b>	
25			
26	<b>100 workdays later</b>		
27	<b>Sunday Monday as weekend</b>	<b>7/30/2011</b>	<b>7/30/2011</b>
28	<b>Sunday as weekend</b>	<b>7/8/2011</b>	<b>7/8/2011</b>
29			
30	<b>Days Between Start and End Date</b>		
31	<b>Sunday Monday as weekend</b>	<b>373</b>	<b>373</b>
32	<b>Sunday as weekend</b>	<b>448</b>	<b>448</b>

**Рис. 7.4.** Примеры использования функций для международных дат

### ❓ Как определить количество рабочих дней между двумя датами?

Нам потребуется функция **ЧИСТРАБДНИ** (**NETWORKDAYS**). Синтаксис этой функции: **ЧИСТРАБДНИ**(нач\_дата;кон\_дата;[праздники]), где **праздники** (**holidays**) — необязательный аргумент, указывающий диапазон ячеек с датами, считающимися праздниками. Функция **ЧИСТРАБДНИ** возвращает количество рабочих дней между начальной и конечной датами, исключая выходные дни и все указанные праздники. В качестве иллюстрации работы функции **ЧИСТРАБДНИ** рассмотрим ячейку `C18` на листе `date format` файла `Dates.xlsx` с формулой `=ЧИСТРАБДНИ(C14;C15)`. По этой формуле рассчитывается количество рабочих дней между 03.01.2003 и 04.08.2003, которое составляет 152 дня. В ячейке `C17` по формуле `=ЧИСТРАБДНИ(C14;C15;F17:F18)` вычисляется количество рабочих дней между 03.01.2003 и 04.08.2003, исключая День Мартина Лютера Кинга и День независимости. В результате получается  $152 - 2 = 150$ .

В Microsoft Excel 2010 появилась функция **ЧИСТРАБДНИ.МЕЖД** (**NETWORKDAYS.INTL**). Аналогично функции **РАБДЕНЬ.МЕЖД**, функция **ЧИСТРАБДНИ.МЕЖД** позволяет задать собственное определение выходных дней. Например, в файле `Dates.xlsx` на листе `date format` в ячейке `C31` по формуле `=ЧИСТРАБДНИ.МЕЖД(C23;C24;2)` вычисляется количество рабочих дней ( 373 ) между 14.03.2011 и 16.08.2012

для случая, когда выходными днями являются воскресенье и понедельник (рис. 7.4). В ячейке D31 те же вычисления выполнены по формуле `=ЧИСТРАБДНИ.МЕЖД(C23;C24;"1000001")`.

**❓ На листе Excel введено 500 дат. Какие формулы нужны, чтобы извлечь из каждой даты месяц, год, день месяца и день недели?**

В файле `Dates.xlsx` на листе `date format` (см. рис. 7.3) в диапазоне ячеек B5:B10 указано несколько дат. В ячейке B5 и ячейках B7:B9 я использовал четыре формата для отображения даты 4 января 2003 г. В столбцах D:G я поместил извлеченные год, месяц, день месяца и день недели для каждой даты. Год из каждой даты я извлек, применив функцию `ГОД (YEAR)` — я скопировал формулу `=ГОД(B5)` из ячейки D5 в ячейки D6:D10. Номер месяца (1 — январь, 2 — февраль и т. д.) из каждой даты я извлек с помощью функции `МЕСЯЦ (MONTH)` путем копирования формулы `=МЕСЯЦ(B5)` из ячейки E5 в ячейки E6:E10. День месяца был извлечен с помощью функции `ДЕНЬ (DAY)` путем копирования формулы `=ДЕНЬ(B5)` из ячейки F5 в ячейки F6:F10. Наконец, день недели для каждой даты был извлечен с помощью функции `ДЕНЬНЕД (WEEKDAY)` путем копирования формулы `=ДЕНЬНЕД(B5;1)` из ячейки G5 в ячейки G6:G10. Обратите внимание: региональные настройки (не для США) могут потребовать внесения в пример некоторых изменений.

Если значение последнего аргумента функции `ДЕНЬНЕД` равно 1, то 1 — это воскресенье, 2 — понедельник и т. д. Если это значение равно 2, то 1 — это понедельник, 2 — вторник и т. д. Если значение равно 3, 0 — это понедельник, 1 — вторник и т. д.

**❓ У меня есть значения года, месяца и дня. Существует ли простой способ восстановления фактической даты?**

Функция `ДАТА (DATE)` с аргументами `ДАТА(год;месяц;день)` возвращает дату для указанного года, месяца и дня. На листе `date format` я восстановил исходные даты, скопировав формулу `=ДАТА(D5;E5;F5)` из ячейки H5 в ячейки H6:H10.

**❓ У меня бизнес по купле-продаже автомобилей. Для некоторых автомобилей известны даты покупки и продажи. Каким образом можно определить, сколько месяцев эти автомобили пробыли на стоянке?**

Количество полных лет, месяцев или дней между двумя датами можно определить с помощью функции `РАЗНДАТ (DATEDIF)`. В файле `Datedif.xlsx` (рис. 7.5) приведен пример покупки автомобиля 15 октября 2016 г и его последующей продажи 10 апреля 2018 г. Сколько полных лет, месяцев или дней стоял на парковке этот автомобиль? Синтаксис функции: `РАЗНДАТ(нач_дата,кон_дата,единицы_времени)`. Если для аргумента `единицы_времени` указать значение "y", функция возвратит количество полных лет между начальной и конечной датами. Если указать значение "m", функция возвратит количество полных месяцев, а если указать значение "d", то будет получено количество полных дней между двумя указанными датами. Таким образом, ввод формулы `=РАЗНДАТ(D4;D5;"y")` в ячейку D6 показывает, что машина хранилась один полный год. Ввод формулы `=РАЗНДАТ(D4;D5;"m")`

в ячейку D7 дает результат: 17 полных месяцев. Введенная в ячейку D8 формула `=РАЗНДАТ(D4;D5;"d")` возвращает число 543 — число полных дней. Кстати, функция `РАЗНДАТ` отсутствует в списке мастера функций.

	C	D	E
4	<b>bought</b>	<b>10/15/2016</b>	
5	<b>sold</b>	<b>4/10/2018</b>	
6	<b>years</b>	<b>1</b>	<code>=DATEDIF(D4,D5,"y")</code>
7	<b>months</b>	<b>17</b>	<code>=DATEDIF(D4,D5,"m")</code>
8	<b>days</b>	<b>542</b>	<code>=DATEDIF(D4,D5,"d")</code>

**Рис. 7.5.** Результат применения функции `РАЗНДАТ`

### ? Как поместить в лист статичную (неизменяемую) дату?

Ввод в ячейке функции `=СЕГОДНЯ()` всегда возвращает сегодняшнюю дату. При создании книги нам может потребоваться, чтобы в книге всегда отображалась дата ее создания. Комбинация клавиш `Ctrl+;` (точка с запятой) поместит сегодняшнюю дату в ячейку, но в отличие от функции `=СЕГОДНЯ()`, эта дата больше меняться не будет. Книга, представленная на рис. 7.6 (см. файл `Staticdate.xls`), была создана 11 июня 2018 г., и эта дата всегда будет отображаться в ячейке F6.

	B	C	D	E
1				
2				
3	<b>6/11/2018</b>			
4		<b>With cell B3 selected</b>		
5		<b>Control+;</b>		
6		<b>placed a static date</b>		
7		<b>(the day we used Control+;)</b>		
8		<b>in cell B3</b>		

**Рис. 7.6.** Создание статичной даты с помощью комбинации `Ctrl+;`

## Задания

1. Как выглядит дата 25 января 2006 г. в числовом формате?
2. Как выглядит дата 14 февраля 1950 г. в числовом формате?
3. Какой фактической дате соответствует дата 4526?
4. Какой фактической дате соответствует дата 45 000?
5. Определите день, который приходится на 74-й рабочий день после сегодняшней даты (включая праздники).



6. Определите день, который приходится на 74-й рабочий день после сегодняшней даты (включая праздники, но исключая Рождество, Новый год и День независимости).
7. Сколько рабочих дней (включая праздники) между 10 июля 2005 г. и 15 августа 2006 г.?
8. Сколько рабочих дней (включая праздники, но исключая Рождество, Новый год и День независимости) между 10 июля 2005 г. и 15 августа 2006 г.?

В файле **Datep.xlsx** содержится несколько сотен дат. Используйте этот файл для выполнения следующей группы заданий.

1. Определите месяц, год, день месяца и день недели для каждой даты.
2. Представьте каждую дату в числовом формате.
3. Проект стартует 4 декабря 2005 г. Он включает этапы 1, 2 и 3. Этап 2 может начаться на следующий день после завершения этапа 1. Этап 3 может начаться на следующий день после завершения этапа 2. Создайте лист, на котором входными данными является продолжительность (в днях) всех трех этапов проекта, а выходными данными — месяц и год, в которые завершится каждый этап.
4. Допустим, вы приобрели акции 29 июля 2005 г. и продали их 30 декабря 2005 г. Фондовая биржа закрыта в День труда, на Рождество и в День благодарения<sup>1</sup>. Создайте список рабочих дней фондового рынка в тот период, когда вы владели акциями.
5. В файле **Machinedates.xlsx** содержатся даты покупки и продажи нескольких автомобилей. Определите, сколько месяцев и лет автомобили провели на стоянке.
6. Для любой даты найдите способ вычисления дня недели первого числа месяца, на который приходится эта дата.
7. Для любой даты найдите способ вычисления последнего дня месяца, на который приходится дата. Подсказка: как ни удивительно, но формула `=ДАТА(2005;13;1)` возвращает 01.01.2006.
8. Как для любой даты вычислить ее порядковый номер от начала года? Например, каким по счету днем года является 15 апреля 2020 г.? Подсказка: формула `=ДАТА(2020;1;0)` возвращает порядковый номер для нулевого дня в январе 2020 г., который в Excel обрабатывается как 31 декабря 2019 г.
9. Во Фредонии выходными днями являются вторник и среда. Какой дате соответствует день через 200 рабочих дней после 03.05.2013?
10. В стране Нижняя Амперия выходными днями являются пятница и суббота. Кроме того, праздничным днем считается День святого Валентина. Сколько рабочих дней там между 10 января 2014 г. и 31 мая 2015 г.?

---

<sup>1</sup> Разумеется, вы можете использовать другие праздничные даты, например 4 ноября — День народного единства. — *Примеч. ред.*

11. Первым кварталом каждого года является январь — март , вторым кварталом — апрель — июнь и т. д. Напишите формулу, которая возвращает первый день квартала для указанной даты.
12. Для кварталов из задания 11 напишите формулу расчета последнего дня предыдущего квартала для любой указанной даты.
13. Создайте таблицу, в которой можно указать дату рождения и получить фактический возраст.
14. День памяти всегда приходится на последний понедельник мая. Создайте таблицу, в которой можно указать год и получить дату Дня памяти.
15. Создайте таблицу, в которой всегда перечислены следующие 50 рабочих дней (праздники не учитываются, и дни с понедельника по пятницу считаются рабочими днями).
16. Грег Уинстон работает на федеральное правительство, и в 2016 году у него было 10 выходных в государственные праздники. (Посмотрите даты этих праздников в интернете.) В каждую не приходящуюся на праздник пятницу Грег также выходной, равно как и в выходные дни, начиная с 8 января 2016 г. Создайте список всех рабочих дней Грега в 2016 г.

## ГЛАВА 8

# Оценка инвестиций по чистой приведенной стоимости

### Обсуждаемые вопросы

- Что такое чистая приведенная стоимость (ЧПС)?
- Как работать с функцией ЧПС в Excel?
- Как рассчитать ЧПС, если денежные потоки приходят в начале или середине года?
- Как рассчитать ЧПС, если денежные потоки приходят через неравные промежутки времени?

Рассмотрим на примере два инвестиционных проекта, денежные потоки которых представлены в файле NPV.xlsx и на рис. 8.1.

	A	B	C	D	E	F
1						
2						
3		r	0.2			
4	NPV	Time	0	1	2	Total cf
5	277.7778	Inv 1 CF	-10000	24000	-14000	0
6	-27.7778	Inv 2 CF	-6000	8000	-1000	1000
7		PV Inv 1	-10000	20000	-9722.22	
8		PV Inv 2	-6000	6666.667	-694.444	
9						
10		PV beginning of Year				
11		Inv 1	\$277.78	\$277.78		
12		Inv 2	(\$27.78)	(\$27.78)		
13		PV End of Year				
14		Inv 1	\$231.48			
15		Inv 2	(\$23.15)			
16		PV Middle of year				
17		Inv 1	\$253.58			
18		Inv 2	(\$25.36)			

**Рис. 8.1.** Для оценки этих двух инвестиций необходимо рассчитать чистую приведенную стоимость

- Инвестиционный проект 1 требует расходов в размере \$10 000 сегодня и \$14 000 через два года. Через год эта инвестиция принесет \$24 000.
- Инвестиционный проект 2 требует расходов в размере \$6000 сегодня и \$1000 через два года. Через год эта инвестиция принесет \$8000.

Какой из проектов представляет собой лучшее вложение средств? Совокупный денежный поток для первого проекта равен \$0, в то время как второй проект принесет \$1000. На первый взгляд, второй инвестиционный проект более эффективен, но не будем спешить с выводами. Большая часть расходов на первый проект будет через два года, в то время как большая часть расходов по второму проекту придется на настоящий момент. Один доллар через два года кажется не такой большой тратой, как один доллар сегодня, так что, вполне возможно, первый инвестиционный проект окажется более привлекательным. Для оценки привлекательности инвестиций необходимо сравнить стоимости денежных потоков, полученные в разные моменты времени. Вот где пригодится понятие чистой приведенной стоимости.

## Ответы на вопросы

### ❓ Что такое чистая приведенная стоимость (ЧПС)?

Чистая приведенная стоимость (ЧПС) потока денежных средств, полученная в разные моменты времени, — это просто стоимость, измеренная в текущих долларах. Допустим, у вас есть один доллар, и вы инвестируете этот доллар под годовую процентную ставку  $r$  процентов. За первый год этот доллар вырастет до  $1 + r$  долларов, за второй год до  $(1 + r)^2$  долларов и т. д. В некотором роде можно сказать, что один доллар сегодня эквивалентен  $1 + r$  долларам через год и  $(1 + r)^2$  долларам через два года. В целом один доллар сегодня эквивалентен  $(1 + r)^n$  долларам через  $n$  лет. Эти расчеты можно записать следующим уравнением:

1 доллар сегодня =  $(1 + r)^n$  долларов через  $n$  лет.

Если разделить обе части этого уравнения на  $(1 + r)^n$ , получим следующий результат:

$$\frac{1}{(1 + r)^n} \text{ доллар сегодня} = 1 \text{ доллар через } n \text{ лет.}$$

Этот результат показывает, как рассчитать ЧПС (в текущих долларах) для любой последовательности денежных потоков. Любой денежный поток можно конвертировать в текущие доллары путем умножения стоимости денежного потока через  $n$  лет ( $n$  может быть дробью) на  $\frac{1}{(1 + r)^n}$ .

Далее для вычисления ЧПС инвестиционного проекта сложите стоимости денежных потоков (в текущих долларах). Пусть  $r$  равно 0,2. ЧПС для данных инвестиционных проектов можно рассчитать следующим образом:

$$\text{ЧПС}_{\text{проект 1}} = -10000 + \frac{24000}{(1+0,20)^1} + \frac{-14000}{(1+0,20)^2} = 277,78 \text{ доллара.}$$

$$\text{ЧПС}_{\text{проект 2}} = -6000 + \frac{8000}{(1+0,20)^1} + \frac{-1000}{(1+0,20)^2} = -27,78 \text{ доллара.}$$

По критерию ЧПС первая инвестиция превосходит вторую. Хотя совокупный денежный поток для второго проекта превышает совокупный денежный поток первого проекта, ЧПС первого проекта больше, поскольку большая часть отрицательного денежного потока первого проекта приходит позже, а критерий ЧПС придает меньший вес денежным потокам, приходящим позже. Если годовая процентная ставка  $r$  составляет 0,02%, то ЧПС второй инвестиции будет больше, поскольку при малых значениях  $r$  более поздние денежные потоки дисконтируются не так сильно, и ЧПС будет соответствовать ранжированию инвестиций по совокупному денежному потоку.

#### ПРИМЕЧАНИЕ

Я выбрал процентную ставку  $r = 0,2$  случайным образом во избежание обсуждения способов определения подходящего значения  $r$ . Чтобы понимать все аспекты определения нужного значения  $r$ , необходимо минимум год посвятить изучению управления финансами. Нужно значение  $r$ , используемое при расчете ЧПС, часто называют *стоимостью капитала компании*. Достаточно сказать, что в большинстве американских компаний годовая стоимость капитала составляет от 0,1 (10%) до 0,2 (20%). Если годовая процентная ставка выбирается в соответствии с общепринятой финансовой практикой, проекты с ЧПС  $> 0$  увеличивают стоимость компании, а проекты с ЧПС  $< 0$  ее уменьшают, в то время как проекты с ЧПС  $= 0$  не изменяют стоимость компании. Компания должна (если располагает неограниченным инвестиционным капиталом) инвестировать в каждый доступный инвестиционный проект с положительной ЧПС.

Для определения ЧПС первой инвестиции в Excel я присвоил диапазону с процентной ставкой (ячейке C3) имя  $r_$ . (Если вы попытаетесь присвоить ячейке имя  $r$ , Excel преобразует его в  $r_$ , см. главу 2.) Затем я скопировал денежный поток нулевого периода из C5 в C7. ЧПС для первого и второго годов я определил, скопировав формулу  $=D5/(1+r_)^D\$4$  из D7 в E7. Символ вставки (^) означает возведение в степень. В ячейке A5 я вычислил ЧПС первой инвестиции, сложив ЧПС первого проекта для каждого года по формуле  $=СУММ(C7:E7)$ . Для определения ЧПС второй инвестиции я скопировал формулы из C7:E7 в C8:E8 и из A5 в A6.

#### ❓ Как работать с функцией ЧПС в Excel?

Синтаксис функции ЧПС (NPV) в Excel: ЧПС(ставка;диапазон\_ячеек). Эта функция определяет ЧПС по заданной ставке для денежных потоков в указанном диапазоне ячеек. При расчете подразумевается, что первый денежный поток отстоит на один период от настоящего момента. То есть для определения ЧПС первой инвестиции нельзя применить формулу  $=\text{ЧПС}(r_;C5:E5)$ . Вместо этого по ней (в ячейке

ке C14) я рассчитал ЧПС следующих денежных потоков: для  $-\$10\,000$  через год от настоящего момента, для  $\$24\,000$  через два года и для  $-\$14\,000$  через три года. Назовем это значение приведенной стоимостью на конец года. ЧПС первого инвестиционного проекта на конец года равна  $\$231,48$ . Для расчета фактической ЧПС первого проекта с учетом денежных потоков в начале года нужно ввести формулу  $=C7+ЧПС(r_;;D5:E5)$  в ячейку C11. По этой формуле денежный отток в нулевой период вообще не учитывается (что правильно, поскольку денежный отток в нулевой период уже выражен в текущих долларах), и сначала денежный поток в ячейке D5 умножается на  $1/1,2$ , а затем денежный поток в ячейке E5 умножается на  $1/1,2^2$ .

Формула в ячейке C11 вычисляет корректную ЧПС первой инвестиции, а именно  $\$277,78$ .

### ❓ Как рассчитать ЧПС, если денежные потоки приходят в начале или середине года?

Чтобы применить функцию ЧПС для расчета чистой приведенной стоимости проекта, денежные потоки которого всегда приходятся на начало года, можно воспользоваться подходом, который я описал для определения ЧПС первой инвестиции. Отделите денежный поток первого года и примените функцию ЧПС к оставшимся денежным потокам. Или учтите то, что для любого  $n$ -го года  $\$1$ , полученный в начале  $n$ -го года, эквивалентен  $\$(1 + r)^n$  в конце  $n$ -го года. Помните: в течение года доллар вырастет в  $1 + r$  раз. Таким образом, если умножить результат, полученный с помощью функции ЧПС, на  $1 + r$ , можно конвертировать ЧПС последовательности денежных потоков на конец года в ЧПС последовательности денежных потоков на начало года. То есть ЧПС первой инвестиции рассчитывается по формуле  $=(1+r)*C14$  (см. ячейку D11). Конечно, вы снова получите  $\$277,78$ . Теперь допустим, что денежные потоки инвестиционного проекта приходят в середине каждого года. Для организации, которая получает ежемесячные взносы в рамках подписки, можно приблизительно считать 12 месячных взносов, полученных за указанный год, единовременной суммой, полученной в середине года. Каким образом с помощью функции ЧПС можно определить ЧПС последовательности денежных потоков в середине года? Для любого  $n$ -го года  $\sqrt{1+r}$  долларов, полученных в конце  $n$ -го года, эквивалентны  $\$1$ , полученному в середине  $n$ -го года, поскольку за полгода  $\$1$  вырастет в  $\sqrt{1+r}$  раз.

Если предположить, что денежные потоки для первой инвестиции приходят в середине года, можно рассчитать ЧПС первого проекта на середину года в ячейке C17 по формуле  $=КОРЕНЬ(1+r)*C14$ . ЧПС равна  $\$253,58$ .

### ❓ Как рассчитать ЧПС, если денежные потоки приходят через неравные промежутки времени?

Деньги часто поступают нерегулярно, что усложняет расчет ЧПС или внутренней ставки доходности (ВСД) этих потоков. К счастью, на помощь нам спешит такая функция Excel, как ЧИСТНЗ (XNPV), вычисляющая ЧПС денежных потоков, спланированных через неравные промежутки времени.

Для функции **ЧИСТНЗ** используется синтаксис **ЧИСТНЗ(ставка;значения;даты)**. Первая дата должна быть самой ранней, а остальные даты не обязательно указывать в хронологическом порядке. Функция **ЧИСТНЗ** вычисляет ЧПС заданных денежных потоков при условии, что текущая дата является первой датой последовательности. Например, если первой указана дата 08.04.13, то ЧПС рассчитывается в стоимости доллара на 8 апреля 2013 г.

Работа функции **ЧИСТНЗ** показана в примере на листе **NPV as of first date** в файле **XNPV.xlsx** (см. также рис. 8.2). Допустим, 8 апреля 2015 г. было выплачено \$900. Позже были получены следующие суммы:

- \$300 15 августа 2015 г.;
- \$400 15 января 2016 г.;
- \$200 25 июня 2016 г.;
- \$100 3 июля 2016 г.

Если годовая процентная ставка составляет 10%, то какова величина ЧПС этих денежных потоков? Введите даты (в формате дат Excel) в D3:D7 и денежные потоки в E3:E7. После ввода формулы **=ЧИСТНЗ(A9;E3:E7;D3:D7)** в ячейку D11 будет рассчитана ЧПС проекта в стоимости доллара на 8 апреля 2015 г., поскольку это первая дата в списке. ЧПС этого проекта в стоимости доллара на 8 апреля 2015 г. равна \$28,64.

	A	B	C	D	E	F	G
1							
2	<b>XNPV Function</b>		<b>Code</b>	<b>Date</b>	<b>Cash Flow</b>	<b>Time</b>	<b>discount factor</b>
3			42102.00	4/8/2015	-\$900.00		1
4			42231.00	8/15/2015	\$300.00	0.353425	0.966876054
5			42384.00	1/15/2016	\$400.00	0.772603	0.92900895
6			42546.00	6/25/2016	\$200.00	1.216438	0.890529581
7			42554.00	7/3/2016	\$100.00	1.238356	0.888671215
8	<b>Rate</b>						
9	0.1						
10				<b>XNPV</b>	<b>Direct</b>		
11				\$28.64	\$28.64		

**Рис. 8.2.** Функция **ЧИСТНЗ** в действии

Функция **ЧИСТНЗ** выполняет следующие вычисления:

1. Вычисляет количество лет после 8 апреля 2015 г, через которое наступит каждая дата в списке (см. столбец F). Например, 15 августа наступит через 0,3534 года после 8 апреля.
2. Дисконтирует денежные потоки по ставке  $1/(1+ставка)^{количество\_лет}$ . Например, денежный поток 15 августа 2015 г. дисконтируется по ставке:  $(1+0.1)^{0.3534}$ .
3. Суммирует совокупные денежные потоки в ячейке E11: (стоимость денежных потоков)\*(коэффициент дисконтирования).

Допустим, сегодня 8 июля 2013 г. Как вычислить ЧПС инвестиции в текущих долларах? Просто добавьте строку с текущей датой и нулевой денежный поток и включите эту строку в диапазон для функции **ЧИСТНЗ**. (См. рис. 8.3 и лист **today**.) ЧПС проекта в текущих долларах составляет \$106,99.

	A	B	C	D	E
1					
2	<b>XNPV Function</b>		<b>Code</b>	<b>Date</b>	<b>Cash Flow</b>
3			41463.00	7/8/2013	\$0.00
4			42102.00	4/8/2015	-\$900.00
5			42231.00	8/15/2015	\$300.00
6			42384.00	1/15/2016	\$400.00
7			42546.00	6/25/2016	\$200.00
8			42554.00	7/3/2016	\$100.00
9			42188.00	7/3/2015	\$100.00
10	<b>Rate</b>				
11	0.1				
12				<b>XNPV</b>	
13				\$106.99	

**Рис. 8.3.** Конвертация ЧПС в текущие доллары

Если денежный поток останется пустым, функция **ЧПС** проигнорирует и денежный поток, и период, а функция **ЧИСТНЗ** возвратит ошибку **#ЧИСЛО!**.

## Задания

1. Игрок НБА получает бонус за подписание контракта в размере \$1 000 000 сейчас и \$2 000 000 через год, через два года и через три года. Предполагая, что  $r = 0,10$ , и игнорируя налоги, стал бы он богаче, если бы получил \$6 000 000 сразу?
2. Ниже приведены денежные потоки проекта:

Сегодня	Через год	Через два года	Через три года
-\$4 млн	\$4 млн	\$4 млн	-\$3 млн

Если стоимость капитала компании составляет 15%, не стоит ли ей отказаться от этой инвестиции?

3. Через месяц клиент будет платить своему интернет-провайдеру \$25 в месяц в течение последующих пяти лет. При условии что все годовые доходы поступают в середине года, рассчитайте ЧПС этих доходов. Примите  $r = 0,15$ .
4. Через месяц клиент будет платить своему интернет-провайдеру \$25 в месяц в течение последующих пяти лет. При условии что все месячные доходы поступают в начале месяца, воспользуйтесь функцией **ЧИСТНЗ** для получения точного значения ЧПС этих доходов. Примите  $r = 0,15$ .



5. Рассмотрите денежные потоки за четыре года в нижеследующей таблице. Определите ЧПС этих денежных потоков, если  $r = 0,15$  и денежные потоки приходят в конце года.

Год	1	2	3	4
	-\$600	\$550	-\$680	\$1000

6. Рассмотрите следующие денежные потоки:

Дата	Денежный поток
15.12.01	-\$1000
11.01.02	\$300
07.04.03	\$600
15.07.04	\$925

Если сегодня 1 ноября 2001 г. и  $r = 0,15$ , какова ЧПС этих потоков?

7. После получения MBA студент начнет получать с 1 сентября 2005 г. \$80 000 в год. Он рассчитывает на ежегодное увеличение оклада на 5%, пока не выйдет на пенсию 1 сентября 2035 г. При условии, что стоимость капитала составляет 8% в год, определите совокупную приведенную стоимость его заработка до уплаты налогов.
8. Рассмотрим 30-летнюю облигацию с выплатой \$50 в конце года с первого года по 29-й и выплатой \$1050 в конце 30-го года. Если дисконтная ставка составляет 5% в год, какая цена является справедливой за эту облигацию?

## ГЛАВА 9

# Внутренняя ставка доходности

### Обсуждаемые вопросы

- Как вычислить внутреннюю ставку доходности денежных потоков?
- Всегда ли проект имеет единственное значение ВСД?
- Существуют ли условия, гарантирующие единственное значение ВСД проекта?
- Если единственное значение ВСД имеют два проекта, как можно использовать значение ВСД для этих проектов?
- Как вычислить ВСД для нерегулярно поступающих денежных потоков?
- Что такое модифицированная внутренняя ставка доходности (МВСД) и как она рассчитывается?

Чистая приведенная стоимость (ЧПС) последовательности денежных потоков зависит от используемой процентной ставки ( $r$ ). Например, анализ денежных потоков первого и второго проектов (см. лист IRR в файле IRR.xlsx и рис. 9.1) показывает, что при  $r = 0,2$  ЧПС второго проекта больше, чем первого, а при  $r = 0,01$  наоборот. При использовании ЧПС для ранжирования инвестиций результат может зависеть от процентной ставки. Человеку свойственно стремление свести все в жизни к единственному показателю. ВСД проекта — это процентная ставка, при которой

	A	B	C	D	E	F	G	H	I	J	K
1		Time	1	2	3	4	5	6	7	NPV $r=0,2$	NPV $r=0,01$
2	Project 1		-400	200	600	-900	1000	250	230	\$268.54	\$918.99
3	Project 2		-200	150	150	200	300	100	80	\$297.14	\$741.07
4	Project 1	IRR Proj 1 No Guess	IRR Proj 2 No Guess								
5	no guess	47.5%	80.1%								
6											
7	guess	Guess Proj 1	Guess proj 2								
8	-0.9	47.5%	80.1%								
9	-0.7	47.5%	80.1%								
10	-0.5	47.5%	80.1%								
11	-0.3	47.5%	80.1%								
12	-0.1	47.5%	80.1%								
13	0.1	47.5%	80.1%								
14	0.3	47.5%	80.1%								
15	0.5	47.5%	80.1%								
16	0.7	47.5%	80.1%								
17	0.9	47.5%	80.1%								

**Рис. 9.1.** Пример использования функции ВСД

величина ЧПС проекта равна 0. Если проект имеет единственное значение ВСД, то у такой ВСД точная трактовка. Например, если ВСД проекта равна 15%, то годовая норма прибыли на вложенные денежные средства составит 15%. В упомянутом примере ВСД первого проекта равна 47,5%. То есть \$400, вложенные в первый год, дадут годовую норму прибыли 47,5%. Но иногда проект имеет несколько ВСД или даже ни одной ВСД. В этих случаях говорить о ВСД проекта не имеет смысла.

## Ответы на вопросы

### ? Как вычислить внутреннюю ставку доходности денежных потоков?

Внутренняя ставка доходности вычисляется с помощью функции ВСД (IRR). Синтаксис функции ВСД: ВСД(диапазон\_денежных\_потоков;[предположение]), где предположение является необязательным аргументом. Если предполагаемое значение для ВСД проекта не указано, Excel начинает вычисления с предположения, что значение ВСД проекта равно 10%, и затем значение ВСД изменяется до тех пор, пока не будет найдена процентная ставка, при которой ЧПС проекта равна 0 (это и есть ВСД проекта). Если процентная ставка, при которой ЧПС проекта равна 0, не найдена, возвращается ошибка #ЧИСЛО!. Я ввел в ячейке B5 листа IRR файла IRR.xlsx формулу =ВСД(C2:I2) для расчета ВСД первого проекта. В результате получилось 47,5%. Таким образом, ЧПС первого проекта равна 0 при годовой процентной ставке 47,5. Аналогично ВСД второго проекта равна 80,1%.

Даже если функция ВСД найдет значение ВСД, у проекта может оказаться несколько ВСД. Чтобы это выяснить, можно варьировать начальное предположение о величине ВСД проекта (например, от -90% до 90%). Предположение о величине ВСД первого проекта можно варьировать, скопировав формулу=ВСД(\$C\$2:\$I\$2;A8) из B8 в B9:B17. Поскольку при всех предположениях о величине ВСД первого проекта в результате получается 47,5%, можно с уверенностью сказать, что первый проект имеет единственную ВСД, и она равна 47,5%. Аналогично можно утверждать, что второй проект имеет единственную ВСД 80,1%.

### ? Всегда ли проект имеет единственное значение ВСД?

В файле IRR.xlsx на листе Multiple IRR и на рис. 9.2 показано, что у третьего проекта (с денежными потоками -20, 82, -60, 2) имеется два значения ВСД. Предположение о величине ВСД третьего проекта варьировалось от -90% до 90% путем копирования формулы =ВСД(\$B\$4:\$E\$4;B8) из ячейки C8 в ячейки C9:C17.

Если предполагаемая величина ВСД не больше 30%, то ВСД проекта равна -9,6%. Для других предполагаемых величин ВСД проекта равна 216,1%. Для обеих этих процентных ставок ЧПС третьего проекта нулевая.

В файле IRR.xlsx на листе No IRR (рис. 9.3) показано, что вне зависимости от предположений о величине ВСД четвертого проекта в результате всегда возвращается ошибка #ЧИСЛО!. Это означает, что четвертый проект не имеет ВСД.

	A	B	C	D	E	F
1	<b>Multiple IRR's</b>					
2						
3		1	2	3	4	
4	<b>Project 3</b>	-20	82	-60	2	
5			plain irr	-9.6%		
6						
7		guess				
8		-0.9	-9.6%		npv at -9.6%	(\$0.01)
9		-0.7	-9.6%		npv at 216.1%	\$0.00
10		-0.5	-9.6%			
11		-0.3	-9.6%			
12		-0.1	-9.6%			
13		0.1	-9.6%			
14		0.3	-9.6%			
15		0.5	216.1%			
16		0.7	216.1%			
17		0.9	216.1%			

**Рис. 9.2.** Проект имеет несколько значений ВСД

Если у проекта несколько ВСД или же она отсутствует, понятие ВСД теряет смысл. Однако несмотря на эту проблему, многие компании все еще используют ВСД в качестве основного инструмента ранжирования инвестиций.

	A	B	C	D
1				
2		<b>No IRR</b>		
3				
4		0	1	2
5	<b>Project 4</b>	10	-30	35
6				
7		guess		
8		-0.9	#NUM!	
9		-0.8	#NUM!	
10		-0.7	#NUM!	
11		-0.6	#NUM!	
12		-0.5	#NUM!	
13		-0.4	#NUM!	
14		-0.3	#NUM!	
15		-0.2	#NUM!	
16		-0.1	#NUM!	
17		0	#NUM!	
18		0.1	#NUM!	
19		0.2	#NUM!	
20		0.3	#NUM!	
21		0.4	#NUM!	
22		0.5	#NUM!	
23		0.6	#NUM!	
24		0.7	#NUM!	
25		0.8	#NUM!	
26		0.9	#NUM!	

**Рис. 9.3.** У этого проекта значение ВСД отсутствует

### ? Существуют ли условия, гарантирующие единственное значение ВСД проекта?

Если в последовательности денежных потоков проекта происходит только одна смена знака, проект гарантированно имеет одно значение ВСД. Например, у второго проекта в ячейках C3:I3 на листе IRR последовательность знаков денежных потоков такая:  $- + + + +$ . Смена знака происходит только один раз (между первым и вторым годом), поэтому у второго проекта должно быть только одно значение ВСД. Последовательность знаков денежных потоков у третьего проекта в ячейках B4:E4 на листе Multiple IRR такова:  $- + - +$ . Поскольку знак денежных потоков меняется три раза, единственное значение ВСД не гарантировано. У четвертого проекта в ячейках B5:D5 на листе No IRR знаки денежных потоков  $(+ - +)$  меняются два раза, в этом случае единственное значение ВСД также не гарантируется. Большинство проектов капиталовложений (таких как строительство завода) начинается с отрицательного денежного потока, за которым следуют положительные денежные потоки. Следовательно, большинство проектов капиталовложений должно иметь единственное значение ВСД.

### ? Если единственное значение ВСД имеют два проекта, как можно использовать значение ВСД для этих проектов?

Если у проекта единственное значение ВСД, то можно утверждать, что проект увеличивает стоимость компании тогда и только тогда, когда ВСД проекта превышает годовую стоимость капитала. Например, если стоимость капитала компании составляет 15%, оба проекта, и первый, и второй, увеличат ее стоимость.

Допустим, что к рассмотрению представлены два проекта (у обоих проектов значение ВСД единственное), но выбрать нужно только один. Кажется, что следует выбрать проект, значение ВСД которого больше. Но принятое на основании этого решение может оказаться неверным, что продемонстрировано на рис. 9.4 и в файле IRR.xlsx на листе Which Project. ВСД пятого проекта 40%, а шестого — 50%. Если проекты ранжируются на основе ВСД и нужно выбрать только один проект, наверное, будет выбран шестой проект. Вспомните, что ЧПС проекта измеряет величину стоимости, которую проект добавляет компании. Очевидно, что пятый проект имеет (практически для любой стоимости капитала) значение ЧПС больше, чем шестой. Следовательно, если требуется выбрать один проект, то это должен быть пятый проект. Основываться только на ВСД опасно, поскольку игнорируется масштаб проекта. Несмотря на то что шестой проект превосходит пятый по доходности на каждый вложенный доллар, благодаря масштабности пятый проект более ценен для компании. ВСД не отражает масштаба проекта, а ЧПС — отражает.

	B	C	D	E
2	Project	Time 0	Time 1	IRR
3	5	-100	140	40%
4	6	-1	1.5	50%

**Рис. 9.4.** Оценка на основе ВСД может привести к выбору не самого выгодного проекта

### ❓ Как вычислить ВСД для нерегулярно поступающих денежных потоков?

Денежные потоки приходятся на реальные даты, а не только на начало или конец года. ВСД для нерегулярных дат вычисляется с помощью функции **ЧИСТВНДОХ** (XIRR). Синтаксис функции **ЧИСТВНДОХ**: **ЧИСТВНДОХ(денежные\_потоки;даты;[предположение])**. Функция **ЧИСТВНДОХ** (XIRR) определяет ВСД последовательности денежных потоков, имеющих место в любом наборе нерегулярных дат. Как и в случае с функцией **ВСД**, аргумент **предположение** является необязательным. См. пример использования функции **ЧИСТВНДОХ** на рис. 9.5 и в файле IRR.xlsx на листе XIRR.

	A	B	C	D	E	F
1					Project 7	
2	<b>XIRR Function</b>			<b>Date</b>	<b>Cash Flow</b>	
3				<b>Code</b>	<b>Date</b>	<b>Cash Flow</b>
4				4/8/2011	4/8/2011	-1500
5				8/15/2011	8/15/2011	300
6				1/15/2012	1/15/2012	400
7				6/25/2012	6/25/2012	200
8				<b>XIRR</b>		
9				-48.69%		

**Рис. 9.5.** Пример использования функция **ЧИСТВНДОХ**

Расчет по формуле **=ЧИСТВНДОХ(F4:F7;E4:E7)** в ячейке D9 показывает, что ВСД седьмого проекта равна -48,69%.

### ❓ Что такое модифицированная внутренняя ставка доходности (МВСД) и как она рассчитывается?

Во многих случаях ставка, по которой компания занимает средства, отличается от ставки, по которой компания реинвестирует средства. При расчете ВСД неявно предполагается, что обе эти ставки равны ВСД. Если известны фактическая ставка займа и ставка реинвестирования, то мы можем, применив функцию **МВСД** (MIRR), рассчитать ставку дисконтирования, при которой ЧПС всех денежных потоков (включая погашение займа и реинвестирование поступлений по займу с заданной ставкой) нулевая. Синтаксис функции **МВСД**: **МВСД(значения\_денежных\_потоков;ставка\_финансирования; ставка\_реинвестирования)**. К счастью, значение **МВСД** всегда единственное. На рис. 9.6 и в файле IRR.xlsx на листе MIRR представлен пример использования функции **МВСД**. Допустим, вы заняли \$120 000 и получили следующие денежные потоки: первый год — \$39 000; второй год — \$30 000; третий год — \$21 000; четвертый год — \$37 000; пятый год — \$46 000. Скажем, вы взяли заем под 10% годовых, а реинвестировали прибыль под 12% годовых.

Введите эти значения на листе **MIRR** в ячейки E7:E12 и рассчитайте **МВСД** в ячейке D15 по формуле **=МВСД(E7:E12;E3,E4)**. Таким образом, **МВСД** этого проекта составляет 12,61%. В ячейке D16 вычислена фактическая ВСД — 13,07%.

	C	D	E	F
1				
2				
3		borrow	0.1	
4		invest	0.12	
5				
6		Year	Amt	
7		0	-120000	
8		1	39000	61367.26
9		2	30000	42147.84
10		3	21000	26342.4
11		4	37000	41440
12		5	46000	46000
13			total	
14				Final Value
15	MIRR	12.6094%		217297.5
16	IRR	13.0736%		
17				MIRR
18				0.126094

Рис. 9.6. Пример использования функции МВСД

Для людей с математическим складом ума в файле MIRR.xlsx (рис. 9.7) представлен детальный расчет МВСД. Как видим, денежный поток отрицателен в периоды 0-й и 1-й и положителен во 2-й и 3-й периоды. Во-первых, обратите внимание, что если вы инвестируете  $x$  \$ в 0-й период и инвестиция приносит  $y$  \$ в 3-й период, то ВСД такой простой инвестиции равна  $(y/x)^{1/3} - 1$ . В данном случае мы рассчитываем  $x$  как нынешнюю стоимость отрицательного денежного потока (ориентируясь

	A	B	C	D	E	F	G	H	I	J	K
1		MIRR	Modified Internal Rate of Return								
2		Borrow at 10% Reinvest at 12%									
3											
4				-1000							
5				-4000							
6				5000	IRR assumes all cash flows invested at IRR						
7				2000							
8											
9	IRR	25.48%	=IRR(D4:D7)								
10	MIRR	17.91%	=MIRR(D4:D7,0.1,0.12)		Find PV of Negative Cash Flows as of beginning at borrowing rate						
11											
12					Find IRR based on these numbers						
13											
14				PV of Negative Cash flows							
15				\$4,636.36 =-D4-D5/1.1							
16											
17				FV of Future Cash Flows							
18				\$7,600.00 =1.12*D6+D7							
19											
20				MIRR							
21				17.91% =(D18/D15)^(1/3)-1							

Рис. 9.7. Расчет МВСД

на ставку по займу) в первый период, а ( $y$  — как будущую стоимость (ориентируясь на ставку реинвестирования) положительного денежного потока в последний период. И тогда выражением для МВСД инвестиции является  $(y/x)^{1/3} - 1$ .

В заключение обращу ваше внимание на то, что при пустом денежном потоке функция ВСД игнорирует и денежный поток, и период и возвращает сообщение об ошибке #ЧИСЛО!.

## Задания

1. Вычислите все виды ВСД для следующего ряда денежных потоков:

Год 1	Год 2	Год 3	Год 4	Год 5	Год 6	Год 7
−\$10 000	\$8000	\$1500	\$1500	\$1500	\$1500	−\$1500

2. Рассмотрите проект с приведенными ниже денежными потоками. Определите ВСД проекта. Если годовая стоимость капитала составляет 20%, следует ли браться за этот проект?

Год 1	Год 2	Год 3
−\$4000	\$2000	\$4000

3. Найдите все виды ВСД для следующего проекта:

Год 1	Год 2	Год 3
\$100	−\$300	\$250

4. Найдите все виды ВСД для проекта с заданными денежными потоками, приходящимися на перечисленные даты:

10.01.2003	10.07.2003	25.05.2004	18.07.2004	20.03.2005	01.04.2005	10.01.2006
−\$1000	\$900	\$800	\$700	\$500	\$500	\$350

5. Рассмотрите следующие два проекта при условии, что стоимость капитала компании составляет 15%. Найдите ВСД и ЧПС каждого проекта. Какие проекты увеличат стоимость компании? Если бы компания могла выбрать только один проект, какой проект ей следовало бы выбрать?

	Год 1	Год 2	Год 3	Год 4
Проект 1	−\$40	\$130	\$19	\$26
Проект 2	−\$80	\$36	\$36	\$36



6. Двадцатипятилетняя Мэг Прайор намерена инвестировать \$10 000 в свой пенсионный фонд в начале каждого года в течение следующих 40 лет. Допустим, в течение следующих 30 лет Мэг заработает 15% на своей инвестиции, а в последние 10 лет перед выходом на пенсию ее инвестиция будет приносить 5%. Определите ВСД, связанную с ее инвестицией, и ее окончательное положение при выходе на пенсию. Будет ли значение ВСД единственным? Как бы вы интерпретировали единственное значение ВСД?
7. Попробуйте объяснить, почему шестой проект (см. в файле IRR.xlsx лист Which Project) имеет ВСД, равную 50%.
8. Исследуйте проект со следующими денежными потоками:

Год 1	Год 2	Год 3
-\$70 000	\$12 000	\$15 000

Попытайтесь найти ВСД этого проекта без предположений. Какая при этом возникает проблема? Какова ВСД этого проекта? Имеет ли проект единственное значение ВСД?

9. Для денежного потока в задании 1 при условии, что можно взять кредит под 12% годовых и инвестировать прибыль под 15% годовых, вычислите МВСД проекта.
10. Допустим, что за облигацию из задания 9 в главе 8 («Оценка инвестиций по чистой приведенной стоимости») сегодня уплачено \$1000. Какова величина ВСД облигации? ВСД облигации часто называют доходностью облигации.

## ГЛАВА 10

# Еще несколько финансовых функций Excel

### Обсуждаемые вопросы

- Вы покупаете копировальный аппарат. Как лучше поступить: заплатить \$11 000 сразу или выплачивать по \$3000 в год в течение пяти лет?
- В конце каждого года в течение следующих 40 лет я собираюсь инвестировать по \$2000 в год вплоть до выхода на пенсию и зарабатывать 8% в год на своих инвестициях. Какую сумму я получу при выходе на пенсию?
- Я занял \$10 000 на срок 10 месяцев под 8% годовых. Каковы мои ежемесячные платежи? Каковы ежемесячные выплаты по основной сумме и по процентам?
- Я собираюсь занять \$80 000 и вносить ежемесячные платежи в течение 10 лет. Максимальный ежемесячный платеж, который я могу себе позволить, составляет \$1000. На какую максимальную процентную ставку я могу рассчитывать?
- Если я возьму займы \$100 000 под 8% годовых и буду вносить платежи в размере \$10 000 в год, через сколько лет я выплачу кредит?
- Я дипломированный бухгалтер и часто работаю со сложными методами оценки амортизации оборудования. Имеются ли в Excel функции для расчета амортизации?

При кредите на покупку автомобиля или жилья всегда возникают сомнения относительно выгоды заключаемой сделки. При откладывании денег в пенсионный период всегда любопытно, насколько велики окажутся сбережения. Подобные финансовые вопросы постоянно возникают в повседневной работе и личной жизни. Ответы на них дает изучение и применение следующих функций Microsoft Excel: ПС (PV), БС (FV), ПЛТ (PMT), ОСПЛТ (PPMT), ПРПЛТ (IPMT), ОБЩДОХОД (CUMPRINC), ОБЩПЛАТ (CUMIPMT), СТАВКА (RATE) и КПЕР (NPER).

## Ответы на вопросы

- ?** **Вы покупаете копировальный аппарат . Как лучше поступить: заплатить \$11 000 сразу или выплачивать по \$3000 в год в течение пяти лет?**

Нам нужно оценить ежегодные выплаты в размере \$3000 в год. Допустим, стоимость капитала составляет 12% в год. Мы можем воспользоваться функцией ЧПС, но функция ПС (PV) в Excel позволяет решить задачу гораздо быстрее. По-

ток денежных средств, состоящий из периодических, равных по сумме платежей (или поступлений), называется *аннуитетом*. Если предположить, что процентная ставка за период постоянна, аннуитет можно рассчитать с помощью функции ПС. Она возвращает приведенную стоимость (в долларах на сегодняшний день) для серии будущих периодических платежей при условии, что платежи и процентная ставка постоянны. Синтаксис функции ПС: ПС(ставка;#кпер[;плт][;бс][;тип]), где плт, бс и тип являются необязательными аргументами.

## ПРИМЕЧАНИЕ

При работе с финансовыми функциями Excel используются следующие соглашения по знакам для аргументов плт (платеж) и бс (будущая стоимость): полученные суммы указываются со знаком «+», а выплаченные — со знаком «-».

- ставка — процентная ставка за период. Например, если кредит берется под 6% годовых, а период равен одному году, то ставка = 0,06. Если период равен одному месяцу, то ставка =  $0,06/12 = 0,005$ .
- #кпер — число периодов аннуитета. Для примера с копирующим аппаратом #кпер = 5. Если вы платите за копирующий аппарат каждый месяц в течение пяти лет, то #кпер = 60. Разумеется, значение аргумента ставка должно соответствовать количеству периодов #кпер. Если значение #кпер указано в месяцах, то необходимо использовать ежемесячную процентную ставку; если в годах — годовую.
- плт — выплата, производимая в каждый период. Для примера с копирующим аппаратом плт = -\$3000. Выплаты указываются со знаком «-», а прибыль — со знаком «+». В вызове функции должен присутствовать по крайней мере один из аргументов, плт или бс.
- бс — остаток денежных средств (или будущая стоимость) после последнего платежа. Для примера с копирующим аппаратом бс = 0. Например, если желаемый остаток средств после последнего платежа равен \$500, то бс = \$500. Если в конце последнего периода требуется внести дополнительный платеж в размере \$500, то бс = -\$500. Если аргумент бс опущен, то его значение считается равным 0.
- тип — число 0 или 1, указывающее, когда производится выплата. Если тип опущен или равен 0, выплаты производятся в конце каждого периода. Если тип = 1, выплаты производятся в начале каждого периода. Обратите внимание, что во всех функциях, описанных в данной главе, вместо 1 можно указать ИСТИНА, а вместо 0 — ЛОЖЬ.

Решение задачи о копирующем аппарате показано на рис. 10.1. (См. лист PV в файле *Excelfinfunctions.xlsx*.)

Приведенная (к текущему моменту) стоимость платежа в размере \$3000 в конце каждого года в течение пяти лет со стоимостью капитала 12% рассчитана в ячейке В3 по формуле =ПС(0,12;5;-3000;0;0). Получившийся результат равен \$10 814,33.

	A	B	C	D	E
1					
2	<b>PMT</b>				
3	Pay \$3000 for 5 years end of year	\$10,814.33	=PV(0.12,5,-3000,0,0)		
4	Pay \$3000 for 5 years beginning of year	\$12,112.05	=PV(0.12,5,-3000,0,1)		
5	Extra \$500 payment end of year 5;end of year payments	\$11,098.04	=PV(0.12,5,-3000,-500,0)		

**Рис. 10.1.** Применение функции ПС

Последние два аргумента можно опустить и по формуле  $\text{=ПС}(0,12;5;-3000)$  получить тот же самый результат. Таким образом, выгоднее платить частями, чем отдать \$11 000 сразу.

Если вы платите за копировальный аппарат по \$3000 в начале каждого года в течение пяти лет, ЧПС платежей будет вычислена в ячейке B4 по формуле  $\text{=ПС}(0,12;5;-3000;0;1)$ . Таким образом, изменяя значение последнего аргумента с 0 на 1, вы переносите расчеты с конца года на его начало. В этом случае ЧПС будет равна \$12 112,05. Следовательно, выгоднее заплатить \$11 000 сразу, чем осуществлять платежи в начале года.

Теперь предположим, что вы платите по \$3000 в конце каждого года и необходимо включить дополнительный платеж в размере \$500 в конце последнего (пятого) года. Для этого случая стоимость всех платежей рассчитана в ячейке B5 с включением будущего платежа в размере \$500 в формулу  $\text{=ПС}(0,12;5;-3000;-500;0)$ . Обратите внимание: денежные потоки \$3000 и \$500 имеют отрицательные знаки, поскольку денежные средства выплачиваются. Приведенная стоимость для всех выплат составляет \$11 098,04.

**?** В конце каждого года в течение следующих 40 лет я собираюсь инвестировать по \$2000 в год вплоть до выхода на пенсию и зарабатывать 8% в год на своих инвестициях. Какую сумму я получу при выходе на пенсию?

В этом случае речь идет о будущей стоимости аннуитета (через 40 лет), а не о стоимости в долларах на сегодняшний день. Это задача для функции БС (будущая стоимость). Функция БС (FV) возвращает будущую стоимость инвестиции при условии, что периодические платежи и процентная ставка постоянны. Синтаксис функции БС:  $\text{БС}(\text{ставка};\text{\#кпер};\text{плт};[\text{пс}];[\text{тип}])$ , где плт, пс и тип являются необязательными аргументами.

- ставка — процентная ставка за период. В данном случае ставка = 0,08.
- #кпер — количество периодов в будущем, для которых требуется вычислить будущую стоимость. #кпер также означает количество периодов выплаты аннуитета. В данном случае #кпер = 40.
- плт — выплата, осуществляемая в каждый период. В данном случае плт = -\$2000. Знак «-» показывает, что мы кладем деньги на счет. Если этот аргумент опущен, то аргумент пс является обязательным.
- пс — сумма денежных средств (в текущих долларах), которую вы взяли в долг сейчас. В данном случае пс равно \$0. Если вы сегодня взяли займы у кого-либо \$10 000, то пс равна \$10 000, потому что кредитор дал вам \$10 000 и вы их получили. Если у вас сегодня было \$10 000 в банке, то пс равна -\$10 000, потому что вы положили на свой счет в банке \$10 000. Если аргумент пс опущен, то по умолчанию его значение равно 0.
- тип — число 0 или 1, обозначающее, когда производятся выплаты или когда денежные средства кладутся на депозит. Если тип = 0 или аргумент тип опущен, выплаты производятся или денежные средства кладутся на депозит в конце периода. В данном примере значение тип равно 0 или опущено. Если значение тип = 1, выплаты производятся или денежные средства кладутся на депозит в начале периода.

В файле *Excelfinfunctions.xlsx* на листе *FV* (рис. 10.2) я ввел в ячейку *B3* формулу  $=\text{БС}(0,08;40;-2000)$  и узнал, что через 40 лет сбережения составят \$518 113,04. Обратите внимание, что для ежегодного платежа необходимо ввести отрицательное значение, потому что \$2000 будут положены на ваш счет. Тот же результат можно получить, указав в формуле  $=\text{БС}(0,08;40;-2000;0;0)$  два последних необязательных аргумента.

	A	B	C	D	E
1					
2	<b>FV</b>				
3	Invest \$2000 end of year for 40 years	\$518,113.04	$=\text{FV}(0.08,40,-2000,0,0)$		
4	Invest \$2000 beginning of year for 40 years	\$559,562.08	$=\text{FV}(0.08,40,-2000,0,1)$		
5	We start with \$30000 and invest \$2000 per year at end of year for 40 years	\$1,169,848.68	$=\text{FV}(0.08,40,-2000,-30000,0)$		

**Рис. 10.2.** Применение функции БС

Если депозит вносится в начале каждого года в течение 40 лет, то по формуле  $=\text{БС}(0,08;40;-2000;0;1)$  в ячейке В4 можно рассчитать результат 40-летних накоплений: \$559 562,08.

Наконец, предположим, что в дополнение к вкладу составляющему \$2000 в конце года в течение 40 лет, вы изначально положили на счет \$30 000. Если вложения приносят 8% в год, какая денежная сумма окажется на вашем счету к моменту выхода на пенсию через 40 лет?

Ответ можно получить, установив в функции БС для аргумента бс значение  $-\$30\,000$ . Знак «-» означает, что вы сделали депозит в размере \$30 000. В ячейке В5 по формуле  $=\text{БС}(0,08;40;-2000;-30000;0)$  вычислена будущая стоимость — \$1 169 848,68.

**❓ Я занял \$10 000 на срок 10 месяцев под 8% годовых. Каковы мои ежемесячные платежи? Каковы ежемесячные выплаты по основной сумме и по процентам?**

Функция ПЛТ (PMT) возвращает сумму периодических платежей по кредиту при условии, что платежи и процентная ставка постоянны. Синтаксис функции ПЛТ:  $\text{ПЛТ}(\text{ставка};\#\text{кпер};\text{пс};[\text{бс}];[\text{тип}])$ , где бс и тип являются необязательными аргументами.

- ставка — процентная ставка по кредиту за период. В данном примере в качестве периода укажите месяц, тогда ставка будет равна  $0,08/12 = 0,006666667$ .
- #кпер — число выплат. В данном случае #кпер = 10.
- пс — приведенная стоимость всех платежей. Иначе говоря, пс — это сумма кредита. В данном случае пс = \$10 000. Значение пс имеет знак «+», поскольку вы получаете \$10 000.
- бс — окончательный баланс ссуды после последнего платежа. В данном случае бс = 0. Если аргумент бс опущен, то его значение принимается равным 0. Допустим, вы взяли «шаровой» кредит, по которому платите в конце каждого месяца, а в конце срока гасите кредит, внося большой одноразовый платеж в размере \$1000. Тогда значение бс равно  $-\$1000$ . Значение бс отрицательное, поскольку платите вы.
- тип — число 0 или 1, показывающее, когда должны быть выплаты. Если этот аргумент равен 0 или опущен, выплаты производятся в конце периода. Если вы собираетесь делать платеж в конце месяца, укажите для аргумента тип значение 0 или ничего не указывайте. Если значение тип равно 1, выплаты производятся или деньги кладутся на хранение в начале периода.

В файле `Excelfinfunctions.xlsx` на листе PMT в ячейке G1 я вычислил по формуле  $=\text{ПЛТ}(0,08/12;10;10000;0;0)$  ежемесячный платеж по 10-месячному кредиту на сумму \$10 000 под 8% годовых и с внесением платежа в конце месяца (рис. 10.3). Ежемесячный платеж составляет \$1037,03. Функция ПЛТ возвращает отрицательное значение, поскольку производятся платежи компании-кредитору.

	C	D	E	F	G	H
1		rate	0.006666667	payment	\$1,037.03	
2		months	10	end of month		
3		loan amount	\$ 10,000.00			
4						
5	Time	Beginning balance	Monthly Payment	Principal	Interest	Ending Balance
6	1	\$ 10,000.00	\$1,037.03	\$970.37	\$66.67	\$9,029.63
7	2	\$9,029.63	\$1,037.03	\$976.83	\$60.20	\$8,052.80
8	3	\$8,052.80	\$1,037.03	\$983.35	\$53.69	\$7,069.45
9	4	\$7,069.45	\$1,037.03	\$989.90	\$47.13	\$6,079.55
10	5	\$6,079.55	\$1,037.03	\$996.50	\$40.53	\$5,083.05
11	6	\$5,083.05	\$1,037.03	\$1,003.15	\$33.89	\$4,079.90
12	7	\$4,079.90	\$1,037.03	\$1,009.83	\$27.20	\$3,070.07
13	8	\$3,070.07	\$1,037.03	\$1,016.56	\$20.47	\$2,053.51
14	9	\$2,053.51	\$1,037.03	\$1,023.34	\$13.69	\$1,030.16
15	10	\$1,030.16	\$1,037.03	\$1,030.16	\$6.87	(\$0.00)
16						
17	NPV of payments	\$10,000.00				
18				cum int months 2-4	cumprinc months 2-4	
19	payment beginning of each month	\$1,030.16		-161.0125862	-2950.083682	
20	monthly payment if we make \$1000 ending payment	\$940.00				

**Рис. 10.3.** Примеры использования функций ПЛТ, ОСПЛТ, ПРПЛТ, ОБЩДОХОД и ОБЩПЛАТ

По необходимости вы можете применить функции ПРПЛТ (IPMT) и ОСПЛТ (PPMT), чтобы рассчитать выплачиваемые проценты и сумму ежемесячного основного платежа (это называется погашением основной суммы) отдельно.

Вычислите, сколько процентов вы платите каждый месяц, применив функцию ПРПЛТ. Синтаксис функции: ПРПЛТ(ставка;период;#кпер;пс;[бс];[тип]), где бс и тип являются необязательными аргументами. Аргумент период указывает номер периода, для которого вычисляется процент. Остальные аргументы аналогичны аргументам функции ПЛТ. Размер ежемесячного платежа в счет погашения основной суммы можно определить с помощью функции ОСПЛТ. Синтаксис функции: ОСПЛТ(ставка;период;#кпер;пс; [бс]; [тип]). Все аргументы этой функции аналогичны аргументам функции ПРПЛТ. Чтобы вычислить размер ежемесячного платежа в счет погашения основной суммы, скопируйте формулу =ОСПЛТ(0,08/12;С6;10;10000;0;0) из ячейки F6 в ячейки F7:F16. Например, за первый месяц выплачивается только \$970,37 в счет погашения основной суммы. Как и ожидалось, размер основного платежа растет каждый месяц. Знак «-» необходим, поскольку основная сумма выплачивается компании-кредитору, и функция ОСПЛТ возвратит отрицательное число. Сумму, ежемесячно выплачиваемую по процентам, вы можете вычислить, скопировав формулу =ПРПЛТ(0,08/12;С6;10;10000;0;0) из ячейки G6 в ячейки G7:G16. Например, за первый месяц необходимо заплатить по процентам \$66,67. Естественно, выплачиваемая по процентам сумма каждый месяц уменьшается.

Обратите внимание, что каждый месяц (платеж по процентам) + (платеж в счет погашения основной суммы) = (ежемесячный платеж). Иногда из общей суммы в результате округления исчезает один цент.

С помощью соотношения (остаток на конец месяца  $t$ ) = (остаток на начало месяца  $t$ ) — (платеж в счет погашения основной суммы за месяц  $t$ ) можно рассчитать в столбце Н остаток на конец каждого месяца. Обратите внимание, что остаток на начало первого месяца равен \$10 000. В столбце D я создал остаток на начало каждого месяца с помощью соотношения (для  $t = 2, 3, \dots, 10$ ) (остаток на начало месяца  $t$ ) = (остаток на конец месяца  $t-1$ ). Естественно, остаток на конец десятого месяца равен \$0 (ячейка H15), как и должно быть. Выплаты по процентам за каждый месяц можно вычислить как (проценты за месяц  $t$ ) = (процентная ставка)  $\times$  (остаток на начало месяца  $t$ ). Например, выплаты по процентам за третий месяц составят  $0,0066667 \times 8052,80 = \$53,69$ .

Очевидно, что ЧПС всех платежей равна точно \$10 000. Я проверил это в ячейке D17, применив формулу =ЧПС(0,08/12;E6:E15). В ячейке D17 приведена формула =ЧПС(E1;E6:E15), а в ячейке E1 — формула =0,08/12 (см. рис. 10.3).

Если платежи производятся в начале каждого месяца, размер каждого такого платежа можно вычислить в ячейке D19 по формуле =ПЛТ(0,08/12;10;10000;0;1). Замена последнего аргумента на 1 перемещает платеж в начало месяца. Поскольку кредитор получает деньги раньше, размер ежемесячного платежа в этом случае меньше, чем при выплате в конце месяца. При выплате в начале месяца ежемесячный платеж составляет \$1030,16.

Наконец, предположим, что в конце десятимесячного периода требуется произвести большой одноразовый платеж в размере \$1000. Если ежемесячные платежи отнесены на конец месяца, то размер такого ежемесячного платежа рассчитан по формуле =ПЛТ(0,08/12;10;10000;-1000;0) в ячейке D20. В этом случае ежемесячный платеж равен \$940. Он меньше исходного ежемесячного платежа, выплачиваемого в размере \$1037,03 в конце месяца, поскольку \$1000 кредита не выплачивается с ежемесячными платежами.

## Функции ОБЩДОХОД и ОБЩПЛАТ

Часто требуется собрать кумулятивные данные о выплатах по процентам или выплатам в счет основной суммы за несколько периодов. С функциями ОБЩДОХОД (CUMPRINC) и ОБЩПЛАТ (CUMIPMT) это проще пареной репы.

Функция ОБЩДОХОД вычисляет выплаты в погашение основной суммы между двумя заданными периодами (включая эти периоды). Синтаксис функции ОБЩДОХОД: ОБЩДОХОД(ставка;#кпер;пс;нач\_период;кон\_период;тип). Аргументы ставка, #кпер, пс и тип аналогичны описанным ранее аргументам.

Функция ОБЩПЛАТ возвращает проценты, выплачиваемые между двумя периодами (включая эти периоды). Синтаксис функции: ОБЩПЛАТ(ставка;#кпер;пс;нач\_период;кон\_период;тип). Аргументы функции аналогичны описанным ранее



аргументам. Например, на листе **PMT** в ячейке **F19** я вычислил по формуле **=ОБЩПЛАТ(0,08/12;10;10000;2;4;0)** сумму процентов, выплачиваемых со второго по четвертый месяц (\$161,01). В ячейке **G19** я вычислил по формуле **ОБЩДОХОД(0,08/12;10;10000;2;4;0)** выплаты в погашение основной суммы со второго по четвертый месяц (\$2950,08).

**?** Я собираюсь занять \$80 000 и вносить ежемесячные платежи в течение 10 лет. Максимальный ежемесячный платеж, который я могу себе позволить, составляет \$1000. На какую максимальную процентную ставку я могу рассчитывать?

Функция **СТАВКА (RATE)** рассчитает для вас процентную ставку, если известны сумма кредита, продолжительность кредита и размер регулярных выплат. Синтаксис функции: **СТАВКА(#кпер;плт;пс;[бс];[тип];[прогноз])**, где **бс**, **тип** и **прогноз** являются необязательными аргументами. Аргументы **#кпер**, **плт**, **пс**, **бс** и **тип** аналогичны ранее описанным аргументам. Аргумент **прогноз** указывает предполагаемый размер ставки. Как правило, аргумент **прогноз** можно опустить. На листе **Rate** (в файле **Excelfinfunctions.xlsx**) в ячейке **D9** я рассчитал по формуле **=СТАВКА(120;-1000;80000;0;0)** размер ежемесячной ставки — 0,7241%. Предполагается, что выплаты производятся в конце каждого месяца (рис. 10.4).

	D	E	F	G	H
6	BORROWING \$80,000				
7	120 MONTHLY PAYMENTS OF \$1000 PER MONTH				
8	WHAT IS MAX RATE YOU CAN HANDLE?				
9	0.72410%	=RATE(120,-1000,80000,0,0)			
10	IF YOU CAN PAY \$10,000 AT END				
11	WHAT IS MAX RATE YOU CAN HANDLE?				
12	0.818%	=RATE(120,-1000,80000,-10000,0,0)			
13					
14					
15	\$80,000.08	=PV(0.007241,120,-1000,0,0)			
16		CHECK!			
17					
18					

**Рис. 10.4.** Пример с функцией **СТАВКА**

Правильно ли рассчитана ставка, можно проверить в ячейке **D15** по формуле **=ПС(0,007241;120;-1000;0;0)**. Результат \$80 000,08 говорит о том, что выплаты по \$1000 в конце каждого месяца в течение 120 месяцев имеют текущую стоимость \$80 000,08.

Если вы готовы выплатить единовременно \$10 000 в 120-м месяце, то формула **=СТАВКА(120;-1000;80000;-10000;0;0)** в ячейке **D12** рассчитает вам максимальную процентную ставку, что вы можете себе позволить. В этом случае она будет равна 0,818%.

**?** Если я возьму займы \$100 000 под 8% годовых и буду вносить платежи в размере \$10 000 в год, через сколько лет я выплачу кредит?

Если известны сумма кредита, размер платежа по периодам и процентная ставка, то количество периодов, необходимых для погашения ссуды, можно рассчитать с помощью функции КПЕР (NPER). Синтаксис функции: КПЕР(ставка;плт;пс;[бс];[тип]), где бс и тип являются необязательными аргументами.

Если мы вносим платеж в конце года, то по формуле=КПЕР(0,08;-10000;100000;0;0) в ячейке D7 листа Nper (в файле Excelfinfunctions.xlsx) получим 20,91 года (рис. 10.5). Таким образом, для выплаты кредита двадцати лет недостаточно, а за двадцать один год возникнет небольшая переплата. Проверим вычисления с помощью функции ПС в ячейках D10 и D11 и увидим, что в случае платежей по \$10 000 в год в течение 20 лет будет возвращено \$98 181,47, а в случае платежей по \$10 000 в течение 21 года будет возвращено \$100 168,03.

Допустим, вы собираетесь внести единовременный платеж в размере \$40 000 в заключительный период оплаты. Через сколько лет будет возвращен кредит? Расчет по формуле =КПЕР(0,08;-10000;100000;-40000;0) в ячейке D14 показывает, что срок погашения кредита составит 15,90 года. Таким образом, для оплаты кредита 15 лет недостаточно, а за 16 лет возникнет небольшая переплата по кредиту.

	C	D
1		
2		
3		<b>Borrow \$100000 8%</b>
4		<b>ANNUAL PAYMENTS OF \$10,000 PER YEAR</b>
5		<b>END OF YEAR PAYMENT</b>
6		<b>HOW MANY YEARS?</b>
7		<b>20.91237188</b>
8		<b>20 YEARS WILL NOT PAY IT OFF; 21 WILL</b>
9		<b>CHECK</b>
10	<b>20 YEARS</b>	<b>\$98,181.47</b>
11	<b>21 YEARS</b>	<b>\$100,168.03</b>
12		
13		<b>IF WE PAY \$40,000 AT END OF PROBLEM</b>
14		<b>15.9012328</b>
15		<b>15 YEARS WILL NOT PAY IT OFF; 16 YEARS WILL</b>

**Рис. 10.5.** Пример с функцией КПЕР

**?** Я дипломированный бухгалтер и часто работаю со сложными методами расчета амортизации оборудования. Имеются ли в Excel функции для расчета амортизации?

*Амортизация* — это снижение стоимости долгосрочных активов вследствие их использования или устаревания. Чаще всего используются следующие три метода расчета амортизации:

- равномерная (линейная) амортизация, функция АПЛ (SLN);
- амортизация по сумме числа лет срока полезного использования (амортизация суммы лет), функция АСЧ (SYD);
- Метод двойного уменьшения остатка (двойной регрессии), функция ДДОБ (DDB).

Пусть стоимость нового оборудования составляет \$15 000, а его остаточная стоимость после пяти лет эксплуатации определена в \$3000. Как различные методы расчета амортизации распределяют  $\$15\,000 - \$3000 = \$12\,000$  по пяти годам эксплуатации?

- Метод линейной амортизации просто ежегодно уменьшает стоимость оборудования на фиксированную сумму (в данном случае  $\$12\,000/5 = \$2400$ ).
- Если срок полезного использования определен в  $N$  лет, метод амортизации суммы лет определяет отчисления за год 1 по формуле:  $(\text{начальная\_стоимость} - \text{остаточная\_стоимость}) \times (5 - 1 + 1) / (N \times (N+1)/2)$ .  $(N \times (N+1)/2)$  — это сумма чисел 1, 2, ...,  $N$ . В нашем примере к году 1 относится  $5/15$  полного износа, к году 2 —  $4/15$  и т. д.
- Если мы определяем балансовую стоимость как  $\text{начальная\_стоимость} - \text{начисленный\_износ}$ , то для износа в течение  $N$  лет эксплуатации метод двойного уменьшения остатка вычисляет износ за год по формуле:  $2 \times \text{балансовая\_стоимость} / N$ . В нашем примере  $N = 5$ . Поэтому за каждый год амортизация должна составлять 40% балансовой стоимости. К сожалению, метод двойного уменьшения остатка может не позволить точно распределить сумму, равную  $(\text{начальная\_стоимость} - \text{остаточная\_стоимость})$ , на амортизационные расходы. В таких случаях метод двойного уменьшения остатка подстраивает амортизацию за последние несколько лет так, чтобы общая амортизация была равна  $\text{стоимость} - \text{остаточная\_стоимость}$ . В нашем примере метод двойного уменьшения остатка распределяет  $0,4 \times \$15\,000 = \$6000$  на год 1,  $0,4 \times \$9000 = \$3600$  на год 2 и  $0,4 \times (\$5400) = \$2160$  на год 3. Если отнести амортизацию  $0,4 \times (\$3240) = \$1296$  на год 4, полная амортизация превысит \$12 000. Поэтому амортизация за год 4 составит  $12\,000 - (6000 + 3600 + 2160) = \$240$  долларов, а за год 5 будет равна 0.

Excel предлагает функции для каждого из этих типов расчета амортизации:

- Для равномерной амортизации функция АПЛ(нач\_стоимость;ост\_стоимость;время\_эксплуатации), вычисляет ежегодную амортизацию.
- Для метода амортизации по сумме лет функция АСЧ(нач\_стоимость;ост\_стоимость;время\_эксплуатации;период) вычисляет амортизацию за указанный период.
- Для метода двойного уменьшения остатка функция ДДОБ(нач\_стоимость;ост\_стоимость;время\_эксплуатации;период;[коэффициент]) вычисляет амортизацию за указанный период.

После создания имен диапазонов для ячеек C2:C4 на основании диапазона B2:B4 (как показано на рис. 10.6 и в файле Depreciationexamples.xlsx) мы можем вычис-

лить амортизацию с использованием каждого из этих трех методов следующим образом:

- Вычислить линейную амортизацию, скопировав из ячейки E8 в F8:J8 формулы = АПЛ(нач\_стоимость;ост\_стоимость;время\_эксплуатации) (=SLN(Cost,Salvage\_Value,Years)).
- Вычислить амортизацию по сумме лет, скопировав из ячейки E9 в F9:J9 формулы =АСЧ(нач\_стоимость;ост\_стоимость;время\_эксплуатации;период) (=SYD(Cost,Salvage\_value,Years,E7)).
- Вычислить амортизацию по методу двойного уменьшения остатка, скопировав из ячейки E10 в F10:J10 формулы =ДДОБ(нач\_стоимость;ост\_стоимость;время\_эксплуатации;период) (=DDB(Cost,Salvage\_value,Years,E7)).

Обратите внимание, что методы АСЧ и ДДОБ переносят большую часть амортизации на первые годы эксплуатации.

	B	C	D	E	F	G	H	I	J
1									
2	Cost	\$15,000.00							
3	Salvage value	\$3,000.00							
4	Years	5							
5									
6				Year					
7				1	2	3	4	5	Total
8	=SLN(Cost,Salvage_value,Years)		Straight-Line	\$2,400.00	\$2,400.00	\$2,400.00	\$2,400.00	\$2,400.00	\$12,000.00
9	=SYD(Cost,Salvage_value,Years,E7)		Sum-of-Years'-Digits	\$4,000.00	\$3,200.00	\$2,400.00	\$1,600.00	\$800.00	\$12,000.00
10	=DDB(Cost,Salvage_value,Years,E7)		Double-Declining-Balance	\$6,000.00	\$3,600.00	\$2,160.00	\$240.00	\$0.00	\$12,000.00

**Рис. 10.6.** Примеры функций Excel для расчета амортизации

## Задания

Все платежи производятся в конце периода, если не указано иное.

1. Вы только что выиграли в лотерею. В конце каждого года в течение следующих 20 лет вы будете получать по \$50 000. Если стоимость капитала составляет 10% в год, какова текущая стоимость выигрыша?
2. Перпетуитет — это бессрочный аннуитет. Если при сдаче дома в аренду я получаю в начале каждого года \$14 000, какова стоимость этого перпетуитета? Допустим, что стоимость капитала равна 10% в год. Подсказка: воспользуйтесь функцией ПС, и пусть периодов будет много!
3. У меня положено \$250 000 на банковский счет. В конце каждого года в течение следующих 20 лет я буду снимать по \$15 000. Если я зарабатываю 8% в год на своем вложении, сколько денег у меня будет через 20 лет?
4. Я буду класть в банк \$2000 в месяц (в конце каждого месяца) в течение следующих 10 лет. Мои вложения принесут мне 0,8% в месяц. Через 10 лет я хотел бы получить 1 млн долл. Сколько денег я должен внести сейчас?

5. Игрок футбольной премьер-лиги будет получать 15 млн долл. в конце каждого года в течение следующих семи лет. На своих вложениях он может зарабатывать 6% в год. Какова текущая стоимость его будущих доходов?
6. В конце каждого года в течение следующих 20 лет я буду получать следующие суммы:

Годы	Сумма
1–5	\$200
6–10	\$300
11–20	\$400

Найдите с помощью функции ПС текущую стоимость денежных потоков при условии, что стоимость капитала составляет 10%. Подсказка: начните с вычисления стоимости получения \$400 в течение 20 лет, а затем вычтите стоимость получения \$100 в течение 10 лет и т. д.

7. Вы взяли ипотеку на 30 лет в размере \$200 000 под 10% годовых. Вычислите ежемесячный платеж, ежемесячные выплаты по процентам и ежемесячные выплаты в счет погашения основной суммы при условии, что выплаты производятся в конце месяца.
8. Выполните задание 7 при условии, что выплаты производятся в начале месяца.
9. С помощью функции БС определите накопленную стоимость для \$100 в течение трех лет при 7% годовых.
10. У вас есть обязательство выплатить \$1 000 000 через 10 лет. Стоимость капитала составляет 10% в год. Какую сумму следует откладывать в конце каждого года в течение 10 лет для выполнения этого обязательства?
11. Вы планируете купить новый автомобиль за \$50 000. Вам предложили две схемы оплаты:
- 10%-ная скидка на продажную цену автомобиля и далее 60 ежемесячных платежей со ставкой 9% в год;
  - скидка на продажную цену автомобиля отсутствует; 60 ежемесячных платежей со ставкой 2% в год.

Если годовая стоимость вашего капитала составляет 9%, какая схема оплаты выгоднее? Считается, что все платежи производятся в конце месяца.

12. У меня \$10 000 на банковском счете. В начале каждого года в течение следующих 20 лет я буду класть на счет \$4000 под 6% годовых. Сколько денег у меня будет через 20 лет?
13. В случае «шаровой» ипотеки необходимо погасить часть кредита в течение определенного периода, а затем произвести единовременный платеж остав-

шегося кредита. Предположим, вы взяли «шаровой» ипотечный кредит в размере \$400 000 сроком на 20 лет с процентной ставкой 0,5% в месяц. Выплаты в конце каждого месяца в течение 20 лет должны погасить \$300 000 из вашего кредита, затем вы должны выплатить остальные \$100 000. Вычислите свои ежемесячные платежи по этому кредиту.

14. Когда вы берете ипотеку с регулируемой процентной ставкой, ежемесячные выплаты привязаны к индексу ставок (например, к ставкам векселей казначейства США). Допустим, что вы взяли ипотеку размером \$60 000 с регулируемой процентной ставкой на 30 лет (360 ежемесячных платежей). Первые 12 платежей регулируются по текущей ставке векселей казначейства 8%. Со второго по пятый годы ежемесячные платежи устанавливаются по ежемесячной ставке векселей казначейства на начало года плюс 2%. Ставки векселей казначейства на начало годов со второго по пятый представлены ниже:

Начало года	Ставка векселей казначейства, %
2	10
3	13
4	15
5	10

Вычислите ежемесячные платежи с первого по пятый годы и остаток на конец периода для каждого года.

15. Предположим, что вы взяли кредит под 14,4% годовых и выплачиваете его ежемесячно. Если вы пропустили четыре платежа подряд, сколько вам придется выплатить в следующий раз?
16. Через 10 лет вы намерены поменять станок на новый, стоимость которого оцениваете приблизительно в \$80 000. При условии, что ежегодно можно заработать 8% на вкладе, сколько денег необходимо откладывать в конце каждого года для покрытия стоимости нового станка?
17. При покупке мотоцикла вы платите \$1500 сразу и \$182,50 ежемесячно в течение трех лет. Если годовая процентная ставка составляет 18%, какова была первоначальная стоимость мотоцикла?
18. Пусть годовая процентная ставка равна 10%. Вы платите \$200 каждый месяц в течение двух лет, \$300 каждый месяц в течение года и \$400 каждый месяц в течение двух лет. Какова текущая стоимость всех платежей?
19. Вы собираетесь вкладывать \$500 в конце каждого шестимесячного периода в течение пяти лет. Если через пять лет необходимо накопить \$6000, какой должна быть годовая норма прибыли по вкладу?
20. Я собираюсь взять кредит в размере \$2000 и выплачивать его ежеквартально в течение двух лет. Годовая процентная ставка равна 24%. Каков размер каждого платежа?

21. Я взял в кредит \$15 000. Мне предстоит выполнить 48 ежемесячных платежей, и годовая процентная ставка равна 9%. Какова общая сумма выплачиваемых процентов по кредиту?
22. Я собираюсь взять кредит в размере \$5000 и планирую погасить его за 36 месяцев, осуществляя платежи каждый месяц. Годовая процентная ставка равна 16,5%. Через год я дополнительно выплачу \$500 и сокращу срок кредита до двух лет в общей сложности. Каким будет мой ежемесячный платеж в течение второго года действия кредита?
23. При ипотеке с переменной ставкой ежемесячные выплаты зависят от процентных ставок в начале каждого года. Вы взяли \$60 000 в ипотеку на 30 лет. В первый год ежемесячные платежи рассчитываются, исходя из текущей годовой ставки векселей казначейства 9%. Со второго по пятый годы ежемесячные платежи базируются на перечисленных далее годовых ставках векселей казначейства +2%.
- второй год: 10%;
  - третий год: 13%;
  - четвертый год: 15%;
  - пятый год: 10%.

Проблема в том, что для ипотек с переменной ставкой существует положение, гарантирующее, что ежемесячные платежи могут расти год от года максимум на 7,5%. Для компенсации кредитору этого условия заемщик регулирует остаток на конец периода в конце каждого года, исходя из разницы между тем, что заемщик фактически заплатил, и тем, что он должен был заплатить. Рассчитайте ежемесячные платежи с первого по пятый годы ипотеки.

24. Вы можете выбрать: получать ли \$8000 ежегодно, начиная с 62 лет и до конца жизни, или получать \$10 000 ежегодно, начиная с 65 лет и до конца жизни. Если можно заработать 8% в год на вкладах, какой выбор будет правильным?
25. Вы только что выиграли в лотерею и будете получать \$50 000 ежегодно в течение 20 лет. При какой процентной ставке эти выплаты будут эквивалентны получению \$500 000 единовременно?
26. Имеется облигация с купонной выплатой в размере \$50 в конце каждого года в течение 30 лет и выплатой номинальной стоимости в размере \$1000 через 30 лет. Если денежные потоки дисконтируются по годовой ставке 6%, какая цена за облигацию будет справедливой?
27. Вы взяли \$100 000 в ипотеку на 40 лет с ежемесячными выплатами. Годовая процентная ставка составляет 16%. Сколько вы выплатите за время ипотеки? Какую сумму останется выплатить за четыре года до окончания ипотеки?
28. Мне необходимо занять \$12 000. Я могу выплачивать \$500 в месяц, а годовая процентная ставка составляет 4,5%. Сколько месяцев потребуются для погашения кредита?

29. Вы собираетесь взять кредит в размере \$50 000 на 180 месяцев. Годовая процентная ставка по кредиту зависит от вашего кредитного балла следующим образом:

Кредитный балл	Годовая процентная ставка, %
740–750	8,15
720–739	8,45
700–719	8,95
670–699	9,725
640–669	11,225
620–639	12,475

Напишите формулу зависимости ежемесячных платежей от кредитного балла.

30. Вы собираетесь взять автокредит в размере \$40 000. Нужно вычислить ежемесячные платежи и итоговые проценты, выплачиваемые для следующих ситуаций:
- кредит на 48 месяцев под 6,85% годовых;
  - кредит на 60 месяцев под 6,59% годовых.
31. Допустим, есть станок, который стоит \$50 000 и остаточная стоимость которого после десятилетнего износа составит \$5000. Вычислите в течение каждого года равномерную амортизацию, амортизацию по сумме числа лет срока полезного использования и амортизацию по методу двойного уменьшения остатка.



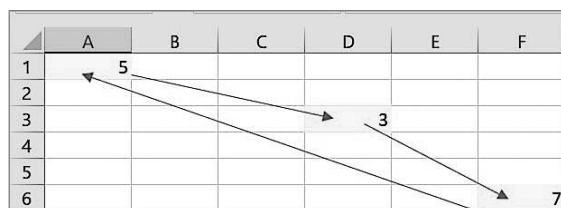
## ГЛАВА 11

# Циклические ссылки

### Обсуждаемые вопросы

- Я часто получаю в Excel сообщение о циклической ссылке. Означает ли это, что я сделал ошибку?
- Как исправить циклическую ссылку?

Если в Microsoft Excel 2019 появляется сообщение о циклической ссылке в книге, это означает, что между двумя или более ячейками книги образовался замкнутый контур — петля, или зависимость. Например, циклическая ссылка возникает если значение в ячейке A1 оказывает влияние на значение в ячейке D3, значение в ячейке D3 влияет на значение в ячейке F6, а значение в ячейке F6 влияет на значение в ячейке A1. На рис. 11.1 представлена схема возникновения циклической ссылки.



**Рис. 11.1.** Возникновение циклической ссылки

Как вы скоро увидите, для устранения циклических ссылок в книге следует выбрать на левой панели вкладки **Файл (File)** команду **Параметры (Options)**, а затем выбрать группу параметров **Формулы (Formulas)** и установить флажок **Включить итеративные вычисления (Enable Iterative Calculation)** в разделе **Параметры вычислений (Calculation Options)**.

## Ответы на вопросы

- ❓ **Я часто получаю в Excel сообщение о циклической ссылке. Означает ли это, что я сделал ошибку?**

Циклическая ссылка обычно образуется на логически непротиворечивом листе, на котором несколько ячеек связаны между собой так, как показано на рис. 11.1.

Рассмотрим элементарный пример, не имеющий простого решения в Excel без создания циклической ссылки.

Небольшая компания с доходом \$1500 и расходами \$1000 собирается потратить 10% своей чистой прибыли на благотворительность. Ставка налога составляет 40%. Сколько денежных средств пойдет на благотворительность? Решение этой задачи представлено в файле *Circular.xlsx* на листе *Sheet1*, показанном на рис. 11.2.

	C	D	E	F	G
1		Charity=10%*After tax profits			
2					
3	Revenues	\$1,500.00			
4	tax rate	0.4			
5	costs	\$1,000.00			
6	charity		0	=0.1*after_tax_profit	
7	before tax profit		500	=Revenues-costs-charity	
8	after tax profit		0	=(1-tax_rate)*before_tax_profit	

**Рис. 11.2.** Циклическая ссылка может возникнуть при расчете налогов

Я начал с того, что присвоил ячейкам D3:D8 соответствующие имена в ячейках C3:C8. В ячейке D3:D5 я ввел размер доходов, налоговой ставки и расходов компании. Для расчета суммы пожертвования, равной 10% от прибыли после уплаты налогов, я ввел формулу `=0,1*после_уплаты_налогов` в ячейку D6. Затем я рассчитал прибыль до уплаты налогов в ячейке D7, вычтя расходы и пожертвование из дохода по формуле `=доход – издержки – благотворительность`. Наконец, я вычислил прибыль после уплаты налогов в ячейке D8 по формуле `=(1 – налоговая_ставка)*до_уплаты_налогов`.

Excel сообщает о циклической ссылке в ячейке D8 (в левом нижнем углу файла *Circular.xlsx*). В чем дело?

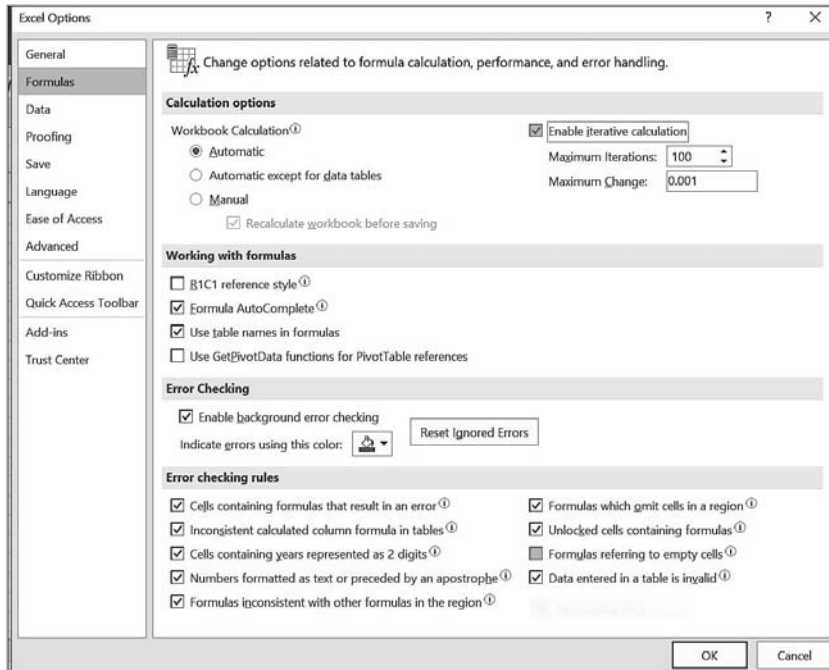
1. Размер пожертвования (ячейка D6) влияет на размер прибыли до уплаты налогов (ячейка D7).
2. Размер прибыли до уплаты налогов (ячейка D7) влияет на размер прибыли после уплаты налогов (ячейка D8).
3. Размер прибыли после уплаты налогов (ячейка D8) влияет на размер пожертвования.

Итак, образовалась петля D6-D7-D8-D6 (показана стрелками на рис. 11.2), вызвавшая сообщение о циклической ссылке. Лист является логически непротиворечивым; все было сделано правильно. Тем не менее, как видно по рис. 11.2, Excel неправильно рассчитывает величину взноса на благотворительность.

### ❓ Как исправить циклическую ссылку?

Исправление циклических ссылок не представляет затруднений. Просто перейдите на вкладку **Файл (File)** в левой части ленты Excel, выберите **Параметры**

(Options) и затем Параметры Excel (Excel Options). На левой панели окна выберите группу параметров Формулы (Formulas) и установите флажок Включить итеративные вычисления (Enable iterative calculation) в разделе Параметры вычислений (Calculation Options), как показано на рис. 11.3.



**Рис. 11.3.** Разрешение циклических ссылок с помощью флажка Включить итеративные вычисления

После установки флажка Включить итеративные вычисления (Enable iterative calculation) программа Excel распознает порожденную циклической ссылкой систему трех уравнений с тремя неизвестными:

$$\text{благотворительность} = 0,1 * (\text{после\_уплаты\_налогов})$$

$$\text{до\_уплаты\_налогов} = \text{доход} - \text{издержки} - \text{благотворительность}$$

$$\text{после\_уплаты\_налогов} = (1 - \text{налоговая\_ставка}) * (\text{до\_уплаты\_налогов})$$

Три неизвестными являются `благотворительность`, `до_уплаты_налогов` и `после_уплаты_налогов`. После установки флажка Включить итеративные вычисления (Enable iterative calculation) Excel выполняет итерации (опыт работы с циклическими ссылками показывает, что потребуется 100 итераций) для поиска решения всех уравнений, порожденных циклической ссылкой. От итерации к итерации значения неизвестных изменяются по сложной математической процедуре (по методу итераций Гаусса — Зейделя). Excel завершает процесс, если изменение значения в любой

ячейке листа от итерации к итерации меньше, чем значение, указанное в поле **Относительная погрешность (Maximum Change)**. По умолчанию оно равно 0,001. При необходимости значение относительной погрешности можно уменьшить, например, до 0,000001. Если его не уменьшить, Excel может, например, присвоить ячейке со значением 5 значение 5,001, что будет неточно. Кроме того, для некоторых сложных листов на разрешение циклической ссылки может потребоваться более 100 итераций. Однако в рассматриваемом примере заикливание было разрешено практически мгновенно (рис. 11.4).

	C	D	E	F
1		<b>Charity=10%*After tax profits</b>		
2				
3	<b>Revenues</b>	<b>\$ 1,500.00</b>		
4	<b>tax rate</b>	<b>0.4</b>		
5	<b>costs</b>	<b>\$ 1,000.00</b>		
6	<b>charity</b>	<b>28.301887</b>		
7	<b>before tax profit</b>	<b>471.69811</b>		
8	<b>after tax profit</b>	<b>283.01887</b>		

**Рис. 11.4.** Разрешение циклической ссылки методом итераций

Размер пожертвования в размере \$28,30 в точности соответствует 10% от прибыли после уплаты налогов (\$283,01). Значения во всех остальных ячейках теперь рассчитываются правильно.

Обратите внимание, что метод итераций в Excel гарантированно работает только при решении *линейных* уравнений. В остальных случаях решение не гарантируется. В примере с налогами разрешение циклических ссылок потребовало решения системы линейных уравнений, поэтому Excel выдал правильный ответ.

Вот еще один пример циклической ссылки. В любой формуле Excel можно обратиться по имени к целому столбцу или к целой строке. Например, по формуле **=СРЗНАЧ(В:В)** рассчитывается средняя величина значений в столбце В. Формула **=СРЗНАЧ(1:1)** вычисляет среднюю величину в строке 1. Это удобно, когда в столбец или строку постоянно добавляются новые данные (например, данные ежемесячных продаж). По этой формуле всегда вычисляется средний объем продаж, и ее не требуется изменять. Проблема заключается в том, что если ввести эту формулу в столбец или строку, на которую она ссылается, то, конечно, создастся циклическая ссылка. Подобные циклические ссылки мгновенно исправляются установкой флажка **Включить итеративные вычисления (Enable iterative calculation)**.

## Задания

1. Компания получила прибыль \$60 000 до выплаты бонусов сотрудникам и уплаты налога штата и федерального налога. Компания выплачивает со-

трудникам бонус в размере 5% от прибыли после уплаты налогов. Налог штата составляет 5% от прибыли (после выплаты бонусов). Федеральный налог составляет 40% от прибыли (после выплаты бонусов и уплаты налога штата). Вычислите суммы, выплаченные в виде бонусов, налога штата и федерального налога.

2. 1 января 2002 г. у меня было \$500. В конце каждого месяца я получаю 2% со своих денег. Проценты за каждый месяц вычисляются от среднего значения баланса на начало и на конец месяца. Сколько денег у меня будет через 12 месяцев?
3. Мой самолет летит по следующему маршруту: Хьюстон — Лос-Анджелес — Сиэтл — Миннеаполис — Хьюстон. На каждом отрезке путешествия расход топлива (в милях на галлон) равен  $40 - 0,02 \cdot (\text{средний расход топлива для маршрута})$ . Здесь средний расход топлива для маршрута равен  $0,5 \cdot (\text{начальное количество топлива для маршрута} + \text{конечное количество топлива для маршрута})$ . Полет начинается в Хьюстоне с 1000 галлонов топлива. Дальность полета на каждом отрезке приведена ниже:

Отрезок	Миля
Хьюстон — Лос-Анджелес	1200
Лос-Анджелес — Сиэтл	1100
Сиэтл — Миннеаполис	1500
Миннеаполис — Хьюстон	1400

Сколько галлонов топлива останется по возвращении в Хьюстон?

4. Распространенным методом распределения затрат для вспомогательных подразделений является метод взаимного распределения затрат. Этот метод может быть реализован с помощью циклических ссылок. Предположим, что в компании Widgetco есть два вспомогательных подразделения: бухгалтерия и отдел консалтинга. Кроме того, в компании есть два производственных подразделения: отдел 1 и отдел 2. Компания приняла решение распределить \$600 000 затрат бухгалтерии и \$116 000 затрат отдела консалтинга между подразделениями компании. Далее указаны соотношения затрат между подразделениями:

	Бухгалтерия	Консалтинг	Отдел 1	Отдел 2
Процент бухгалтерской работы, выполненной для других подразделений компании	0	20	30	50
Процент консультационной работы, выполненной для других подразделений компании	10	0	80	10

Какая часть затрат бухгалтерии и отдела консалтинга должна быть распределена между другими подразделениями компании? Необходимо определить две величины: совокупные затраты, присвоенные бухгалтерии, и совокупные затраты, присвоенные отделу консалтинга. Совокупные затраты, присвоенные бухгалтерии, равны  $\$600\,000 + 0,1 * (\text{совокупные затраты, присвоенные отделу консалтинга})$ , поскольку 10% всей консалтинговой работы было выполнено для бухгалтерии.

Аналогичное уравнение можно написать для совокупных затрат, присвоенных отделу консалтинга. Теперь вы сможете правильно рассчитать распределение затрат бухгалтерии и отдела консалтинга между другими подразделениями компании.

5. Мы начали год с \$200 и в течение года получили еще \$100. Мы также получили в конце года 10% прибыли на основе баланса на начало года. Мы закончили год с \$320. Вычислите баланс на конец года, если проценты начисляются в зависимости от среднего значения баланса на начало и на конец года.

## ГЛАВА 12

# Функции ЕСЛИ, ЕСЛИОШИБКА, ЕСЛИМН, ВЫБОР и ПЕРЕКЛЮЧ

### Обсуждаемые вопросы

- При заказе 500 единиц продукции необходимо заплатить \$3,00 за единицу продукции. При заказе от 501 до 1200 единиц цена за единицу падает до \$2,70. При заказе от 1201 до 2000 единиц стоимость единицы продукции составляет \$2,30. При заказе свыше 2000 единиц — \$2,00. Как написать формулу, выражающую затраты на приобретение как функцию количества приобретенных единиц продукции?
- Я купил 100 акций по цене \$55 за акцию. Для хеджирования риска снижения стоимости акций я купил 60 шестимесячных европейских пут-опционов. Каждый опцион имеет цену исполнения \$45 и стоит \$5 долларов. Как разработать таблицу, в которой указана процентная доходность моего инвестиционного портфеля через шесть месяцев для ряда возможных будущих цен?
- Многие аналитики фондового рынка считают, что торговые правила для торговли по скользящим средним могут опережать рынок. Общее рекомендуемое правило для торговли по скользящим средним гласит: покупай, когда цена акции превысит средний показатель за последние 15 месяцев, и продавай, когда цена акции упадет ниже среднего показателя за последние 15 месяцев. Как это торговое правило может быть сформулировано для биржевого индекса Standard&Poor's 500 (S&P)?
- В игре в кости брошены две кубика. Если сумма чисел при первом броске равна 2, 3 или 12 — вы проиграли. Если сумма чисел при первом броске равна 7 или 11 — вы выиграли. Во всех остальных случаях игра продолжается. Как создать формулу, определяющую состояние игры после первого броска?
- В большинстве гипотетических финансовых отчетов в качестве средства уравнивания активов и обязательств используются денежные средства. Я знаю, что более реалистично использовать для этого долгосрочные обязательства. Как настроить гипотетический отчет с долгосрочными обязательствами в качестве средства уравнивания?
- При копировании формулы с функцией ВПР для вычисления зарплаты сотрудников я получаю множество сообщений об ошибке #Н/Д. Затем при расчете средней зарплаты сотрудников я не могу получить числовой ответ из-за ошибки #Н/Д. Можно ли заменить ошибки #Н/Д пробелом и рассчитать, таким образом, среднюю зарплату?
- Лист содержит ежеквартальные доходы компании Walmart. Как вычислить доход за каждый год и внести его в строку, содержащую продажи в первом квартале соответствующего года?

- Формула с функцией ЕСЛИ может оказаться довольно длинной. Сколько вложенных функций ЕСЛИ можно поместить в ячейку? Каково максимально допустимое количество символов в формуле Excel?
- Вложенные функции ЕСЛИ могут оказаться слишком длинными и сложными. Как новая функция ЕСЛИМН облегчает написание вложенных функций ЕСЛИ?
- Как работает функция ВЫБОР?
- Как работает новая функция ПЕРЕКЛЮЧ?

Перечисленные выше проблемы на первый взгляд имеют мало общего (или вообще не имеют). Тем не менее, создание в Microsoft Excel 2019 модели для каждой из этих ситуаций требует применить функцию ЕСЛИ (IF). Функция ЕСЛИ (и новая функция ЕСЛИМН) — вероятно, самые распространенные функции в Excel. С ее помощью можно проверять условия для значений и формул, подражая (в какой-то степени) условной логике языков программирования, таких как C, C++ и Java.

Первым аргументом функции ЕСЛИ является условие, например  $A1 > 10$ . Если условие истинно, функция возвращает первое из двух приведенных в ней значений; в противном случае возвращается второе значение. Самый простой способ показать эффективность функций ЕСЛИ — это применить их в ответах на все вопросы данной главы.

## Ответы на вопросы

**❓ При заказе 500 единиц продукции необходимо заплатить \$3,00 за единицу продукции. При заказе от 501 до 1200 единиц цена за единицу падает до \$2,70. При заказе от 1201 до 2000 единиц стоимость единицы продукции составляет \$2,30. При заказе свыше 2000 единиц — \$2,00. Как написать формулу, выражающую затраты на приобретение как функцию количества приобретенных единиц продукции?**

Решение этой задачи можно найти на листе Quantity Discount в файле Ifstatement.xlsx. Этот лист показан на рис. 12.1.

Пусть в ячейке A9 содержится объем заказа. Стоимость заказа как функцию от объема заказа можно вычислить, реализовав следующую логику:

- если значение в ячейке A9 меньше или равно 500, стоимость заказа равна  $3 * A9$ ;
- если значение в ячейке A9 больше или равно 501 и меньше 1200, стоимость заказа равна  $2,70 * A9$ ;
- если значение в ячейке A9 больше или равно 1201 и меньше 2000, стоимость заказа равна  $2,30 * A9$ ;
- если значение в ячейке A9 больше 2000, стоимость заказа равна  $2 * A9$ .



	A	B	C	D	E	F
1		cutoff	price			
2	cut1	500	\$ 3.00	price1		
3	cut2	1200	\$ 2.70	price2		
4	cut3	2000	\$ 2.30	price3		
5			\$ 2.00	price4		
6		=IF(A9<=_cut1,price1*A9,IF(A9<=_cut2,price2*A9,IF(A9<=_cut3,price3*A9,price4*A9)))				
7						
8	order quantity	cost	per unit cost			
9	450	\$ 1,350.00	\$ 3.00			
10	900	\$ 2,430.00	\$ 2.70			
11	1450	\$ 3,335.00	\$ 2.30			
12	2100	\$ 4,200.00	\$ 2.00			

**Рис. 12.1.** Моделирование оптовых скидок с помощью функции ЕСЛИ

Сперва присвоим имена из ячеек A2:A4 ячейкам B2:B4 и имена из ячеек D2:D5 ячейкам C2:C5. Затем реализуем указанную выше логику в ячейке B9 в следующей формуле:

=ЕСЛИ(A9<=опт1;цена1\*A9;ЕСЛИ(A9<=опт2;цена2\*A9;ЕСЛИ(A9<=опт3;цена3\*A9;цена4\*A9))).

Для понимания того, как в Excel вычисляется значение по этой формуле, напомним, что функции ЕСЛИ обрабатываются слева направо. Если величина объема заказа меньше или равна 500 (опт1), стоимость определяется как цена1\*A9. Если величина объема заказа больше 500, то проверяется, меньше она или равна 1200. Если это так, то было заказано от 501 до 1200 единиц, и стоимость заказа вычисляется как цена2\*A9. Затем, если объем заказа больше 1200, формула проверяет, меньше он или равен 2000. Если это так, то было заказано от 1201 до 2000 единиц, и стоимость заказа рассчитывается как цена3\*A9. Наконец, если стоимость пока не вычислена, считается, что она равна цена4\*A9. В каждом случае функция ЕСЛИ возвращает правильную стоимость заказа. Обратите внимание, что я ввел в ячейки A10:A12 три других объема заказа и скопировал формулу расчета стоимости в ячейки B10:B12. Для любого объема заказа возвращается правильная общая стоимость.

Для каждой функции ЕСЛИ внутри формулы требуется одна закрывающая скобка.

Формула, содержащая несколько функций ЕСЛИ, называется *формулой с вложенными функциями ЕСЛИ*.

**?** Я купил 100 акций по цене \$55 за акцию. Для хеджирования риска снижения стоимости акций я купил 60 шестимесячных европейских пут-опционов. Каждый опцион имеет цену исполнения \$45 и стоит \$5 долларов. Как разработать таблицу, в которой указана процентная доходность моего инвестиционного портфеля через шесть месяцев для ряда возможных будущих цен?

Перед тем как приступить к решению, рассмотрим некоторые основные понятия из мира финансов. Европейский пут-опцион означает право (но не обязанность) продать акцию в указанный момент времени в будущем (в данном случае через

шесть месяцев) по оговоренной цене (в данном случае \$45). Если цена акции через шесть месяцев составит \$45 или больше, то опцион не будет иметь внутренней стоимости. Предположим, однако, что цена акции через шесть месяцев упала ниже \$45. Тогда можно получить прибыль, если купить акции и тут же продать за \$45. Например, если через шесть месяцев акция продается за \$37, можно получить прибыль  $\$45 - \$37 = \$8$  за акцию, использовав опцион продавца. Таким образом, пут-опционы защищают от падения цен на акции. В этом случае, если через шесть месяцев цена на акцию опустится ниже \$45, опцион накопит некоторую внутреннюю стоимость. Это смягчает снижение стоимости портфеля при снижении стоимости входящих в него акций. Отметим также, что процентная доходность портфеля (при допущении отсутствия выплат дивидендов по акциям в портфеле) вычисляется путем деления величины изменения стоимости портфеля (конечная стоимость портфеля – начальная стоимость портфеля) на начальную стоимость портфеля.

Теперь посмотрим, как процентная доходность портфеля, состоящего из 60 пут-опционов и 100 акций, изменится через шесть месяцев в зависимости от изменения цены на акцию с \$20 до \$65. Решение вы найдете на листе **Hedging** в файле **Ifstatement.xlsx** (рис. 12.2).

	A	B	C	D	E
1					
2	<b>Nputs</b>	<b>60</b>			
3	<b>Nshares</b>	<b>100</b>			
4	<b>exprice</b>	<b>\$ 45.00</b>			
5	<b>pricenow</b>	<b>\$ 55.00</b>			
6	<b>putcost</b>	<b>\$ 5.00</b>			
7	<b>startvalue</b>	<b>\$ 5,800.00</b>			
	<b>final stock price</b>	<b>final put value</b>	<b>final share value</b>	<b>percentage return hedged</b>	<b>percentage return unhedged</b>
8					
9	\$ 20.00	\$ 1,500.00	\$ 2,000.00	-39.7%	-63.6%
10	\$ 25.00	\$ 1,200.00	\$ 2,500.00	-36.2%	-54.5%
11	\$ 30.00	\$ 900.00	\$ 3,000.00	-32.8%	-45.5%
12	\$ 35.00	\$ 600.00	\$ 3,500.00	-29.3%	-36.4%
13	\$ 40.00	\$ 300.00	\$ 4,000.00	-25.9%	-27.3%
14	\$ 45.00	\$ -	\$ 4,500.00	-22.4%	-18.2%
15	\$ 50.00	\$ -	\$ 5,000.00	-13.8%	-9.1%
16	\$ 55.00	\$ -	\$ 5,500.00	-5.2%	0.0%
17	\$ 60.00	\$ -	\$ 6,000.00	3.4%	9.1%
18	\$ 65.00	\$ -	\$ 6,500.00	12.1%	18.2%

**Рис. 12.2.** Пример хеджирования с использованием функции ЕСЛИ

Присвойте имена из ячеек A2:A7 ячейкам B2:B7. Начальная стоимость портфеля ( $100 \times \$55 + 60 \times \$5 = \$5800$ ) указана в ячейке B7. Для расчета конечной стоимости опционов скопируйте формулу `=ЕСЛИ(A9<цена_исполнения;цена_исполнения-A9;0)*число_опционов` из ячейки B9 в ячейки B10:B18. Если цена акции через шесть

месяцев меньше цены исполнения, то стоимость каждого опциона равна цене исполнения – цене акции через шесть месяцев. В противном случае через шесть месяцев стоимость каждого опциона будет равна \$0. Для расчета конечной стоимости акций скопируйте формулу  $\text{=количество\_акций} \times \text{A9}$  из ячейки C9 в ячейки C10:C18. Для расчета процентной доходности хеджированного портфеля скопируйте формулу  $\text{=((C9+B9) - начальная\_стоимость)/начальная\_стоимость}$  из ячейки D9 в ячейки D10:D18. Для расчета процентной доходности портфеля без хеджирования (то есть без пут-опционов) скопируйте формулу  $\text{=(C9 - число\_акций} \times \text{цена\_акции)/ (число\_акций} \times \text{цена\_акции)}$  из ячейки E9 в ячейки E10:E18.

Как видно из рис. 12.2, при цене акции ниже \$45 хеджированный портфель принесет более высокий ожидаемый доход, чем нехеджированный. Обратите внимание, что если цена акции не снизится, более высокий ожидаемый доход принесет нехеджированный портфель. Именно поэтому покупку пут-опционов часто называют *страхованием портфеля*.

**?** Многие аналитики фондового рынка считают, что торговые правила для торговли по скользящим средним могут опережать рынок. Общее рекомендуемое правило для торговли по скользящим средним гласит: покупай, когда цена акции превысит средний показатель за последние 15 месяцев, и продавай, когда цена акции упадет ниже среднего показателя за последние 15 месяцев. Как это торговое правило может быть сформулировано для биржевого индекса Standard&Poor's 500 (S&P)?

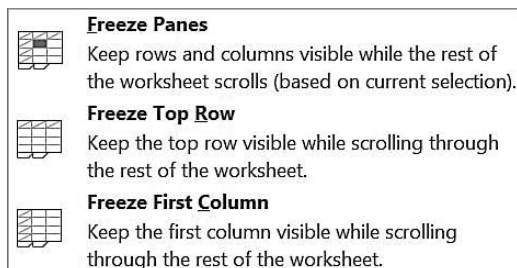
В этом примере я сравниваю эффективность правила для торговли по скользящим средним (в отсутствие транзакционных издержек при покупке и продаже акций) с эффективностью стратегии долгосрочного инвестирования. Преимущество правила для торговли по скользящим средним заключается в том, что оно дает возможность отслеживать тенденции рынка. Правило для торговли по скользящим средним позволяет подниматься на «рынке быков» и продавать акции до того, как на вас обрушится «рынок медведей».

Файл *Matradingrule.xlsx* содержит ежемесячное значение биржевого индекса S&P 500 с января 1871 г. по октябрь 2002 г. Для отслеживания эффективности стратегии торговли по скользящим средним мне необходимо каждый месяц учитывать следующее:

- ☐ Какова средняя величина индекса S&P 500 за последние 15 месяцев?
- ☐ Есть ли у меня акции в начале каждого месяца?
- ☐ Покупаю ли я акции в течение месяца?
- ☐ Продаю ли я акции в течение месяца?
- ☐ Каков денежный поток за месяц (при продаже акции — положительный, при покупке акции — отрицательный, иначе — нулевой)?

В процессе решения этой задачи придется прокрутить много строк. При прокрутке хотелось бы сохранить видимость столбцов A и B, а также заголовков в строке 8.

Для этого в файле `Matradingruletemp.xlsx` поставьте курсор в ячейку `C9`, откройте на ленте вкладку **Вид (View)** и в группе **Окно (Window)** выберите **Закрепить области (Freeze Panes)** (рис. 12.3).



**Рис. 12.3.** Часть группы **Закрепить области**

Строка **Закрепить области (Freeze Panes)** из выпадающего меню **Freeze Panes** сохранит видимыми столбцы **A** и **B** и строки с 6-й по 8-ю при прокрутке остальной части листа. Строки 6–8 будут оставаться видимыми независимо от того, сколько вы прокрутили строк. При прокрутке вправо всегда будут видны столбцы **A** и **B**. **Закрепив первую строку (Freeze Top Row)**, вы будете держать в поле зрения первую строку, прокручивая всю остальную часть листа. Например, закрепив 6-ю строку, вы всегда будете видеть 6-ю строку, неважно насколько вы ушли вниз при прокрутке. Если выбрать в списке строку **Закрепить первый столбец (Freeze First Column)**, при прокрутке листа всегда будет виден крайний слева столбец. Выбрав в списке строки **Снять закрепление областей (Unfreeze Panes)**, вы вернетесь к обычному режиму просмотра.

Файл `Matradingrule.xlsx` (рис. 12.4) содержит формулы для отслеживания эффективности стратегии торговли по скользящим средним. Для решения этой задачи нам нужны несколько функций **ЕСЛИ** и, возможно, логическая функция **И (AND)**. Например, я покупаю акции в течение месяца тогда и только тогда, когда не владею акциями на начало месяца и цена акции в текущем месяце выше, чем скользящее среднее значение цены акции за 15 месяцев. Первым месяцем, для которого можно вычислить скользящее среднее за 15 месяцев, является апрель 1872 г., так что расчеты можно начинать со строки 24.

Предположим, что акция приобретена в апреле 1872 г., поэтому я ввожу **Да** в ячейку `C24`.

- Скопировав формулу `=CPЗНАЧ(B9:B23)` из `D24` в `D25:D1590`, я получил скользящее среднее за 15 месяцев для каждого месяца.
- Скопировав формулу `=ЕСЛИ(И(C24="Нет";B24>D24);"Да";"Нет")` из `E24` в `E25:E1590`, я определил тем самым для каждого месяца, была ли акция S&P приобретена в течение месяца. Помните, что акция покупается тогда и только тогда, когда она не находится во владении в начале месяца и текущая стоимость акции S&P выше ее скользящего среднего значения за 15 месяцев. Обратите

	A	B	C	D	E	F	G
6		<b>S&amp;P</b>				<b>MA profit</b>	<b>\$1,319.75</b>
7		<b>Comp.</b>				<b>Buy and hold profit</b>	<b>\$849.45</b>
8	<b>Date</b>	<b>P</b>	<b>Own?</b>	<b>MA</b>	<b>Buy?</b>	<b>Sell?</b>	<b>Cash flow</b>
18	1871.1	4.59					
19	1871.11	4.64					
20	1871.12	4.74					
21	1872.01	4.86					
22	1872.02	4.88					
23	1872.03	5.04					
24	1872.04	5.18	Yes	4.739	no	no	0
25	1872.05	5.18	Yes	4.788	no	no	0
26	1872.06	5.13	Yes	4.833	no	no	0
27	1872.07	5.1	Yes	4.868	no	no	0
28	1872.08	5.04	Yes	4.892	no	no	0
29	1872.09	4.95	Yes	4.904	no	no	0
30	1872.1	4.97	Yes	4.913	no	no	0
31	1872.11	4.95	Yes	4.929	no	no	0
32	1872.12	5.07	Yes	4.939	no	no	0
33	1873.01	5.11	Yes	4.955	no	no	0
34	1873.02	5.15	Yes	4.989	no	no	0
35	1873.03	5.11	Yes	5.023	no	no	0
36	1873.04	5.04	Yes	5.048	no	yes	5.04

**Рис. 12.4.** Стратегия торговли по скользящим средним превосходит стратегию долгосрочного инвестирования

внимание на аргумент И в формуле. Она содержит два условия (допускается более двух условий), разделенных точкой с запятой. Если оба условия выполнены, то функция возвращает значение **Да**; в противном случае она возвращает **Нет**. Текст в функции ЕСЛИ необходимо заключить в кавычки (" ").

- Скопировав формулу =ЕСЛИ(И(C24="Да";B24<D24);"Да";"Нет") из F24 в F25:F1590, я определил тем самым для каждого месяца, продана ли акция S&P. Акция продается тогда и только тогда, когда она находится во владении в начале месяца и текущая стоимость акции S&P ниже ее скользящего среднего за 15 месяцев. Акция S&P продана впервые в апреле 1873 г.

## ПРИМЕЧАНИЕ

Проще всего скопировать формулу из D24 в D25:D1590 так: поместите курсор на правый нижний угол ячейки D24 (курсор выглядит как перекрестие) и затем дважды щелкните левой кнопкой мыши. По двойному щелчку формула копируется во все ячейки столбца со значениями в столбце слева. Этим же приемом можно воспользоваться для копирования формул в несколько столбцов.

- Для любого месяца до октября 2002 г. в случае покупки акции S&P в течение месяца денежный поток эквивалентен стоимости купленной акции S&P со знаком «минус». При продаже акции S&P в течение месяца денежный поток эквивалентен стоимости S&P. Иначе денежный поток равен 0. В октябре 2002 г. акция S&P, если она находилась во владении, продается для получения

суммы ее стоимости. Следовательно, копирование формулы =ЕСЛИ(E24="Да";-B24;ЕСЛИ(F24="Да";B24;0)) из G24 в G25:G1589 даст денежный поток для всех месяцев до октября 2002 г. Формула =ЕСЛИ(C1590="Да";B1590;0) в ячейке G1590 вычисляет сумму, вырученную от продажи любой акции, которой я обладал в начале последнего месяца.

- В ячейке G6 я вычислил общую прибыль от применения стратегии торговли по скользящим средним по формуле=СУММ(G24:G1590). Стратегия торговли по скользящим средним принесла прибыль в размере \$1319,75.
- Прибыль от покупки по стратегии долгосрочного инвестирования равна стоимости S&P в октябре 2002 г. минус стоимость S&P в апреле 1872 г. Я вычислил прибыль от применения стратегии долгосрочного инвестирования в ячейке G7 по формуле =B1590 – B24.

Как видите, прибыль при долгосрочном инвестировании составляет \$849,45, что намного меньше, чем прибыль при использовании правила торговли по скользящим средним. Конечно, это решение не учитывает транзакционные издержки при купле-продаже акций. Если транзакционные издержки велики, они могут свести на нет сверхприбыль от применения стратегии торговли по скользящим средним.

**❓ В игре в кости брошены две кубика. Если сумма чисел при первом броске равна 2, 3 или 12 — вы проиграли. Если сумма чисел при первом броске равна 7 или 11 — вы выиграли. Во всех остальных случаях игра продолжается. Как создать формулу, определяющую состояние игры после первого броска?**

Факт проигрыша в кости, если сумма чисел при первом броске равна 2, 3 или 12, можно смоделировать, объединив функцию ИЛИ с функцией ЕСЛИ. В файле Ifstatement.xlsx на листе Craps в ячейку B5 (рис. 12.5) я ввел формулу =ЕСЛИ(ИЛИ(A5=2; A5=3; A5=12); "проигрыш"; ЕСЛИ(ИЛИ(A5=7;A5=11); "выигрыш"; "продолжать")). Затем я скопировал эту формулу из B5 в B6:B7. При вводе в ячейку A5 значений 2, 3 или 12 по формуле вычисляется значение проигрыш. При вводе 7 или 11 вычисляется значение выигрыш, при вводе любого другого значения вычисляется значение продолжайте.

	A	B	C
1			
2	Craps		
3			
4	Toss	Result	
5	3	lose	=IF(OR(A5=2,A5=3,A5=12),"lose",IF(OR(A5=7,A5=11),"win","keep going"))
6	7	win	
7	9	keep going	

**Рис. 12.5.** Моделирование первого броска при игре в кости с помощью функции ЕСЛИ

**❓ В большинстве гипотетических финансовых отчетов в качестве средства уравнивания активов и обязательств используются денежные средства. Я знаю, что более реалистично использовать для этого долгосрочные обя-**



# зательства. Как настроить гипотетический отчет с долгосрочными обязательствами в качестве средства уравнивания?

Гипотетический отчет является предсказанием финансового будущего компании. Для прогнозирования создается гипотетический баланс и отчет о прибылях и убытках. Балансовый отчет представляет собой срез активов и финансовых обязательств компании на любой момент времени; отчет о прибылях и убытках содержит данные о финансовом положении компании на любой момент времени. Гипотетические отчеты помогают компании определять будущие долгосрочные обязательства; они также являются ключевыми частями моделей, используемых фондовыми аналитиками для определения справедливой цены на акции. В файле Proforma.xlsx для компании я сгенерировал свободные денежные потоки на следующие четыре года. Балансовый отчет представлен на рис. 12.6, а отчет о прибылях и убытках — на рис. 12.7.

	B	C	D	E	F	G	H
1							
2							
3	Sales growth	SG	0.02		CA/Sales		0.15
4	Initial sales	IS	1,000.00		CL/Sales		0.07
5	Interest rate on debt	IRD	0.10		NFA/Sales		0.60
6	Dividend payout	DIV	0.05		GFA/Sales		0.90
7	Tax rate	TR	0.53				
8	COGS/Sales	COGS	0.75				
9	Depreciation rate	DEP	0.10				
10	Liquid asset interest rate	LAIR	0.09				
11	Balance sheet						
12			0.00	1.00	2.00	3.00	4.00
13	Cash and marketable securities			0.00	0.00	0.00	52.56
14	Current assets		150.00	153.00	156.06	159.18	162.36
15	Gross fixed assets		900.00	1,001.33	1,115.29	1,243.29	1,386.93
16	Acc. dep.		300.00	400.13	511.66	635.99	774.68
17	Net fixed assets		600.00	601.20	603.62	607.30	612.24
18	Total assets		750.00	754.20	759.68	766.48	827.17
19							
20	Current liabilities		70.00	71.40	72.83	74.28	75.77
21	Debt		180.00	118.96	59.33	1.80	0.00
22	Stock		400.00	400.00	400.00	400.00	400.00
23	Retained earnings		100.00	163.84	227.52	290.39	351.40
24	Equity		500.00	563.84	627.52	690.39	751.40
25	Total liabilities		750.00	754.20	759.68	766.48	827.17
26							

**Рис. 12.6.** Гипотетический балансовый отчет

В столбце D содержится информация об актуальном состоянии компании (в год 0). Основные положения таковы:

- рост продаж (sales growth, SG) составляет 2% в год;
- первоначальные продажи составляют \$1000;
- процентная ставка по долговым обязательствам (interest rate on debt, IRD) равна 10%;

	B	C	D	E	F	G	H
27	Income statement		0.00	1.00	2.00	3.00	4.00
28	Sales		1,000.00	1,020.00	1,040.40	1,061.21	1,082.43
29	Cost of goods sold		700.00	765.00	780.30	795.91	811.82
30	Depreciation			100.13	111.53	124.33	138.69
31	Operating income			154.87	148.57	140.97	131.92
32	Interest income			0.00	0.00	0.00	4.73
33	Interest expense			11.90	5.93	0.18	0.00
34	Income before taxes			142.97	142.64	140.79	136.65
35	Taxes			75.77	75.60	74.62	72.42
36	Net income			67.20	67.04	66.17	64.22
37							
38	Beg. retained earnings			100.00	163.84	227.52	290.39
39	Dividends			3.36	3.35	3.31	3.21
40	Ending retained earnings			163.84	227.52	290.39	351.40

**Рис. 12.7.** Гипотетический отчет о прибылях и убытках

- выплачиваемые дивиденды (DIV) составляют 5% от чистой прибыли;
- налоговая ставка (TR) равна 53%;
- себестоимость реализованной продукции (cost of goods sold, COGS) составляет 75% от объема продаж;
- норма амортизации (DEP) равна 10% от общей суммы неликвидных активов;
- процентная ставка по ликвидным активам (liquid asset interest rate, LAIR) равна 9%;
- текущие активы составляют 15% от объема продаж;
- текущие обязательства составляют 7% от объема продаж.

Я присвоил имена из ячеек C3:C10 ячейкам диапазона D3:D10. Для каждого года  $t$  основы финансовых вычислений и бухучета подразумевают отношения, реализованные в следующих формулах.

- **Формула 12.1.** Объем продаж за год  $(t + 1) = (\text{объем продаж за год } t) \times (1 + SG)$ . Для вычисления объема продаж за каждый год я скопировал формулу  $=D28*(1+SG)$  из E28 в F28:H28.
- **Формула 12.2.**  $COGS \text{ в год } t = COGS \times (\text{объем продаж за год } t)$ . COGS для каждого года я вычислил по формуле  $=COGS*E28$ , скопировав ее из E29 в F29:H29.
- **Формула 12.3.** Если величина активов за год  $t >$  величины обязательств за год  $t$ , величина долгосрочных обязательств за год  $t$  должна быть установлена равной (совокупные активы за год  $t$  – текущие обязательства за год  $t$  – собственный капитал компании за год  $t$ ). В противном случае величина долгосрочных обязательств за год  $t$  равна 0. Величину долгосрочных обязательств для каждого года я рассчитал, скопировав формулу  $=\text{ЕСЛИ}(E18>E20+E24;E18-E20-E24;0)$  из E21 в F21:H21. Если величина совокупных активов за год  $t$  больше, чем общая сумма обязательств за год  $t$ , то величина долгосрочных обязательств за год  $t$  должна быть установлена равной (совокупные активы за год  $t$  – текущие обязательства за год  $t$  – собственный капитал компании за год  $t$ ). Это



уравнивает, или уравнивает, активы и обязательства. В противном случае величина долгосрочных обязательств за год  $t$  принимается равной 0. Тогда для уравнивания активов и обязательств используются суммы в виде денежных средств и ликвидных ценных бумаг за год  $t$ .

- **Формула 12.4.** Текущие обязательства за год  $t = (CL/\text{продажи}) \times (\text{объем продаж за год } t)$ . В ячейке E20 я использовал формулу `=H$4*E28`. Чтобы вычислить величину текущих обязательств для каждого года, я скопировал ее из E20 в F20:H20.
- **Формула 12.5.** Собственный капитал компании за год  $t = \text{акционерный капитал за год } t + \text{нераспределенная прибыль за год } t$ . Для расчета собственного капитала компании я скопировал формулу `=СУММ(E22:E23)` из E24 в F24:H24.
- **Формула 12.6.** Если величина долгосрочных обязательств за год  $t$  больше 0, то сумма денежных средств и ликвидных ценных бумаг для года  $t$  принимается равной 0. В противном случае эта сумма для года  $t$  равна МАКС (0; общая сумма обязательств для года  $t$  – текущие активы для года  $t$  – балансовая стоимость основного капитала для года  $t$ ). В E13:H13 я вычислил сумму денежных средств и ликвидных ценных бумаг для каждого года, скопировав формулу `=ЕСЛИ(E21>0;0;МАКС(0;E25-E14-E17))` из E13 в F13:H13. Если величина долгосрочных обязательств за год  $t$  больше 0, то нет необходимости в использовании денежных средств и ликвидных ценных бумаг для уравнивания активов и обязательств для года  $t$ . В этом случае я устанавливаю сумму денежных средств и ликвидных ценных бумаг для года  $t$  равной 0. Иначе я устанавливаю эту сумму для года  $t$  равной (общая сумма обязательств для года  $t$  – текущие активы для года  $t$  – балансовая стоимость основного капитала для года  $t$ ). Это сбалансирует активы и обязательства, если активы для года  $t$  (без денежных средств и ликвидных ценных бумаг) меньше обязательств для года  $t$ . Если долгосрочные обязательства не уравнивают активы и обязательства, в качестве средства уравнивания активов и обязательств создаются легко-реализуемые активы.
- **Формула 12.7.** Затраты на выплату процентов для года  $t = (\text{величина долгосрочных обязательств за год } t) \times \text{IRD}$ . Я ввел в E33 формулу `=IRD*E21` для расчета затрат на выплату процентов в первом году и скопировал ее в F33:H33 для расчета процентов во 2–4-й годы.
- **Формула 12.8.** Доход в виде процентов для года  $t = (\text{сумма денежных средств и ликвидных ценных бумаг для года } t) \times \text{LAIR}$ . Я скопировал формулу `=LAIR*E13` из E32 в F32:H32, чтобы вычислить доход от процентов.
- **Формула 12.9.** Операционный доход для года  $t = \text{продажи для года } t - \text{себестоимость реализованной продукции для года } t - \text{амортизационные отчисления для года } t$ . Для расчета операционного дохода я скопировал формулу `=E28-E29-E30` из E31 в F31:H31.
- **Формула 12.10.** Дивиденды для года  $t = (\text{чистая прибыль для года } t) \times \text{DIV}$ . Для расчета дивидендов за каждый год я скопировал формулу `=E36*DIV` из E39 в F39:H39.

- **Формула 12.11.** Нераспределенная прибыль на начало года  $t + 1$  = нераспределенная прибыль на конец года  $t$ . Я вычислил нераспределенную прибыль на начало каждого года, скопировав формулу =E40 из F38 в G38:H38.
- **Формула 12.12.** Нераспределенная прибыль на конец года  $t$  = нераспределенная прибыль на начало года  $t$  + чистая прибыль за год  $t$  – дивиденды за год  $t$ . Я вычислил нераспределенную прибыль на конец каждого года, скопировав формулу =E38+E36-E39 из E40 в F40:H40.
- **Формула 12.13.** Прибыль до уплаты налогов за год  $t$  = операционный доход за год  $t$  – затраты на выплату процентов для года  $t$  + доход в виде процентов для года  $t$ . Я вычислил прибыль до уплаты налогов, скопировав формулу =E31-E33+E32 из E34 в F34:H34.
- **Формула 12.14.** Налоги за год  $t$  = (прибыль до уплаты налогов за год  $t$ )  $\times$  TR. Я вычислил налоги за каждый год по формуле =TR\*E34, скопировав ее из E35 в F35:H35.
- **Формула 12.15.** Чистая прибыль за год  $t$  = (прибыль до уплаты налогов за год  $t$ ) – (налоги за год  $t$ ). Для вычисления чистой прибыли за каждый год я скопировал формулу =E34-E35 из E36 в F36:H36.
- **Формула 12.16.** Общая сумма неликвидных активов за год  $t$  = балансовая стоимость основного капитала за год  $t$  + накопленная амортизация за год  $t$ . Я вычислил общую сумму неликвидных активов, скопировав формулу =E17+E16 из E15 в G15:H15.
- **Формула 12.17.** Амортизационные отчисления за год  $t$  = (балансовая стоимость основного капитала за год  $t$ )  $\times$  DEP. Я вычислил амортизационные отчисления за каждый год, скопировав формулу =DEP\*E15 из E30 в F30:H30.
- **Формула 12.18.** Накопленная амортизация за год  $t$  = накопленная амортизация за год  $(t - 1)$  + амортизационные отчисления за год  $t$ . Я вычислил накопленную амортизацию для каждого года, скопировав формулу =D16+E30 из E16 в F16:H16.
- **Формула 12.19.** Балансовая стоимость основного капитала за год  $t$  = общая сумма неликвидных активов за год  $t$  – накопленная амортизация за год  $t$ . В строке 17 для расчета балансовой стоимости основного капитала скопируйте формулу =E15-E16 из ячейки E17 в ячейки F17:H17.
- **Формула 12.20.** Совокупные активы за год  $t$  = денежные средства и ликвидные активы за год  $t$  + текущие активы за год  $t$  + балансовая стоимость основного капитала за год  $t$ . Сложив ликвидные активы, текущие активы и балансовую стоимость основного капитала, я вычислил совокупные активы, скопировав формулу =СУММ(E13;E14;E17) из E18 в F18:H18.
- **Формула 12.21.** Общая сумма обязательств за год  $t$  = текущие обязательства за год  $t$  + долгосрочные обязательства за год  $t$  + собственный капитал компании за год  $t$ . Я вычислил общую сумму обязательств для каждого периода, скопировав формулу =СУММ(E20;E21;E24) из E25 в F25:H25. Для каждого года баланс сойдется благодаря долгосрочным обязательствам и ликвидным активам.

В формулах 12.3 и 12.6 необходимо использовать функцию **ЕСЛИ**. На этом листе также содержатся циклические ссылки. (О том, как разрешать циклические ссылки, сказано в главе 11.) Например, циклическая ссылка создается следующими зависимостями:

- денежные средства для года  $t$  влияют на совокупные активы для года  $t$ ;
- совокупные активы для года  $t$  влияют на долгосрочные обязательства для года  $t$ ;
- долгосрочные обязательства для года  $t$  влияют на денежные средства для года  $t$ .

Для разрешения циклических ссылок на вкладке **Файл (File)** выберите команду **Параметры (Options)**, затем **Формулы (Formulas)** и установите флажок **Включить итеративные вычисления (Enable Iterative Calculations)**. Как показано в главе 11, после этого циклические ссылки разрешаются автоматически. Обратите внимание, что для каждого года  $t$  совокупные активы в строке 18 равны общей сумме обязательств в строке 25. Этот пример демонстрирует, насколько эффективны функции **ЕСЛИ** и циклические ссылки при совместной работе.

**❓ При копировании формулы с функцией ВПР для вычисления зарплаты сотрудников я получаю множество сообщений об ошибке #Н/Д. Затем при расчете средней зарплаты сотрудников я не могу получить числовой ответ из-за ошибки #Н/Д. Можно ли заменить ошибки #Н/Д пробелом и рассчитать, таким образом, среднюю зарплату?**

В файле **Errortrap.xlsx** (рис. 12.8), в ячейках D3:E7 на Лист 1, хранятся зарплаты и имена пяти сотрудников. В ячейках D11:D15 находится другой список из пяти имен сотрудников (некоторые повторяются), и я вычисляю их зарплаты, копируя формулу **=ВПР(D11;\$D3:\$E7;2;ЛОЖЬ)** из E11 в E12:E15. К сожалению, при этом в ячейках E13 и E14 появляется сообщение об ошибке **#Н/Д**. Сокращение **Н/Д** означает «значение недоступно». Excel возвращает сообщение об ошибке **#Н/Д**, если по формуле не может быть вычислен соответствующий результат. Поскольку имена **JR** и **Josh** не указаны в списке зарплат, функция **ВПР** не может вернуть для них значения зарплат. Аналогично сообщение об ошибке **#Н/Д** появляется в ячейке E16 при вычислении средней зарплаты по формуле **=СРЗНАЧ(E11:E15)**. Многие вручную заменяют ошибки **#Н/Д** пробелами для правильного расчета среднего значения с помощью функции **СРЗНАЧ**, которая игнорирует пробелы. Но есть лучший способ! Можно заменить ошибки каким-нибудь символом (например, пробелом или нулем) с помощью функции **ЕСЛИОШИБКА (IFERROR)**. Синтаксис функции **ЕСЛИОШИБКА**: **ЕСЛИОШИБКА(значение;значение\_если\_ошибка)**. Далее я покажу вам, как это делается.

Первым аргументом функции является формула, которую требуется вычислить, второй аргумент — это значение, записываемое в ячейку в случае, если формула возвращает ошибку. (О других частых ошибках, таких как **#ДЕЛ/0**, **#ИМЯ?**, **#ЧИСЛО!**, **#ССЫЛКА!**, **#ЗНАЧ!**, см. далее в этом разделе.) Таким образом, по формуле **=ЕСЛИОШИБКА(ВПР(D11;\$D3:\$E7;2;ЛОЖЬ);" ")**, скопированной из F11 в F12:F15,

для каждого фактического сотрудника зарплата будет вычислена правильно, а для людей, не являющихся фактическими сотрудниками, в соответствующие ячейки будет записан пробел. (Поэтому ячейки F13:F14 выглядят как пустые. Вы можете проверить формулы в этих ячейках.) Теперь средняя зарплата для всех сотрудников из списка вычисляется по формуле =СРЗНАЧ(F11:F15) правильно.

	C	D	E	F	G	H	I	J
1								
2			salary					
3		Jane	40					
4		Jack	60					
5		Jill	70					
6		Erica	34					
7		Adam	120					
8								
9								
10		Name	Salary	errortrapped				
11		Erica	34	34	=IFERROR(VLOOKUP(D11,\$D\$3:\$E\$7,2,FALSE)," ")			
12		Adam	120	120				
13		JR	#N/A					
14		Josh	#N/A					
15		Jill	70	70				
16		average	#N/A	74.66667				
17								
18	aggregate		74.66667	=AGGREGATE(1,6,E11:E15)				
19	function							
20	works!							

**Рис. 12.8.** Перехват ошибок в формуле с помощью функции ЕСЛИОШИБКА

Начиная с Microsoft Excel 2010 в Excel представлена функция АГРЕГАТ (AGGREGATE), которая при расчетах игнорирует строки с ошибками. Синтаксис функции: АГРЕГАТ(номер\_функции;параметры;массив). Аргумент номер\_функции — это код между 1 и 19, присвоенный функциям в вычислениях. Например 1 = СРЗНАЧ, а 9 = СУММ. Полный список доступных функций доступен в Справка по этой функции (Help On This Function) в окне Вставка функции (Function Wizard). Аргумент параметры указывает типы ячеек, которые должны игнорироваться при вычислении (см. таблицу ниже). Если значение аргумента параметры равно 6, то игнорируются ячейки с сообщениями об ошибках. Аргумент массив — это диапазон ячеек, используемых в вычислениях. Как показано в ячейке E18, средняя зарплата для перечисленных зарплат рассчитана по формуле =АГРЕГАТ(1;6;E11:E15) правильно (74,66667).

Параметр	Содержимое игнорируемых ячеек
0 или опущен	Вложенные функции ПРОМЕЖУТОЧНЫЕ.ИТОГИ и АГРЕГАТ
1	Скрытые строки и вложенные функции ПРОМЕЖУТОЧНЫЕ.ИТОГИ и АГРЕГАТ
2	Ошибки и вложенные функции ПРОМЕЖУТОЧНЫЕ.ИТОГИ и АГРЕГАТ

Параметр	Содержимое игнорируемых ячеек
3	Скрытые строки, ошибки и вложенные функции ПРОМЕЖУТОЧНЫЕ. ИТОГИ и АГРЕГАТ
4	Нет
5	Скрытые строки
6	Сообщения об ошибках
7	Скрытые строки и сообщения об ошибках

	B	C	D	E	F	G
1	Error Examples					
2						
3	5	0	#DIV/0!	=C3/D3		
4						
5						
6	5	jack	#VALUE!	=c6+d6		
7			#NAME?	=sum(sales)		
8			#NUM!	=sqrt(-1)		
9			#REF!	=SUM(A1:A3);DELETED COLUMN A		

**Рис. 12.9.** Примеры сообщений Excel об ошибках

В файле Errortypes.xlsx, показанном на рис. 12.9, содержатся примеры сообщений о других частых ошибках.

- В ячейке D3 при вычислении формулы `=C3/B3` появилось сообщение об ошибке `#ДЕЛ/0`, поскольку была попытка деления на 0.
- В ячейке D6 при вычислении формулы `=B6+C6` появилось сообщение об ошибке `#ЗНАЧ!`, поскольку значение `jack` имеет тип данных, не соответствующий введенной формуле (`jack` — это текст).
- В ячейке D7 при вычислении формулы `=СУММ(sales)` появилось сообщение об ошибке `#ИМЯ?`, указывающее на то, что имя диапазона `sales`, на которое есть ссылка в формуле, не определено.
- В ячейке D8 при вычислении формулы `=КОРЕНЬ(-1)` появилось сообщение об ошибке `#ЧИСЛО!`, сигнализирующее о вводе неподходящего аргумента в функцию. Excel выводит это сообщение об ошибке при попытке извлечь квадратный корень из отрицательного числа.
- В ячейке C9 я ввел формулу `=СУММ(A1:A3)` и удалил столбец A — это вызвало ошибку `#ССЫЛКА!`, поскольку указанные в формуле ячейки (`A1:A3`) отсутствуют на листе.

Любое из этих сообщений об ошибке можно заменить с помощью функции ЕСЛИОШИБКА каким-либо числом или текстовой строкой.

❓ Лист содержит ежеквартальные доходы компании Walmart. Как вычислить доход за каждый год и внести его в строку , содержащую продажи в первом квартале соответствующего года?

В файле `Walmartrev.xlsx` содержатся данные о ежеквартальных доходах (в млн долл.) компании Walmart (рис. 12.10). В строках 6, 10, 14 и т. д. приведены доходы за первый квартал каждого года. В каждой из этих строк в столбце Е я хочу рассчитать суммарный доход за год. В остальных строках ячейки столбца Е должны быть пустыми. Можно ввести в ячейку Е6 формулу `СУММ(D6:D9)` и скопировать ее в Е10, Е14, Е18 и т. д., но можно поступить умнее. Объединив функцию ЕСЛИ с двумя чудесными функциями Excel — СТРОКА (ROW) и ОСТАТ (MOD) — я могу ввести формулу лишь однажды. Функция СТРОКА(ссылка на ячейку) возвращает номер строки указанной ячейки. Функция `=СТРОКА(A6)` возвращает значение 6; если текущей является строка 6, функция `=СТРОКА()` также возвращает значение 6. Функция ОСТАТ(число;делитель) возвращает остаток от деления значения аргумента число на делитель. Например, формула `=ОСТАТ(9;4)` возвращает 1, а формула `=ОСТАТ(6;3)` возвращает 0. Обратите внимание, что я хочу , чтобы формула работала только в строках, номера которых при делении на 4 дают в остатке 2. Таким образом, формула `=ЕСЛИ(ОСТАТ(СТРОКА());4)=2;СУММ(D6:D9);"` ), скопированная из Е6 в Е7:Е57, гарантирует, что доход за текущий год будет добавлен только в строки, номера которых при делении на 4 дают в остатке 2. Это означает , что ежегодный доход будет указан только в первом квартале каждого года, что и требовалось сделать.

Walmart Revenues			
Year	Quarter	revenue	Annual revenues
1991	1	9,281	43886.76096
1991	2	10,340	
1991	3	10,628	
1991	4	13,639	
1992	1	11,649	55483.59296
1992	2	13,028	
1992	3	13,684	
1992	4	17,122	
1993	1	13,920	67344.29593
1993	2	16,237	
1993	3	16,827	
1993	4	20,361	
1994	1	17,686	82493.89093
1994	2	19,942	
1994	3	20,418	
1994	4	24,448	
1995	1	20,440	93627
1995	2	22,723	
1995	3	22,914	

Рис. 12.10. Данные о годовом доходе компании Walmart

**?** Формула с функцией ЕСЛИ может оказаться довольно длинной. Сколько вложенных функций ЕСЛИ можно поместить в ячейку? Каково максимально допустимое количество символов в формуле Excel?

В Microsoft Excel 2016 и 2019 можно поместить в ячейку до 64 вложенных функций ЕСЛИ. В предыдущих версиях Excel можно было использовать максимум семь вложенных функций ЕСЛИ. В Microsoft Excel 2010 (и Microsoft Excel 2007) ячейка может содержать до 32 000 символов.

**?** Вложенные функции ЕСЛИ могут оказаться слишком длинными и сложными. Как новая функция ЕСЛИМН облегчает написание вложенных функций ЕСЛИ?

Длинные формулы с вложенными функциями ЕСЛИ крайне неудобны в записи, так как, много раз вводя слово ЕСЛИ, приходится постоянно держать в голове, сколько же вам понадобится закрывающих скобок. К счастью, если у вас стоит Office 365 или Excel 2019, то у вас есть доступ к замечательной функции ЕСЛИМН, которая облегчает ввод длинных формул с ЕСЛИ. В файле IFSfinal.xlsx (рис. 12.11 и 12.12) показано применение функции ЕСЛИМН. В сущности, она позволяет вам избежать многократного ввода слова ЕСЛИ и закрыть формулу единственным набором скобок.

Функция ЕСЛИМН применяется, чтобы определить результат бросания костей после первого встряхивания. Обратите внимание, что нам потребуется слово ИСТИНА перед последней вероятностью из списка. В ячейке E15 мы задаем результат бросания костей после первого встряхивания формулой =ЕСЛИМН(ИЛИ(D15=2;D15=3;D15=12);"ПРОИГРЫШ";ИЛИ(D15=7;D15=11);"ВЫИГРЫШ";ИСТИНА;"продолжайте").

	A	B	C	D	E	F	G	H
2								
3				Shirt Code	1st hyphen	2nd hyphen	Size	What I Want
4	XXL	Really Big		1023-XL-1539	5	8	XL	Extra Large
5	XL	Extra Large		982-L-555	4	6	L	Large
6	L	Large		1257-XXL-423	5	9	XXL	Really Big
7	M	Medium		1072-M-863	5	7	M	Medium
8	S	Small		464-S-1526	4	6	S	Small
9	XS	Extra Small		746-XS-791	4	7	XS	Extra Small
10	XXS	Really Small		791-XL-136	4	7	XL	Extra Large
11				831-L-376	4	6	L	Large
12				1139-XXL-108	5	9	XXL	Really Big

**Рис. 12.11.** Применение функции ЕСЛИМН для определения результата бросания костей после первого встряхивания.

Если выпадает 2, 3 или 12, то результат — Проигрыш; если 7 или 11 — Выигрыш, в другом случае слово ИСТИНА заставит выдать последний результат в списке: про-



должайте. Заметьте, что эти формулы потребовали только одно слово **ЕСЛИ** и одну закрывающую скобку.

Рисунок 2.12 показывает, как использование функции **ЕСЛИМН** упростит первую формулу с **ЕСЛИ**, приведенную в этой главе. Чтобы определить цену за единицу продукции с помощью **ЕСЛИМН**, скопируем из D9 в D10:D12 формулу **=ЕСЛИМН(A9<=\_опт1; цена1, A9<=\_опт2; цена2; A9<=\_опт3; цена3; ИСТИНА; цена4)**.

Если объем заказа меньше или равен 500, формула возвратит значение цены за единицу 3\$, если объем заказа находится между 501 и 1200, возвратится значение 2,7\$ за единицу, если объем заказа между 1201 и 2000, возвращается цена 2,3\$ за единицу, в ином случае (из-за слова **ИСТИНА**) возвращается значение цены в 2\$. Опять же, заметьте: слово **ЕСЛИ** мы употребляем один раз, и нам нужна только одна закрывающая скобка. Также обратите внимание на то, что в старой формуле с **ЕСЛИ**, когда ни одно из заданных условий не является истинным, по умолчанию возвращается последний результат, а в формуле **ЕСЛИМН** для этого нужно ввести **ИСТИНА** в формулу — тогда она возвратит последний результат.

	A	B	C	D
1	IFSFUNCTION	cutoff	price	
2	cut1	500	\$	3.00 price1
3	cut2	1200	\$	2.70 price2
4	cut3	2000	\$	2.30 price3
5			\$	2.00 price4
6				
7				
8	order quantity	UNIT COST		
9	450	\$ 3.00	=IF(A9<=_cut1,price1,IF(A9<=_cut2,price2,IF(A9<=_cut3,price3,price4)))	\$3,00
10	900	\$ 2.70	=IF(A10<=_cut1,price1,IF(A10<=_cut2,price2,IF(A10<=_cut3,price3,price4)))	\$2.70
11	1450	\$ 2.30	=IF(A11<=_cut1,price1,IF(A11<=_cut2,price2,IF(A11<=_cut3,price3,price4)))	\$2.30
12	2100	\$ 2.00	=IF(A12<=_cut1,price1,IF(A12<=_cut2,price2,IF(A12<=_cut3,price3,price4)))	\$2.00
13				
14				
15				
16			=IFS(A9<=_cut1,price1,A9<=_cut2,price2,A9<=_cut3,price3,TRUE,price4)	

**Рис. 12.12.** Применение функции **ЕСЛИМН** для определения цены за единицу продукции

## ❓ Как работает функция **ВЫБОР**?

Функция **ВЫБОР** возвращает значение из листа значений на основании заданной позиции. Синтаксис функции **ВЫБОР**: **ВЫБОР (индекс\_числ;значение1;[значение2];...)**.

Значения здесь могут быть диапазонами. Если индекс равняется *n*, то выбирается значение *n*. В файле **Choosenfinal.xlsx** (рис. 12.13–12.15) вы найдете три примера применения функции **ВЫБОР**. Великий математик Эдвард Торп придумал, как считать карты в игре в блек-джек. Когда из колоды раздаются карты 2–7, это выгодно игроку, и игрок засчитывает себе за эти карты по очку: +1. Когда из колоды раздаются карты 9, 10 или туз (номер первый в списке), это невыгодно игроку и за них снимается по очку: –1. Восмерка нейтральна и засчитывается как 0. На рабочем листе **Blackjack** (рис. 12.13 файла **Choosenfinal.xlsx**) показано, как использо-



вать функцию **ВЫБОР**, чтобы рассчитать кумулятивный итог по очкам по формуле **=ВЫБОР(G3;-1;1;1;1;1;1;1;0;-1;-1)**, где дан список значений очков по текущей карте. Как нам и хотелось бы, сдача карт 2–7 приносит по одному очку: +1, сдача карт 1, 9 или 10 забирает по одному очку: –1, и сдача восьмерки имеет значение 0 очков. В колонке I записывается кумулятивный итог по очкам. Например, после того как были сданы 7 карт, итог составляет +3, что довольно благоприятно для игрока.

	D	E	F	G	H	I	J
1					<b>=CHOOSE(G3,-1,1,1,1,1,1,1,0,-1,-1)</b>		
2			Card Number	Card	Current Points	Total Points	
3	1	-1	Card 1	10	-1	-1	
4	2	1	Card 2	5	1	0	
5	3	1	Card 3	1	-1	-1	
6	4	1	Card 4	5	1	0	
7	5	1	Card 5	5	1	1	
8	6	1	Card 6	7	1	2	
9	7	1	Card 7	6	1	3	
10	8	0	Card 8	10	-1	2	
11	9	-1	Card 9	10	-1	1	
12	10	-1	Card 10	10	-1	0	

**Рис. 12.13.** Применение функции **ВЫБОР** к подсчету карт при игре в блекджек

На рабочем листе **Quarters** содержатся еще два примера использования функции **ВЫБОР** (рис. 12.14 и 12.15). Предположим, 1–3-й месяцы года — это второй квартал, 4–6-й месяц — это третий квартал, 7–9-й месяцы — четвертый квартал, а 10–12-й месяцы — это первый квартал. На основе месяца в ячейке **H5**

	D	E	F	G	H	I	J	K
1								
2								
3								
4	1	2			Month	Quarter		
5	2	2			7	4		
6	3	2			10	1		
7	4	3						
8	5	3						
9	6	3			<b>=CHOOSE(H5,2,2,2,3,3,3,4,4,4,1,1,1)</b>			
10	7	4						
11	8	4			<b>=VLOOKUP(H6,D4:E15,2)</b>			
12	9	4						
13	10	1						
14	11	1						
15	12	1						

**Рис. 12.14.** Применение функции **ВЫБОР** для определения квартала года

формула в ячейку I5 возвращает номер квартала: **=ВЫБОР(H5;2;2;2;3;3;3;4;4;4;1;1;1)**. Конечно, мы могли бы получить точно такой же результат с помощью формулы **=ВПР(H6;D4:E15;2)**.

В ячейке M2 на рабочем листе **Quarters** (рис. 12.15) мы вводим квартал года, а затем в ячейке N2 нам нужен итог по продажам за введенный квартал. Эту цель мы достигаем, введя в ячейку N2 формулу **=СУММ(ВЫБОР(M2;M6:M8;N6:N8;O6:O8;P6:P8))**.

Например, когда в M2 содержится 3, функция **ВЫБОР** выбирает диапазон O6:O8. Затем функция **СУММ** складывает цифры в этом диапазоне; в сумме получается 120.

	M	N	O	P	Q	R
1	Quarter	Total				
2	3	120				
3		<b>=SUM(CHOOSE(M2,M6:M8,N6:N8,O6:O8,P6:P8))</b>				
4						
5	Quarter 1	Quarter 2	Quarter 3	Quarter 4		
6	10	20	30	40		
7	15	30	40	50		
8	20	40	50	60		

**Рис. 12.15.** Применение функции **ВЫБОР** для определения итоговых продаж по кварталу

### ❓ Как работает новая функция **ПЕРЕКЛЮЧ**

Если у вас Excel 2019 или Office 365, то вы располагаете функцией **ПЕРЕКЛЮЧ** (**SWITCH**). Функция **ПЕРЕКЛЮЧ** оценивает одно значение (называемое выражением, *expression*) против списка значений и возвращает результат, соответствующий первому совпадающему значению. Если нет совпадений, то может возвращаться выбранное значение по умолчанию. Функция **ПЕРЕКЛЮЧ** может оценить до 126 совпадающих значений и результатов. В файле **Switchfinal.xlsx** (рис. 12.16) показано применение функции **ПЕРЕКЛЮЧ**. На основании кодов рубашек (**Shirt Code**) в столбце D, в столбце G перечислены коды размеров для каждой конкретной рубашки. Нам нужно, чтобы код размера преобразовывался в определение в столбце B, соответствующее коду каждой рубашки. Это делается путем копирования формулы из H4 в H5:H27 — код рубашки преобразовывается в связанное с ним определение.

**=ПЕРЕКЛЮЧ(G4;"XXL";"На самых полных";"XL";"Очень большой";"L";  
"Большой";"M";"Средний";"S";"Маленький";"XS";"Очень маленький";"XXS";  
"Самый маленький";"Неправильный код")**

Обратите внимание, что если в колонке G не содержится кода, поддающегося расшифровке, то наша формула с функцией **ПЕРЕКЛЮЧ** вернет в ячейке H27 сообщение «Неправильный код».

	A	B	C	D
1	IFSFUNCTION	cutoff	price	
2	cut1	500	\$	3.00 price1
3	cut2	1200	\$	2.70 price2
4	cut3	2000	\$	2.30 price3
5			\$	2.00 price4
6				
7				
8	order quantity	UNIT COST		
9	450	\$ 3.00	=IF(A9<= _cut1,price1,IF(A9<= _cut2,price2,IF(A9<= _cut3,price3,price4)))	\$3.00
10	900	\$ 2.70	=IF(A10<= _cut1,price1,IF(A10<= _cut2,price2,IF(A10<= _cut3,price3,price4)))	\$2.70
11	1450	\$ 2.30	=IF(A11<= _cut1,price1,IF(A11<= _cut2,price2,IF(A11<= _cut3,price3,price4)))	\$2.30
12	2100	\$ 2.00	=IF(A12<= _cut1,price1,IF(A12<= _cut2,price2,IF(A12<= _cut3,price3,price4)))	\$2.00
13				
14				
15				
16			=IFS(A9<= _cut1,price1,A9<= _cut2,price2,A9<= _cut3,price3,TRUE,price4)	

**Рис. 12.16.** Применение функции ПЕРЕКЛЮЧ для преобразования кода рубашки в конкретный размер

## Задания

- Допустим, цена продукта будет меняться так, как указано в таблице:

Дата	Цена
До 15 февраля 2004 г. включительно	\$8
С 16 февраля 2004 г. до 10 апреля 2005 г.	\$9
С 11 апреля 2005 г. до 15 января 2006 г.	\$10

Напишите формулу, вычисляющую цену продукта по дате его продажи.

- Вместимость самолета компании Blue Yonder Airlines, делающей рейсы из Сиэтла в Нью-Йорк, составляет 250 человек. Авиакомпания продала 270 билетов на рейс по цене \$300 за билет. Билеты не подлежат возврату. Переменные затраты на полет одного пассажира (в основном это расходы на продукты питания и топливо) составляют \$30. Если на посадку явятся более 250 человек, то мест не хватит, и компании Blue Yonder придется выплатить компенсацию за избыточное бронирование в размере \$350 каждому пассажиру, оказавшемуся без места. Разработайте таблицу, в которой вычисляется прибыль компании Yonder в зависимости от количества клиентов, изъявивших желание лететь.
- Крупная фармацевтическая компания пытается вычислить производственную мощность завода для нового препарата. Затраты на строительство для единицы годовой производительности составляют \$10. Каждая единица товара продается за \$12, и переменные затраты составляют \$2. Препарат будет продаваться в течение 10 лет. Разработайте таблицу, в которой вычисляется прибыль компании за 10 лет с учетом выбранного уровня годовой произво-

длительности и годовой потребности в препарате. Допустим, что спрос на препарат всегда одинаков. Кроме того, при решении этой задачи можно проигнорировать стоимость денег с учетом фактора времени.

4. Упомянутая выше фармацевтическая компания производит новый препарат. Компания предполагает, что:
- в первый год будет продано 1 100 000 единиц товара;
  - продажи будут расти в течение трех лет, а затем снижаться в течение семи лет;
  - в период роста продажи будут расти на 15% в год. В период снижения продажи будут падать со скоростью 10% в год.

Создайте таблицу, в которой для вычисления объемов продаж с первого по одиннадцатый годы используется объем продаж за первый год, продолжительность периода роста, продолжительность периода спада, темпы роста и темпы спада.

5. Вы принимаете участие в тендере на строительный проект. Предложение с самой привлекательной ценой победит. Вы оцениваете стоимость проекта в \$10 000. Помимо вас в тендере принимают участие еще четыре компании. Подготовка заявки стоит \$400. Напишите формулу (с учетом вашей цены и цен четырех конкурентов), вычисляющую вашу прибыль (или потери, если проиграете тендер).
6. Мы участвуем в аукционе, где выставлена ценная картина. Картину приобретет тот, кто предложит самую высокую цену. Мы оценили картину в \$10 000. В аукционе участвуют еще четыре компании. Подготовка заявки стоит \$400. Напишите формулу, определяющую, получим ли мы картину (с учетом нашей ставки и ставок четырех конкурентов).
7. Фармацевтическая компания рассчитывает, что отгрузит в 2004 г. 10 000 единиц своего нового препарата. Она ожидает выхода на рынок двух конкурентов. В год выхода на рынок первого конкурента компания прогнозирует потерю 30% от своей доли рынка. В год выхода на рынок второго конкурента компания прогнозирует потерю 15% от своей доли рынка. Рынок растет на 10% в год. Учитывая годы выхода на рынок двух конкурентов, создайте таблицу для вычисления годового объема продаж в 2004–2013 гг.
8. Магазин одежды заказал 100 000 купальников. Производство одного купальника обходится в \$22. Магазин планирует продавать купальники до 31 августа по цене \$40 за штуку, а потом снизить цену до \$30. Создайте лист для вычисления прибыли от этого заказа с учетом величины спроса до 31 августа и после 31 августа.
9. Для каждого броска при игре в кости после первого броска правила таковы: если игра не закончилась и при броске выпало столько очков, сколько при первом броске, вы выиграли. Если игра не закончилась и выпало 7, вы про-

играли. Во всех остальных случаях игра продолжается. Создайте лист , на котором указано положение в игре после четырех бросков (с учетом данных первых четырех бросков).

10. Пусть в примере с куплей-продажей акций S&P по скользящим средним вы по-прежнему покупаете акции, когда текущая цена превышает скользящее среднее за 15 месяцев, и продаете, когда текущая цена опускается ниже скользящего среднего за 5 месяцев. Является ли эта стратегия более выгодной, чем продажа акции, когда ее текущая цена ниже скользящего среднего за 15 месяцев?
11. Европейский колл-опцион дает право купить акцию в определенный момент времени по цене исполнения. *Спред «бабочка»* включает покупку одного колл-опциона с низкой ценой исполнения, покупку одного колл-опциона с высокой ценой исполнения и продажу двух колл-опционов с ценой исполнения где-то посередине между низкой и высокой ценами. Пример спреда «бабочка»: текущая цена акции составляет \$60. Вы покупаете 54-долларовый шестимесячный европейский колл-опцион за \$9 и 66-долларовый шестимесячный европейский колл-опцион за \$4, а продаете два \$60 европейских колл-опциона по \$6. Вычислите прибыль (в долларах, а не в процентах) от этой операции как функцию цены за акцию через шесть месяцев в диапазоне от \$40 до \$80. Когда трейдер покупает спред «бабочка», на какой тип движения цены акции через шесть месяцев он делает ставку?
12. Пусть акция продается по \$32. Вы покупаете шестимесячный европейский колл-опцион с ценой исполнения \$30 за \$2,50 и шестимесячный европейский колл-опцион с ценой исполнения \$35 за \$1. Вычислите прибыль (в долларах) от применения этой стратегии как функцию цены акции через шесть месяцев в диапазоне от \$25 до \$45. Почему такая стратегия называется спредом «быков»? Как бы вы изменили эту стратегию, чтобы получить спред «медведей»?
13. Вернемся к примеру с гипотетическим отчетом. Пусть процентная ставка для долгосрочных обязательств зависит от нашего финансового благосостояния. Точнее говоря, если размер прибыли до уплаты процентов и налогов (EBIT, earnings before interest and taxes) отрицательный, то процентная ставка для долгосрочных обязательств равна 16%. Если затраты на выплату процентов больше 10% от величины EBIT и величина EBIT положительная, процентная ставка для долгосрочных обязательств равна 13%. Иначе процентная ставка равна 10%. Измените гипотетический отчет с учетом переменной процентной ставки.
14. Выполните это задание независимо от задания 13. Допустим, наша компания хочет, чтобы соотношение размера задолженности и собственного капитала составляло 50% для каждого года. Каким образом следует изменить гипотетический отчет? Подсказка: для каждого года необходимо поддерживать величину акционерного капитала неотрицательной и использовать для уравнивания активов и обязательств акционерный капитал, а также денежные средства и ликвидные ценные бумаги.

15. День Мартина Лютера Кинга приходится на третий понедельник января. Напишите формулу, определяющую (по заданному году) дату Дня Мартина Лютера Кинга. Подсказка: сначала вычислите день недели для 1 января заданного года.
16. День благодарения приходится на четвертый четверг ноября. Напишите формулу для определения (по заданному году) даты Дня благодарения. Подсказка: сначала вычислите день недели для 1 ноября заданного года.
17. К первому кварталу года относятся месяцы с января по март, ко второму кварталу — месяцы с апреля по июнь, к третьему кварталу — месяцы с июля по сентябрь и к четвертому кварталу — месяцы с октября по декабрь. Создайте формулу, определяющую (для любой заданной даты) квартал года.
18. Напишите формулу, вычисляющую возраст по дате рождения.
19. День труда приходится на первый понедельник сентября. Напишите формулу, определяющую дату Дня труда для заданного года.
20. В файле **Nancybonds.xlsx** содержится рейтинг для нескольких облигаций в предыдущий и текущий месяцы. Необходимо подсчитать рациональным способом, у какого количества облигаций рейтинг понизился. К сожалению, каждая компания представлена в нескольких строках. При условии, что данные отсортированы по названиям компаний, как можно узнать количество облигаций с понизившимся рейтингом?
21. В файле **Addresses.xlsx** имена и фамилии людей указаны в одной строке, часть адреса с названием улицы в другой строке, а город, штат и почтовый индекс — в следующей строке. Каким образом можно указать информацию о каждом человеке в одной строке?
22. В файле **FormattingDDAnum.xlsx** содержится несколько текстовых строк, таких как **DDA : D** в ячейке **C4**, **DDA1250045** в ячейке **C17** и т. д. Ячейка считается правильно отформатированной, если первыми тремя символами являются **DDA**, а последними семью символами — число не меньше миллиона. Определите, какие ячейки имеют правильное форматирование.
23. Пусть число членов первой группы указано в ячейке **B1**, число членов второй группы — в ячейке **B2** и число членов третьей группы — в ячейке **B3**. Общее количество членов во всех группах всегда равно 100. Допустим, в первой группе 50 членов, во второй группе 30 членов и в третьей группе 20 членов. Поместите рациональным способом в столбец **D** значение 1 для каждого члена первой группы, значение 2 для каждого члена второй группы и значение 3 для каждого члена третьей группы. Таким образом, в столбце **D** (для указанных данных) должно быть записано: 1 в ячейках **D1:D50**, 2 в ячейках **D51:D80** и 3 в ячейках **D81:D100**.
24. В файле **Dividebyprice.xlsx** содержатся данные о количестве проданных единиц каждого продукта и общая выручка. Необходимо определить среднюю цену для каждого продукта и отобразить ее в столбце **H**. Учтите, что если количество проданных единиц равно 0, то средняя цена отсутствует. Отловите ошибку.

ки в файле `Dividebyprice.xlsx` таким образом, чтобы для всех продуктов с нулевым объемом продаж было указано сообщение `No Sales`, а не сообщение об ошибке `#ДЕЛ/0!`.

25. В школе изящных искусств имеется 100 шкафчиков, пронумерованных от 1 до 100. В настоящее время все шкафчики открыты. Начните с того, что закройте шкафчики, номера которых кратны 3. Затем переключите (то есть откройте закрытый шкафчик и закройте открытый шкафчик) каждый шкафчик, номер которого кратен 4, потом переключите шкафчики, номера которых кратны 5, ..., и, наконец, переключите все шкафчики, номера которых кратны 100. Сколько шкафчиков теперь открыто?
26. В файле `Matchlist.xlsx` содержится список покупателей, приобретших ваш продукт в феврале, и список покупателей, приобретших его в марте. Определите, сколько клиентов, купивших продукт в феврале, купили его и в марте.
27. Настройте лист календаря, который по указанному месяцу и году вычисляет день недели, на который приходится каждый день месяца.
28. В файле `Problem28data.xlsx` некоторые строки столбца `C` содержат слово `and`. Введите в столбец `D` формулу, которая записывает `X` в строку столбца `D` тогда и только тогда, когда столбец `C` в этой строке содержит слово `and`.
29. Вы подбросили два 20-сторонних кубика. Возможных результатов 400: (1, 1), (1, 2), ..., (20, 20). Примените функцию `ЕСЛИ` для генерирования всех 400 возможных комбинаций.
30. Для данных в файле `Catsanddogs.xlsx` определите, сколько раз встречается текстовая строка `cat` и текстовая строка `dog`.
31. Файл `Footballdata.xlsx` содержит записи о победах/поражениях команд НФЛ за сезон 2015 года. Введите в ячейке `H13` формулу, которую можно скопировать в `H14:H31` и которая вычисляет процент побед для каждой команды и не выдает сообщений об ошибке.
32. В начале года 1 стоимость полиса автострахования на следующий год составляет \$300. При отсутствии страховых случаев в течение года стоимость полиса на следующий год понижается на 10%, при наличии одного страхового случая в течение года стоимость полиса на следующий год повышается на 10%, а при наличии двух страховых случаев в течение года стоимость полиса на следующий год повышается на 30%. Составьте таблицу расчета стоимости страхового полиса на годы 1–30.
33. Продажи нашего продукта зависят от нашей цены и цены конкурента следующим образом: если наша цена выше цены конкурента не менее чем на \$3, мы продаем 500 единиц товара; если наша цена ниже цены у конкурента не менее чем на \$3, мы продаем 1500 единиц товара; в противном случае мы продаем 1000 единиц товара. Предположив, что все цены — это целые числа от 1 до 10 долларов включительно, напишите формулы, которые вычисляют наши продажи этого продукта для всех возможных сочетаний цен.



34. Необходимо рассчитать годовые доходы компании, исходя из информации в ячейках C1:D5 файла **Problem34data.xlsx**. Продажи продукта растут с заданной скоростью до выхода на рынок конкурента; после этого уровень продаж остается на достигнутом уровне. Каждый год цена на продукт повышается с учетом темпов инфляции. Напишите формулы в ячейках, выделенных желтым цветом, правильно вычисляющие цену, продажи и доход за годы 1–10.
35. В ячейке D2 файла **Problem35data.xlsx** содержится значение размера рынка за год 1. Ячейка D3 представляет долю на рынке (неизменную для каждого года), ячейка D4 содержит цену продукта (неизменную для каждого года), а D5 содержит коэффициент прибыльности (неизменный для каждого года). В ячейке D6 указан рост размера рынка для каждого года, когда есть рост, в D7 представлены годы роста рынка. Напишите формулы, вычисляющие ежегодный доход и совокупный доход за 10 лет.
36. Президентский день отмечается в третий понедельник февраля. Создайте таблицу, в которой по введенному году определяется дата празднования Президентского дня.
37. Вам нужно написать формулы, вычисляющие прибыль от продажи поздравительных открыток. Текущая информация представлена в ячейках B1:B7 файла **Problem37data.xlsx**. Мы заказали открытки по цене \$3,50 за штуку. Полная цена открытки составляет \$6. Заказ на открытки по полной цене будет вводиться в ячейку B7. Первые 40 открыток из распродажи продаются за \$2. Оставшиеся открытки распродаются по \$0,50. Напишите формулы, которые всегда возвращают правильную прибыль при вводе любых неотрицательных данных.
38. В столбец B в файле **Problem38data.xlsx** в нечетные строки внесены имена продавцов, а прямо под именем — количество проданных единиц продукции. Напишите формулы, необходимые, чтобы в нечетных строках столбца C появились имена продавцов вместе с количеством проданных единиц продукции (через запятую). Подсказка: для определения нечетности номера можно применить функцию **ЧЁТН (ISEVEN)**, а функция **СТРОКА (ROW)** возвращает номер строки текущей ячейки.
39. В файле **Problem39data.xlsx**, представлены данные по клиентам аудиторской компании **QXZ Inc.** У вас на руках информация о том, большая это фирма или маленькая и прошли они или провалили июньский аудит. Если фирма крупная и провалила июньский аудит, следующий аудит ей предстоит в июле. Если фирма крупная и прошла июньский аудит, следующий аудит ей предстоит в сентябре. Если фирма маленькая, ее следующий аудит в августе. Напишите формулы, помещающие в столбец D месяц следующего аудита.



## ГЛАВА 13

# Время и функции времени

### Обсуждаемые вопросы

- Как указать время в Excel?
- Как указать время и дату в одной ячейке?
- Как в Excel обрабатываются значения времени?
- Как отобразить на листе текущее время?
- Как создать значения времени с помощью функции ВРЕМЯ?
- Как использовать функцию ВРЕМЗНАЧ для преобразования строки из текстового формата в формат времени?
- Как извлечь часы, минуты и секунды из заданного значения времени?
- Как определить, сколько часов отработал сотрудник, если известно время начала и окончания работы?
- Я суммирую общее время, отработанное сотрудником, но у меня никогда не получается больше 24 часов. Где ошибка?
- Как создать последовательность равномерно распределенных временных интервалов?
- Как поместить в книгу статичное временное значение?

В главе 7 я упоминал о том, что в Excel 2019 дате 1 января 1900 г. соответствует порядковое число 1; дате 2 января 1900 г. — число 2 и т. д. Кроме того, в Excel моментам времени соответствуют десятичные дроби (как доли 24-часового дня). Начальной точкой является полночь, поэтому трем часам ночи (3:00) соответствует число 0,125, полдню — число 0,5, а шести часам вечера (18:00) — число 0,75 и т.д. При объединении даты и времени в одной ячейке в числовом формате число складывается из порядкового номера дня после 1 января 1900 г. плюс десятичная дробь, связанная с указанным временем. Таким образом, дата 01.01.2007 дает в результате (в формате Общий) число 39083, а дата и время 01.01.2007 6:00 — число 39083,25.

## Ответы на вопросы

### ❓ Как указать время в Excel?

Для этого необходимо ввести двоеточие ( :) после количества часов и еще одно двоеточие перед количеством секунд. Например, в файле Time.xlsx (рис. 13.1) вре-

мя 8:30 А.М. в ячейке C2 введено как 8:30 AM. В ячейке C3 время 8:30 Р.М. введено как 8:30 PM. Как показано в ячейке D3, можно также ввести время 8:30 Р.М. в 24-часовом формате как 20:30. В ячейке A4 введена формула =ВРЕМЯ(15,10,30), что приносит результат 3:10:30 PM. Это момент времени через 30 секунд после 3:10 Р .М.

	A	B	C	D	E
1					
2	8:30 AM	=TIME(8,30,0)	8:30 AM	8:30	
3	8:30 PM	=time(20,30,0)	8:30 PM	20:30	
4	3:10:30 PM	=TIME(15,10,30)	HOUR(A4)	Minute(a4)	SECOND(A4)
5	1:10:30 AM	=TIME(25,10,30)	15.00	10	30
6					
7	0.354166667	=TIMEVALUE("8:30")			
8					
9			start	finish	
10		Jane	9:00 PM	6:00 AM	
11		Jack	7:00 AM	3:30 PM	
12		elapsed time			
13		Jane	9.00		
14		Jack	8.50		
15					
16					
17	Start	5/12/2006 8:12	5/12/2006 8:12		
18	Finish	6/10/2006 12:30	6/10/2006 7:30		
19		29.18	28.97		

**Рис. 13.1.** Примеры ввода времени в разных форматах

### ❓ Как указать время и дату в одной ячейке?

Просто вставьте пробел после даты и введите время. В файле Time.xlsx в ячейку F13 я ввел 1 января, 2007 5:35. Разумеется, это 5:35 утра 1 января 2007 г. Excel немедленно переформатировал эту запись к виду 1/1/2007 5:35:00, и то же самое отражается в строке формул.

### ❓ Как в Excel обрабатываются значения времени?

В Excel результат вычислений со значениями времени зависит от формата, используемого в ячейке. Различные форматы времени в Excel показаны на рис. 13.2.

	C	D	E	F	G	H	I
1					=NOW()	ASTIME	ASNUMBER
2	8:30 AM	8:30			6/17/2018 6:46	6:46 AM	0.28
3	8:30 PM	20:30					
4	HOUR(A4)	Minute(a4)	SECOND(A4)				
5	15.00	10	30	0.50	=C3-C2	12:00 PM	
6				0.50	=d3-d2	12:00	
7				#####	=D2-D3 PROBLEM!		
8				-0.50	CHANGE TO NUMBER FORMAT OK		

**Рис. 13.2.** Различные форматы времени в Excel

В файле *Time.xlsx* (см. рис. 13.2) я указал разницу между 8:30 Р .М. и 8:30 А.М. в ячейках F5 и H5 с помощью формулы `=C3-C2`. Если вы не меняли формат ячейки, Excel считает, что между этими двумя моментами времени 12 часов, и в ячейке H5 отображается результат **12:00 PM**. В большинстве случаев требуется отобразить эту разницу как 0,5 дня. (Умножив эту разницу во времени на 24, получим ее значение в часах.) Для получения в ячейке F5 значения 0,5 необходимо присвоить ячейке F5 формат Числовой (Number).

В ячейке F7 я попытался вычесть ранний момент времени из позднего по формуле `=D2-D3`. Поскольку ячейке не был присвоен соответствующий формат, Excel отобразил жуткие #####. Если ячейке с формулой присвоить формат Числовой (Number), как у ячейки F8, разница во времени отобразится правильно — -0,5 дня.

В ячейках B17 и C17 указано время начала двух рабочих проектов, а в ячейках B18 и C18 — время окончания этих проектов (рис. 13.3). Вы можете вычислить, сколько часов занимает выполнение каждого проекта, по формуле `=B18-B17`, скопировав ее из B19 в C19 и присвоив ячейкам формат Числовой (Number). Таким образом, первый проект длится 29,18 дня, а второй — 28,97 дня.

	A	B	C
17	Start	5/12/2006 8:12	5/12/2006 8:12
18	Finish	6/10/2006 12:30	6/10/2006 7:30
19		29.18	28.97

**Рис. 13.3.** Определите время, необходимое для завершения проектов

### ❓ Как отобразить на листе текущее время?

В Excel дату и текущее время можно получить с помощью функции ТДАТА (NOW) по формуле `=ТДАТА()`. Например, в файле *Time.xlsx* в ячейке G2 (рис. 13.4) после ввода формулы `=ТДАТА()` отображаются текущая дата и время, 1/9/2019 14:12 (я сделал скриншот в 14:12 9 января 2019 года). Если вы будете редактировать файл *Time.xlsx*, то в ячейке G2 отобразятся текущая дата и время. Для вычисления текущего времени можно ввести в ячейки H2 или I2 формулу `=ТДАТА()-СЕГОДНЯ()`. Ячейка H2 отформатирована для показа времени (2:12 PM), а ячейка I2 — для показа числа (0,59 дня). Это означает, что время 2:12 PM составляет 59% от промежутка между полночью одного дня и полночью следующего дня.

	G	H	I
1	=NOW()	ASTIME	ASNUMBER
2	1/9/2019 14:12	2:12 PM	0.59
3		=NOW()-TODAY()	

**Рис. 13.4.** Функции ТДАТА и СЕГОДНЯ

### ? Как создать значения времени с помощью функции ВРЕМЯ?

Функция ВРЕМЯ (TIME) имеет следующий синтаксис: ВРЕМЯ(часы;минуты;секунды). По указанным значениям часов, минут и секунд функция ВРЕМЯ возвращает время дня. Функция ВРЕМЯ никогда не возвращает значение, превышающее 24 часа.

В ячейке A2 (см. рис. 13.1) по формуле =ВРЕМЯ(8;30;0) выведено 8:30 AM. В ячейке A3 по формуле =ВРЕМЯ(20;30;0) выведено 8:30 PM. В ячейке A4 по формуле =ВРЕМЯ(15;10;30) выведено 3:10:30 PM. Наконец, обратите внимание, что в ячейке A5 по формуле =ВРЕМЯ(25;10;30) значение 25 обрабатывается как  $25 - 24 = 1$ , и выведено 1:10:30 AM.

Если в ячейке не отображается количество секунд, просто выберите в диалоговом окне Формат ячеек (Format Cells) формат времени, который отображает секунды.

### ? Как использовать функцию ВРЕМЗНАЧ для преобразования строки из текстового формата в формат времени?

Функция ВРЕМЗНАЧ (TIMEVALUE) имеет следующий синтаксис: =ВРЕМЗНАЧ(время\_как\_текст), где время\_как\_текст — текстовая строка значения времени в правильном формате. Функция ВРЕМЗНАЧ возвращает время в виде десятичной дроби от 0 до 1. (Это означает, что функция ВРЕМЗНАЧ игнорирует любую дату в строке время\_как\_текст.) Например, в ячейке A7 (см. рис. 13.1) формула =ВРЕМЗНАЧ("8:30") дает в результате 0,354166667, поскольку 8:30 AM составляет 35,4% от промежутка между полночью одного дня и полночью следующего дня.

### ? Как извлечь часы, минуты и секунды из заданного значения времени?

В Excel из ячейки, содержащей время, временные единицы извлекаются при помощи функций ЧАС (HOUR), МИНУТЫ (MINUTE) и СЕКУНДЫ (SECOND). Например, как показано на рис. 13.1, формула =ЧАС(A4) в ячейке C5 дает 15,0; формула =МИНУТЫ(A4) в ячейке D5 дает 10, а формула =СЕКУНДЫ(A4) в ячейке E5 дает 30.

### ? Как определить, сколько часов отработал сотрудник, если известно время начала и окончания работы?

Я ввел время начала и окончания работы Джейн и Джека в ячейки C10:D11 (рис. 13.5). Нам нужно вычислить, сколько времени проработал каждый. Проблема в том, что Джейн закончила работу на следующий день после того, как приступила к ней, поэтому простое вычитание двух цифр не даст желаемого результата. Нам поможет формула =ЕСЛИ(D10>C10;24\*(D10-C10);24+24\*(D10-C10)), скопированная из C13 в C14. Этим ячейкам я присвоил Числовой (Number) формат. Если время окончания *больше* времени начала, то количество отработанных часов равно разности времени начала и времени окончания, умноженной на 24. Если время окончания *меньше* времени начала, то  $24 * (\text{время\_окончания} - \text{время\_начала})$  дает отрицательное количество часов. В таком случае нужно добавить 24 часа, при условии что окончание смены пришлось на следующий день. Таким образом, Джейн отработала 9 часов, а Джек — 8,5 часа.

	B	C	D
9		start	finish
10	Jane	9:00 PM	6:00 AM
11	Jack	7:00 AM	3:30 PM
12	elapsed time		
13	Jane	9.00	
14	Jack	8.50	
15			
16			
17	5/12/2006 8:12	5/12/2006 8:12	
18	6/10/2006 12:30	6/10/2006 7:30	
19	29.18	28.97	

**Рис. 13.5.** Вычисление продолжительности рабочего дня сотрудников

❓ Я суммирую общее время, отработанное сотрудником, но у меня никогда не получается больше 24 часов. Где ошибка?

В ячейках C31:D35 (рис. 13.6) я указал количество часов (в формате ч:мм), отработанное сотрудником в каждый из дней недели. В ячейке D36 по формуле =СУММ(D31:D35) я вычислил общее количество часов, отработанное за неделю. Excel выдал 14:48. Очевидно, что это ошибка! В формате ч:мм никогда не отображается значение, превышающее 24 часа. Необходимо выбрать формат, отображающий значения больше 24 часов (38:48:00), как показано в ячейке D38 (седьмой в списке). Тогда отработанные часы просуммируются корректно (38 часов 48 минут).

	C	D
31	mon	9:23
32	tues	8:30
33	wed	7:20
34	thur	9:40
35	fri	3:55
36	total	14:48
37	38 hours 48 min	
38	total reformatted	38:48:00

**Рис. 13.6.** Вычисление общего количества отработанных за неделю часов

❓ Как создать последовательность равномерно распределенных временных интервалов?

Допустим, врач принимает с 8:00 до 17:00, тратя на каждого пациента 20 минут. Как ввести список возможных записей на определенное время в разные строки? Просто воспользуйтесь функцией Excel Автозаполнение (AutoFill) (рис. 13.7). Введите в ячейки L15:L16 первые два значения времени (8:00 AM и 8:20 AM). Теперь выделите ячейки L15:L16 и двигайте курсор в правый нижний угол ячейки L16 до тех пор, пока курсор не превратится в черный крестик. Перемещайте указатель вниз, пока не увидите значение 5:00 PM (последнее время приема). По содержи-

мому ячеек L15:L16 Excel автоматически (правильно!) определил, что необходимо указать двадцатиминутные интервалы. Если ввести в ячейку **Понедельник**, а в ячейку под ней **Вторник**, то функция **Автозаполнение** создаст последовательность **Понедельник, Вторник, Среда, ...**, продолжением которой снова станет **Понедельник**. Введите 01.01.2007 в одну ячейку, затем 02.01.2007 в другую, выделите обе ячейки и с помощью функции **Автозаполнение** получите последовательность дат 01.01.2007, 02.01.2007, 03.01.2007 и т. д.

	L
14	sequence of times
15	8:00 AM
16	8:20 AM
17	8:40 AM
18	9:00 AM
19	9:20 AM
20	9:40 AM
21	10:00 AM
22	10:20 AM
23	10:40 AM
24	11:00 AM
25	11:20 AM
26	11:40 AM
27	12:00 PM
28	12:20 PM
29	12:40 PM
30	1:00 PM
31	1:20 PM
32	1:40 PM
33	2:00 PM
34	2:20 PM
35	2:40 PM
36	3:00 PM

**Рис. 13.7.** Ввод последовательности временных интервалов

### ❓ Как поместить в книгу статичное временное значение?

Допустим, вы создаете книгу и хотите, чтобы на листе всегда отображалось точное время его создания. Встаньте на пустую ячейку, используйте комбинацию **Ctrl+Shift+;** (точка с запятой) и всегда будете видеть, когда лист был создан. Например, на рис. 13.8 показано, что лист создан в 7:10 AM.

	C
1	
2	
3	
4	
5	7:10 AM

**Рис. 13.8.** Вставка в ячейку статичного времени

## Задания

1. Напишите формулу, вычисляющую время через 18 часов после текущего времени.
2. В файле **Marathon.xlsx** хранятся данные о том, за сколько времени четыре марафонца прошли дистанцию. Вычислите среднее время прохождения дистанции марафонцами. Выполните с этими данными еще три задания.
  - Насколько Джон опередил Джил?
  - Сколько минут бежал каждый марафонец?
  - Сколько секунд бежал каждый марафонец?
3. В файле **Jobshop.xlsx** содержатся даты и время начала различных видов работ, а также время, необходимое для их выполнения. Вычислите время завершения для каждого вида работ.

## ГЛАВА 14

# Команда Специальная вставка

### Обсуждаемые вопросы

- Как переместить результаты расчетов (но не формулы) в другую часть листа?
- В столбце имеется список имен. Как поместить список в строку вместо столбца?
- Я скачал с сайта в Excel процентные ставки по казначейским векселям США. В полученных данных отображается 5, если процентная ставка составляет 5%, 8 — если процентная ставка составляет 8%, и т. д. Как быстро разделить все результаты на 100, чтобы, например, пятипроцентная ставка отображалась как 0,05?

Команда Специальная вставка (Paste Special) в Microsoft Excel 2019 позволяет легко обрабатывать данные на листе. В этой главе я покажу, как применить данную команду для следующих операций:

- вставка в ячейки только значений (но не формул) на различных частях листа;
- транспонирование столбцов данных в строки и наоборот;
- преобразование диапазона чисел путем сложения, вычитания, деления или умножения каждого числа в диапазоне и заданной константы.

## Ответы на вопросы

### ❓ Как переместить результаты расчетов (но не формулы) в другую часть листа?

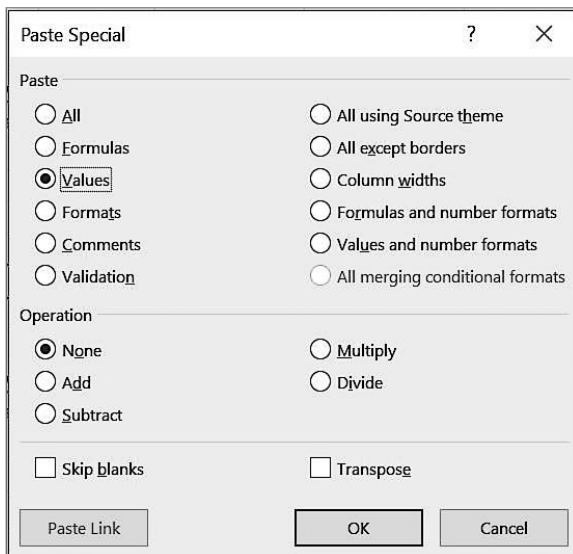
В файле Pastespecial.xlsx на листе Paste Special Value в ячейках E4:H9 содержатся имена и фамилии, количество сыгранных игр, общее количество заработанных очков и среднее количество заработанных очков за игру для пяти баскетболистов 10–11 лет из Блумингтона, штат Индиана. В диапазоне ячеек H5:H9 я использовал данные из ячеек F5:G9, чтобы вычислить среднее количество заработанных очков за игру для каждого ребенка (рис. 14.1).

Предположим, нам необходимо скопировать эти данные и вычисленное количество очков за игру — но не формулы, по которым были выполнены вычисления, — в другой диапазон ячеек (например, E13:H18). Просто выделите диапазон E4:H9, нажмите Ctrl+C и поставьте курсор в левую верхнюю ячейку целевого диапазона, в который требуется скопировать данные (в данном примере это ячейка E13). Затем щелкните правой кнопкой мыши, выберите в контекстном меню Специальная



	E	F	G	H
4		<b>Games</b>	<b>Points</b>	<b>Points/game</b>
5	<b>Dan</b>	4	28	7.00
6	<b>Gabe</b>	4	28	7.00
7	<b>Gregory</b>	5	35	7.00
8	<b>Christian</b>	6	22	3.67
9	<b>Max</b>	6	15	2.50
10				
11				
12				
13		<b>Games</b>	<b>Points</b>	<b>Points/game</b>
14	<b>Dan</b>	4	28	7
15	<b>Gabe</b>	4	28	7
16	<b>Gregory</b>	5	35	7
17	<b>Christian</b>	6	22	3.666666667
18	<b>Max</b>	6	15	2.5

**Рис. 14.1.** Вставка только значений с помощью команды Специальная вставка



**Рис. 14.2.** Диалоговое окно Специальная вставка с отмеченным пунктом значения. В этом случае вставляются только значения, но не формулы

вставка... (Paste Special), выберите значения (Values), как показано на рис. 14.2. Нажмите OK. Мы перенесли в диапазон E13:H18 данные, но не формулы, из диапазона E4:H9. Это можно проверить, щелкнув на ячейке H16. Появится значение (7), а не формула, которая использовалась для вычисления среднего количества очков за игру, заработанных Грегори. Имейте в виду, что если выбрать значения (Values), как описано выше, а затем вставить данные в тот же диапазон, из которого они были скопированы, формулы исчезнут с листа.

### ? В столбце имеется список имен. Как поместить список в строку вместо столбца?

Чтобы перенести данные из столбца в строку (и наоборот), скопируйте их и используйте диалоговое окно **Специальная вставка (Paste Special)**, в котором установите флажок **транспонировать (Transpose)**. По существу, транспонирование переворачивает выбранные ячейки таким образом, что первая строка скопированного диапазона становится первым столбцом диапазона, в который вставляются данные, и наоборот. Например, см. в файле *Pastespecial.xlsx* лист *Paste Special Transpose* (рис. 14.3).

	E	F	G	H	I	J
4		<b>Games</b>	<b>Points</b>	<b>Points/game</b>		
5	<b>Dan</b>	4	28	7.00		
6	<b>Gabe</b>	4	28	7.00		
7	<b>Gregory</b>	5	35	7.00		
8	<b>Christian</b>	6	22	3.67		
9	<b>Max</b>	6	15	2.50		
10						
11						
12						
13	<b>Dan</b>	<b>Gabe</b>	<b>Gregory</b>	<b>Christian</b>	<b>Max</b>	
14						
15						
16						
17		<b>Dan</b>	<b>Gabe</b>	<b>Gregory</b>	<b>Christian</b>	<b>Max</b>
18	<b>Games</b>	4	4	5	6	6
19	<b>Points</b>	28	28	35	22	15
20	<b>Points/game</b>	7.00	7.00	7.00	3.67	2.50

**Рис. 14.3.** Транспонирование столбца данных в строку или строки данных в столбец с помощью флажка «транспонировать» в диалоговом окне **Специальная вставка**

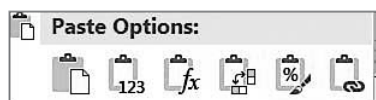
Допустим, вы хотите указать имена игроков в одной строке (начиная с ячейки E13). Выделите диапазон ячеек E5:E9 и нажмите **Ctrl+C**. Щелкните правой кнопкой мыши на ячейке E13, выберите в контекстном меню **Специальная вставка (Paste Special)** и в диалоговом окне **Специальная вставка (Paste Special)** установите флажок **транспонировать (Transpose)**. Нажмите **OK** и готово.

Теперь нам нужно транспонировать содержимое таблицы из диапазона E4:H9 в целевой диапазон с началом в ячейке E17. Выделите диапазон E4:H9 и нажмите **Ctrl+C**. Теперь переместите курсор в левый верхний угол диапазона, в который требуется перенести информацию (E17). Щелкните правой кнопкой мыши, выберите **Специальная вставка (Paste Special)**, затем в диалоговом окне **Специальная вставка (Paste Special)** выберите **транспонировать (Transpose)** и нажмите **OK**. Содержимое ячеек E4:H9 транспонируется, как показано на рис. 14.3. Следует отметить, что в ячейках F20:J20 Excel автоматически заменил в формуле расчета среднего количества очков за игру данные из строки данными из столбца для каждого игрока.

**ПРИМЕЧАНИЕ**

Если в диалоговом окне Специальная вставка (Paste Special) снять флажок транспонировать (Transpose) и вместо ОК нажать Вставить связь (Paste Link), находящаяся в нижнем левом углу, то данные будут скопированы без транспонирования, но целевые ячейки окажутся связанными с исходными ячейками, и любые изменения в исходных данных будут отражаться в скопированных данных. При замене значения в ячейке F5 на 7 значение в ячейке F18 также изменится на 7, и среднее количество очков, заработанных Дэном за игру, отображаемое в ячейке F20, будет пересчитано (4 очка за игру).

Если выбрать диапазон, нажать Ctrl+C и щелкнуть правой кнопкой мыши, то перед вами появится сокращенное меню панели **Параметры вставки**, представленное на рис. 14.4.



**Рис. 14.4.** Панель Параметры вставки

Как видите, слева направо вам предложены шесть действий:

- ☐ просто вставить все;
- ☐ вставить только значения;
- ☐ вставить только формулы;
- ☐ вставить с транспонированием;
- ☐ вставить только форматы;
- ☐ вставить связь.

**?** Я скачал с сайта в Excel процентные ставки по казначейским векселям США. В полученных данных отображается 5, если процентная ставка составляет 5%, 8 — если процентная ставка составляет 8%, и т. д. Как быстро разделить все результаты на 100, чтобы, например, пятипроцентная ставка отображалась как 0,05?

В файле Pastespecial.xlsx на листе Paste Special Divide Before (рис. 14.4) содержатся данные годовой ставки процентов, выплаченных по трехмесячным казначейским векселям США, для каждого месяца между январем 1970 г. и февралем 1987 г. В январе 1970 г. годовая ставка по трехмесячным казначейским векселям составляла 8,01%. Нам требуется вычислить заработанный годовой процент на \$1, вложенный по текущей ставке векселей казначейства. Формула для расчета ставки:  $(1 + (\text{годовая ставка})/100)$ . Заработанный процент вычислить проще, если данные в столбце с годовыми процентными ставками разделить на 100.

Область **Операция (Operation)** в диалоговом окне **Специальная вставка (Paste Special)** позволяет вам складывать, вычитать, умножать и делить — то есть производить арифметические действия с каждым числом в диапазоне и заданным числом. Мы хотим разделить на 100 каждое число в столбце **D**. Для начала введите делитель (100). Его можно ввести в любом месте листа, я выбрал **F5** (рис. 14.7). Выделите ячейку **F5**, нажмите **Ctrl+C**. Вы увидите «бегущих муравьев» вокруг ячейки **F5**. Выберите диапазон чисел, которые хотите изменить. Для выделения всех данных в столбце **D** щелкните на ячейке **D10**, нажмите **Ctrl+Shift** и затем **↓**. Этой комбинацией клавиш удобно выбирать большой диапазон в столбце. (Для выбора большого диапазона ячеек в строке перейдите в первую ячейку данных, нажмите **Ctrl+Shift** и затем **→**.) Щелкните правой кнопкой мыши и выберите **Специальная вставка (Paste Special)**, затем установите переключатель **Операция (Operation)** в положение **разделить (Divide)**, как показано на рис. 14.6.

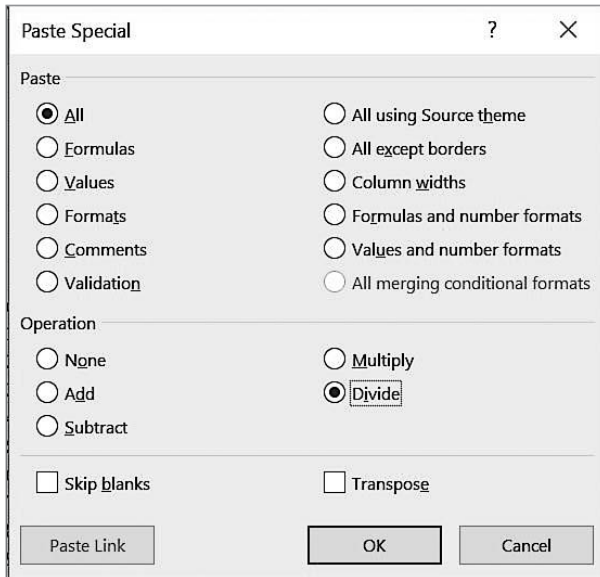
	C	D
9	date	3mo
10	1.1970	8.01
11	2.1970	7.01
12	3.1970	6.48
13	4.1970	7.03
14	5.1970	7.04
15	6.1970	6.52
16	7.1970	6.43
17	8.1970	6.38
18	9.1970	6.03
19	10.1970	5.96
20	11.1970	5.07
21	12.1970	4.9
22	1.1971	4.17
23	2.1971	3.43
24	3.1971	3.64
25	4.1971	4.04
26	5.1971	4.38
27	6.1971	5.12
28	7.1971	5.31

**Рис. 14.5.** Данные для применения операции разделить из меню **Специальная вставка** над диапазоном данных для деления их на константу

Нажмите **OK**, и Excel разделит все выделенные числа в столбце **D** на 100. Результаты вычисления представлены на рис. 14.6. Если выбрать **сложить (Add)**, в ячейке **D10** отобразится значение 108,01. Если выбрать **вычесть (Subtract)**, в **D10** отобразится значение -91,99. И наконец, если выбрать **умножить (Multiply)**, в **D10** отобразится 801.

После того как вы нажмете **OK**, Excel выполнит деление каждой выбранной цифры в столбце **D** на 100. Результат показан на рис. 14.7. Если бы было выбрано **Добавить**, то в ячейке **D10** отобразилось бы число 108,01; если бы было выбрано **Вычесть**, то

значение в D10 стало бы **-99,99**; а если бы было выбрано **Умножить**, то в D10 отразилось бы число **801**.



**Рис. 14.6.** Применение к диапазону ячеек операции из диалогового окна Специальная вставка

	C	D	E	F	G
3					
4					
5				<b>100</b>	
6					
7					
8					
9	<b>date</b>	<b>3mo</b>			
10	<b>1.1970</b>	<b>0.0801</b>			
11	<b>2.1970</b>	<b>0.0701</b>			
12	<b>3.1970</b>	<b>0.0648</b>			
13	<b>4.1970</b>	<b>0.0703</b>			
14	<b>5.1970</b>	<b>0.0704</b>			
15	<b>6.1970</b>	<b>0.0652</b>			
16	<b>7.1970</b>	<b>0.0643</b>			
17	<b>8.1970</b>	<b>0.0638</b>			
18	<b>9.1970</b>	<b>0.0603</b>			

**Рис. 14.7.** Данные после деления на константу с помощью установки переключателя в диалоговом окне Специальная вставка

## Задания

1. В файле **Mavs.xlsx** содержатся статистические данные баскетбольной команды Dallas Mavericks за 2002–2003 гг. Имена игроков перечислены в столбце **A**, статистические категории указаны в строке **3**, а статистические результаты в строках **4–20**. Объедините имена всех игроков в одну строку, а все статистические категории в один столбец.
2. Процент реализации бросков с игры, процент реализации штрафных бросков и процент реализации трехочковых бросков указан в виде десятичных дробей. Например, у Стива Нэша (Steve Nash) процент реализации штрафных бросков равен 91,9%, и на листе указано 0,919. Замените на листе все значения для процентов реализации бросков числами от 1 до 100.
3. В файле **Productpaste.xlsx** содержатся данные о квартальных продажах четырех продуктов. Скопируйте эти данные в другой диапазон и расположите данные о продажах по горизонтали, а не по вертикали. Свяжите скопированные данные с исходными данными таким образом, чтобы изменение данных в строках **5–8** отражалось на расчете данных годовых продаж.
4. В файле **Productsalespaste.xlsx** содержатся данные о продажах продуктов **X**, **Y** и **Z** (в тысячах единиц). Преобразуйте эти данные в данные о продажах в фактических единицах. Подсказка: помните, что для выбора несмежных ячеек можно использовать клавишу **Ctrl** (как было объяснено в главе 2 «Имена диапазонов»).

## ГЛАВА 15

# Трехмерные формулы и гиперссылки

### Обсуждаемые вопросы

- Существует ли простой способ создания многостраничной книги, в которой все листы имеют одинаковую структуру? Можно ли создать формулы со ссылками на ячейки, расположенные на разных листах?
- Я работаю с 200-страничной книгой. Как организовать простую навигацию по страницам?

Из этой главы вы узнаете, как создавать книги с отдельными листами, имеющими одинаковую структуру. Также мы обсудим трехмерные формулы, позволяющие выполнять расчеты с ячейками, расположенными на разных листах. И я покажу, как использовать гиперссылки для удобной навигации по страницам.

## Ответы на вопросы

**❓ Существует ли простой способ создания многостраничной книги, в которой все листы имеют одинаковую структуру? Можно ли создать формулы со ссылками на ячейки, расположенные на разных листах?**

Пусть требуется создать книгу с отдельными листами для отслеживания продаж в разных регионах США (восток, юг, средний запад и запад). Кроме того, необходимо вычислить общий объем продаж на сводном листе. На каждом листе должна отслеживаться цена продукта, себестоимость единицы продукции и количество проданных единиц продукции, а также постоянные расходы и доходы. На сводном листе должна отслеживаться общая прибыль и количество проданных единиц продукции. Региональные листы должны выглядеть так, как на рис. 15.1.

Для создания такой структуры внесите цену продукта в ячейку C3, себестоимость единицы продукции — в ячейку C4, количество проданных единиц продукции — в ячейку C5 и постоянные затраты — в ячейку C6. Затем в ячейке C7 вычислите прибыль по региону Восток по формуле  $=(C3-C4)*C5-C6$ . Теперь вам нужна такая же структура листов для других регионов. Для этого нужно ввести заголовки и формулу в ячейки только на одном листе, а затем Excel автоматически скопирует их на листы для других регионов.

	A	B	C	D	E
1	East				
2					
3		Price	10		
4		Unit cost	5		
5		Units sold	35		
6		Fixed cost	100		
7		Profit	75	=(C3-C4)*C5-C6	

**Рис. 15.1.** Объемы продаж на востоке

Сначала откройте пустую книгу, содержащую по умолчанию один лист. Щелкнув на значке **Новый лист** ("+" ) справа от последнего из названных листов (или нажав Shift+F11), вставьте четыре новых листа (всего в книге должно получиться пять листов). Назовите четыре первых листа **Восток**, **Юг**, **Средний Запад** и **Запад**. Последний лист назовите **Сводка**. Общий объем продаж будет подсчитан на листе **Сводка**. (Кстати, если на вкладке **Файл** (File) выбрать **Параметры** (Options), то в группе **Общие** (General) можно изменить количество листов, включаемых в книгу по умолчанию, введя нужное значение в поле **Число листов** (This Many Sheets) в разделе **При создании новых книг** (When Creating New Workbooks).)

Для заполнения данных по регионам выберите первый лист ( **Восток**); удерживая клавишу **Shift**, выберите лист последнего региона (**Запад**). Теперь всё, что вводится на лист **Восток**, дублируется на листах для остальных регионов. Введите **Цена** в ячейку B3, **Себестоимость** в ячейку B4, **Продано** в ячейку B5, **Постоянные затраты** в ячейку B6 и **Прибыль** в ячейку B7. Наконец, введите формулу  $=(C3-C4)*C5-C6$  в ячейку C7. Затем для завершения процедуры ввода данных щелкните на имени листа последнего региона. На листах для всех регионов отобразились одинаковые заголовки в столбце B и правильная формула вычисления прибыли в ячейке C7.

Теперь вы готовы использовать трехмерные формулы для вычисления общего количества проданных единиц продукции и общей прибыли. Напомню, что для каждого региона количество проданных единиц указывается в ячейке C5. На листе **Сводка**, на котором будет вычислено общее количество проданных единиц продукции, поставьте курсор в ячейку C5. Начните вводить формулу **=СУММ** и затем переместите курсор в первую ячейку, значение которой требуется включить в итоги (ячейка C5 на листе **Восток**). Далее, удерживая Shift, щелкните на имени последнего листа, данные с которого включаются в итоги (лист **Запад**). Наконец, введите в строке формул закрывающую скобку, и в ячейке C5 листа **Сводка** появится формула **=СУММ(восток:запад!C5)**. Это трехмерная формула. Большинство формул Excel оперирует в двух измерениях (строки и столбцы). Трехмерная формула действует в трех измерениях: сквозь листы. По этой формуле суммируются значения из ячеек C5 на всех листах, начиная с листа **Восток** и заканчивая листом **Запад**. Конечно, если хотите, вы можете просто ввести эту формулу в ячейку C5 листа **Сводка**. Теперь для вычисления общей прибыли скопируйте эту формулу и вставьте ее в ячейку C7 листа **Сводка** (рис. 15.2). После этого вам останется только ввести заголовки в лист **Сводка** и значения в листы отдельных регионов. В файле **Threedim.xlsx** представлен конечный результат.



	A	B	C	D
1	Summary			
2				
3				
4				
5		Units sold	140	=SUM(East:West!C5)
6				
7		Profit	300	=SUM(East:West!C7)

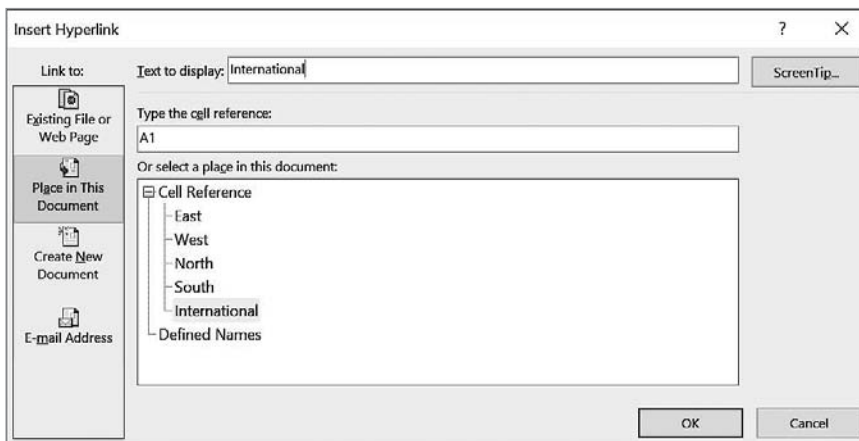
**Рис. 15.2.** Суммирование количества проданных единиц продукции и общей прибыли

В следующих главах вы узнаете о еще четырех способах суммирования данных с разных листов или из разных книг:

- Сводные таблицы и срезы для описания данных (глава 45).
- Модель данных (глава 46).
- Power Pivot (глава 47).
- Консолидация данных (глава 52).

### ❓ Я работаю с 200-страничной книгой. Как организовать простую навигацию по страницам?

Один из простых способов организовать навигацию между листами — использовать гиперссылки. Книга *Hyperlinktemp.xlsx* (в папке *Template* этой главы) содержит пять листов. Согласитесь, было бы удобно иметь на первом листе ссылку, активируемую щелчком мыши, которая позволила бы перейти в ячейку **A1** любого листа? Чтобы создать ссылку на лист *International* в ячейке **A10** первого листа, поместите курсор в ячейку **A10** и выберите **Вставка (Insert), Гиперссылка (Hyperlink)** (или нажмите **Ctrl+K**). Затем заполните поля в диалоговом окне, как показано на рис. 15.3.



**Рис. 15.3.** Создание гиперссылки

Если щелкнуть на гиперссылке *International* в ячейке A10 (рис. 15.4), мы немедленно перейдем в ячейку A1 листа *International*. Аналогично можно создать ссылки на ячейку A1 на других листах (см. файл *Hyperlinks.xlsx* и рис. 15.4). Так же легко вы можете создать гиперссылку на веб-страницу, позицию в другом файле или на адрес электронной почты.

	F	G
6	<u>East</u>	
7	<u>West</u>	
8	<u>North</u>	
9	<u>South</u>	
10	<u>International</u>	

**Рис. 15.4.** Гиперссылки для связи с другими листами в книге

Функцию ГИПЕРССЫЛКА (HYPERLINK) можно использовать для облегчения задачи создания большого количества ссылок. Как — показано в файле *Hyperlinkfunction.xlsx* и на рис. 15.5. Синтаксис функции ГИПЕРССЫЛКА: ГИПЕРССЫЛКА(расположение ссылки, читабельное имя для ссылки). Так, в ячейках D3:D5 содержатся веб-адреса трех сайтов. В ячейках C3:C5 помещены ярлыки для попадания на эти сайты. Скопировав формулу ГИПЕРССЫЛКА (D3,C3) из B3 в B4 :B5, мы создадим «ярлыковые» ссылки на желаемые веб-сайты. В главе 23 «Функция ДВССЫЛ» рассказывается, как использовать функции ГИПЕРССЫЛКА и ДВССЫЛ для автоматизации процесса создания оглавления, содержащего гиперссылки на каждый лист в рабочей книге.

	A	B	C	D
1				
2				
3	=HYPERLINK(D3,C3)	Hyperlink	Hyperlink	<a href="https://support.office.com/en-us/article/HYPERLINK-function-333c7ce6-c5ae-4164-9c47-7de9b76f577f">https://support.office.com/en-us/article/HYPERLINK-function-333c7ce6-c5ae-4164-9c47-7de9b76f577f</a>
4	=HYPERLINK(D4,C4)	Flash fill	Flash fill	<a href="https://support.office.com/en-us/article/Video-Use-AutoFill-and-Flash-Fill-2e79a709-c814-4b27-8bc2-c4dc84d49464">https://support.office.com/en-us/article/Video-Use-AutoFill-and-Flash-Fill-2e79a709-c814-4b27-8bc2-c4dc84d49464</a>
5	=HYPERLINK(D5,C5)	Excel support	Excel support	<a href="https://support.office.com/en-us/excel">https://support.office.com/en-us/excel</a>

**Рис. 15.5.** Пример использования функции ГИПЕРССЫЛКА

Далее приведены некоторые другие способы, помогающие перемещаться по многостраничным книгам:

- нажатие **Ctrl+Page Down** перемещает вас на следующий лист книги, а **Ctrl+Page Up** — на предыдущий лист;
- по щелчку правой кнопкой мыши на стрелках в левой нижней части экрана открывается список листов текущей книги, как показано на рис. 15.6. Для перехода на нужный лист достаточно щелкнуть по его названию;
- предположим, ваша книга содержит много листов (см. файл *Multiplework sheetstemp.xlsx*). Если вы находитесь недалеко от первого листа и наведете курсор на стрелки слева от закладки первого листа, а затем нажмете **Ctrl+щелчок** левой кнопкой мыши, то вы перейдете на последний лист. Если вы находитесь

недалеко от последнего листа и наведете курсор на стрелки слева от закладки первого листа, а затем нажмете Ctrl+щелчок левой кнопкой мыши, то перейдете на первый лист.

## Задания

1. Вы владелец шести кондитерских. Выручка и количество клиентов для каждой кондитерской представлены в таблице ниже.

Кондитерская	Выручка	Количество клиентов
1	\$8000	1950
2	\$7000	1800
3	\$9000	2200
4	\$8400	2000
5	\$5900	1400
6	\$10 100	2500

Создайте книгу для ввода суммы выручки и количества клиентов по каждой кондитерской, а также сводный лист (с помощью трехмерных формул), на котором вычисляются общая выручка за неделю и общее количество клиентов.

2. Поместите в файле `Multipleworksheetstemp.xlsx` гиперссылки на ячейку A5 в листах Лист 3, Лист 5 и Лист 7, а также на веб-сайт [Officeblog.com](http://Officeblog.com).

## ГЛАВА 16

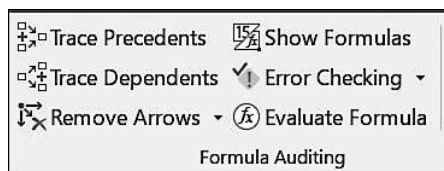
# Инструменты проверки зависимостей и надстройка Inquire

### Обсуждаемые вопросы

- Я только что получил лист с 5000 строк, на котором вычисляется чистая приведенная стоимость (ЧПС) нового автомобиля. На этом листе мой финансовый аналитик сделал предположение о годовом проценте роста цены продукта. Каких ячеек на листе касается это предположение?
- Мне кажется, что мой финансовый аналитик сделал ошибку при вычислении прибыли до уплаты налогов за первый год. Какие ячейки на модели листа использовались в этом расчете?
- Как инструменты проверки зависимостей работают с данными на нескольких листах или в нескольких книгах?
- Что такое надстройка Inquire и как ее установить?
- Как применить надстройку Inquire для сравнения книг?
- Как применить надстройку Inquire для анализа структуры книги?
- Как применить надстройку Inquire для анализа связей между листами и книгами?
- Как применить надстройку Inquire для анализа влияющих и зависимых ячеек для конкретной ячейки?
- Как с помощью Inquire очистить ячейку от избыточного форматирования?

Когда мы слышим слово «структура», на ум приходит образ каких-то построений. Применительно к структуре модели листа мы говорим о способе, каким входные предположения (данные, например, объем, цена и себестоимость единицы продукции) используются при расчетах представляющих интерес результатов, например ЧПС, прибыли или стоимости. В Microsoft Excel 2019 инструмент проверки зависимостей предоставляет простой способ выявления структуры листа, что упрощает понимание логики содержащихся на нем сложных моделей документирования структуры листа или книги.

Для просмотра возможностей инструментов проверки зависимостей в Excel перейдите на вкладку **Формулы (Formulas)** ленты Excel и откройте группу **Зависимости формул (Formula Auditing)** (рис. 16.1). В данной главе я подробно рассмотрю большинство инструментов. Инструмент **Вычислить формулу (Evaluate Formula)** подробно рассмотрен в главе 22 («Функция СМЕЩ»).



**Рис. 16.1.** Инструментальная панель Зависимости формул

Значок Показать формулы (Show Formulas) служит переключателем между режимом просмотра в ячейках формул и режимом просмотра значений, полученных в результате вычисления этих формул. Переключиться между режимом просмотра формул и режимом просмотра результатов можно также нажатием клавиш **Ctrl+~**. Кроме того, формулы можно просмотреть в сводной таблице с помощью новой функции, появившейся в Microsoft Excel 2013, — **Ф.ТЕКСТ (FORMULATEXT)**. В файле **ISFORMULA\_TEXT.xlsx** (рис. 16.2) показано, как применять эту функцию.

	A	B	C	D	E
			Column A ISFORMULA?	Column B ISFORMULA?	FORMULATEXT
3	x	y			
4	1	5	FALSE	TRUE	=5*A4
5	2	10	FALSE	TRUE	=5*A5
6	3	15	FALSE	TRUE	=5*A6

**Рис. 16.2.** Применение функций ЕФОРМУЛА и Ф.ТЕКСТ

В столбец А мы ввели числа 1, 2 и 3, а в столбце В умножили их на 5. После копирования формулы **=Ф.ТЕКСТ(B4)** из Е4 в Е5:Е6 в столбце Е появятся формулы из столбца В.

В столбцах С и D показано применение другой функции, впервые появившейся в Microsoft Excel 2013, — **ЕФОРМУЛА (ISFORMULA)**. Эта функция возвращает значение **ИСТИНА**, если ячейка содержит формулу, и **ЛОЖЬ** в противном случае. После копирования формулы **=ЕФОРМУЛА(A4)** из ячейки С4 в ячейки С4:D6 становится понятно, что в столбце В содержатся формулы, а в столбце А их нет.

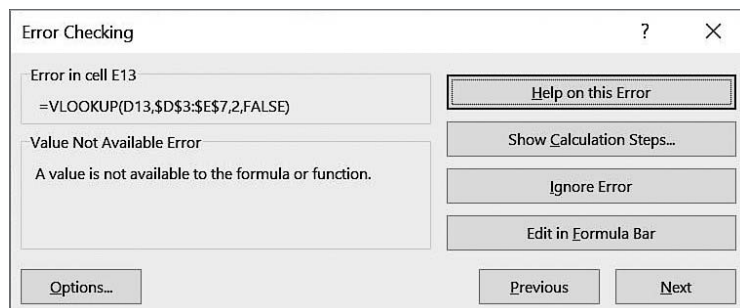
С помощью инструмента Проверка наличия ошибок (Error Checking) лист проверяется на наличие ошибок, и по имеющимся ошибкам выдается подсказка. Для иллюстрации возможностей Excel по перехвату ошибок рассмотрим еще раз файл **Errortrap.xlsx** из главы 12 (рис. 16.3). (Файл **Errortrap.xlsx** также лежит в папке **Practice Files** главы 16, но это тот же самый файл.)

В ячейке Е13 вы можете нажать на стрелку значка ошибки и открыть меню возможных действий. Или, в группе Зависимости формул, вы можете нажать на стрелку значка проверки наличия ошибок и вызвать три инструмента: Проверка наличия ошибок (Error Checking), Источник ошибки (Trace Errors) и Циклические ссылки (Circular

References). Наша книга не содержит циклических ссылок, поэтому инструмент Циклические ссылки (Circular References) недоступен. Если в книге есть циклические ссылки, этот инструмент тоже будет выделен. Если вы выберете Проверка наличия ошибок (Error Checking), Excel дойдет до первой ячейки с ошибкой (первая ошибка будет найдена в ячейке E13) и появится диалоговое окно, показанное на рис. 16.4. Для получения дополнительной информации об ошибке нажмите кнопку Справка по этой ошибке (Help On This Error). Кнопка Показать этапы вычисления (Show Calculation Steps) позволяет произвести вычисления по формуле пошагово. Нажатие кнопки Пропустить ошибку (Ignore Error) закрывает диалоговое окно. Кнопка Изменить в строке формул (Edit In Formula Bar) предназначена для редактирования формулы. При нажатии кнопки Назад (Previous) происходит возврат к предыдущей найденной ошибке, а при нажатии кнопки Далее (Next) — переход в диалоговое окно следующей ошибки.

	C	D	E	F	G
1					
2			salary		
3		Jane	40		
4		Jack	60		
5		Jill	70		
6		Erica	34		
7		Adam	120		
8					
9					
10		Name	Salary	errortrapped	
11		Erica	34	34	
12		Adam	120	120	
13		JR	#N/A		
14		Josh	#N/A		
15		Jill	70	70	
16		average	#N/A	74.66667	
17					
18	aggregate		74.66667		
19	function				
20	works!				

**Рис. 16.3.** Применение инструментов Проверка наличия ошибок и Источник ошибки



**Рис. 16.4.** Диалоговое окно Контроль ошибок

Если при выборе инструмента **Источник ошибки (Trace Errors)** активная ячейка содержит ошибку, ячейки, необходимые для расчета в активной ячейке, обозначаются синими стрелками или синим прямоугольником. Например, на рис. 16.3 для расчета в ячейке **E13** необходимы ячейка **D13** и диапазон ячеек **D3:E7**. (В ячейке **E13** содержится сообщение об ошибке **#Н/Д**.)

При использовании инструмента **Окно контрольного значения (Watch Window)** вы можете выделить ячейку или ячейки и затем независимо от своего местонахождения в книге следить за изменением значений в выбранных ячейках. Например, в диалоговом окне **Окно контрольного значения (Watch Window)** можно выбрать инструмент **Добавить контрольное значение (Add Watch)** и добавить формулу (скажем, в ячейку **C3** на листе **Лист1**). Теперь даже если вы будете работать в другой книге, в диалоговом окне **Окно контрольного значения** можно будет следить за изменением значения в ячейке **C3** на листе **Лист1**.

Кнопки **Влияющие ячейки (Trace Precedents)** и **Зависимые ячейки (Trace Dependents)** на панели проверки зависимостей являются инструментами поиска и просмотра влияний и зависимостей для ячеек или формул на листе. *Влияющей* является любая ячейка, значение в которой необходимо для расчета значения выделенной формулы. Например, анализируя кампанию прямой почтовой рассылки, вы можете сделать предположения о количестве отправленных писем и проценте ответивших. Затем можно вычислить общее количество ответов как **(процент\_ответивших\*количество\_отправленных\_писем)**. В этом случае **процент\_ответивших** и **количество\_отправленных\_писем** оказывают влияние на ячейку с формулой расчета общего количества ответов. *Зависимой* является любая ячейка, содержащая формулу, значение которой не может быть вычислено, если неизвестно значение выделенной ячейки. В предыдущем примере ячейка, содержащая общее количество ответов, является зависимой от ячейки, содержащей процент ответивших. При использовании этих инструментов проверки зависимостей Excel автоматически помечает влияющие и зависимые ячейки синими стрелками. Инструмент **Убрать стрелки (Remove Arrows)** на панели проверки зависимостей убирает отображение стрелок. Чтобы найти влияющие или зависимые ячейки, нужно, чтобы была отображена вся панель проверки зависимостей. Для переключения между ее компактным и развернутым видом можно воспользоваться комбинацией **Ctrl+F1**.

Давайте применим инструменты проверки зависимостей к решению некоторых практических задач.

## Ответы на вопросы

❓ Я только что получил лист с 5000 строк, на котором вычисляется чистая приведенная стоимость (ЧПС) нового автомобиля. На этом листе мой финансовый аналитик сделал предположение о годовом проценте роста цены продукта. Каких ячеек на листе касается это предположение?

В файле NPVaudit.xlsx на листе Original Model содержится расчет ЧПС на основе прибыли после уплаты налогов для нового автомобиля, который, как ожидается, поступит от производителя через пять лет (рис. 16.5). Цена и себестоимость указаны в тысячах долларов. Значения параметров, принятые для анализа, представлены в ячейках C1:C8 (с соответствующими именами в ячейках B1:B8). Предположим, что цена продукта растет на 3% в год. Какие ячейки листа зависят от этого предположения?

	A	B	C	D	E	F
1		taxrate	0.4			
2		Year1sales	10000			
3		Sales growth	0.1			
4		Year1price	\$ 9.00			
5		Year1cost	\$ 6.00			
6		intrate	0.15			
7		costgrowth	0.05			
8		pricegrowth	0.03			
9	Year	1	2	3	4	5
10	Unit Sales	10000	11000	12100	13310	14641
11	unit price	\$ 9.00	\$ 9.27	\$ 9.55	\$ 9.83	\$ 10.13
12	unit cost	\$ 6.00	\$ 6.30	\$ 6.62	\$ 6.95	\$ 7.29
13	Revenues	\$ 90,000.00	\$ 101,970.00	\$ 115,532.01	\$ 130,897.77	\$ 148,307.17
14	Costs	\$ 60,000.00	\$ 69,300.00	\$ 80,041.50	\$ 92,447.93	\$ 106,777.36
15	Before Tax Profits	\$ 30,000.00	\$ 32,670.00	\$ 35,490.51	\$ 38,449.83	\$ 41,529.81
16	Tax paid	\$ 12,000.00	\$ 13,068.00	\$ 14,196.20	\$ 15,379.93	\$ 16,611.92
17	Aftertax Profits	\$ 18,000.00	\$ 19,602.00	\$ 21,294.31	\$ 23,069.90	\$ 24,917.89
18						
19	NPV	\$70,054.34				

**Рис. 16.5.** Использование инструментов проверки зависимости для отслеживания формул в сложных таблицах

Для ответа на этот вопрос выделите ячейку C8 (ячейку с предположением о 3%-ном росте цены) и затем на вкладке **Формулы (Formulas)** в группе **Зависимости формул (Formula Auditing)** нажмите кнопку **Зависимые ячейки (Trace Dependents)**. Появится набор стрелок (рис. 16.6), указывающий на ячейки, непосредственно зависящие от ячейки C8.

При однократном нажатии на кнопку **Зависимые ячейки (Trace Dependents)** Excel указывает на ячейки, находящиеся в прямой зависимости от предположения о росте цены. На рис. 16.6 показано, что от этого предположения непосредственно зависит только цена единицы продукции для годов 2–5. При многократном нажатии кнопки **Зависимые ячейки (Trace Dependents)** стрелками будут отмечены все формулы, в которых для расчета требуется значение ежегодного роста цены (рис. 16.7).

Как вы видите, предположение о росте цены влияет не только на цену единицы продукции в годы 2–5, но и на выручку в годы 2–5, прибыль до уплаты налогов, сумму налогов, прибыль после уплаты налогов и ЧПС. Стрелки можно удалить с помощью инструмента **Убрать стрелки (Remove Arrows)**.



	A	B	C	D	E	F
1		taxrate	0.4			
2		Year1sales	10000			
3		Sales growth	0.1			
4		Year1price	\$ 9.00			
5		Year1cost	\$ 6.00			
6		intrate	0.15			
7		costgrowth	0.05			
8		pricegrowth	0.03			
9	Year					
10	Unit Sales	10000	11000	12100	13310	14641
11	unit price	\$ 9.00	\$ 9.27	\$ 9.55	\$ 9.83	\$ 10.13
12	unit cost	\$ 6.00	\$ 6.30	\$ 6.62	\$ 6.95	\$ 7.29
13	Revenues	\$ 90,000.00	\$ 101,970.00	\$ 115,532.01	\$ 130,897.77	\$ 148,307.17
14	Costs	\$ 60,000.00	\$ 69,300.00	\$ 80,041.50	\$ 92,447.93	\$ 106,777.36
15	Before Tax Profits	\$ 30,000.00	\$ 32,670.00	\$ 35,490.51	\$ 38,449.83	\$ 41,529.81
16	Tax paid	\$ 12,000.00	\$ 13,068.00	\$ 14,196.20	\$ 15,379.93	\$ 16,611.92
17	Aftertax Profits	\$ 18,000.00	\$ 19,602.00	\$ 21,294.31	\$ 23,069.90	\$ 24,917.89
18						
19	NPV	\$70,054.34				

Рис. 16.6. Отслеживание непосредственно зависящих ячеек

	A	B	C	D	E	F
1		taxrate	0.4			
2		Year1sales	10000			
3		Sales growth	0.1			
4		Year1price	\$ 9.00			
5		Year1cost	\$ 6.00			
6		intrate	0.15			
7		costgrowth	0.05			
8		pricegrowth	0.03			
9	Year					
10	Unit Sales	10000	11000	12100	13310	14641
11	unit price	\$ 9.00	\$ 9.27	\$ 9.55	\$ 9.83	\$ 10.13
12	unit cost	\$ 6.00	\$ 6.30	\$ 6.62	\$ 6.95	\$ 7.29
13	Revenues	\$ 90,000.00	\$ 101,970.00	\$ 115,532.01	\$ 130,897.77	\$ 148,307.17
14	Costs	\$ 60,000.00	\$ 69,300.00	\$ 80,041.50	\$ 92,447.93	\$ 106,777.36
15	Before Tax Profits	\$ 30,000.00	\$ 32,670.00	\$ 35,490.51	\$ 38,449.83	\$ 41,529.81
16	Tax paid	\$ 12,000.00	\$ 13,068.00	\$ 14,196.20	\$ 15,379.93	\$ 16,611.92
17	Aftertax Profits	\$ 18,000.00	\$ 19,602.00	\$ 21,294.31	\$ 23,069.90	\$ 24,917.89
18						
19	NPV	\$70,054.34				

Рис. 16.7. При многократном нажатии кнопки Зависимые ячейки проявятся все ячейки, зависящие от предположения о росте цены

При нажатии Ctrl+] синим цветом отмечаются все ячейки, непосредственно зависящие от активной ячейки; при нажатии клавиш Ctrl+Shift+] синим цветом отмечаются все ячейки, зависящие от активной ячейки.

❓ Мне кажется, что мой финансовый аналитик сделал ошибку при вычислении прибыли до уплаты налогов за первый год. Какие ячейки на модели листа использовались в этом расчете?

Здесь нам нужно найти влияющие ячейки для B15, необходимые для расчета прибыли за год 1 до уплаты налогов. Выберите ячейку B15 и нажмите кнопку Влияющие ячейки (Trace Precedents) один раз. Появятся стрелки, как на рис. 16.8.

Как видите, для расчета прибыли за год 1 до уплаты налогов необходимы значения в ячейках с выручкой за год 1 и затратами за год 1. (Прибыль за год 1 до уплаты

	A	B	C	D	E	F
1		taxrate	0.4			
2		Year1sales	10000			
3		Sales growth	0.1			
4		Year1price	\$ 9.00			
5		Year1cost	\$ 6.00			
6		intrate	0.15			
7		costgrowth	0.05			
8		pricegrowth	0.03			
9	Year	1	2	3	4	5
10	Unit Sales	10000	11000	12100	13310	14641
11	unit price	\$ 9.00	\$ 9.27	\$ 9.55	\$ 9.83	\$ 10.13
12	unit cost	\$ 6.00	\$ 6.30	\$ 6.62	\$ 6.95	\$ 7.29
13	Revenues	\$ 90,000.00	\$ 101,970.00	\$ 115,532.01	\$ 130,897.77	\$ 148,307.17
14	Costs	\$ 60,000.00	\$ 69,300.00	\$ 80,041.50	\$ 92,447.93	\$ 106,777.36
15	Before Tax Profits	\$ 30,000.00	\$ 32,670.00	\$ 35,490.51	\$ 38,449.83	\$ 41,529.81
16	Tax paid	\$ 12,000.00	\$ 13,068.00	\$ 14,196.20	\$ 15,379.93	\$ 16,611.92
17	Aftertax Profits	\$ 18,000.00	\$ 19,602.00	\$ 21,294.31	\$ 23,069.90	\$ 24,917.89
18						
19	NPV	\$70,054.34				

**Рис. 16.8.** Ячейки, непосредственно влияющие на ячейку с величиной прибыли за год 1 до уплаты налогов

	A	B	C	D	E	F
1		taxrate	0.4			
2		Year1sales	10000			
3		Sales growth	0.1			
4		Year1price	\$ 9.00			
5		Year1cost	\$ 6.00			
6		intrate	0.15			
7		costgrowth	0.05			
8		pricegrowth	0.03			
9	Year	1	2	3	4	5
10	Unit Sales	10000	11000	12100	13310	14641
11	unit price	\$ 9.00	\$ 9.27	\$ 9.55	\$ 9.83	\$ 10.13
12	unit cost	\$ 6.00	\$ 6.30	\$ 6.62	\$ 6.95	\$ 7.29
13	Revenues	\$ 90,000.00	\$ 101,970.00	\$ 115,532.01	\$ 130,897.77	\$ 148,307.17
14	Costs	\$ 60,000.00	\$ 69,300.00	\$ 80,041.50	\$ 92,447.93	\$ 106,777.36
15	Before Tax Profits	\$ 30,000.00	\$ 32,670.00	\$ 35,490.51	\$ 38,449.83	\$ 41,529.81
16	Tax paid	\$ 12,000.00	\$ 13,068.00	\$ 14,196.20	\$ 15,379.93	\$ 16,611.92
17	Aftertax Profits	\$ 18,000.00	\$ 19,602.00	\$ 21,294.31	\$ 23,069.90	\$ 24,917.89
18						
19	NPV	\$70,054.34				

**Рис. 16.9.** Для просмотра всех ячеек, влияющих на величину прибыли за первый год до уплаты налогов, нажмите кнопку Влияющие ячейки несколько раз

налогов равна выручке за год 1 минус затраты за год 1.) Нажимайте кнопку **Влияющие ячейки (Trace Precedents)** до тех пор, пока не появятся все ячейки, влияющие на ячейку с величиной прибыли за год 1 до уплаты налогов (рис. 16.9). Здесь на прибыль за год 1 до уплаты налогов влияют продажи, цена и себестоимость единицы продукции в год 1.

### ? Как инструменты проверки зависимостей работают с данными на нескольких листах или в нескольких книгах?

Рассмотрим простую модель листа в файле `Audittwosheets.xlsx`, показанную на рис. 16.10. На листе **Profit** по формуле (количество проданных единиц\*(цена – переменные затраты) – постоянные затраты) вычисляется прибыль компании на основе данных на листе **Data**.

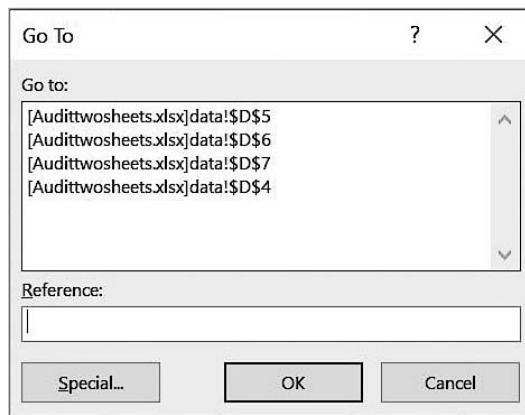
Предположим, вы хотите найти влияющие ячейки для ячейки с формулой прибыли. На листе **Profit** выделите ячейку D7 и на вкладке **Формулы (Formulas)** в группе **Зависимости формул (Formula Auditing)** нажмите кнопку **Влияющие ячейки (Trace Precedents)**. Появится пунктирная линия, стрелка и значок листа (рис. 16.11).

	C	D
1		
2		
3		
4	Fixed cost	10000
5	Unit Sales	3000
6	Price	\$ 7.50
7	Variable cost	\$ 3.20

**Рис. 16.10.** Применение инструментов проверки зависимостей к данным на нескольких листах

	B	C	D
1			
2			
3			
4			
5			
6			Profit
7			\$2,900.00

**Рис. 16.11.** Результат отслеживания влияющих ячеек для данных на нескольких листах



**Рис. 16.12.** Проверка зависимостей на нескольких листах с помощью диалогового окна **Переход**

Значок листа показывает, что влияющие ячейки для формулы прибыли находятся на другом листе. После двойного щелчка на пунктирной линии откроется диалоговое окно **Переход (Go To)**, как на рис. 16.12.

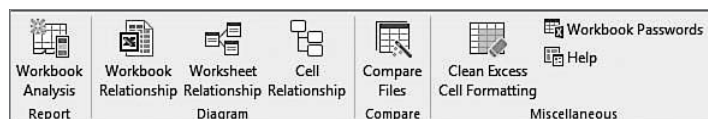
В нем можно выбрать любую влияющую ячейку (из списка ячеек D4:D7 на листе Data) и нажать **OK** для перехода к выбранной влияющей ячейке.

### ? Что такое надстройка Inquire и как ее установить?

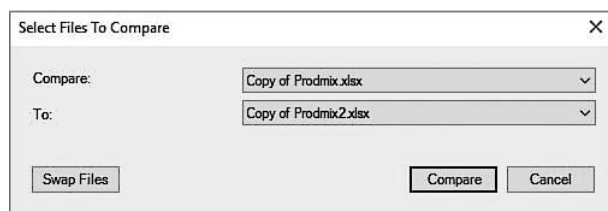
Надстройка **Inquire** может быть использована самыми разнообразными путями для изучения структуры листов и книг, их сравнения и изучения взаимосвязей между ними. Надстройка **Inquire** имеется только в Office Professional Plus и Office 365 Plus. Чтобы установить надстройку **Inquire**, щелкните **Настройки** в меню **Файл** и затем выберите **Надстройки (Add-Ins)**. В диалоговом окне **Управление** выберите **Com.Add-ins**, затем **Перейти**, вслед за чем поставьте флажок с надстройкой **Inquire (Запрос)**. Если вы не видите надпись **Inquire**, значит, в вашей версии Excel она не установлена.

### ? Как применить надстройку Inquire для сравнения книг?

Предположим, Джилл создала файл **Copy-of-Prodmi.xlsx**, а Джеймс изменил его и переименовал в **Copy-of-Prodmix2.xlsx**. Дипломированному бухгалтеру фирмы Джоан необходимо знать, какие конкретно изменения Джеймс внес в этот файл. Операция **Сравнить файлы** в **Inquire** — это как раз то, что ей нужно. Щелкнув по панели инструментов **Inquire** (рис. 16.13), вы увидите имеющиеся в вашем распоряжении инструменты. Чтобы применить операцию **Сравнить файлы**, откройте те два файла, которые вы хотите сравнить (рис. 16.14).



**Рис. 16.13.** Панель инструментов **Inquire**



**Рис. 16.14.** Диалоговое окно **Выбрать файлы для сравнения**

Мы видим (рис. 16.15), что во втором файле добавился столбец **J**, а также были изменены данные в ячейке **D4** и формулы в **D14** и **D15**.

Sheet	Cell	Value 1	Value 2	Change Description	
optimal				Added Column J.	
optimal	D14	=SUMPRODU...	=SUMPRODU...	Formula Changed.	
optimal	D15	=SUMPRODU...	=SUMPRODU...	Formula Changed.	
optimal	D4	6	7	Entered Value Changed.	
optimal	D14	4499.9999999...	4509.9999999...	Calculated Value Changed.	
optimal	D15	1236.1333333...	1251.1333333...	Calculated Value Changed.	

Рис. 16.15. Обнаруженные ключевые различия

### ❓ Как применить надстройку Inquire для анализа структуры книги?

В файле Salessummary.xlsx использованы данные из файла Salesdata.xlsx для расчета кумулятивного объема продаж за каждый квартал. Открыв файл Salessummary.xlsx, вы, возможно, захотите собрать много разной информации о структуре этого файла. На Панели инструментов Inquire выберите Отчет об анализе книги. Появится диалоговое окно для формирования Отчета об анализе книги, где вы сможете отметить, что вас интересует. Это может быть схема связей книги, сведения обо всех формулах, скрытых строках и столбцах, поименованных объектах и поименованных объектах с ошибками (Linked Workbooks, All Formulas, Hidden Rows And Columns, Named Items, Named Items With Errors). На рис. 16.16, например, показаны все формулы, имеющиеся в книге Salessummary.xlsx. При желании полученная информация может быть экспортирована в Excel (см. файл Salesananalysissummary.xlsx).

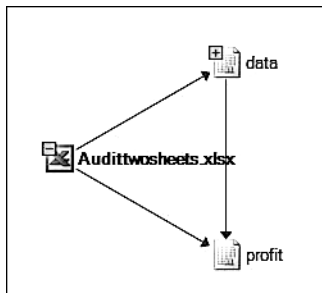
	A	B	C	D
1	<b>All Formulas (12 total)</b>			
2	E:\salessummary.xlsx			
3				
4	Sheet Name	Cell Address	Formula	Value
5	Q1	C2	=SUMIF(E:\[Salesdata.xlsx]Q1 data!C\$6:C\$20,B2,E:\[Salesdata.xlsx]Q1 data!D\$6:D\$20)	442.3333333
6	Q1	C3	=SUMIF(E:\[Salesdata.xlsx]Q1 data!C\$6:C\$20,B3,E:\[Salesdata.xlsx]Q1 data!D\$6:D\$20)	319.3333333
7	Q1	C4	=SUMIF(E:\[Salesdata.xlsx]Q1 data!C\$6:C\$20,B4,E:\[Salesdata.xlsx]Q1 data!D\$6:D\$20)	355.3333333
8	Q2	C2	=SUMIF(E:\[Salesdata.xlsx]Q2data!\$C\$6:\$C\$17,Q2!B2,E:\[Salesdata.xlsx]Q2data!\$D\$6:\$D\$17)*Q1!C2	724.3333333
9	Q2	C3	=SUMIF(E:\[Salesdata.xlsx]Q2data!\$C\$6:\$C\$17,Q2!B3,E:\[Salesdata.xlsx]Q2data!\$D\$6:\$D\$17)*Q1!C3	597.3333333
10	Q2	C4	=SUMIF(E:\[Salesdata.xlsx]Q2data!\$C\$6:\$C\$17,Q2!B4,E:\[Salesdata.xlsx]Q2data!\$D\$6:\$D\$17)*Q1!C4	647.3333333
11	Q3	C2	=SUMIF(E:\[Salesdata.xlsx]Q3data!\$D\$7:\$D\$18,B2,E:\[Salesdata.xlsx]Q3data!\$E\$7:\$E\$18)*Q2!C2	980.3333333
12	Q3	C3	=SUMIF(E:\[Salesdata.xlsx]Q3data!\$D\$7:\$D\$18,B3,E:\[Salesdata.xlsx]Q3data!\$E\$7:\$E\$18)*Q2!C3	818.3333333
13	Q3	C4	=SUMIF(E:\[Salesdata.xlsx]Q3data!\$D\$7:\$D\$18,B4,E:\[Salesdata.xlsx]Q3data!\$E\$7:\$E\$18)*Q2!C4	893.3333333
14	Q4	C2	=SUMIF(E:\[Salesdata.xlsx]Q4data!\$C\$5:\$C\$16,B2,E:\[Salesdata.xlsx]Q4data!\$D\$5:\$D\$16)*Q3!C2	1254.3333333
15	Q4	C3	=SUMIF(E:\[Salesdata.xlsx]Q4data!\$C\$5:\$C\$16,B3,E:\[Salesdata.xlsx]Q4data!\$D\$5:\$D\$16)*Q3!C3	1156.3333333
16	Q4	C4	=SUMIF(E:\[Salesdata.xlsx]Q4data!\$C\$5:\$C\$16,B4,E:\[Salesdata.xlsx]Q4data!\$D\$5:\$D\$16)*Q3!C4	1150.3333333

Рис. 16.16. Перечень формул из рабочей книги Salesdata.xlsx

### ❓ Как применить надстройку Inquire для анализа связей между листами и книгами?

Заглянем еще раз в книгу Audittwosheets.xlsx. Если мы хотим понять, как взаимосвязаны в ней листы, то нужно открыть книгу и, находясь в любой ячейке, кликнуть иконку Взаимосвязи листов (Worksheet Relationship) на панели инструментов Inquire (рис. 16.17). Эта красивая диаграмма показывает что данные для листа Data

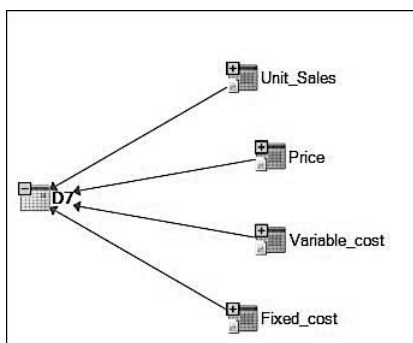
содержатся на листе **Profit**. Если вы откроете сразу несколько книг, то сможете получить информацию о связях между ними всеми, щелкнув значок **Отчет о взаимосвязях книг**.



**Рис. 16.17.** Отчет о взаимосвязях листов

### ? Как применить надстройку **Inquire** для анализа влияющих и зависимых ячеек для конкретной ячейки?

Предположим, вы хотите узнать, какая информация нужна для того, чтобы рассчитать прибыль на листе **Profit** книги **Audittwosheets.xlsx**. Просто поместите курсор в ячейку **D7** листа расчета прибыли и выберите **Взаимосвязи между ячейками (Cell Relationship)** на вкладке **Inquire**. Затем выберите **Влияющие ячейки (Trace Precedents)**, и появится чудесная диаграмма, показанная на рис. 16.18.



**Рис. 16.18.** Три ячейки на листе данных, используемые в расчете прибыли

### ? Как с помощью **Inquire** очистить ячейку от избыточного форматирования?

В Excel 2013 или более поздней версии в вашем распоряжении имеется до 64 000 форматов ячеек. Выбрав **Удалить лишнее форматирование ячеек (Clean Excel Cell Formatting)** на вкладке **Inquire**, вы удалите с листа ячейки, следующие за по-

следней непустой ячейкой. Например, если 10 000-я строка является последней на листе, то все ячейки ниже строки 10 000 будут удалены, и это может весьма благотворно отразиться на числе форматов, используемых в вашей книге.

## Задания

1. В примере с ЧПС нового автомобиля определите следующее:
  - для ячейки с процентной ставкой — непосредственно зависимые ячейки и все зависимые ячейки;
  - для ячейки со ставкой налога — непосредственно зависимые ячейки и все зависимые ячейки;
  - для ячейки с количеством проданных единиц на год 4 — непосредственно влияющие ячейки и все влияющие ячейки;
  - для ячейки с затратами на год 3 — непосредственно влияющие ячейки и все влияющие ячейки.
2. Для файла **Prodmix.xlsx** определите:
  - ячейки, влияющие на ячейку с величиной прибыли;
  - ячейки, зависимые от ячейки с количеством произведенного продукта 1.

## ГЛАВА 17

# Анализ чувствительности с помощью таблиц данных

### Обсуждаемые вопросы

- Я подумываю открыть магазин в местном торговом центре и продавать натуральный лимонад. Я хотел бы знать до открытия магазина, как прибыль, доход и переменные затраты будут зависеть от цены и себестоимости единицы продукции.
- Я собираюсь построить новый дом. Сумма кредита (с периодом погашения 15 лет) зависит от цены, по которой я продам свой нынешний дом. Я также не уверен в годовой процентной ставке, которую получу по завершении процесса продажи. Можно ли определить, как ежемесячные платежи будут зависеть от суммы кредита и годовой процентной ставки?
- Крупная интернет-компания подумывает о покупке еще одного магазина. Текущий годовой доход от этого магазина составляет \$100 млн, а расходы — \$150 млн. Текущие прогнозы показывают, что доходы магазина ежегодно растут на 25%, а расходы — на 5%. Известно, однако, что прогнозы могут быть ошибочными, и хотелось бы знать с учетом различных предположений о годовом росте доходов и расходов, через сколько лет магазин начнет приносить прибыль.
- Как создать диаграмму на основе таблицы данных?

Большинство моделей листа содержат предположения о конкретных параметрах или вводных данных модели. В примере с продажей лимонада к таким вводным относятся:

- цена одного стакана лимонада;
- себестоимость производства стакана лимонада;
- чувствительность спроса к цене лимонада;
- годовые постоянные затраты на эксплуатацию точки для продажи лимонада.

На основе предположений об исходных параметрах можно рассчитать интересные нас результаты. Для примера с лимонадом такими результатами могут быть:

- годовая прибыль;
- годовой доход;
- годовые переменные затраты.



Несмотря на все приложенные усилия, предположения о вводных данных могут оказаться ошибочными. Например, наиболее правдоподобное предположение о переменных затратах на производство стакана лимонада составляет \$0,45, но оно не обязательно верно. В электронной таблице изменение результатов в ответ на изменения вводных данных определяется с помощью *анализа чувствительности*. Например, вам нужно отследить, как изменение цены продукта влияет на годовую прибыль, доход и переменные затраты. В таблице данных Microsoft Excel 2019 можно легко поменять значения одного или двух исходных параметров и выполнить анализ чувствительности. С помощью *однонаправленной* таблицы данных можно узнать, как изменение значений одного исходного параметра отразится на любом количестве результатов. С помощью *двухнаправленной* таблицы можно определить, как изменение значений двух исходных параметров отразится на одном результате. На трех примерах в этой главе мы покажем, как просто использовать таблицу данных для получения значимых результатов.

## Ответы на вопросы

**?** Я подумываю открыть магазин в местном торговом центре и продавать натуральный лимонад. Я хотел бы знать до открытия магазина, как прибыль, доход и переменные затраты будут зависеть от цены и себестоимости единицы продукции.

Необходимые для такого анализа данные находятся в файле **Lemonade.xlsx** (рис. 17.1, 17.2, 17.4). Предположительные исходные данные указаны в диапазоне D1:D4. Допустим, что годовой спрос на лимонад равен  $65000 - 9000 \cdot \text{цена}$  (см. формулу в ячейке D2). (Оценка кривой спроса рассматривается в главе 87.) Имена в ячейках C1:C7 соответствуют ячейкам D1:D7.

Я рассчитал годовой доход в ячейке D5 по формуле  $\text{спрос} \cdot \text{цена}$ . В ячейке D6 я рассчитал по формуле  $\text{себестоимость} \cdot \text{спрос}$  переменные затраты за год. Наконец в ячейке D7 по формуле  $\text{доход} - \text{постоянные\_затраты} - \text{переменные\_затраты}$  я вычислил прибыль.

	C	D
1	price	\$ 4.00
2	demand	29000
3	unit cost	\$ 0.45
4	fixed cost	\$ 45,000.00
5	revenue	\$ 116,000.00
6	variable cost	\$ 13,050.00
7	profit	\$ 57,950.00

**Рис. 17.1.** Входные данные, влияющие на рентабельность магазина по продаже лимонада

Допустим, я хочу узнать, как изменение цены (например, с \$1 до \$4 с шагом \$0,25) повлияет на годовую прибыль, доход и переменные затраты. Поскольку изменяется только один исходный параметр, задача решается с помощью однонаправленной таблицы данных. Эта таблица данных представлена на рис. 17.2.

	C	D	E	F
9		profit	revenue	variable cost
10		57950	116000	13050
11	\$ 1.00	\$ (14,200.00)	\$ 56,000.00	\$ 25,200.00
12	\$ 1.25	\$ (2,000.00)	\$ 67,187.50	\$ 24,187.50
13	\$ 1.50	\$ 9,075.00	\$ 77,250.00	\$ 23,175.00
14	\$ 1.75	\$ 19,025.00	\$ 86,187.50	\$ 22,162.50
15	\$ 2.00	\$ 27,850.00	\$ 94,000.00	\$ 21,150.00
16	\$ 2.25	\$ 35,550.00	\$ 100,687.50	\$ 20,137.50
17	\$ 2.50	\$ 42,125.00	\$ 106,250.00	\$ 19,125.00
18	\$ 2.75	\$ 47,575.00	\$ 110,687.50	\$ 18,112.50
19	\$ 3.00	\$ 51,900.00	\$ 114,000.00	\$ 17,100.00
20	\$ 3.25	\$ 55,100.00	\$ 116,187.50	\$ 16,087.50
21	\$ 3.50	\$ 57,175.00	\$ 117,250.00	\$ 15,075.00
22	\$ 3.75	\$ 58,125.00	\$ 117,187.50	\$ 14,062.50
23	\$ 4.00	\$ 57,950.00	\$ 116,000.00	\$ 13,050.00

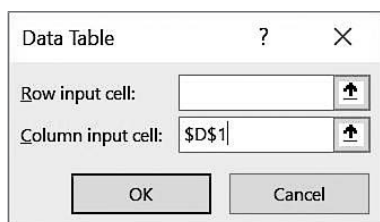
**Рис. 17.2.** Данные для создания однонаправленной таблицы с различными значениями цены

Создание однонаправленной таблицы данных начинается с перечисления исходных параметров в столбце. В диапазоне C11:C23 я указал интересующие меня цены (в диапазоне от \$1 до \$4 с шагом \$0,25). Строкой выше списка исходных параметров я разместил формулы, необходимые для расчета в таблице данных. Формулу для вычисления прибыли я ввел в ячейку D10, формулу вычисления величины дохода — в ячейку E10 и формулу для вычисления величины переменных затрат — в ячейку F10. Не вводите просто цифры — в Excel нужны формулы!

Теперь выделим диапазон для таблицы (C10:F23). Таблица начинается строкой выше первого исходного параметра, а ее последняя строка — это строка, которая содержит последний исходный параметр. Первый столбец таблицы — это столбец с исходными параметрами, а ее последний столбец — это последний из столбцов с результатами. После выделения таблицы откройте вкладку Данные (Data) на ленте Excel. В группе Работа с данными (Data Tools) в раскрывающемся списке Анализ "что, если" (What-If Analysis) выберите Таблица данных (Data Table). В диалоговом окне Таблица данных (Data Table) заполните поле Подставляя значения по столбцам в (Column input cell), как показано на рис. 17.3.

В этом поле следует указать ячейку значению которой будут присваиваться значения из списка вводных данных, то есть значения, перечисленные в первом столбце таблицы. Так как список вводных данных состоит из цен, в поле я выбрал ячейку D1. Когда вы нажмете ОК, Excel создаст однонаправленную таблицу (рис. 17.4).

В диапазоне ячеек D11:F11 вычисляются прибыль, доход и переменные затраты для цены \$1. В ячейках D12:F12 прибыль, доход и переменные затраты вычисля-



**Рис. 17.3.** Создание таблицы данных

	C	D	E	F
1	price	\$ 4.00		
2	demand	29000		
3	unit cost	\$ 0.45		
4	fixed cost	\$ 45,000.00		
5	revenue	\$ 116,000.00		
6	variable cost	\$ 13,050.00		
7	profit	\$ 57,950.00		
8				
9		profit	revenue	variable cost
10		57950	116000	13050
11	\$ 1.00	\$ (14,200.00)	\$ 56,000.00	\$ 25,200.00
12	\$ 1.25	\$ (2,000.00)	\$ 67,187.50	\$ 24,187.50
13	\$ 1.50	\$ 9,075.00	\$ 77,250.00	\$ 23,175.00
14	\$ 1.75	\$ 19,025.00	\$ 86,187.50	\$ 22,162.50
15	\$ 2.00	\$ 27,850.00	\$ 94,000.00	\$ 21,150.00
16	\$ 2.25	\$ 35,550.00	\$ 100,687.50	\$ 20,137.50
17	\$ 2.50	\$ 42,125.00	\$ 106,250.00	\$ 19,125.00
18	\$ 2.75	\$ 47,575.00	\$ 110,687.50	\$ 18,112.50
19	\$ 3.00	\$ 51,900.00	\$ 114,000.00	\$ 17,100.00
20	\$ 3.25	\$ 55,100.00	\$ 116,187.50	\$ 16,087.50
21	\$ 3.50	\$ 57,175.00	\$ 117,250.00	\$ 15,075.00
22	\$ 3.75	\$ 58,125.00	\$ 117,187.50	\$ 14,062.50
23	\$ 4.00	\$ 57,950.00	\$ 116,000.00	\$ 13,050.00

**Рис. 17.4.** Однонаправленная таблица данных с различными значениями цены

ются для цены \$1,25 и далее для всего диапазона цен. Максимально увеличивающая прибыль цена — это цена \$3,75. Эта цена принесет ежегодную прибыль в размере \$58 125,00, ежегодный доход в размере \$117 187,50 и ежегодные переменные затраты в размере \$14 062,50.

Допустим, я хочу узнать, как изменится годовая прибыль при изменении цены с \$1,50 до \$5,00 (с шагом \$0,25) и при изменении себестоимости единицы продукции с \$0,30 до \$0,60 (с шагом \$0,05). Поскольку в этом случае изменяются два исходных параметра, нужна двунаправленная таблица данных (рис. 17.5). Я указал значения для одного исходного параметра в первом столбце таблицы (используйте диапазон H11:H25 для цены) и значения для второго исходного параметра в первой строке таблицы. (В данном случае диапазон I10:O10 включает список значений себестоимости единицы продукции.) В двунаправленной таблице может быть только одна ячейка с результатом, и формула для расчета результата должна быть помещена в левый верхний угол таблицы. Поэтому в ячейку H10 я вписал формулу расчета прибыли.

Я выделил таблицу (ячейки H10:O25) и открыл вкладку Данные (Data). В группе Прогноз (Forecast) я нажал Анализ "что, если" (What-If Analysis) и затем Таблица данных (Data Table). Ячейка D1 (цена) — это столбец ввода значений (Column input cell), а ячейка D3 (себестоимость единицы продукции) — это строка ввода значений (Row input cell). Такая настройка таблицы означает, что значения в первом столбце таблицы используются как цены, а значения в первой строке таблицы — как себе-

стоимости единицы продукции. После нажатия ОК появится таблица, показанная на рис. 17.5. В качестве примера в ячейке K19 при цене \$3,50 и себестоимости единицы продукции \$0,40 указана годовая прибыль, равная \$58 850,00. Для каждой себестоимости единицы продукции цена, максимизирующая прибыль, выделена цветом. Обратите внимание, что при росте себестоимости единицы продукции растет цена, максимизирующая прибыль, поскольку часть выросшей себестоимости передается покупателю. Конечно, можно гарантировать только то, что цена, максимизирующая прибыль, в таблице данных находится в пределах \$0,25 от фактической максимизирующей прибыли цены. После знакомства с инструментом Excel Поиск решения (Solver) в главе 88 «Ценообразование продуктов с сопутствующими товарами» вы сможете определить точную (до цента) максимизирующую прибыль цену.

	H	I	J	K	L	M	N	O
9		unit cost						
10	57950	\$ 0.30	\$ 0.35	\$ 0.40	\$ 0.45	\$ 0.50	\$ 0.55	\$ 0.60
11	\$ 1.50	\$16,800.00	\$14,225.00	\$11,650.00	\$ 9,075.00	\$ 6,500.00	\$ 3,925.00	\$ 1,350.00
12	\$ 1.75	\$26,412.50	\$23,950.00	\$21,487.50	\$19,025.00	\$16,562.50	\$14,100.00	\$11,637.50
13	\$ 2.00	\$34,900.00	\$32,550.00	\$30,200.00	\$27,850.00	\$25,500.00	\$23,150.00	\$20,800.00
14	\$ 2.25	\$42,262.50	\$40,025.00	\$37,787.50	\$35,550.00	\$33,312.50	\$31,075.00	\$28,837.50
15	\$ 2.50	\$48,500.00	\$46,375.00	\$44,250.00	\$42,125.00	\$40,000.00	\$37,875.00	\$35,750.00
16	\$ 2.75	\$53,612.50	\$51,600.00	\$49,587.50	\$47,575.00	\$45,562.50	\$43,550.00	\$41,537.50
17	\$ 3.00	\$57,600.00	\$55,700.00	\$53,800.00	\$51,900.00	\$50,000.00	\$48,100.00	\$46,200.00
18	\$ 3.25	\$60,462.50	\$58,675.00	\$56,887.50	\$55,100.00	\$53,312.50	\$51,525.00	\$49,737.50
19	\$ 3.50	\$62,200.00	\$60,525.00	\$58,850.00	\$57,175.00	\$55,500.00	\$53,825.00	\$52,150.00
20	\$ 3.75	\$62,812.50	\$61,250.00	\$59,687.50	\$58,125.00	\$56,562.50	\$55,000.00	\$53,437.50
21	\$ 4.00	\$62,300.00	\$60,850.00	\$59,400.00	\$57,950.00	\$56,500.00	\$55,050.00	\$53,600.00
22	\$ 4.25	\$60,662.50	\$59,325.00	\$57,987.50	\$56,650.00	\$55,312.50	\$53,975.00	\$52,637.50
23	\$ 4.50	\$57,900.00	\$56,675.00	\$55,450.00	\$54,225.00	\$53,000.00	\$51,775.00	\$50,550.00
24	\$ 4.75	\$54,012.50	\$52,900.00	\$51,787.50	\$50,675.00	\$49,562.50	\$48,450.00	\$47,337.50
25	\$ 5.00	\$49,000.00	\$48,000.00	\$47,000.00	\$46,000.00	\$45,000.00	\$44,000.00	\$43,000.00

**Рис. 17.5.** Двухнаправленная таблица для расчета величины прибыли как функции цены и себестоимости единицы продукции

Вот еще несколько замечаний к данной задаче.

- При изменении исходных значений на листе также изменяются значения, вычисляемые в таблице данных. Например, если увеличить постоянные затраты на \$10 000, все значения для прибыли в таблице данных уменьшатся на \$10 000.
- Невозможно удалить или отредактировать часть таблицы данных. Если нужно сохранить значения в таблице, выделите табличный диапазон, скопируйте значения, а затем щелкните правой кнопкой мыши и выберите инструмент **Специальная вставка (Paste Special)**. В диалоговом окне **Специальная вставка (Paste Special)** установите переключатель в положение **значения (Values)**. Однако после этого изменения исходных данных больше не будут вызывать обновления результатов в таблице данных.

- При создании двунаправленной таблицы данных будьте внимательны и не перепутайте исходные значения, размещаемые в столбце и в строке. Ошибка приведет к бессмысленным результатам.
- Многие устанавливают для расчетов автоматический режим. В этом случае любые изменения на листе вызывают пересчет всех таблиц данных. Как правило, это и требуется, но если таблицы данных большие, автоматический пересчет может идти крайне медленно. Если постоянный пересчет таблиц данных замедляет работу, откройте вкладку **Файл (File)** на ленте, выберите команду **Параметры (Options)** и затем в диалоговом окне выберите **Формулы (Formulas)**. Установите переключатель в положение **Автоматически, кроме таблиц данных (Automatic Except For Data Tables)**. После этого все таблицы данных будут пересчитываться только при нажатии **F9** (пересчет). Или же на вкладке **Формулы (Formulas)** в группе **Вычисление (Calculation)** в раскрывающемся списке инструмента **Параметры вычислений (Calculation Options)** выберите **Автоматически, кроме таблиц данных (Automatic Except For Data Tables)**.

❓ **Я собираюсь построить новый дом. Сумма кредита (с периодом погашения 15 лет) зависит от цены, по которой я продам свой нынешний дом. Я также не уверен в годовой процентной ставке, которую получу по завершении процесса продажи. Можно ли определить, как ежемесячные платежи будут зависеть от суммы кредита и годовой процентной ставки?**

Реальная мощь таблиц данных становится очевидной при объединении таблиц данных с одной из функций Excel. В данном примере используется двунаправленная таблица с варьированием двух исходных параметров (суммы кредита и годовой процентной ставки) для функции **ПЛТ (PMT)**. С ее помощью рассчитывается изменение ежемесячного платежа при изменении этих исходных параметров. (Функция **ПЛТ** подробно рассматривалась в главе 10 «Еще несколько финансовых функций Excel».) Данные к этому примеру находятся в файле **Mortgagedt.xlsx** (рис. 17.6).

	C	D	E	F	G	H	I	J
2	Amt Borrowed	\$400,000.00						
3	Number of Months	180						
4	Annual int rate	6%						
5	Monthly Payment	\$3,375.43						
6			Annual	Interest	Rate			
7	\$3,375.43	5.0%	5.5%	6.0%	6.5%	7.0%	7.5%	8.0%
8	\$ 300,000.00	\$2,372.38	\$2,451.25	\$2,531.57	\$2,613.32	\$2,696.48	\$2,781.04	\$2,866.96
9	\$ 350,000.00	\$2,767.78	\$2,859.79	\$2,953.50	\$3,048.88	\$3,145.90	\$3,244.54	\$3,344.78
10	\$ 400,000.00	\$3,163.17	\$3,268.33	\$3,375.43	\$3,484.43	\$3,595.31	\$3,708.05	\$3,822.61
11	\$ 450,000.00	\$3,558.57	\$3,676.88	\$3,797.36	\$3,919.98	\$4,044.73	\$4,171.56	\$4,300.43
12	\$ 500,000.00	\$3,953.97	\$4,085.42	\$4,219.28	\$4,355.54	\$4,494.14	\$4,635.06	\$4,778.26
13	\$ 550,000.00	\$4,349.36	\$4,493.96	\$4,641.21	\$4,791.09	\$4,943.56	\$5,098.57	\$5,256.09
14	\$ 600,000.00	\$4,744.76	\$4,902.50	\$5,063.14	\$5,226.64	\$5,392.97	\$5,562.07	\$5,733.91
15	\$ 650,000.00	\$5,140.16	\$5,311.04	\$5,485.07	\$5,662.20	\$5,842.38	\$6,025.58	\$6,211.74

**Рис. 17.6.** Таблица данных для определения изменения ежемесячного платежа по ипотеке в зависимости от изменения суммы заимствования и ставки процента



Допустим, вы взяли ипотеку на 15 лет и платите в конце каждого месяца. Введите сумму кредита в ячейку D2, количество месяцев ипотеки (180) в ячейку D3 и годовую процентную ставку в D4. Присвойте имена из ячеек C2:C4 ячейкам D2:D4. Исходя из этих данных, вычислите размер ежемесячного платежа в ячейке D5 по формуле

=-ПЛТ(годовая\_ставка\_процента/12;количество\_месяцев;сумма\_заимствования).

Допустим, сумма кредита варьирует (в зависимости от цены продажи нынешнего дома) от \$300 000 до \$650 000, а годовая процентная ставка составляет от 5 до 8%. Введите суммы кредита в ячейки C8:C15 и значения возможной процентной ставки в ячейки D7:J7. В C7 содержится результат, который требуется пересчитать для различных комбинаций исходных данных. Поэтому установите значение в C7 равным значению в D5. Выделите таблицу (C7:J15), откройте вкладку Данные (Data), в группе Работа с данными (Data Tools), в Анализ "что, если" (What-If Analysis) выберите Таблица данных (Data Table). Поскольку числа в первом столбце таблицы являются суммами кредита, в поле Подставлять значения по столбцам в (Column input cell) укажите D2. Числа в первой строке таблицы представляют годовые процентные ставки, поэтому в поле Подставлять значения по строкам в (Row input cell) укажите D4. После нажатия ОК появится таблица данных, представленная на рис. 17.6. Данные в этой таблице показывают, например, что при кредите \$400 000 под годовую ставку 6% ежемесячные платежи составят чуть более \$3375. Данные в таблице также показывают, что при более низкой процентной ставке (например, 5%) увеличение суммы кредита на \$50 000 повышает ежемесячный платеж примерно на \$395, тогда как при более высокой процентной ставке (например, 8%) увеличение суммы кредита на \$50 000 повышает ежемесячный платеж примерно на \$478.

❓ **Крупная интернет-компания подумывает о покупке еще одного магазина. Текущий годовой доход от этого магазина составляет \$100 млн, а расходы — \$150 млн. Текущие прогнозы показывают, что доходы магазина ежегодно растут на 25%, а расходы — на 5%. Известно, однако, что прогнозы могут быть ошибочными, и хотелось бы знать с учетом различных предположений о годовом росте доходов и расходов, через сколько лет магазин начнет приносить прибыль.**

Следует определить период времени, необходимый для достижения безубыточности, на основе ежегодных темпов роста доходов от 10 до 50% и ежегодных темпов роста расходов от 2 до 20%. Предположим также, что если компания не достигнет безубыточности за 13 лет, то не сможет его достичь в принципе. Данные к этой задаче находятся в файле Bezos.xlsx и показаны на рис. 17.7.

Я скрыл столбцы A и B и строки 16–18. Чтобы скрыть столбцы A и B, сначала выделите любые ячейки в столбцах A и B (или заголовки столбцов) и затем перейдите на вкладку Главная (Home). В группе Ячейки (Cells) в раскрывающемся списке Формат (Format) выберите Скрыть или отобразить (Hide & Unhide) и укажите Скрыть

столбцы (Hide Columns). Чтобы скрыть строки 16–18, выделите любые ячейки в этих строках (или заголовки строк) и, повторив предыдущие действия, укажите Скрыть строки (Hide Rows). Конечно, параметры | Видимости (Visibility) также позволяют вам Отобразить строки (Unhide Rows) и Отобразить столбцы (Unhide Columns). Если на листе много скрытых столбцов и строк и требуется отобразить их быстро, выделите весь лист, нажав Выделить все (Select All) на пересечении заголовков столбцов и строк. Затем выберите Отобразить строки (Unhide Rows) или Отобразить столбцы (Unhide Columns), и на листе появятся все скрытые прежде строки или столбцы.

В строке 11 планируем доходы компании на 13 лет вперед (на основе предполагаемых ежегодных темпов роста доходов в ячейке E7) путем копирования формулы  $=E11*(1+\$E\$7)$  из F11 в G11:R11. В строке 12 планируем расходы компании на 13 лет вперед (на основе предполагаемых ежегодных темпов роста расходов в ячейке E8) путем копирования формулы  $=E12*(1+\$E\$8)$  из F12 в G12:R12 (рис. 17.7).

	C	D	E	F	G	H	P	Q	R	S	T	U	V	W	X
6			growth												
7	Revenue	1E+08	0.25												
8	Expenses	1.50E+08	0.05												
9															
10			0	1	2	3	11	12	13						
11	Revenues	1E+08	1.25E+08	1.56E+08	1.95E+08	1.16E+09	#####	1.82E+09							
12	Expenses	2E+08	1.58E+08	1.65E+08	1.74E+08	2.57E+08	#####	2.83E+08							
13	Breakeven		0	0	3	0	0	0							
14															
15	Total		3												
19				Expense	Growth										
20			3	0.02	0.03	0.04	0.12	0.13	0.14	0.15	0.16	0.17	0.18	0.19	0.2
21			0.1	6	7	8	No BE	No BE	No BE	No BE	No BE	No BE	No BE	No BE	No BE
22	Revenue		0.11	5	6	7	No BE	No BE	No BE	No BE	No BE	No BE	No BE	No BE	No BE
23	growth		0.12	5	5	6	No BE	No BE	No BE	No BE	No BE	No BE	No BE	No BE	No BE
24			0.13	4	5	5	No BE	No BE	No BE	No BE	No BE	No BE	No BE	No BE	No BE
25			0.14	4	4	5	No BE	No BE	No BE	No BE	No BE	No BE	No BE	No BE	No BE
55			0.44	2	2	2	2	2	2	2	2	2	3	3	3
56			0.45	2	2	2	2	2	2	2	2	2	2	3	3
57			0.46	2	2	2	2	2	2	2	2	2	2	2	3
58			0.47	2	2	2	2	2	2	2	2	2	2	2	2
59			0.48	2	2	2	2	2	2	2	2	2	2	2	2
60			0.49	2	2	2	2	2	2	2	2	2	2	2	2
61			0.5	2	2	2	2	2	2	2	2	2	2	2	2

**Рис. 17.7.** Таблица данных для расчета периода достижения безубыточности (диапазон E19:R61)

Мы будем использовать двустороннюю таблицу данных для определения того, как разные темпы роста доходов и расходов влияют на количество лет, необходимых для достижения безубыточности. Нам нужна одна ячейка, в которой будет указываться количество лет. Поскольку уровень безубыточности может быть достигнут в течение любого из следующих 13 лет, это может показаться довольно трудной задачей.

Начнем с применения функции ЕСЛИ для каждого года в строке 13 с целью определить, достигнут ли уровень безубыточности в течение этого года. Функция ЕСЛИ возвращает порядковый номер года в случае достижения уровня безубы-

точности или 0 в противном случае. В ячейке E15 определяем год, в который достигнут уровень безубыточности, складывая все числа в строке 13. Наконец, ячейку E15 можно использовать как ячейку с конечным результатом для двунаправленной таблицы.

Скопируйте формулу `=ЕСЛИ(И(E11<E12;F11>F12);F10;0)` из F13 в G13:R13. Эта формула показывает, что безубыточность в первый раз будет достигнута тогда и только тогда, когда в предыдущем году доходы были меньше расходов и в течение текущего года доходы превышают расходы. В этом случае в строке 13 записывается порядковый номер года; в противном случае записывается 0.

Теперь в ячейке E15 можно определить год достижения безубыточности по формуле `=ЕСЛИ(СУММ(F13:R13)>0;СУММ(F13:R13);"Нет")`. Если безубыточность не будет достигнута даже через 13 лет, в результате вычисления будет получена текстовая строка "Нет".

Затем введите ежегодные темпы роста доходов (от 10 до 50%) в ячейки E21:E61. Введите ежегодные темпы роста расходов (от 2 до 20%) в ячейки F20:X20. Убедитесь, что формула безубыточности скопирована в ячейку E20 как `=E15`. Выделите таблицу E20:X61, на вкладке Данные (Data) в группе Прогноз (Forecast) щелкните в раскрывающемся списке Анализ "что, если" (What-If Analysis) и выберите Таблица данных (Data Table). Укажите ячейку E7 (темпы роста доходов) в поле Подставлять значения по столбцам в (Column input cell) и ячейку E8 (темпы роста расходов) в поле Подставлять значения по строкам в (Row input cell). При таких настройках получится двунаправленная таблица (рис. 17.7).

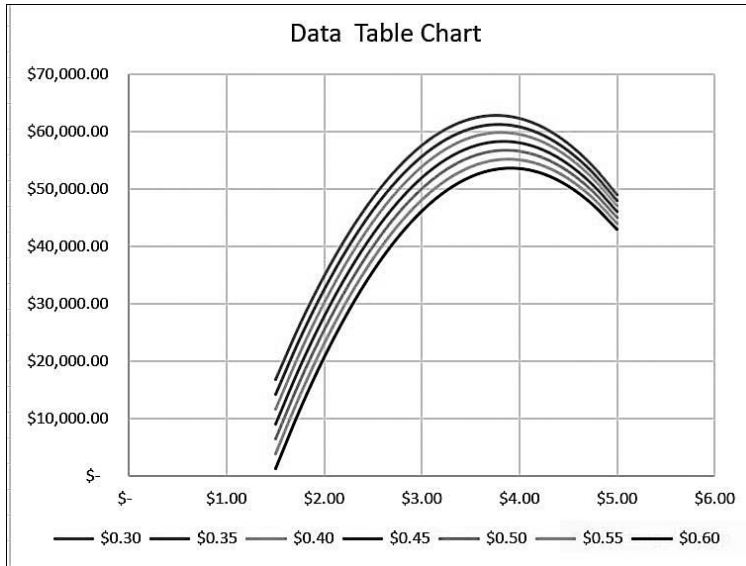
Отметим, например, что если расходы растут на 4% в год, то при 10%-ном ежегодном росте доходов безубыточность будет достигнута через 8 лет, в то время как при 50%-ном ежегодном росте доходов она будет достигнута уже через 2 года! Также отметим, что если расходы растут на 12% в год, а доходы — на 14% в год, то безубыточность не будет достигнута к концу 13-го года.

### Как создать диаграмму на основе таблицы данных?

Таблица данных — это всего лишь ворох цифр. Диаграмма же, основанная на таблице данных, зачастую дает аналитику куда большее понимание ситуации. Чтобы проиллюстрировать, как двунаправленная таблица может послужить основой для построения диаграммы, снова взглянем на двунаправленную таблицу данных в файле *Lemonade.xlsx* (рис. 17.8).

Сначала мы скопировали данные из диапазона H10:O24 в H28:O43, а затем удалили запись в ячейке H28. Выбрав на вкладке Вставка (Insert) диапазон H28:O43, выбираем третий тип диаграммы (тот, который только с линиями), Точечную диаграмму (X-Y (Scatter) Chart). Получаем диаграмму, показанную на рис. 17.8. Как видим, самая верхняя кривая, как и следовало ожидать, отражает минимальную себестоимость за единицу. Также видно, что рост цены влечет за собой рост прибыли, пока цена за единицу растет в интервале от \$3,75 до \$4. При этом с ростом себестоимости за единицу кривые дохода сближаются друг с другом.





**Рис. 17.8.** Диаграмма, отражающая зависимость прибыли от цены и себестоимости единицы продукции

## Задания

1. Вам поручили проанализировать, насколько прибыльным делом будет издать автобиографию Билла Клинтона. Были сделаны следующие предположения:
  - Билл получает единовременно роялти в размере \$12 млн;
  - фиксированные расходы на производство книги в твердой обложке составят \$1 млн;
  - переменные расходы на производство одного экземпляра книги в твердой обложке составляют \$4;
  - чистый доход издательства от продажи книги составляет \$15 за каждый проданный экземпляр в твердой обложке;
  - издатель рассчитывает продать 1 млн экземпляров книги в твердой обложке;
  - фиксированные расходы на производство книги в мягкой обложке составляют \$100 000;
  - переменные расходы на производство одного экземпляра книги в мягкой обложке составляют \$1;
  - чистый доход издательства от продажи книги составляет \$4 за каждый проданный экземпляр книги в мягкой обложке;

- объем продаж книги в мягкой обложке в 2 раза превысит объем продаж книги в твердой обложке.

Используйте эту информацию при ответе на следующие вопросы:

- Как будет меняться прибыль издательства до уплаты налогов при продажах книги в твердой обложке в интервале от 100 тыс. до 1 млн экземпляров?
  - Как будет меняться прибыль издателя до уплаты налогов при продажах книги в твердой обложке в интервале от 100 тыс. до 1 млн экземпляров и при изменении отношения объемов продаж книги в мягкой обложке к объему продаж книги в твердой обложке от 1 до 2,4?
2. Ежегодный спрос на продукт составляет  $500 - 3p + 10a \cdot 5$ , где  $p$  — цена продукта в долларах,  $a$  — сотни долларов, потраченные на рекламу продукта. Ежегодные фиксированные расходы на продажу продукта составляют \$10 000, а переменные расходы на производство единицы продукции — \$12. Определите цену (с шагом \$10) и сумму, затраченную на рекламу (с шагом \$100), которые максимизируют прибыль.
  3. Вернемся к примеру с хеджированием из главы 12 «Функции ЕСЛИ, ЕСЛИОШИБКА, ЕСЛИМН, ВЫБОР и ПЕРЕКЛЮЧ». Определите процентную доходность портфеля для цены акции через шесть месяцев, варьирующей от \$20 до \$65, и количества проданных пут-опционов, изменяющегося от 0 до 100 (с шагом 10).
  4. Допустим, в примере с ипотечным кредитом известна годовая процентная ставка — 5,5%. Создайте таблицу, в которой показана разница в платежах для 15-, 20- и 30-летней ипотеки с суммой кредита от \$300 000 до \$600 000 (с шагом \$50 000).
  5. На данный момент вы продаете 40 000 единиц продукции по \$45 за каждую единицу. Переменные расходы на производство единицы продукции составляют \$5. Вы собираетесь снизить цену продукта на 30%, поскольку уверены, что это приведет к увеличению объема продаж с 10 до 50%. Выполните анализ чувствительности и покажите, как меняется размер прибыли в зависимости от роста продаж в процентах. Не учитывайте фиксированные расходы.
  6. Допустим, в конце каждого года из следующих 40 лет вы будете откладывать одну и ту же сумму в свой пенсионный фонд по одной и той же ежегодной процентной ставке. Покажите, каким образом изменится сумма при выходе на пенсию при изменении ежегодного взноса с \$5000 до \$25 000 и при изменении процентной ставки с 3 до 15%.
  7. Срок окупаемости проекта — это период времени, необходимый для покрытия первоначальной инвестиции будущей прибылью от проекта. Для проекта требуется инвестиция в размере 300 млн долл. в нулевой год. Проект будет приносить прибыль в течение 10 лет, и по итогам первого года денежный поток составит от 30 до 100 млн долл. Ежегодные темпы роста денежного потока составят 5–25%. Как окупаемость проекта зависит от денежного потока по итогам первого года и от темпов роста денежного потока?

8. Компания, занимающаяся разработкой программного обеспечения, собирается локализовать программный продукт на суахили. В настоящее время в год продается 200 000 единиц продукта по цене \$100 за единицу. Переменные расходы на производство продукта составляют \$20. Фиксированные расходы на перевод составят \$5 млн. Локализация продукта на суахили приведет к росту продаж в каждом году из последующих трех лет на неизвестный процент (текущий объем продаж — 200 000 единиц). Покажите, как изменение прибыли после локализации зависит от процентного увеличения объема продаж продукта. В расчетах можно не учитывать временную стоимость денег и налоги.
9. В файле *Citydistances.xlsx* приведены данные о широте и долготе для нескольких городов США. Также в нем есть формула, вычисляющая расстояние между двумя городами по заданной широте и долготе. Создайте таблицу, в которой вычисляются расстояния между любой парой указанных городов.
10. Вы начали копить деньги на высшее образование своего ребенка. Вы намерены вносить на счет по \$5000 ежегодно, и вам необходимо узнать, какая сумма окажется на вашем счету через 10–15 лет при годовой ставке доходности инвестиции от 4 до 12%.
11. Если вам начисляются проценты по процентной ставке  $r$  каждый год и сложные проценты  $n$  раз в год, тогда через  $y$  лет \$1 вырастет до  $(1 + (r/n))^{ny}$  долларов. При условии, что годовая процентная ставка составляет 10%, создайте таблицу, показывающую коэффициент роста \$1 через 5–15 лет при ежедневном, ежемесячном, ежеквартальном и полугодовом начислении сложных процентов.
12. На вашем счету в банке лежат \$100. Ежегодно вы можете снять  $x\%$  (от 4 до 10) от первоначальной суммы. Определите для годового темпа роста от 3 до 10%, сколько пройдет лет до того, как закончатся деньги. Подсказка: следует воспользоваться функцией ЕСЛИОШИБКА (IFERROR), рассмотренной в главе 12, поскольку если годовой темп роста превышает темпы изъятия, деньги на счету никогда не закончатся.
13. Если вам начисляются проценты по годовой ставке  $x\%$  в год, через  $n$  лет \$1 будет равен  $(1 + x)^n$  долларов. Для годовой процентной ставки от 1 до 20% определите точно, когда (в годах) произойдет удвоение \$1.
14. Вы заняли \$200 000 и вносите платеж в конце каждого месяца. Для годовой процентной ставки от 5 до 10% и срока возврата кредита 10, 15, 20, 25 и 30 лет определите итоговые проценты, выплаченные по кредиту.
15. Вы начали копить деньги на высшее образование своего ребенка. Вы кладете на счет одну и ту же сумму в конце каждого года. Ваша цель — накопить \$100 000. Определите размер ежегодного взноса при ежегодной прибыли по инвестициям от 4 до 15% и сроке инвестирования от 5 до 15 лет.
16. В файле *Antitrustdata.xlsx* содержатся данные о годах начала и завершения судебных дел. Определите количество активных судебных дел в течение каждого года.

17. Вы можете выйти на пенсию в 62 года и получать \$8000 в год или выйти на пенсию в 65 лет и получать \$10 000 в год. Какова разница (в долларах на сегодняшний день) между этими двумя вариантами при изменении годовой процентной ставки, по которой дисконтируются денежные потоки, от 2 до 10% и при изменении возраста смерти от 70 до 84 лет?
18. Вы собираетесь открыть ресторан, в котором будет шесть столиков на четыре персоны каждый. Каждый день работы ресторана в период ланча обеспечивается две рассадки посетителей и в период ужина — три рассадки.
- Ресторан закрыт по понедельникам.
  - Средний чек за ланч составляет \$20, и коэффициент прибыльности по чекам за ланч составляет 40 %.
  - Средний чек за ужин составляет \$40, и коэффициент прибыльности по чекам за ужин составляет 50%.
  - Исходите из предположения, что постоянные затраты на содержание ресторана составляют \$400 000 в год.

Взяв в расчет 364 дня в году, создайте таблицу данных, отражающую изменение годового дохода в зависимости от изменения количества занятых мест в обоих периодах подачи блюд между 10 и 100 процентами.

19. Вы инвестируете в новую бродвейскую пьесу «Бабочка на свету». У вас есть следующая информация о постановке:
- Постоянные затраты на первое представление пьесы составляют 5 млн долл.
  - Средняя цена билета — \$100.
  - Театр вмещает 2000 человек, и представления идут 365 дней в году.

Создайте таблицу данных, в которой определяется, как общий доход от постановки изменяется при продолжительности проката пьесы от 1 года до 5 лет и при средней наполненности зала, меняющейся от 70 до 90%.

20. В настоящее время журнал имеет 5000 подписчиков. Определите, как может измениться число подписчиков на конец года 6, если число новых подписчиков меняется от 1000 до 5000 и процент сохраняющих подписку ежегодно меняется между 60 и 90%.
21. Когда покупатель заказывает товар из каталога, наша прибыль составляет \$50 (стоимость рассылки каталога не включена). Создайте таблицу данных для определения, как изменяется общий доход от рассылки, если процент ответа на рассылку колеблется между 1 и 10%.
22. Сегодня вы купили пять акций Пакета 1 по цене \$30 за акцию и три акции Пакета 2 по цене \$25 за акцию. Через месяц вы собираетесь продать все акции. Создайте таблицу данных, отслеживающую, как меняется общий доход при изменении цены каждой акции от -\$5 до \$5.

## ГЛАВА 18

# Инструмент Подбор параметра

### Обсуждаемые вопросы

- Сколько стаканов лимонада в год по определенной цене должен продать магазин для достижения безубыточности?
- Мы собираемся выплатить ипотеку за 15 лет. Годовая процентная ставка составляет 6%. Банк сообщил, что максимальный ежемесячный платеж для нас составляет \$2000. Какую сумму мы можем взять в кредит?
- Мне всегда плохо давались сюжетные задачи по алгебре в средней школе. Может ли Excel упростить решение таких задач?

В Microsoft Excel 2019 вы можете использовать **Подбор параметра (Goal Seek)** для вычисления значения исходного параметра листа, при котором значение заданной формулы соответствует указанной цели. Например, предположим, что в примере с магазином по продаже лимонада из главы 17 «Анализ чувствительности с помощью таблиц данных» имеются постоянные накладные расходы, фиксированные расходы на производство единицы продукции и фиксированная цена продажи. С учетом этих данных можно с помощью инструмента **Подбор параметра (Goal Seek)** вычислить количество стаканов лимонада, которое требуется продать для достижения безубыточности. По сути инструмент **Подбор параметра (Goal Seek)** представляет собой встроенный решатель уравнений. Для использования инструмента **Подбор параметра (Goal Seek)** в Excel вам нужно указать следующую информацию.

- **Установить в ячейке (Set Cell)** — указывает, что ячейка содержит формулу, вычисляющую искомое значение. В примере с лимонадом это поле должно содержать ссылку на ячейку с формулой расчета прибыли.
- **Значение (To Value)** — указывает целевое числовое значение, которое должно быть вычислено в ячейке с формулой. В примере с лимонадом, поскольку требуется определить объем продаж, необходимый для достижения безубыточности, в это поле следует ввести значение 0.
- **Изменяя значение ячейки (By Changing Cell)** — указывает ячейку с исходным параметром, значение которого автоматически изменяется до тех пор, пока в ячейке с формулой не будет получено целевое числовое значение, определенное в поле **Значение (To Value)**. В примере с лимонадом это поле должно содержать ссылку на ячейку с годовым объемом продаж лимонада.

## Ответы на вопросы

### ? Сколько стаканов лимонада в год по определенной цене должен продать магазин для достижения безубыточности?

Данные к этому разделу находятся в файле `Lemonadegs.xlsx` (рис. 18.1). Как и в главе 17, предположим, что годовые фиксированные расходы составляют \$45 000,00, а переменные расходы на производство единицы продукции — \$0,45. Допустим также, что цена стакана лимонада равна \$3,00. Вопрос в том, сколько стаканов лимонада необходимо продавать каждый год для достижения безубыточности.

Для начала введите любое число для величины спроса в ячейку D2. На вкладке Данные (Data) в группе Прогноз (Forecast) в раскрывающемся списке Анализ "что, если" (What-If Analysis) выберите Подбор параметра (Goal Seek). В открывшемся диалоговом окне заполните поля, как показано на рис. 18.2.

	C	D
1	price	\$ 3.00
2	demand	17647.05882
3	unit cost	\$ 0.45
4	fixed cost	\$ 45,000.00
5	revenue	\$ 52,941.18
6	variable cost	\$ 7,941.18
7	profit	\$ -

**Рис. 18.1.** Входные данные инструмента Подбор параметра

**Рис. 18.2.** Диалоговое окно Подбор параметра с данными для анализа безубыточности

Данные, указанные в этом диалоговом окне, означают, что требуется изменять значение в ячейке D2 (годовой спрос или продажи) до тех пор, пока значение в ячейке D7 (прибыль) не будет равно 0. После нажатия ОК появится результат, показанный на рис. 18.1. Уровень безубыточности будет достигнут при продаже около 17 647 стаканов лимонада в год (или 48 стаканов в день). Для поиска целевого значения величина спроса в ячейке D2 изменялась (чередованием между верхней и нижней границей) до тех пор, пока не было найдено значение, при котором прибыль равна \$0. Даже если задача имеет несколько решений, инструмент Подбор параметра (Goal Seek) отобразит только одно из них.

### ? Мы собираемся выплатить ипотеку за 15 лет. Годовая процентная ставка составляет 6%. Банк сообщил, что максимальный ежемесячный платеж для нас составляет \$2000. Какую сумму мы можем взять в кредит?

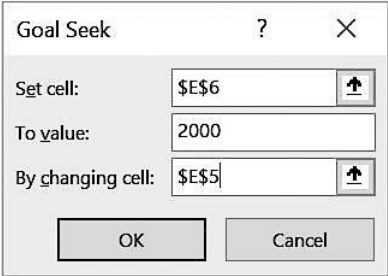
Для ответа на этот вопрос начнем создавать таблицу для вычисления размера ежемесячных платежей по 15-летней ипотеке (при условии выполнения плате-

жей в конце месяца) как функции годовой процентной ставки и пробной суммы кредита. См. проделанную работу в файле **Paymentgs.xlsx** и на рис. 18.3.

В ячейку E6 введена формула расчета ежемесячного платежа **=-ПЛТ(годовая\_процентная\_ставка/12;месяцы;сумма\_кредита)**, связанного с суммой кредита, которая вносится в ячейку E5. Заполнение полей в диалоговом окне **Подбор параметра (Goal Seek)**, как показано на рис. 18.4, приведет к вычислению суммы кредита при заданном размере ежемесячных платежей \$2000. С лимитом \$2000 для ежемесячных платежей можно взять в кредит до \$237 007,03.

	D	E
3	Years	180
4	Annual int rate	0.06
5	Amount borrowed	\$ 237,007.03
6	Monthly payment	\$2,000.00

**Рис. 18.3.** Данные для определения суммы заимствования с помощью инструмента Подбор параметра на основе заданного ежемесячного платежа



**Рис. 18.4.** Диалоговое окно Подбор параметра с данными примера с ипотекой

### ? Мне всегда плохо давались сюжетные задачи по алгебре в средней школе. Может ли Excel упростить решение таких задач?

На уроках алгебры в средней школе в большинстве сюжетных задач требовалось найти переменную (обычно ее называли  $x$ ), решив уравнение. Инструмент **Подбор параметра (Goal Seek)** представляет собой своеобразный решатель уравнений, так что он идеально подходит для решения сюжетных задач. Вот типичная школьная задача по алгебре:

*Мария и Эдмунд поссорились во время медового месяца в Сизтле. Мария вскочила в свою Mazda Miata и со скоростью 64 мили в час направилась к дому своей матери в Лос-Анджелесе. Два часа спустя Эдмунд запрыгнул в свой BMW и отправился за ней со скоростью 80 миль в час. Сколько миль проедет каждый из них до того, как Эдмунд догонит Марию?*

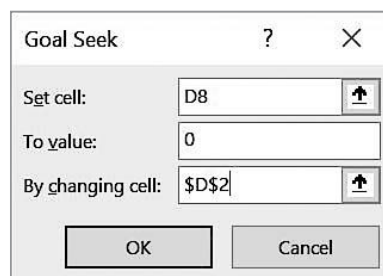
Решение можно найти в файле **Maria.xlsx**, показанном на рис. 18.5.

В поле **Установить в ячейке (Set Cell)** следует указать ячейку с разностью между расстоянием, которое успела проехать Мария, и расстоянием, которое проехал Эдмунд. В этой ячейке должно находиться значение 0, которое связано со временем, проведенным в пути Марией. Разумеется, Эдмунд провел в пути на два часа меньше, чем Мария.

В ячейку D2 я ввел пробное количество часов, проведенное Марией в пути. Затем я присвоил имена из ячеек C2:C8 ячейкам D2:D8. Поскольку Эдмунд ехал на два часа меньше, я ввел в ячейку D4 формулу `=время_в_пути_Марии - 2`. В ячейках D6 и D7, вспомнив, что `расстояние = скорость*время`, вычислим путь, который проехали Мария и Эдмунд. Разность между этими величинами вычисляется в ячейке D8 по формуле `=длина_пути_Марии - длина_пути_Эдмунда`. Теперь можно заполнить поля в диалоговом окне Подбор параметра (Goal Seek), как показано на рис. 18.6.

	C	D
2	Time Maria drives	10
3	Maria speed	64
4	Time Edmund drives	8
5	Edmund speed	80
6	Maria distance	640
7	Edmund distance	640
8	Difference	0

**Рис. 18.5.** Решение текстовых задач с помощью инструмента Подбор параметра



**Рис. 18.6.** Диалоговое окно Подбор параметра с данными для решения задачи по алгебре

Я менял значение в ячейке D2 (время в пути Марии) до тех пор, пока разность (в ячейке D8) не стала равна 0. Как видите, Мария проведет за рулем 10 часов до того, как ее догонит Эдмунд, бывший в пути 8 часов, и они оба преодолеют расстояние в 640 миль.

## Задания

1. По данным из задания 1 в главе 17 определите, сколько книг в твердой обложке должно быть продано для достижения безубыточности.
2. Для примера с чистой приведенной стоимостью (ЧПС) автомобиля в главе 16 «Инструменты проверки зависимостей и надстройка Inquire» определите, каким должен быть ежегодный рост продаж для получения ЧПС в размере 1 млн долл.?
3. Для этого же примера из главы 16 определите, какая себестоимость единицы продукции в год 1 увеличит ЧПС до 1 млн долл.?
4. Пусть в примере с ипотекой необходимо взять кредит в размере \$200 000 на 15 лет. Если максимальный ежемесячный платеж составляет \$2000, какую годовую процентную ставку я потяну?
5. Как с помощью инструмента Подбор параметра определить внутреннюю ставку доходности (ВСД) проекта?



6. В конце каждого года из следующих 40 лет вы собираетесь откладывать \$20 000 в пенсионный фонд. Какова должна быть ставка доходности вашей инвестиции, которая позволит выйти через 40 лет на пенсию с 2 млн долл. на счету?
7. Вы собираетесь зарабатывать по 10% в год на инвестициях в пенсионный фонд. В конце каждого года в течение 40 лет вы собираетесь класть одну и ту же сумму в свой инвестиционный портфель. Какую сумму необходимо класть на счет каждый год для накопления 2 млн долл. к моменту выхода на пенсию?
8. Рассмотрим два проекта с денежными потоками, указанными в таблице ниже.

	Год 1	Год 2	Год 3	Год 4
Проект 1	–\$1000	\$400	\$350	\$400
Проект 2	–\$900	\$100	\$100	\$1000

При какой процентной ставке у первого проекта ЧПС будет больше? (Подсказка: найдите процентную ставку, при которой оба проекта имеют одинаковую ЧПС.)

9. Я организую конференцию в своем колледже. Фиксированные расходы составляют \$15 000. Каждому спикеру, которых будет 10, я должен заплатить по \$700. Колледжу я должен заплатить \$300 за питание и проживание каждого участника конференции. С каждого участника, но не докладчика, конференции я беру взнос в размере \$900, которые пойдут на организационный взнос и расходы на питание и проживание. Сколько потребуется платных участников конференции для достижения безубыточности?
10. Я собираюсь купить 40 фунтов конфет. Некоторые конфеты продаются по цене \$10 за фунт, а другие — по \$6 за фунт. Сколько фунтов конфет по каждой из цен необходимо купить для получения средней цены \$7 за фунт?
11. Три электрика делают электропроводку в новом доме. Первый электрик выполнил бы эту работу самостоятельно за 11 дней. Второму электрику для этого необходимо 5 дней. Третий электрик выполнит работу самостоятельно за 9 дней. Если все три электрика будут работать вместе, сколько времени потребуется для завершения работ?
12. В ознаменование экспедиции Льюиса и Кларка я пройду на каноэ 40 миль вверх по течению и затем 40 миль вниз по течению. Скорость течения реки составляет 5 миль в час. Если весь поход занял 5 часов, с какой скоростью я плыл, если не учитывать скорость течения реки?
13. В главе 8, «Оценка инвестиций по чистой приведенной стоимости» в файле **NPV.xlsx** было показано, что при высоких процентных ставках ЧПС больше у первого проекта, а при низких процентных ставках — у второго. При какой процентной ставке оба проекта будут иметь одинаковое значение ЧПС?

14. Пусть вы взяли в кредит \$500 000 на 20 лет и вносите ежегодные платежи в начале каждого года. Какая годовая процентная ставка будет соответствовать этим условиям?
15. Сегодня 15 апреля, и к этому дню в этом месяце мы продали товара на \$35 000. План продаж на апрель составляет \$72 500. Если мы установим цену  $p$  за наш продукт, то продадим  $800 - 4p$  единиц товара за оставшуюся часть месяца. Какая цена позволит нам выполнить план по доходам за апрель?
16. Текущая годовая процентная ставка составляет 5%, и вы хотите взять кредит на 20 лет с ежемесячной выплатой в конце месяца. Если вы готовы ежемесячно выплачивать \$3000, какова может быть максимальная сумма кредита?
17. Грег Уинстон обанкротился и занялся продажей лимонада. Он вложил \$400 в оборудование по производству лимонада и продает стакан лимонада за \$4. Себестоимость производства стакана лимонада — \$2,50. Сколько стаканов лимонада должен продать Грег, чтобы получить доход в \$300?
18. Мой доход в 1-й год составил 80 000 \$. Проработав 35 лет, я уйду на пенсию. Предположим, что мой доход станет возрастать на 5% в год на протяжении всего этого времени, и я буду получать его в начале года. Предположим, что после получения дохода  $x\%$  его уходит на потребление. Затем моя оставшаяся наличность приносит 10% дохода. Какое значение  $x$  принесет в последний год наличность, равную 1,5 миллиона долларов?

## ГЛАВА 19

# Анализ чувствительности с помощью Диспетчера сценариев

### Обсуждаемый вопрос

- Я хочу создать для компании наиболее благоприятный, наименее благоприятный и наиболее вероятный сценарии продаж модели автомобиля, изменяя значения объема продаж за год 1, годового роста продаж и розничной цены в год 1. В таблице данных для анализа чувствительности можно изменять только один или два исходных параметра, поэтому для решения этой задачи таблица данных не подходит. Существует ли в Excel другой инструмент, который при анализе чувствительности позволяет изменять более двух параметров?

Диспетчер сценариев (Scenario Manager) поможет нам выполнить анализ чувствительности с возможностью изменения до 32 исходных значений. Сначала с помощью инструмента Диспетчер сценариев (Scenario Manager) определите набор ячеек с исходными данными, которые будут меняться. Присвойте имена сценариям и введите значения для каждой исходной ячейки в каждом сценарии. Наконец, выберите ячейки для выходных данных (также называемые *ячейками результата*), значения в которых хотите отслеживать. Тогда Диспетчер сценариев (Scenario Manager) создаст отчет, содержащий исходные данные и значения в ячейках выходных данных для каждого сценария.

### Ответ на вопрос

**?** Я хочу создать для компании наиболее благоприятный, наименее благоприятный и наиболее вероятный сценарии продаж модели автомобиля, изменяя значения объема продаж за год 1, годового роста продаж и розничной цены в год 1. В таблице данных для анализа чувствительности можно изменять только один или два исходных параметра, поэтому для решения этой задачи таблица данных не подходит. Существует ли в Excel другой инструмент, который при анализе чувствительности позволяет изменять более двух параметров?

Пусть необходимо создать следующие три сценария, связанные с чистой приведенной стоимостью (ЧПС) автомобиля, для данных в примере из главы 16.

	Объем продаж за год 1	Рост годового объема продаж	Розничная цена в год 1
Наиболее благоприятный вариант	\$20 000	20%	\$10,00
Наиболее вероятный вариант	\$10 000	10%	\$7,50
Наименее благоприятный вариант	\$5000	2%	\$5,00

Для каждого сценария требуется отследить ЧПС компании и прибыль за каждый год после уплаты налогов. Данные к этому примеру находятся в файле NPVauditscenario.xlsx. На рис. 19.1 представлена модель листа (на листе Original Model), а на рис. 19.2 приведен отчет по трем сценариям (на листе Scenario Summary).

	A	B	C	D	E	F
1		taxrate	0.4			
2		Year1sales	12000			
3		Sales growth	0.05			
4		Year1price	\$ 7.50			
5		Year1cost	\$ 6.00			
6		intrate	0.15			
7		costgrowth	0.05			
8		pricegrowth	0.03			
9	Year	1	2	3	4	5
10	Unit Sales	12000	12600	13230	13891.5	14586.075
11	unit price	\$ 7.50	\$ 7.73	\$ 7.96	\$ 8.20	\$ 8.44
12	unit cost	\$ 6.00	\$ 6.30	\$ 6.62	\$ 6.95	\$ 7.29
13	Revenues	\$ 90,000.00	\$ 97,335.00	\$ 105,267.80	\$ 113,847.13	\$ 123,125.67
14	Costs	\$ 72,000.00	\$ 79,380.00	\$ 87,516.45	\$ 96,486.89	\$ 106,376.79
15	Before Tax Profits	\$ 18,000.00	\$ 17,955.00	\$ 17,751.35	\$ 17,360.24	\$ 16,748.88
16	Tax	\$ 7,200.00	\$ 7,182.00	\$ 7,100.54	\$ 6,944.10	\$ 6,699.55
17	Aftertax Profits	\$ 10,800.00	\$ 10,773.00	\$ 10,650.81	\$ 10,416.15	\$ 10,049.33
18						
19	NPV	\$35,492.08				

**Рис. 19.1.** Данные, на которых основаны сценарии

Для определения наиболее благоприятного сценария откройте вкладку Данные (Data) и в группе Прогноз (Forecast) в раскрывающемся списке Анализ "что, если" (What-If Analysis) выберите инструмент Диспетчер сценариев (Scenario Manager), кнопку Добавить (Add) и заполните поля в диалоговом окне, как показано на рис. 19.3.

Введите имя сценария (Наиболее благоприятный) и выберите ячейки C2:C4 как ячейки с исходными данными, определяющие сценарий. Нажмите ОК в диалоговом окне Добавить сценарий и заполните поля входными значениями в диалоговом окне Значения ячеек сценария (Scenario Values), определяющими наиболее благоприятный вариант (рис. 19.4).

	B	C	D	E	F	G
1						
2	Scenario Summary					
3	Current Values:		Best	Most Likely	Wors	
5	Changing Cells:					
6	Year1sales	12000	20000	10000	5000	
7	Sales_growth	0.05	0.2	0.1	0.02	
8	Year1price	\$ 7.50	\$ 10.00	\$ 7.50	\$ 5.00	
9	Result Cells:					
10	\$B\$17	\$ 10,800.00	\$ 48,000.00	\$ 9,000.00	\$ (3,000.00)	
11	\$C\$17	\$ 10,773.00	\$ 57,600.00	\$ 9,405.00	\$ (3,519.00)	
12	\$D\$17	\$ 10,650.81	\$ 69,016.32	\$ 9,741.10	\$ (4,090.33)	
13	\$E\$17	\$ 10,416.15	\$ 82,560.80	\$ 9,980.12	\$ (4,718.50)	
14	\$F\$17	\$ 10,049.33	\$ 98,588.50	\$ 10,087.17	\$ (5,408.35)	
15	\$B\$19	\$35,492.08	\$226,892.67	\$32,063.83	(\$13,345.75)	
16	Notes: Current Values column represents values of changing cells at					
17	time Scenario Summary Report was created. Changing cells for each					
18	scenario are highlighted in gray.					

Рис. 19.2. Отчет по сценариям

**Edit Scenario**

Scenario name: Best

Changing cells: \$C\$2:\$C\$4

Ctrl+click cells to select non-adjacent changing cells.

Comment: Created by Owner on 6/23/2018

Protection: ☒ Prevent changes, ☐ Hide

OK Cancel

Рис. 19.3. Исходные данные для наиболее благоприятного сценария

**Scenario Values**

Enter values for each of the changing cells.

1: Year1sales 20000

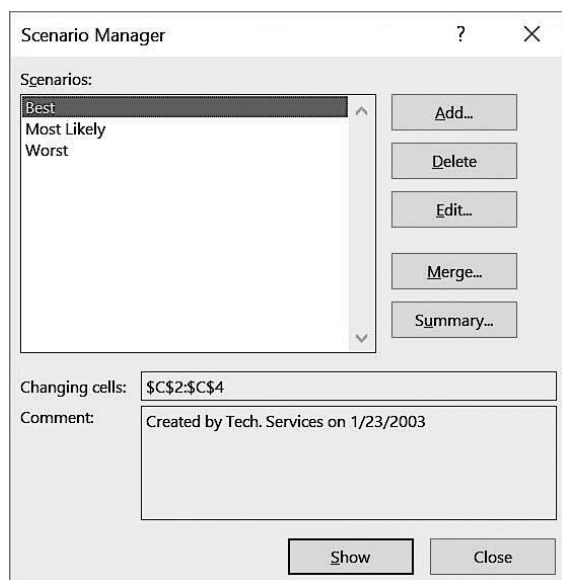
2: Sales\_growth .2

3: Year1price .10

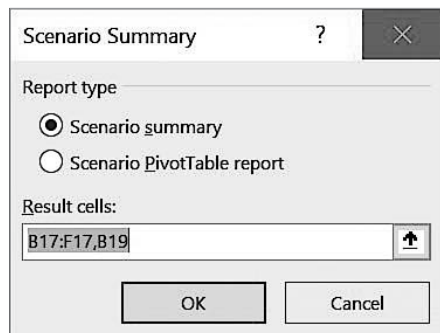
Add OK Cancel

Рис. 19.4. Определение исходных значений для наиболее благоприятного сценария

Для наиболее благоприятного сценария введите 2000 в качестве объема продаж за год 1 (Year1sales), 0,2 как значение роста продаж (Sales\_growth) и 10 как розничную цену в год 1 (Year1price). В диалоговом окне **Значения ячеек сценария (Scenario Values)** нажмите **Добавить (Add)**, недоступную, пока вы редактируете сценарий, и аналогичным образом введите данные для наиболее вероятного и наименее благоприятного сценариев. После ввода данных для всех трех сценариев в диалоговом окне **Значения ячеек сценария (Scenario Values)** нажмите **OK**. В диалоговом окне **Диспетчер сценариев (Scenario Manager)**, показанном на рис. 19.5, перечислены созданные сценарии.



**Рис. 19.5.** Диалоговое окно Диспетчера сценариев со списком созданных сценариев



**Рис. 19.6.** Диалоговое окно Отчет по сценарию для выбора в отчет ячеек с результатами

В диалоговом окне **Диспетчер сценариев (Scenario Manager)** нажмите **Отчет (Summary)** и выберите ячейки результата, которые должны отображаться в отчетах по сценариям. На рис. 19.6 показано, как в диалоговом окне **Отчет по сценарию (Scenario Summary)** я выбрал для отслеживания значение прибыли за каждый год после уплаты налогов (ячейки **B17:F17**) и значение ЧПС (ячейка **B19**).

Так как ячейки результата **B17:F17** и **B19** находятся в несмежных диапазонах, я разделил их точкой с запятой. (Т акже несколько диапазонов ячеек можно выбрать и внести, нажав **Ctrl.**) Установите переключатель **Тип отчета (Report type)** в положение **структура (Scenario Summary)**, а не **сводная таблица (Scenario PivotTable Report)**, и нажмите **ОК**. Excel создаст замечательный отчет **Структура сценария (Scenario Summary)**, показанный ранее на рис. 19.2. Обратите внимание, что в отчет включен столбец, помеченный как **Текущие значения (Current Values)**, для изначально указанных на листе значений. При наименее благоприятном сценарии компания несет убытки (в размере \$13 345,75), при наиболее благоприятном — получает прибыль (в размере \$226 892,67). Так как при наименее благоприятном сценарии цена ниже переменных расходов, компания теряет деньги каждый год.

## Замечания

- При установке в диалоговом окне **Отчет по сценарию (Scenario Summary)** переключателя в положение **сводная таблица (Scenario PivotTable Report)** результаты по сценариям представляются в формате сводных таблиц.
- Допустим, вы выбрали сценарий в диалоговом окне **Диспетчер сценариев (Scenario Manager)** и затем нажали **Вывести (Show)**. На листе появятся значения входных ячеек для выбранного сценария, и все формулы будут автоматически пересчитаны. Этот инструмент отлично подходит для презентации слайд-шоу сценариев.
- С помощью инструмента **Диспетчер сценариев (Scenario Manager)** сложно создать много сценариев, поскольку приходится вводить значения отдельно для каждого сценария. Большое количество сценариев можно создать с помощью моделирования по методу Монте-Карло (см. главу 77 «Введение в моделирование по методу Монте-Карло»). Метод Монте-Карло позволяет узнать, например, вероятность того, что ЧПС денежных потоков проекта является неотрицательной. Это важный показатель, поскольку такая вероятность показывает, повышает ли проект стоимость компании.
- При нажатии на знак «–» в строке 5 отчета **Структура сценария (Scenario Summary)** строки с предполагаемыми значениями скрываются, а отображаются только результаты. При нажатии на знак «+» отчет восстанавливается в полном объеме.
- Вы отправили файл нескольким людям, и каждый из них добавил собственный сценарий. После того как каждый человек вернет файл со сценариями, можно объединить все сценарии в одну книгу следующим образом: откройте каждую версию книги, нажмите в исходной книге в диалоговом окне **Диспетчер**

сценариев (Scenario Manager) кнопку **Объединить (Merge)** и затем выберите книги, содержащие сценарии, которые требуется объединить. Выбранные сценарии будут автоматически объединены в исходной книге.

## Задания

1. Удалите наиболее благоприятный сценарий и запустите еще один отчет по сценариям.
2. Добавьте сценарий с именем **Высокая цена (High Price)**, в котором цена в год 1 равна \$15, а для двух других исходных величин оставьте наиболее вероятные значения.
3. Для примера с продажей лимонада из главы 17 «Анализ чувствительности с помощью таблиц данных» с помощью инструмента **Диспетчер сценариев (Scenario Manager)** создайте отчет, суммирующий прибыль для следующих сценариев:

Сценарий	Цена	Себестоимость	Фиксированные расходы
Высокая себестоимость/высокая цена	\$5,00	\$1,00	\$65 000,00
Средняя себестоимость/средняя цена	\$4,00	\$0,75	\$45 000,00
Низкая себестоимость/низкая цена	\$2,50	\$0,40	\$25 000,00

4. Для примера с ипотекой из главы 17 с помощью инструмента **Диспетчер сценариев (Scenario Manager)** создайте отчет, сводящий в таблицу ежемесячные платежи для следующих сценариев:

Сценарий	Сумма кредита	Годовая процентная ставка	Количество ежемесячных платежей
Самый низкий платеж	\$300 000	4%	360
Наиболее вероятный платеж	\$400 000	6%	240
Самый высокий платеж	\$550 000	8%	180



## ГЛАВА 20

# Функции СЧЁТЕСЛИ, СЧЁТЕСЛИМН, СЧЁТ, СЧЁТЗ и СЧИТАТЬПУСТОТЫ

### Обсуждаемые вопросы

- Предположим, что у меня есть список исполненных по радио песен. Для каждой песни известны исполнитель, дата исполнения и продолжительность звучания. Как ответить на следующие вопросы о песнях из этого списка:
  - Сколько песен исполнил каждый певец?
  - Сколько песен было исполнено не Эминемом?
  - Сколько песен длились не меньше четырех минут?
  - Сколько песен длились дольше, чем средняя продолжительность звучания песен из списка?
  - Сколько песен было исполнено певцами, фамилии которых начинаются с S?
  - Сколько песен было исполнено певцами, фамилии которых состоят из шести букв?
  - Сколько песен прозвучало после 15 июня 2005 г.?
  - Сколько песен прозвучало до 2009 г.?
  - Сколько было исполнено песен, длящихся ровно четыре минуты?
  - Сколько песен, длящихся ровно четыре минуты, было исполнено Брюсом Спрингстином?
  - Сколько песен, длящихся от трех до четырех минут, было исполнено Мадонной?
- В более общем случае, как можно выполнить следующие операции:
  - подсчитать в диапазоне количество ячеек, содержащих числа;
  - подсчитать количество пустых ячеек в диапазоне;
  - подсчитать количество непустых ячеек в диапазоне.

Часто возникает необходимость подсчитать в диапазоне число ячеек, отвечающих заданному критерию. Например, если на листе содержатся данные о продажах косметики, вам может понадобиться информация о количестве сделок, заключенных продавцом по имени Дженнифер, или о количестве сделок, совершенных

после 10 июня. Количество ячеек, отвечающих заданным критериям, определенным на основе одной строки или столбца, можно подсчитать с помощью функции СЧЁТЕСЛИ (COUNTIF).

Синтаксис функции СЧЁТЕСЛИ: СЧЁТЕСЛИ(диапазон;критерий), где

- **диапазон** — диапазон ячеек, в котором необходимо подсчитать ячейки, отвечающие заданному критерию;
- **критерий** — число, дата или выражение, определяющее необходимость подсчета данной ячейки в диапазоне.

Синтаксис функции СЧЁТЕСЛИМН (COUNTIFS): СЧЁТЕСЛИМН(диапазон\_условия1; условие1[;диапазон\_условия2;условие2]...)

Функция СЧЁТЕСЛИМН подсчитывает число строк, в которых ячейки из первого диапазона соответствуют первому условию, ячейки из второго диапазона соответствуют второму условию и т. д. Таким образом, функция СЧЁТЕСЛИМН позволяет проверять условия для нескольких столбцов или несколько условий для одного столбца. Другие функции, учитывающие несколько критериев, рассматриваются в главах 21 «Функции СУММЕСЛИ, СРЗНА ЧЕСЛИ, СУММЕСЛИМН, СРЗНАЧЕСЛИМН, МАКСЕСЛИ и МИНЕСЛИ» и 50 «Обработка данных с помощью статистических функций для баз данных».

Ключом к успешному применению функции СЧЁТЕСЛИ (и других подобных функций) является знание всего спектра критериев, допустимых в Microsoft Excel. Типы используемых критериев проще всего объяснить на примерах. В дополнение к примерам с функцией СЧЁТЕСЛИ в этой главе представлены примеры использования функций СЧЁТ (COUNT), СЧЁТЗ (COUNTA) и СЧИТАТЬПУСТОТЫ (COUNTBLANK).

- Функция СЧЁТ возвращает число ячеек диапазона, содержащих числа.

	D	E	F	G
5				3.483701
6	<b>Song Number</b>	<b>Singer</b>	<b>Date</b>	<b>Minutes</b>
7	1	Eminem	5/21/2004	4
8	2	Eminem	4/15/2004	2
9	3	Cher	1/28/2005	2
10	4	Eminem	1/28/2005	4
11	5	Moore	11/5/2004	2
12	6	Cher	9/18/2004	4
13	7	Spears	4/15/2004	3
14	8	Spears	3/17/2005	3
15	9	Manilow	1/16/2005	4
16	10	Eminem	4/10/2005	4
17	11	Madonna	2/15/2004	3
18	12	Eminem	1/10/2004	4
19	13	Springsteen	4/10/2005	2
20	14	Spears	4/15/2004	4

**Рис. 20.1.** Часть песенной базы данных для примеров с функцией СЧЁТЕСЛИ

- Функция СЧЁТЗ возвращает число непустых ячеек в диапазоне.
- Функция СЧИТАТЬПУСТОТЫ возвращает число пустых ячеек в диапазоне.

Проиллюстрируем применение этих функций на примере базы данных (файл Rock.xlsx), в которой для каждой песни, прозвучавшей на радиостанции WKRP, содержатся следующие данные:

- исполнитель песни;
- дата исполнения песни;
- продолжительность звучания песни.

На рис. 20.1 показано подмножество данных.

## Ответы на вопросы

### ? Сколько песен исполнил каждый певец?

Сначала я выделил первую строку базы данных, диапазон D6:G6. Затем я выделил всю базу данных, нажав комбинацию Ctrl+Shift+↓. На вкладке **Формулы (Formulas)** в группе **Определенные имена (Defined Names)** я щелкнул на **Создать из выделенного (Create From Selection)** и выбрал в строке выше **(Top Row)**. Диапазон D7:D957 получил имя **Номер песни**, диапазон E7:E957 — имя **Исполнитель**, диапазон F7:F957 — имя **Дата** и диапазон G7:G957 — имя **Минуты**. Для определения числа песен, исполненных каждым певцом, скопируйте формулу `=СЧЁТЕСЛИ(Исполнитель;B5)` из ячейки C5 в ячейки C6:C12. В ячейке C5 теперь отображается число ячеек в диапазоне **Исполнитель**, значение в которых соответствует значению в ячейке B5 (**Eminem**). В базе находится 114 песен, исполненных Эминемом. Шер спела 112 песен и т. д. (рис. 20.2). Также можно подсчитать число песен, исполненных Эминемом, по формуле `=СЧЁТЕСЛИ(Исполнитель; "Eminem")`. Обратите внимание, что текст (например, **Eminem**) необходимо заключать в кавычки ("" ) и что в критериях не учитывается регистр.

	В	С
2	count	9
3	counta	18
4	countblank	2
5	Eminem	114
6	Cher	112
7	Moore	131
8	Spears	129
9	Mellencamp	115
10	Madonna	133
11	Springsteen	103
12	Manilow	114
13	Total	951

**Рис. 20.2.** Применение функции СЧЁТЕСЛИ для подсчета количества песен

### ? Сколько песен было исполнено не Эминемом?

Для решения этой задачи необходимо знать, что в Excel комбинация символов <> интерпретируется как «не равно». Формула =СЧЁТЕСЛИ(Исполнитель;"<>Eminem"), введенная в ячейку C15, вычислит, что 837 песен в базе данных были исполнены не Эминемом (рис. 20.3). Символы <>Eminem необходимо заключить в кавычки, поскольку в Excel комбинация символов «не равно» (<>) интерпретируется как текст, и значение Eminem, конечно, тоже является текстом. Тот же результат можно получить по формуле =СЧЁТЕСЛИ(Исполнитель;"<>"&B5), в которой знак «&» используется для объединения ссылки на ячейку B5 и оператора <>.

	В	С
15	Not by Eminem	837
16	Songs >= 4 minutes	477
17	Songs longer than average	477
18	Singer begins with S	232
19	Singer has six letters in name	243
20	Songs after 6/15/2005	98
21	Songs Before 2009	951
22	Songs exactly 4 minutes	247
23	Songs exactly 5 minutes	230
24	Springsteen songs 4 minutes	24
25	Madonna songs 3 or 4 minutes	70

**Рис. 20.3.** Функция СЧЁТЕСЛИ в комбинации с оператором «не равно» (<>)

### ? Сколько песен длились не меньше четырех минут?

Количество песен, длившихся не менее четырех минут, можно подсчитать по формуле =СЧЁТЕСЛИ(Минуты;">=4") в ячейке C16. Символы >=4 необходимо заключить в кавычки, поскольку комбинация символов «больше или равно» (>=), как и комбинация <>, обрабатывается как текст. Как видно из рис. 20.3, дольше четырех минут звучало 477 песен.

### ? Сколько песен длились дольше, чем средняя продолжительность звучания песен из списка?

Для ответа на этот вопрос сначала вычислите среднюю продолжительность звучания песен по формуле =СРЗНАЧ(Минуты) в ячейке G5. Затем в ячейке C17 подсчитайте число песен с продолжительностью звучания больше средней по формуле =СЧЁТЕСЛИ(Минуты;">"&G5). Символ & используется для ссылки в критерии на другую ячейку (в данном случае, G5). Как видно из того же рисунка, дольше среднего звучали 477 песен, что соответствует числу песен, звучавших не менее четырех минут. Эти результаты совпадают, поскольку подразумевается, что продолжительность звучания каждой песни является целым числом. Таким образом, если песня звучала не меньше 3,48 минут, значит, она звучала не меньше 4 минут.

**❓ Сколько песен было исполнено певцами, фамилии которых начинаются с S?**

Для ответа на этот вопрос воспользуйтесь в критерии подстановочным знаком «\*». Звездочка означает любую последовательность символов. Таким образом, по формуле `=СЧЁТЕСЛИ(Исполнитель;"S*")` в ячейке C18 подсчитываются все песни, исполненные певцами, чьи фамилии начинаются с S. (В критериях регистр не учитывается.) Певцы, фамилии которых начинаются с S, исполнили 232 песни. Это просто общее число песен, исполненных Брюсом Спрингстином и Бритни Спирс ( $103 + 129 = 232$ ).

**❓ Сколько песен было исполнено певцами, фамилии которых состоят из шести букв?**

Для этой задачи следует воспользоваться подстановочным знаком «?», который означает любой символ. Следовательно, число песен, исполненных певцами с фамилиями из шести букв, можно вычислить в ячейке C19 по формуле `=СЧЁТЕСЛИ(Исполнитель;"?????")`. В результате получится 243 песни. (У двух исполнителей фамилии состоят из шести букв у Бритни Спирс и у Эминема, которые в сумме исполнили 243 песни ( $129 + 114 = 243$ ).)

**❓ Сколько песен прозвучало после 15 июня 2005 г.?**

В критериях, используемых в функциях СЧЁТЕСЛИ, даты обрабатываются на основе порядкового номера. (Более поздняя дата считается больше, чем более ранняя дата.) В ячейке C20 по формуле `=СЧЁТЕСЛИ(Дата;">15/6/2005")` подсчитано, что после 15 июня 2005 г. прозвучало 98 песен.

**❓ Сколько песен прозвучало до 2009 г.?**

Здесь критерий должен учесть все даты до 31 декабря 2008.включительно. Я ввел в ячейку C21 формулу `=СЧЁТЕСЛИ(Дата;"<=31/12/2008")`. Как показано на рис. 20.3, до начала 2009 г. была исполнена 951 песня (как оказалось, это все песни).

**❓ Сколько было исполнено песен, длящихся ровно четыре минуты?**

Вычислите в ячейке C22 количество песен, длящихся ровно четыре минуты, по формуле `=СЧЁТЕСЛИ(Минуты;4)`. По этой формуле вычисляется количество ячеек в диапазоне G7:G957, содержащих значение 4. Как видно из рисунка, продолжительность звучания четыре минуты имеют 247 песен. Аналогичным образом в ячейке C23 я подсчитал, что 230 песен имеют продолжительность звучания пять минут.

**❓ Сколько песен, длящихся ровно четыре минуты, было исполнено Брюсом Спрингстином?**

Нам нужно подсчитать все строки, у которых в столбце Исполнитель находится значение Springsteen, а в столбце Минуты — значение 4. Поскольку для решения этой

задачи одного критерия мало, воспользуемся функцией СЧЁТЕСЛИМН (COUNTIFS). Введите в ячейку C2 формулу: .

=СЧЁТЕСЛИМН(Исполнитель;"Springsteen";Минуты;4).

По этой формуле учитывается каждая строка, в которой указаны исполнитель Springsteen и продолжительность звучания 4 минуты. Брюс Спрингстин спел 24 песни, звучащие ровно 4 минуты. Моей любимой песней Спрингстина является композиция «Thunder Road», но она длится более четырех минут. Воспользуйтесь мастером функций для ввода формул с функцией СЧЁТЕСЛИМН. И не забывайте, что можно указать диапазон в формуле с помощью клавиши F3.

### **❓ Сколько песен, длящихся от трех до четырех минут, было исполнено Мадонной?**

Поскольку речь идет о нескольких критериях, снова воспользуемся функцией СЧЁТЕСЛИМН. В ячейку C25 введите формулу =СЧЁТЕСЛИМН(Исполнитель;"Madonna"; Минуты;"<=4"; Минуты;">=3") для подсчета всех строк, в которых указано, что Мадонна исполняла песни, звучащие от трех до четырех минут. Это те самые строки, которые требуется подсчитать. Оказывается, Мадонна исполнила 70 песен продолжительностью от трех до четырех минут. (Моя любимая песня — «Crazy for You!».)

### **❓ Как подсчитать в диапазоне количество ячеек, содержащих числа?**

Количество ячеек с числовым значением можно подсчитать с помощью функции СЧЁТ (COUNT). Например, после ввода формулы =СЧЁТ(B5:C14) в ячейке C2 отображается значение 9, поскольку в диапазоне B5:C14 числа содержат 9 ячеек (C5:C13). (См. рис. 20.2.)

### **❓ Как подсчитать количество пустых ячеек в диапазоне?**

Для подсчета пустых ячеек в диапазоне предназначена функция СЧИТАТЬПУСТОТЫ (COUNTBLANK). Например, после ввода формулы =СЧИТАТЬПУСТОТЫ(B5:C14) в ячейке C4 отображается значение 2, поскольку в диапазоне B5:C14 две ячейки пустые (B14 и C14).

### **❓ Как подсчитать количество непустых ячеек в диапазоне?**

Количество непустых ячеек в диапазоне возвращает функция СЧЁТЗ (COUNTA). Например, после ввода формулы =СЧЁТЗ(B5:C14) в ячейке C3 отображается значение 18, поскольку в диапазоне B5:C14 18 непустых ячеек.

## **Замечания**

В оставшейся части книги обсуждаются альтернативные методы решения задач с изменением двух или более критериев (таких, например, как «Сколько песен Бритни Спирс прозвучало до 10 июня 2005 г.?»):

- в главе 50 рассматриваются статистические функции для баз данных;
- в главе 91 «Формулы массива и функции, возвращающие массив» рассматриваются формулы для обработки массивов.

## Задания

Данные к заданиям находятся в файле **Rock.xlsx**.

1. Сколько песен было исполнено не Бритни Спирс?
2. Сколько песен прозвучало до 15 июня 2004 г.?
3. Сколько песен прозвучало между 1 июня 2004 г. и 4 июля 2006 г.? Подсказка: возьмите разность значений, возвращенных двумя функциями СЧЁТЕСЛИ.
4. Сколько песен было исполнено певцами, чьи фамилии начинаются с М?
5. Сколько песен было исполнено певцами, в именах которых есть буква «Е»?
6. Создайте формулу, которая всегда возвращает количество песен, исполненных сегодня. Подсказка: используйте функцию СЕГОДНЯ().
7. В диапазоне ячеек **D4:G15** подсчитайте ячейки, содержащие числовые значения. Сосчитайте пустые ячейки. Сосчитайте непустые ячейки.
8. Сколько песен Барри Манилоу прозвучало в 2004 г.?
9. Сколько песен, длящихся не менее четырех минут и исполненных Мэнди Мур, прозвучало в 2007 г. или ранее?
10. Сколько песен в исполнении Бритни Спирс, длящихся три минуты, прозвучало позднее 2004 г.?
11. Файл **NBA.xlsx** содержит следующие данные:
  - в столбцах **A** и **B** находятся названия команд НБА и их кодовые номера. Например, команда с номером 1 представляет Атланту;
  - в столбце **C** указана команда хозяев поля для каждой игры;
  - в столбце **D** указана команда гостей для каждой игры;
  - в столбце **E** указаны очки, набранные командой хозяев поля;
  - в столбце **F** указаны очки, набранные командой гостей.

На основе этих данных подсчитайте количество игр, в которых участвовала каждая команда.

12. В файле **Matchthesecond.xlsx** находится список имен. Некоторые имена встречаются несколько раз. Определите строку, в которой встречается, например, второе вхождение имени «Dave». Создайте лист, на котором можно ввести имя и целое положительное число (например, *n*) и подсчитать строку, в которой введенное имя встречается в *n*-й раз.

13. В файле **Numbers.xlsx** содержится набор чисел. Подсчитайте, сколько этих чисел находится в интервале от 1 до 12 включительно.
14. В файле **Problem14data.xlsx** содержится информация о 2000 крупнейших компаниях в мире. Определите, сколько из них являются банками и сколько из этих банков находятся в США.
15. В файле **Problem14data.xlsx** содержится информация о 2000 крупнейших компаниях в мире. Определите, у скольких банков доходы от продаж равны или превышают 100 млрд долл., равны или превышают 90 млрд долл., равны или превышают 10 млрд долл. и т. д.



## ГЛАВА 21

# Функции СУММЕСЛИ, СРЗНАЧЕСЛИ, СУММЕСЛИМН, СРЗНАЧЕСЛИМН, МАКСЕСЛИ и МИНЕСЛИ

### Обсуждаемые вопросы

- Я работаю менеджером по продажам в компании, торгующей косметикой, и суммирую данные по каждой торговой сделке: продавец, дата продажи, число проданных (или возвращенных) единиц товара и полученная общая сумма (или уплаченная за возвраты сумма). Мне необходимо ответить на следующие вопросы:
  - Каков объем продаж в долларах для каждого продавца?
  - Сколько единиц товара было возвращено?
  - Каков объем продаж в долларах начиная с 2005 г.?
  - Сколько штук блеска для губ было продано? Какова выручка от продажи блеска для губ?
  - Каков объем продаж в долларах для всех продавцов, за исключением Джен?
  - Каково среднее количество единиц проданного товара по всем сделкам, совершенным конкретным продавцом?
  - Каков объем продаж помады в долларах у Джен?
  - Каково среднее количество проданной помады (в единицах товара) по всем продажам, совершенным Зарет?
  - Каково среднее количество проданной помады (в единицах товара) по всем сделкам, совершенным Зарет, для сделок по крайней мере с 50 единицами товара?
  - Каков объем продаж помады в долларах по всем сделкам, совершенным конкретным продавцом, для сделок на сумму свыше \$100? Каков объем продаж в долларах для сделок на сумму менее \$100?
  - Может ли Excel искать максимумы и минимумы по условию?

Можно ли применить Excel для расчета максимумов и минимумов по условию? Если вы хотите просуммировать все записи в одном столбце (или строке), которые отвечают критерию, зависящему от другого столбца (или строки), вам может функция СУММЕСЛИ (SUMIF). Синтаксис функции: СУММЕСЛИ(диапазон; критерий[;диапазон\_суммирования]), где:

- **диапазон** — диапазон ячеек, которые необходимо оценить в соответствии с критерием;
- **критерий** — это число, дата или выражение, определяющее, суммируется ли соответствующая ячейка в диапазоне суммирования;
- **диапазон суммирования** — диапазон ячеек, значения которых суммируются. Если диапазон суммирования опущен, то считается, что диапазон суммирования и диапазон совпадают.

Правила для критериев, указываемых в функции СУММЕСЛИ, идентичны правилам для функции СЧЁТЕСЛИ. Для получения дополнительной информации о функции СЧЁТЕСЛИ см. главу 20 «Функции СЧЁТЕСЛИ, СЧЁТЕСЛИМН, СЧЁТ, СЧЁТЗ и СЧИТАТЬ-ПУСТОТЫ».

Функция СРЗНАЧЕСЛИ (AVERAGEIF) возвращает среднее значение для отвечающих критерию ячеек диапазона. Синтаксис функции: СРЗНАЧЕСЛИ(диапазон; условие[;диапазон\_усреднения]). В Microsoft Excel 2019 имеется три функции (появившиеся еще в Microsoft Excel 2007), с помощью которых можно отметить строки, отвечающие нескольким критериям: СЧЁТЕСЛИМН (рассматривалась в главе 20), СУММЕСЛИМН (SUMIFS) и СРЗНАЧЕСЛИМН (AVERAGEIFS). Другие функции, допускающие вычисления по нескольким критериям, рассматриваются в главе 50 «Обработка данных с помощью статистических функций для баз данных». Кроме того, вычисления по нескольким критериям можно выполнять с помощью функций для обработки массивов (см. главу 91 «Формулы массива и функции, возвращающие массив»).

Синтаксис функции СУММЕСЛИМН: СУММЕСЛИМН(диапазон\_суммирования; диапазон\_условий1; условия1; диапазон\_условий2; условия2;... диапазон\_условийN; условияN). Функция СУММЕСЛИМН суммирует каждую запись из диапазона суммирования, удовлетворяющую всем указанным условиям: первому условию (на основе первого диапазона условий), второму условию (на основе второго диапазона) и т. д. Аналогично функция СРЗНАЧЕСЛИМН имеет синтаксис: СРЗНАЧЕСЛИМН(диапазон\_суммирования; диапазон\_условий1; условия1; диапазон\_условий2; условия2;... диапазон\_условийN; условияN). Функция СРЗНАЧЕСЛИМН возвращает среднее значение для записей в диапазоне суммирования, удовлетворяющих всем указанным условиям: первому условию (на основе первого диапазона условий), второму условию (на основе второго диапазона) и т. д.

В Office 365 и Excel 2019 имеются новые функции **МАКСЕСЛИ (MAXIFS)** и **МИНЕСЛИ (MINIFS)**, позволяющие вам искать максимумы и минимумы по условию. Функция **МАКСЕСЛИ** имеет синтаксис **МАКСЕСЛИ(макс\_диапазон; диапазон\_условий1; условия1; диапазон\_условий2; условия2;... диапазон\_условийN; условияN)**.

Функция возвращает максимальное значение, найденное среди всех записей, удовлетворяющих заданному критерию. Может быть задан только один критерий. Функция МИНЗНАЧЕСЛИМН работает схожим образом.

## Ответы на вопросы

### ? Каков объем продаж в долларах для каждого продавца?

Данные для примеров, рассмотренных в этой главе, находятся в файле Makeup2007.xlsx. Подмножество этих данных представлено на рис. 21.1.

	G	H	I	J	K	L
4	Trans Number	Name	Date	Product	Units	Dollars
5	1	Betsy	4/1/2004	lip gloss	45	\$ 137.20
6	2	Hallagan	3/10/2004	foundation	50	\$ 152.01
7	3	Ashley	2/25/2005	lipstick	9	\$ 28.72
8	4	Hallagan	5/22/2006	lip gloss	55	\$ 167.08
9	5	Zaret	6/17/2004	lip gloss	43	\$ 130.60
10	6	Colleen	11/27/2005	eye liner	58	\$ 175.99
11	7	Cristina	3/21/2004	eye liner	8	\$ 25.80
12	8	Colleen	12/17/2006	lip gloss	72	\$ 217.84
13	9	Ashley	7/5/2006	eye liner	75	\$ 226.64
14	10	Betsy	8/7/2006	lip gloss	24	\$ 73.50
15	11	Ashley	11/29/2004	mascara	43	\$ 130.84
16	12	Ashley	11/18/2004	lip gloss	23	\$ 71.03
17	13	Emilee	8/31/2005	lip gloss	49	\$ 149.59
18	14	Hallagan	1/1/2005	eye liner	18	\$ 56.47
19	15	Zaret	9/20/2006	foundation	-8	\$ (21.99)
20	16	Emilee	4/12/2004	mascara	45	\$ 137.39
21	17	Colleen	4/30/2006	mascara	66	\$ 199.65
22	18	Jen	8/31/2005	lip gloss	88	\$ 265.19

**Рис. 21.1.** Данные для примеров с функцией СУММЕСЛИ

Как обычно, начнем с присвоения данным в столбцах с G по L соответствующих имен из ячеек G4:L4. Например, имя диапазона Продукт соответствует диапазону J5:J1904. Для расчета общей суммы продаж, совершенных каждым продавцом (рис. 21.2), я просто скопировал формулу =СУММЕСЛИ(Имя;A5;Доллары) из B5 в B6:B13. Эта формула суммирует все значения в столбце Доллары для Emilee в столбце Имя и показывает, что Эмили продала косметики на сумму \$25 258,87. Безусловно, формула =СУММЕСЛИ(Имя;"Emilee";Доллары) даст тот же результат.

### ? Сколько единиц товара было возвращено?

В ячейке B16 по формуле =СУММЕСЛИ(Единицы\_товара;"<0";Единицы\_товара) суммируются все отрицательные числа в столбце Единицы товара (столбец K). Ре-

зультат равен –922. Если поставить в начале формулы знак «–», то мы получим количество возвращенных единиц товара (922). Напомню, что если аргумент **диапазон\_суммирования** в вызове функции **СУММЕСЛИ** опущен, подразумевается, что он совпадает с аргументом **диапазон**. Таким образом, формула **=СУММЕСЛИ(Единицы\_товара;"<0")** также возвратит 922.

	A	B	C
4	Name	Dollar Volume	
5	Emilee	\$ 25,258.87	=SUMIF(Name,A5,Dollars)
6	Hallagan	\$ 28,705.16	=SUMIF(Name,A6,Dollars)
7	Ashley	\$ 25,947.24	=SUMIF(Name,A7,Dollars)
8	Zaret	\$ 26,741.31	=SUMIF(Name,A8,Dollars)
9	Colleen	\$ 24,890.66	=SUMIF(Name,A9,Dollars)
10	Cristina	\$ 23,849.56	=SUMIF(Name,A10,Dollars)
11	Betsy	\$ 28,803.15	=SUMIF(Name,A11,Dollars)
12	Jen	\$ 29,050.53	=SUMIF(Name,A12,Dollars)
13	Cici	\$ 27,590.57	=SUMIF(Name,A13,Dollars)
14			
15			
16	Units returned	922	=SUMIF(Units,"<0",Units)
17	Total dollars sold 2005 or later	\$ 157,854.32	=SUMIF(Date,">=1/1/2005",Dollars)
18	Units of lip gloss sold	16333	=SUMIF(Product,"lip gloss",Units)
19	\$s of lip gloss sold	\$ 49,834.64	=SUMIF(Product,"lip gloss",Dollars)
20	\$s sold not by Jen	\$ 211,786.51	=SUMIF(Name,"<>Jen",Dollars)
21	Lipstick \$s by Jen	3953	=SUMIFS(Dollars,Name,"Jen",Product,"lipstick")
22	avg units lipstick by Zaret	33	=AVERAGEIFS(Units,Name,"Zaret",Product,"lipstick")
23	avg units lipstick Zaretunits >=50	68	=AVERAGEIFS(Units,Name,"Zaret",Product,"lipstick",Units,">=50")
24	Lipstick \$s >=\$100 Jen	3583	=SUMIFS(Dollars,Name,"Jen",Product,"lipstick",Dollars,">=100")
25	Lipstick \$s <\$100 Jen	370	check
26	average units by Jen	43.548	
27		=AVERAGEIF(Name,"Jen",Units)	=SUMIF(Name,"Jen",Units)/COUNTIF(Name,"Jen")

**Рис. 21.2.** Результаты вычислений посредством функции **СУММЕСЛИ**

### ❓ Каков объем продаж в долларах начиная с 2005 г.?

В ячейке B17 по формуле **=СУММЕСЛИ(Дата;">=1/1/2005";Доллары)** суммируются все записи в столбце **Доллары** (столбец L), содержащие дату не ранее 1/1/2005 в столбце **Дата**. Вычисления показывают, что начиная с 2005 г. было продано косметики на \$157 854,32.

### ❓ Сколько штук блеска для губ было продано? Какова выручка от продажи блеска для губ?

В ячейке B18 по формуле **=СУММЕСЛИ(Продукт;"lip gloss";Единицы\_товара)** суммируются значения во всех ячейках столбца **Единицы товара**, для которых в столбце **Продукт** (столбец J) содержится текст "lip gloss". Как мы видим, было продано 16 333 шт. блеска для губ. Это чистый объем продаж: сделки, по которым блеск для губ был возвращен, учитываются как отрицательные продажи.

Аналогично, в ячейке B19 по формуле **=СУММЕСЛИ(Продукт;"lip gloss";Доллары)** вычислена чистая сумма продаж блеска для губ — \$49 834,64. Здесь суммы, связанные с возвратом товара, учтены как отрицательная выручка.

**? Каков объем продаж в долларах для всех продавцов, за исключением Джен?**

В ячейке B20 по формуле =СУММЕСЛИ(Имя;"<>Jen";Доллары) складываются суммы в долларах для всех сделок, для которых в столбце Имя не указано имя Jen. Объем продаж косметики без учета продаж, совершенных Джен, составил \$211 786,51.

**? Каково среднее количество единиц проданного товара по всем сделкам, совершенным конкретным продавцом?**

Это работа для функции СРЗНАЧЕСЛИ. Введите формулу =СРЗНАЧЕСЛИ(Имя;"Jen";Единицы\_товара) в ячейку B26 и получите среднее значение для всех записей в столбце Единицы товара, содержащих в столбце Имя значение Jen. Средний размер сделки Джен составляет 43,548 единиц товара. Это можно проверить в ячейке C26 по формуле =СУММЕСЛИ(Имя;"Jen";Единицы\_товара)/СЧЁТЕСЛИ(Имя;"Jen").

**? Каков объем продаж помады в долларах у Джен?**

На этот раз в расчете участвуют два критерия (Имя="Jen" и Продукт="lipstick"). Таким образом, искомый объем продаж вычисляется в ячейке B21 по формуле =СУММЕСЛИМН(Доллары;Имя;"Jen";Продукт;"lipstick"). Всего Джен продала помады на \$3953.

**? Каково среднее количество проданной помады (в единицах товара) по всем продажам, совершенным Зарет?**

Для расчета потребуется функция СРЗНАЧЕСЛИМН. Я вычислил количество в ячейке B22 по формуле =СРЗНАЧЕСЛИМН(Единицы\_товара;Имя;"Zaret";Продукт;"lipstick"). В сделках Зарет по продаже помады среднее количество проданных единиц товара составило 33 шт.

**? Каково среднее количество проданной помады (в единицах товара) по всем сделкам, совершенным Зарет, для сделок по крайней мере с 50 единицами товара?**

Здесь я снова использовал функцию СРЗНАЧЕСЛИМН, но добавил еще один критерий, обеспечивающий учет только тех сделок, для которых количество проданных штук составляет не менее 50. В ячейку B23 я ввел формулу =СРЗНАЧЕСЛИМН(Единицы\_товара;Имя;"Zaret";Продукт;"lipstick";Единицы\_товара;">=50"). Во всех сделках Зарет с помадой, в которых продано не менее 50 единиц товара, количество помады составило 68 шт.

**? Каков объем продаж помады в долларах по всем сделкам, совершенным конкретным продавцом, для сделок на сумму свыше \$100? Каков объем продаж в долларах для сделок на сумму менее \$100?**

Поскольку в дополнение к критериям (Имя="Jen" и Продукт="lipstick") здесь имеются некоторые утверждения о размере сделок в долларах, необходима функция

СУММЕСЛИМН. В ячейке B24 я вычислил общую сумму сделок, по которым Джен продала помаду и в которых сумма составила не меньше \$100, по формуле `=СУММЕСЛИМН(Доллары;Имя;"Jen";Продукт;"lipstick";Доллары;">=100")`. В таких сделках Джен продала помады на \$3583. Для сделок с помадой на сумму менее \$100 результат вычислен в ячейке B25 по формуле `=СУММЕСЛИМН(Доллары;Имя;"Jen";Продукт;"lipstick";Доллары;"<100")` и составляет \$370. Обратите внимание, что сумма  $370 + 3583$  соответствует общей долларовой выручке Джен от продаж помады, вычисленной в ячейке B21.

### ? Может ли Excel искать максимумы и минимумы по условию?

В Office 365 и Excel 2019 имеются функции `МАКСЕСЛИ(MAXIFS)` и `МИНЗНАЧЕСЛИ(MINIFS)`, применяемые для расчета максимумов и минимумов, заданных определенными условиями. Применение этих функций показано в файле `Maxifsminifs.xlsx` (рис. 21.3). В файле содержатся имя, команда, позиция на поле и число очков, заработанных каждым игроком НБА в сезон игр 2016–2017 года. Нам нужно узнать максимальное и минимальное количество очков для каждого сочетания позиции и команды. Данные перечислены в столбцах I–K, а имена в строке 1. На рабочем листе `Maxifsminifs.xlsx` мы увидим максимальное число заработанных очков по месту в таблице для каждой команды, скопировав формулу `=МАКСЕСЛИ(PTS;Tm;$A4;Pos;B$3)` из ячейки B4 в диапазон B4:F33. Например, мы видим, что максимальное число очков заработано хьюстонским подвижным защитником (SG) и составляет 2085 очков. Заработал их Джеймс Харден. На рабочем

	A	B	C	D	E	F
1		MAX POINTS BY TEAM			MAXIFS	
2		AND POSITION				
3		SG	PF	C	SF	PG
4	ATL	972	1174	854	740	1237
5	BOS	812	450	830	843	1916
6	BRK	863	641	1350	783	484
7	CHI	1062	639	731	1543	459
8	CHO	1006	746	749	633	1588
9	CLE	516	961	572	1667	1586
10	DAL	912	1393	437	329	857
11	DEN	789	521	1006	981	634
12	DET	944	1174	1028	1004	741
13	GSW	1517	680	377	1494	1715
14	HOU	1067	907	697	816	2085
15	IND	546	675	1019	1443	1079
16	LAC	1034	1108	901	182	911
17	LAL	1078	825	401	640	846
18	MEM	572	889	1370	383	1200
19	MIA	729	809	1127	294	1266
20	MIL	693	1025	828	1609	509
21	MIN	889	687	1720	1607	649
22	NOP	628	584	1854	477	881

**Рис. 21.3.** Пример применения функции `МАКСЕСЛИ`

листе **Minifs** показано применение функции **МИНЕСЛИ**(**MINIFS**). В главе 91 будет показано применение формул обработки массивов для расчета условных максимумов и минимумов.

## Задания

1. Для каждого продукта определите общее количество проданных единиц товара и объем продаж в долларах.
2. Определите общую выручку, полученную до 10 декабря 2005 г.
3. Определите общее количество единиц товара, проданное работниками, чьи фамилии начинаются с С.
4. Определите общую выручку для продавцов, имена которых состоят из пяти букв.
5. Сколько единиц товара было продано всеми продавцами, за исключением Коллин?
6. Сколько единиц косметики было продано с 15 января 2004 г. до 15 февраля 2005 г.?
7. Файл **NBA.xlsx** содержит следующие данные:
  - в столбцах **A** и **B** находятся названия команд НБА и их кодовые номера. Например, команда с номером 1 представляет Атланту;
  - в столбце **C** указана команда хозяев поля для каждой игры;
  - в столбце **D** указана команда гостей для каждой игры;
  - в столбце **E** указаны очки, набранные командой хозяев поля;
  - в столбце **F** указаны очки, набранные командой гостей.Исходя из этих данных, вычислите для каждой команды среднее количество набранных очков за игру и среднее количество проигранных очков.
8. В файле **Toysrus.xlsx** содержится объем продаж (в млн долл.) для каждого квартала 1997–2001 гг. и первых двух кварталов 2002 г. На основе этих данных вычислите индекс сезонности для каждого квартала. Например, если средний объем продаж в первом квартале составил 80% от общего среднего объема продаж за квартал, индекс сезонности для первого квартала будет равен 0,8.
9. В файле **Sumifrows.xlsx** содержатся данные о продажах за несколько зимних, весенних, летних и осенних кварталов. Определите средний объем продаж для зимних, весенних, летних и осенних кварталов.
10. И снова на основе данных из файла **Makeup2007.xlsx** (используйте его для заданий 10–16) определите объем выручки от сделок по продаже не менее 50 единиц косметики.



11. Сколько штук блеска для губ продала Сиси в 2004 г.?
12. Каково среднее количество проданной Эмили основы под макияж?
13. Каков средний размер в долларах сделок Бетси по продаже основы под макияж после конца 2004 г.?
14. Какова общая сумма сделок Эшли, по которым было продано не менее 40 штук помады?
15. Создайте таблицу, содержащую объемы продаж каждого продукта каждым продавцом.
16. Создайте таблицу, в которой при вводе года отображаются объемы продаж по каждому продукту для каждого продавца в этом году.
17. На основе данных из файла **Numbers.xlsx** найдите общую сумму всех чисел не меньше 5 и не больше 15.
18. На основе данных из файла **Numbers.xlsx** найдите среднее значение всех чисел не меньше 10 и не больше 25.
19. В файле **Problem19data.xlsx** приведены ежемесячные данные о температуре в США в 1895–2015 гг. Вычислите среднюю температуру по каждому из месяцев года за весь этот срок.
20. В файле **Problem20data.xlsx** указано расстояние в ярдах, засчитанное в рывках и пасах командам НФЛ. Определите среднее число ярдов, набранных за игру каждой командой в рывках и пасах.
21. В файле **Problem21data.xlsx** представлена статистика по квотербекам НФЛ. Узнайте общее количество забитых мячей и передач для каждой команды НФЛ. Учтите, что в некоторых командах есть несколько квотербеков!
22. Опираясь на данные в файле **Problem22data.xlsx**, определите итоговое число единиц каждого продукта, сбытое каждым менеджером продаж.
23. В файле **Problem23\_24data.xlsx** собраны данные по 2000 крупнейших компаний в мире. Определите количество банковских организаций, средний доход на каждую фирму из перечисленных банковских организаций и долю всего банковского дохода, заработанного банками, чей доход составляет от 20 млрд долл.
24. В файле **Problem23\_24data.xlsx** собраны данные по 2000 крупнейших компаний в мире. Определите доход и название крупнейшей французской фирмы розничной торговли.



## ГЛАВА 22

# Функция СМЕЩ

### Обсуждаемые вопросы

- Как создать ссылку на диапазон ячеек, который отстоит от ячейки или другого диапазона на заданное количество строк и столбцов?
- Как можно выполнить операцию поиска при помощи крайнего правого столбца таблицы вместо крайнего левого?
- Я часто скачиваю данные о продажах программных продуктов, сгруппированные по странам/регионам. Я должен отслеживать для Ирана выручку, затраты и количество проданных программных продуктов, но данные по Ирану не всегда находятся в одной и той же части листа. Можно ли создать формулу, в которой выручка, затраты и количество проданных программных продуктов выбираются всегда правильно?
- Каждый разрабатываемый компанией препарат проходит три этапа разработки. У меня есть список затрат по месяцам на каждый препарат, и я знаю продолжительность в месяцах каждого этапа разработки. Можно ли создать формулу, вычисляющую общие затраты для каждого препарата на каждом этапе разработки?
- Я владелец небольшого видеопроката. На рабочем листе мой бухгалтер указал названия всех фильмов и количество копий на складе. К сожалению, он объединил информацию для каждого фильма в одной ячейке. Как мне переписать данные о количестве копий на складе для каждого фильма в отдельную ячейку?
- Как работает в Excel инструмент Вычислить формулу?
- Как написать формулу, всегда возвращающую последнее число в столбце?
- Как создать диапазон, автоматически включающий новые данные?
- Я ежемесячно составляю диаграммы по объемам продаж в штуках для продукта компании. Каждый месяц я загружаю новые данные. Хорошо бы диаграммы обновлялись автоматически. Есть ли простой способ добиться этого?

Для создания ссылки на диапазон, отстоящий на заданное число строк и столбцов от ячейки или диапазона ячеек, предназначена функция **СМЕЩ (OFFSET)**. По сути, при создании ссылки на диапазон ячеек сначала определяется опорная ячейка. Затем указывается число строк и столбцов от опорной ячейки до требуемой ячейки, начиная с которой должен быть создан диапазон. Например, при помощи функции **СМЕЩ** можно создать ссылку на диапазон ячеек, содержащий две строки и три столбца и начинающийся на два столбца правее и на одну строку выше текущей

ячейки. Число строк и столбцов, которые требуется отсчитать от опорной ячейки, можно вычислять при помощи других функций Microsoft Excel.

Синтаксис функции СМЕЩ: СМЕЩ(ссылка;смещ\_по\_строкам;смещ\_по\_столбцам[;высота][;ширина]), где

- **ссылка** — опорная ячейка или диапазон ячеек, откуда отсчитывается смещение. Если указывается диапазон ячеек, ячейки должны быть смежные. **смещ\_по\_строкам** — число строк от опорной ячейки или диапазона до строки, в которой должен начинаться смещенный диапазон (строка, в которой находится левая верхняя ячейка в смещенном диапазоне). Если указано отрицательное число, отсчет идет вверх от опорной ячейки, если положительное — вниз. Например, если для аргумента **ссылка** указана ячейка C5 и для аргумента **смещ\_по\_строкам** указано значение -1, смещение произойдет в строку 4. Если для аргумента **смещ\_по\_строкам** указано значение +1, то смещение произойдет в строку 6. Если значение для **смещ\_по\_строкам** равно 0, смещенный диапазон начинается в строке 5;
- **смещ\_по\_столбцам** — число столбцов от опорной ячейки или диапазона до столбца, в котором должен начинаться смещенный диапазон. Отрицательное число означает смещение влево от опорной ячейки, положительное — вправо. Например, если для аргумента **ссылка** указана ячейка C5 и если для аргумента **смещ\_по\_столбцам** указано значение -1, смещение произойдет в столбец B. Если значение для **смещ\_по\_столбцам** равно +1, смещение произойдет в столбец D. Если значение для **смещ\_по\_столбцам** равно 0, смещенный диапазон начнется в столбце C;
- **высота** и **ширина** — необязательные аргументы, задающие количество строк и столбцов в смещенном диапазоне. Если аргумент **высота** или **ширина** опущен, функция СМЕЩ возвращает ссылку на диапазон, высота или ширина которого та же, что в аргументе **ссылка**.

## Ответы на вопросы

- ❓ **Как создать ссылку на диапазон ячеек, который отстоит от ячейки или другого диапазона на заданное количество строк и столбцов?**

Файл *Offsetexample.xlsx* (рис. 22.1) содержит несколько примеров с функцией СМЕЩ в действии.

Например, в ячейку B10 я ввел формулу, показанную в ячейке A10: =СУММ(СМЕЩ(B7;-1;1;2;1)). Вычисление начнется с опорной ячейки B7. Смещение произойдет на одну строку вверх и один столбец вправо, то есть в ячейку C6. Затем формула выделит диапазон, состоящий из двух строк и одного столбца, а именно C6:C7. Функция СУММ возвратит сумму чисел в этом диапазоне:  $2 + 6 = 8$ . Два других примера (рис. 22.1) обрабатываются аналогичным образом. Далее я покажу вам, как применить функцию СМЕЩ для решения некоторых задач, присланных мне бывшими студентами, а ныне сотрудниками крупных американских компаний.

	A	B	C	D	E	F	G	H	I	J	K
1											
2											
3		Offset examples									
4											
5											
6		1	2	3	4			1	2	3	4
7		5	6	7	8			5	6	7	8
8		9	10	11	12			9	10	11	12
9											
10	=SUM(OFFSET(B7,-1,1,2,1))	8					=SUM(OFFSET(H6,0,1,3,2))	39			
11											
12											
13											
14		1	2	3	4						
15		5	6	7	8						
16		9	10	11	12						
17											
18	=SUM(OFFSET(E16,-2,-3,2,3))	24									

**Рис. 22.1.** Функция СМЕЩ в комбинации с функцией СУММ

**?** Как можно выполнить операцию поиска при помощи крайнего правого столбца таблицы вместо крайнего левого?

На рис. 22.2 (см. книгу Lefthandlookup.xlsx) я привел список членов баскетбольной команды НБА «Даллас Маверикс» сезона 2002–2003 гг. и указал процент бросков с игры для каждого игрока. Если бы меня попросили найти игрока с определенным процентом бросков с игры, я легко справился бы с задачей при помощи функции ВПР (VLOOKUP). Но здесь требуется *найти слева* процент бросков с игры для игрока по его имени. Функция ВПР не может выполнить поиск слева, но ее можно заменить комбинацией функций ПОИСКПОЗ (MATCH) и СМЕЩ.

	B	C	D	E
5		Left-hand lookup		
6			Name	FG %age
7	FG%	Player	Walt Williams	0.397
8	45.8%	Dirk Nowitzki		=OFFSET(B7,MATCH(D7,\$C\$8:\$C\$22,0),0,1,1)
9	41.8%	Michael Finley		
10	46.3%	Steve Nash		
11	39.5%	Nick Van Exel		
12	53.5%	Raef LaFrentz		
13	60.2%	Eduardo Najera		
14	51.2%	Shawn Bradley		
15	39.7%	Walt Williams		
16	44.4%	Adrian Griffin		
17	48.4%	Avery Johnson		
18	47.6%	Raja Bell		
19	66.7%	Evan Eschmeyer		
20	41.0%	Popeye Jones		
21	40.0%	Mark Strickland		
22	23.5%	Adam Harrington		

**Рис. 22.2.** Поиск слева при помощи функций ПОИСКПОЗ и СМЕЩ

Сначала я ввел имя игрока в ячейку D7. Затем я воспользовался опорной ячейкой B7 (заголовок Процент бросков с игры) в функции СМЕЩ. Для поиска процента бросков с игры указанного игрока мне нужно сместиться вниз под строку 7, там отображаются имена игроков. Здесь необходима функция ПОИСКПОЗ. Часть формулы =СМЕЩ(B7;ПОИСКПОЗ(D7;\$C\$8:\$C\$22;0);0), содержащая функцию ПОИСКПОЗ, вызовет смещение в строку с указанным именем игрока, а для смещения по столбцам задано значение 0. Поскольку для аргумента ссылка указана одна ячейка, а аргументы высота и ширина опущены, функция СМЕЩ также возвращает одну ячейку. И это ячейка с процентом бросков с игры для указанного игрока. См. ячейку E7.

**?** Я часто скачиваю данные о продажах программных продуктов, сгруппированные по странам/регионам. Я должен отслеживать для Ирана выручку, затраты и количество проданных программных продуктов, но данные по Ирану не всегда находятся в одной и той же части листа. Можно ли создать формулу, в которой выручка, затраты и количество проданных программных продуктов выбираются всегда правильно?

В файле Asiansales.xlsx (рис. 22.3) содержатся данные о количестве проданных программных продуктов, выручке и переменных расходах по нескольким странам в Азии и на Ближнем Востоке. Каждый раз при загрузке ежемесячных финансовых отчетов местоположение данных на листе для каждой страны меняется, так что необходима формула, которая всегда возвращает (для заданной страны) правильные значения количества проданных программных продуктов, выручки и переменных затрат.

	C	D	E	F	G
6	Country/Region	Units sold	Revenue	Variable Cost	
7	India	541	\$ 4,328	\$ 1,082	
8	China	1000	\$ 5,000	\$ 2,000	
9	Iran	577	\$ 2,308	\$ 1,731	
10	Israel	454	\$ 3,632	\$ 1,362	
11	Japan	141	\$ 705	\$ 423	
12	Taiwan	221	\$ 1,105	\$ 663	
13	Thailand	223	\$ 1,115	\$ 446	
14	Indonesia	524	\$ 2,620	\$ 1,048	
15	Malaysia	328	\$ 1,968	\$ 984	
16	Vietnam	469	\$ 2,814	\$ 1,407	
17	Cambodia	398	\$ 1,990	\$ 796	
18					
19		Units sold	Revenue	Variable Cost	
20	Country/Region	1	2	3	
21	Iran	577	2308	1731	
22					
23		=OFFSET(\$C\$6,MATCH(\$C\$21,\$C\$7:\$C\$17,0),D20)			

**Рис. 22.3.** Функция СМЕЩ для работы с данными, отображаемыми в разных частях листа

Вы получите нужный результат, скопировав формулу `=СМЕЩ($C$6;ПОИСКПОЗ($C21; $C$7:$C$17;0);D20)` из D21 в E21:F21. В этой формуле для аргумента ссылка установлена ячейка C6 (в которой содержатся слова Страна/регион) и указано смещение вправо на один столбец (значение в ячейке D20) для поиска количества проданных программных продуктов и вниз до строки, содержащей название страны, указанной в ячейке C21. В ячейке E21 вместо ссылки на ячейку D20 теперь ссылка на ячейку E20, в которой записано значение 2, следовательно, происходит смещение на два столбца вправо от столбца C к столбцу с выручкой. В ячейке E21 вместо ссылки на D20 теперь ссылка на ячейку F20 со значением 3, следовательно, происходит смещение на три столбца вправо от столбца C к столбцу с переменными расходами.

**?** Каждый разрабатываемый компанией препарат проходит три этапа разработки. У меня есть список затрат по месяцам на каждый препарат и я знаю продолжительность в месяцах каждого этапа разработки. Можно ли создать формулу, вычисляющую общие затраты для каждого препарата на каждом этапе разработки?

В файле `Offsetcost.xlsx` находятся данные о ежемесячных затратах на разработку пяти лекарственных препаратов. Для каждого лекарства указано количество месяцев, которое отведено на каждый этап разработки. Подмножество этих данных представлено на рис. 22.4.

	B	C	D	E	F	G	H
1		Dur Phase 1	2	3	9	12	6
2		Dur Phase 2	2	8	5	4	12
3		Dur Phase 3	2	11	4	11	15
4		Phase 1 Cost	110	313	795	1167	615
5		Phase 2 Cost	142	789	465	397	1096
6		Phase 3 Cost	234	876	401	1135	1588
7							
8							
9	Code	Month	Drug 1	Drug 2	Drug 3	Drug 4	Drug 5
10	1	Jan-98	52	135	131	121	69
11	2	Feb-98	58	120	77	60	68
12	3	Mar-98	80	58	66	52	113
13	4	Apr-98	62	56	78	61	146
14	5	May-98	130	126	98	118	94
15	6	Jun-98	104	102	64	117	125
16	7	Jul-98	121	59	115	112	137
17	8	Aug-98	107	123	56	102	77
18	9	Sep-98	80	88	110	85	93
19	10	Oct-98	51	111	72	118	89
20	11	Nov-98	74	124	82	143	66
21	12	Dec-98	76	107	99	78	66
22	13	Jan-99	97	97	129	77	142

**Рис. 22.4.** Вычисление затрат на разработку на этапах 1–3 с помощью функции СМЕЩ

Цель — определить для каждого препарата общие затраты на каждом этапе разработки. В ячейках D4:D6 я вычислил общие затраты на разработку препарата 1 для этапов 1–3. Я вычислил затраты для препарата 1 на этапе 1, указав в качестве ссылки опорную ячейку D10 и значение 0 для смещения по строкам и по столбцам. Для охвата всех затрат на первом этапе я указал для высоты ячейку с числом месяцев на этапе 1, а для ширины — значение 1. Таким образом, затраты на этапе 1 для первого препарата вычисляются в ячейке D4 по формуле =СУММ(СМЕЩ(D10;0;0;D1;1)). Затем в ячейке D5 я вычислил общие затраты для препарата 1 на втором этапе по формуле =СУММ(СМЕЩ(D10;D1;0;D2;1)). Обратите внимание, что я указал ссылку на ячейку D10 (первый месяц затрат) и сместился вниз на количество строк, равное продолжительности этапа 1, к началу этапа 2. Для включения всех затрат этапа 2 нужно указать для аргумента **высота** значение в ячейке D2. Наконец, в ячейке D6 я вычислил затраты на разработку препарата 1 на этапе 3 по формуле =СУММ(СМЕЩ(D10;D1+D2;0;D3;1)). В ней сначала указывается первый месяц затрат и смещение вниз на число строк, равное суммарной продолжительности этапов 1 и 2, к началу этапа 3. В ячейке D3 хранится число строк с затратами для этапа 3. Далее, скопировав формулы из D4:D6 в E4:H6, я могу вычислить общие затраты для этапов 1–3 на разработку препаратов 2–5. Например, для препарата 2 общие затраты на этапе 1 составили \$313, на этапе 2 — \$789 и на этапе 3 — \$876.

❓ Я владелец небольшого видеопроката. На рабочем листе мой бухгалтер указал названия всех фильмов и количество копий на складе. К сожалению, он объединил информацию для каждого фильма в одной ячейке. Как мне переписать данные о количестве копий на складе для каждого фильма в отдельную ячейку?

На листе Movies в книге Movies.xlsx, приведенном на рис. 22.5, содержатся названия всех фильмов и число их копий на складе.

	A	B	C	D	E	F	G	H	I
1	count words	Number	Movie and Number of Copies						
2	2	40	Seabiscuit 40	Seabiscuit	40				
3	4	12	Lara Croft Tombrailer 12	Lara	Croft	Tombrailer	12		
4	6	36	Raiders of the Lost Ark 36	Raiders	of	the	Lost	Ark	36
5	3	5	Annie Hall 5	Annie	Hall	5			
6	2	4	Manhattan 4	Manhattan	4				
7	3	112	Star Wars 112	Star	Wars	112			
8	4	128	How to Deal 128	How	to	Deal	128		
9	4	1	The Matrix Reloaded 1	The	Matrix	Reloaded	1		
10	3	1040	Johnny English 1040	Johnny	English	1040			
11	3	12	Rosemary's Baby 12	Rosemary's	Baby	12			
12	3	1002	High Noon 1002	High	Noon	1002			

**Рис. 22.5.** Функция СМЕЩ в примере с фильмами

Мне нужно извлечь количество копий каждого фильма в отдельную ячейку. Если бы количество копий в ячейке было указано слева от названия фильма, задача

была бы простой; в этом случае следовало бы применить функцию **НАЙТИ** (FIND) и определить местоположение первого пробела, а затем при помощи функции **ЛЕВСИМВ** (LEFT) возвратить все символы слева от первого пробела. (См. в главе 6 способы применения функций **НАЙТИ** и **ЛЕВСИМВ**, а также других текстовых функций для работы с текстом.) К сожалению, этот метод неприменим, если количество копий указано справа от названия фильма. Например, для названия, состоящего из одного слова, количество копий находится справа от первого пробела, а для названия, состоящего из четырех слов, — справа от четвертого пробела.

Можно подойти к решению задачи следующим образом: выбрать на вкладке **Данные** (Data), в группе **Работа с данными** (Data Tools) щелкнуть по значку **Текст по столбцам** (Text To Columns), чтобы открыть **Мастер распределения текста по столбцам** (Convert Text To Columns Wizard, см. указания, как им пользоваться, ниже) и поместить каждое слово из названия фильма и количество копий в отдельные столбцы. Затем при помощи функции **СЧЁТЗ** (COUNTA) подсчитать число слов в названии, включая число копий как слово, для каждого фильма. Затем при помощи функции **СМЕЩ** определить местонахождение числа копий.

Сначала вставьте достаточное количество столбцов справа от данных для извлечения каждого слова названия фильма и числа копий в отдельные столбцы. Я использовал шесть столбцов (для фильма **Raiders of the Lost Ark**), добавив столбцы **D-I**, как показано на рис. 22.5. Затем я выделил диапазон **C2:C12** и щелкнул на вкладке **Данные** (Data) кнопку **Текст по столбцам** (Text To Columns). В диалоговом окне **Мастер распределения текста по столбцам** (Convert Text To Columns Wizard) я выбрал формат **с разделителями** (Delimited), затем нажал **Далее** (чтобы перейти к шагу 2 в Мастере), снял флажок **знак табуляции** и указал в качестве символа разделителя **пробел**. Затем я нажимаю **Далее**, указываю в поле **Поместить в целевую ячейку** **D2** и нажимаю **Готово**. Результат отображен на рис. 22.5.

Затем я сосчитал количество слов (включая цифры) в каждой ячейке с полным названием фильма (пусть число копий тоже считается словом), скопировав формулу **=СЧЁТЗ(D2:I2)** из **A2** в **A3:A12**.

Наконец, для определения количества копий каждого фильма на складе я скопировал формулу **=СМЕЩ(C2;0;A2)** из **B2** в **B3:B12**. По этой формуле от опорной ячейки с названием фильма происходит смещение на число столбцов, равное числу слов в ячейке с названием. Поскольку в этой функции **СМЕЩ** для аргумента **ссылка** указана только одна ячейка, а не диапазон, аргументы **высота** и **ширина** можно опустить. Таким образом, функция возвращает только ячейку, содержащую последнюю «слово» в ячейке с названием, а именно число копий.

Кроме того, как показано на рис. 22.6 и на листе **Flash Fill**, для извлечения числа копий каждого фильма можно применить инструмент **Мгновенное заполнение** (Flash Fill). Я ввел **40** в ячейку **D2** и **12** в ячейку **D3**; нажал **Enter**, а затем **Ctrl+E** — и сразу появляется число копий других фильмов. Обратите внимание, что если вы захотите попробовать применить то же самое к листу **Flash Fill**, то придется удалить данные в ячейках **D4:D12**, чтобы не получить ошибку.

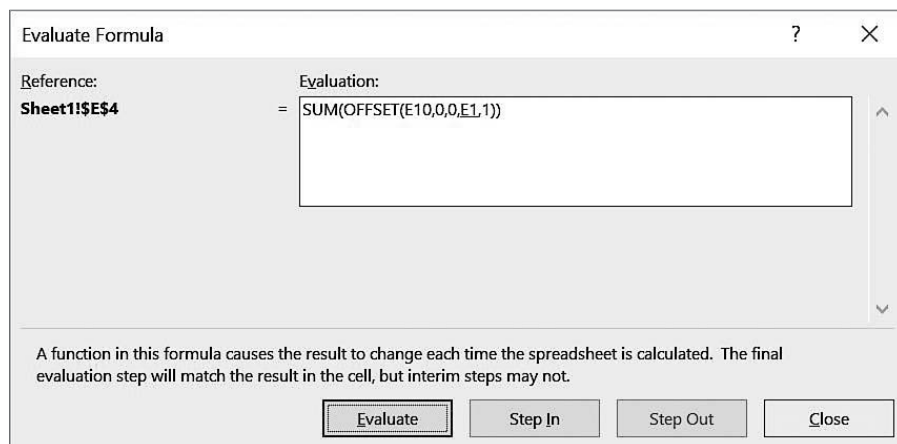


	C	D
1	Movie and Number of Copies	Number
2	Seabiscuit 40	40
3	Laura Croft Tombraider 12	12
4	Raiders of the Lost Ark 36	36
5	Annie Hall 5	5
6	Manhattan 4	4
7	Star Wars 112	112
8	How to Deal 128	128
9	The Matrix Reloaded 1	1
10	Johnny English 1040	1040
11	Rosemary's Baby 12	12
12	High Noon 1002	1002

**Рис. 22.6.** Извлечение числа копий на складе с помощью инструмента Мгновенное заполнение

### ❓ Как работает в Excel инструмент Вычислить формулу?

Если выделить в ячейке любую часть формулы и затем нажать F9, Excel отобразит значение, вычисленное по этой части формулы. Если вы так сделали, обязательно нажмите Esc, иначе потеряете формулу. Этот прием упрощает отлаживание и восприятие сложных формул. Таким образом, если применить этот прием к любой из формул, описанных в данной главе, возможно, будет проще понять, что происходит в части формулы с функцией СМЕЩ. Например, в файле Offsetcost.xlsx в ячейке E4 генерируются общие затраты на этапе 1 по формуле `=СУММ(СМЕЩ(E10;0;0;E1;1))`. Если выделить часть формулы `СМЕЩ(E10;0;0;E1;1)` и нажать F9, в строке формул появится `=СУММ({135;120;58})`, следовательно, в части формулы с функцией СМЕЩ в ячейке D4 используются правильные ячейки (D10:D12). Удостоверьтесь, что нажали Esc, иначе формула будет изменена!



**Рис. 22.7.** Диалоговое окно Вычисление формулы



Еще один способ увидеть, как работает сложная формула, — воспользоваться инструментом **Вычислить формулу** (Evaluate Formula). Встаньте в ячейку E4 и откройте вкладку **Формулы** (Formulas). В группе **Зависимости формул** (Formula Auditing) выберите **Вычислить формулу** (Evaluate Formula). Откроется диалоговое окно **Вычисление формулы** (Evaluate Formula) — рис. 22.7. Нажимайте кнопку **Вычислить** (Evaluate), и формула будет вычисляться пошагово, пока не появится окончательный результат. После второго нажатия формула будет выглядеть как `=СУММ($E$10:E$12)`, что позволяет удостовериться в правильности выбора в ячейке E4 ячеек с данными для этапа 1 разработки препарата 2. После третьего нажатия вы увидите конечный результат — 313.

### ? Как написать формулу, всегда возвращающую последнее число в столбце?

Если часто приходится загружать новые данные на лист, сможете ли вы написать формулу, которая всегда будет возвращать продажи за последний месяц? (См. файл *Mostrecent.xlsx* и рис. 22.8.)

	B	C	D
3			Most recent
4			110
5			=OFFSET(B6,COUNT(B:B),0,1,1)
6	Sales		
7	20		
8	3		
9	40		
10	50		
11	60		
12	90		
13	110		

**Рис. 22.8.** Поиск последнего числа в столбце

Просто введите в ячейку D4 формулу `=СМЕЩ(В6;СЧЁТ(В:В);0;1;1)`.

Формула начнет работать с ячейки B6 и будет двигаться вниз на число строк, равное количеству числовых значений в столбце B. Она приведет вас к самому последнему месяцу, ячейке B13, которая выбирается, поскольку указанные значения для аргументов (1, 1) означают возврат одной ячейки. Ячейка D4 возвращает значение ячейки B13 — 110.

### ? Как создать диапазон, автоматически включающий новые данные?

Пользователи Excel часто добавляют строки или столбцы данных в диапазон данных, используемый для создания сводных таблиц или для выполнения другого типа анализа. Как правило, они просто обновляют диапазон ячеек, на который ссылается формула, и затем перезапускают анализ. Однако вы можете использовать имена динамических диапазонов, и тогда обновлять диапазон данных, на который есть ссылка в формуле или сводной таблице, не придется. Такой диапазон обновится автоматически. Вот вам пример.

В файле `Dynamicrange.xlsx` находятся записи из базы данных отдела кадров (рис. 22.9).

	A	B	C	D	E	F	G
1	<b>Name</b>	<b>Salary</b>	<b>Exp</b>	<b>Gender</b>			
2	John	35500	3	M			
3	Jack	42300	4	M			
4	Jill	53426	5	F			
5	Erica	56000	6	F			
6	JR	62000	8	M			
7	Bianca	49000	10	F			
8	Francis	52000	5	M			
9	Roger	56000	7	M			
10	Maggie	42000	4				
11							448278

**Рис. 22.9.** Пример динамического диапазона

Мы создали динамический диапазон, самопроизвольно расширяющийся по мере добавления строк и столбцов данных к диапазону `A10:D10`. На этот момент данные содержатся в девяти строках и четырех столбцах. Неплохо было бы создать диапазон, автоматически включающий новые строки и столбцы при добавлении новых сотрудников или полей данных в базу данных.

Для создания динамического диапазона откройте вкладку **Формулы (Formulas)** и в группе **Определенные имена (Defined Names)** выберите инструмент **Диспетчер имен (Name Manager)**. Нажмите кнопку **Создать (New)** и определите диапазон (рис. 22.10).

**New Name** ? X

**Name:** data

**Scope:** Workbook ▼

**Comment:**

**Refers to:** =offset(data!\$A\$1,0,0,counta(\$a:\$a),counta(\$1:\$1)) ⬆

OK Cancel

**Рис. 22.10.** Создание динамического диапазона

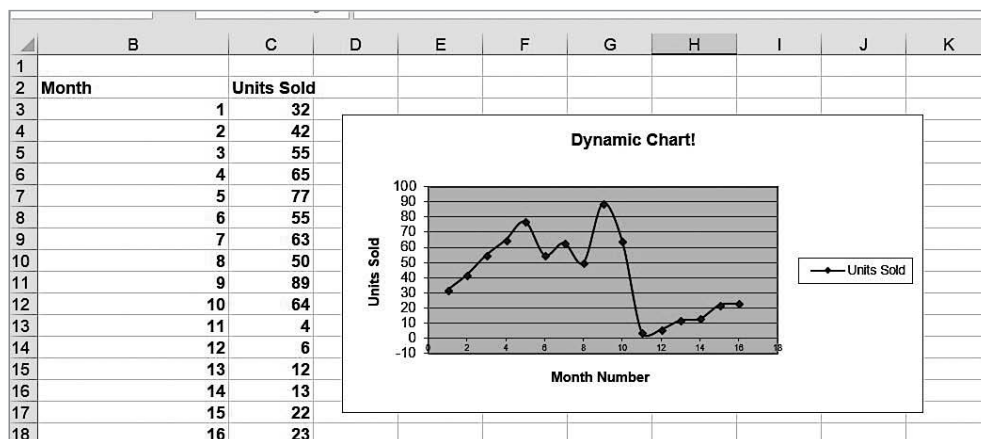
Вы можете выбрать имя диапазона **data**, которое я уже создал. В нем определена формула `=СМЕЩ(data!$A$1,0,0,СЧЕТЗ(data!$A:$A),СЧЕТЗ(data!$1:$1))`. Диапазон начинается в левом верхнем углу: ячейка **A1**. Поскольку ячейка **A1** должна быть левым верхним углом именованного диапазона, укажите для аргументов **смещ\_по\_строкам** и **смещ\_по\_столбцам** значение **0**. В выделенном диапазоне число строк равно числу непустых записей в столбце **A**, а число столбцов — числу непустых записей в строке **1**. Таким образом, при добавлении сотрудников или полей данных формула будет охватывать их автоматически. Знаки доллара (\$) необходимы, поскольку в этом случае определяемый диапазон не сдвинется при перемещении по листу.

Для проверки введите в ячейку **G11** формулу `=СУММ(data)`. К настоящему моменту по этой формуле все числа в диапазоне **A1:D10** дают в сумме \$448 278.

Теперь добавьте имя **Meredith** в строку **11**, введите заработную плату \$10 000 в ячейку **B11**, введите переменную для ошибок (добавьте слово **Ошибки**) в ячейку **E1** и введите **1000** в ячейку **E11**. Формула `=СУММ(data)` теперь охватывает значения **10 000** и **1000**, и значение в ячейке автоматически обновляется до \$459 278 долларов.

**?** Я ежемесячно составляю диаграммы по объемам продаж в штуках для продукта компании. Каждый месяц я загружаю новые данные. Хорошо бы диаграммы обновлялись автоматически. Есть ли простой способ добиться этого?

Книга **Chartdynamicrange.xlsx** (рис. 22.11) содержит данные об объемах продаж продукта компании в штуках. Как видно из рисунка, по этим данным составляется диаграмма разброса (XY-диаграмма).

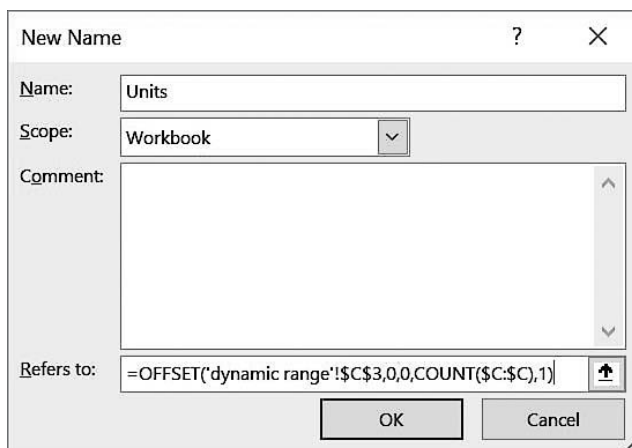


**Рис. 22.11.** Динамическое обновление диаграммы с помощью функции СМЕЩ

Новые данные о продажах загружаются начиная со строки **19**. Есть ли простой способ убедиться в автоматическом включении в диаграмму новых данных?

Для этого используйте функцию СМЕЩ для создания имен динамических диапазонов для столбцов Месяц (Month) и Объем продаж в шт (Units sold). При вводе новых данных динамический диапазон для объема продаж будет автоматически включать все данные о продажах, а динамический диапазон для месяцев будет содержать номер каждого месяца. После создания этих диапазонов можно изменить диаграмму, заменив используемые в диаграмме диапазоны данных новыми динамическими диапазонами. Тогда при вводе новых данных диаграмма также будет обновляться.

Сначала на вкладке Формулы (Formulas) в группе Определенные имена (Defined Names) выберите Присвоить имя (Define Name) для открытия диалогового окна Создание имени (New Name). Создайте диапазон с именем Объем\_продаж, заполнив поля, как показано на рис. 22.12.



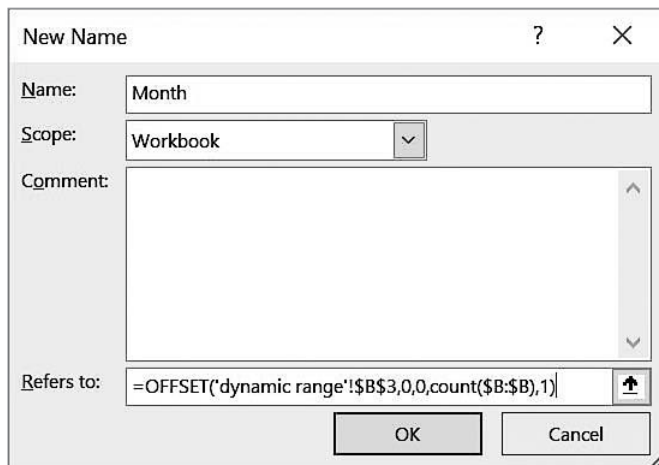
**Рис. 22.12.** Создание динамического диапазона для объема продаж

В поле Диапазон (Refers to) введите =СМЕЩ('лист1'!\$C\$3;0;0;СЧЁТ(!\$C:\$C);1). Будет создан диапазон шириной в один столбец с началом в ячейке С3, которая содержит первую точку данных об объемах продаж. В диапазоне будет столько чисел, сколько в столбце С, поскольку он порожден частью формулы СЧЁТ('лист1'!\$C:\$C). Введенные в столбец С новые данные будут автоматически включены в диапазон Объем\_продаж.

Затем для месяцев, введенных в столбец В, создайте динамический диапазон с именем Месяц. Формула представлена на рис. 22.13.

Теперь нажмите любую точку на диаграмме. В строке формул появится формула =РЯД(лист1!\$C\$2;лист1!\$B\$3:\$B\$18;лист1!\$C\$3:\$C\$18;1) с функцией РЯД (SERIES). Таким способом в Excel извлекаются данные, необходимые для построения диаграммы. Замените диапазоны \$B\$3:\$B\$18 и \$C\$3:\$C\$18 именами динамических диапазонов следующим образом:

=РЯД(лист1!\$C\$2;Chartdynamicrange.xlsx!Месяц;  
Chartdynamicrange.xlsx!Объем\_продаж;1)



The image shows the 'New Name' dialog box in Microsoft Excel. It has a title bar with a question mark and a close button. Inside, there are four main sections: 'Name' with a text box containing 'Month'; 'Scope' with a dropdown menu showing 'Workbook'; 'Comment' with a large empty text area and a vertical scrollbar; and 'Refers to' with a text box containing the formula '=OFFSET('dynamic range'!\$B\$3,0,0,count(\$B:\$B),1)' and a small icon to the right. At the bottom right are 'OK' and 'Cancel' buttons.

**Рис. 22.13.** Формула для создания динамического диапазона с именем Месяц

Этот метод неприменим, если между любыми старыми и новыми данными есть пробел. Введите какие-либо новые данные и посмотрите, как они будут включены в диаграмму.

## Замечания

В Excel автоматическое включение новых данных в диаграммы и формулы обеспечивает инструмент **Таблица (Table)**. Этот инструмент рассмотрен в главе 26 «Таблицы».

## Задания

1. В файле C21p1.xlsx содержатся данные об объемах продаж в штуках для 11 продуктов с 1999 по 2003 г. Напишите формулу с функциями **ПОИСКПОЗ** и **СМЕЩ**, вычисляющую продажи заданного продукта в заданном году. Как решить эту задачу без использования функций **ПОИСКПОЗ** и **СМЕЩ**?
2. Широко используемое правило торговли по скользящим средним звучит так: покупай акцию, когда ее цена превысит средний показатель за последние **D** месяцев, и продавай ее, когда цена упадет ниже среднего показателя за последние **D** месяцев. В главе 12 я показал, что для **D = 15** это торговое правило значительно превзошло по эффективности долгосрочное инвестирование согласно биржевому индексу Standard&Poog's 500. Объединив однонаправлен-

ную таблицу данных с функцией **СМЕЩ**, определите значение **D**, максимизирующее торговую прибыль (без учета транзакционных издержек). Нужные данные находятся в файле **Matradingrule.xlsx**.

3. Широко используемое правило торговли по скользящим средним звучит так: покупай акцию, когда ее цена превысит средний показатель за последние **B** месяцев, и продавай ее, когда цена упадет ниже среднего показателя за последние **S** месяцев. В главе 13 я показал, что для  $B = S = 15$  это торговое правило значительно превзошло по эффективности долгосрочное инвестирование согласно биржевому индексу Standard&Poor's 500. Объединив двунаправленную таблицу данных с функцией **СМЕЩ**, определите значения **B** и **S**, максимизирующие торговую прибыль (без учета операционных издержек). Нужные данные находятся в файле **Matradingrule.xlsx**.
4. В файле **Lagged.xlsx** содержатся данные о количестве рекламных объявлений в журнале, размещенных вербовочным пунктом армии США в течение 60 идущих подряд месяцев. Для каждого месяца количество рекламных объявлений с  $k$ -месячной задержкой определяется как количество объявлений, размещенных  $k$  месяцев назад. Для месяцев с 7-го по 60-й необходимо вычислить значения для количества объявлений с задержкой 1–6 месяцев. Для эффективного вычисления этих значений используйте функцию **СМЕЩ**.
5. В файле **Verizondata.xlsx** хранятся данные о продажах телефонов Verizon в пяти регионах. Найдите эффективный способ ввода в одной строке региона, типа телефона и объемов продаж каждого типа телефона для каждой из 20 комбинаций «регион — продукт».
6. Это сложная задача. В файле **Agingdata.xlsx** находятся данные о прогнозируемом количестве заявлений о страховом случае, которые могут поступать ежедневно, и о количестве свободных сотрудников страховой компании. За один день сотрудник может обработать до 30 заявлений. В первую очередь сотрудники обрабатывают самые старые заявления в системе. В файле **s22\_6.xlsx** из папки **Solution Files** ячейки **H6:AL6** содержат количество заявлений, находящихся в системе на 1 января, до поступления новых заявлений. Создайте лист для отслеживания сроков давности заявлений. Сколько 1-дневных, 2-дневных, ..., 30-дневных и более давних заявлений находится в системе?
7. Каждая строка в файле **DVDsales.xlsx** содержит ежемесячные продажи **D VD**. Напишите формулу, вычисляющую объем продаж для каждого **D VD** в течение первых шести месяцев после выхода на рынок.
8. Для расчета гандикапа игрока в гольф берется среднее значение 10 самых слабых раундов из последних 20 раундов. Затем из среднего значения очков вычитается 80, и результат округляется до ближайшего целого числа. Таким образом, если 10 самых слабых раундов из последних 20 принесли 864 очка, гандикап будет равен 6. В файле **Golfddata.xlsx** содержатся набранные игроком в гольф очки. Начиная со строки 24 вычислите гандикап игрока в гольф после каждого раунда. Предполагается, что если десятый лучший результат среди последних 20 раундов встречается несколько раз, все раунды с этим числом

очков включаются в расчет гандикапа. Отметим также, что в Excel функция ОКРУГЛ (ROUND) округляет  $x$  до ближайшего целого числа ( $=\text{ОКРУГЛ}(x;0)$ ).

9. Каждая строка файла **Carsumdata.xlsx** содержит данные о продажах продукта (автомобиля, поезда или самолета) с января по июль. Вы вводите на листе месяц и продукт. Напишите формулу, вычисляющую общий объем продаж этого продукта в заданном месяце.
10. Файл **Verizon.xlsx** содержит ежемесячные отчетные данные по операциям на складе Verizon. Используя функцию СМЕЩ, извлеките все результаты операций за январь в один столбец, за февраль в другой столбец и т. д.
11. В файле **Casesensitive.xlsx** содержатся коды и цены продуктов. Обратите внимание, что при обработке кодов продуктов необходимо учитывать регистр. Например, DAG32 и dag32 — это разные продукты. Напишите формулу, предоставляющую цену продукта по его коду. Подсказка: возможно, потребуется функция СОВПАД (EXACT). Формула  $=\text{СОВПАД}(\text{текст1};\text{текст2})$  возвращает значение ИСТИНА, если текст1 и текст2 совпадают. Функция СОВПАД различает прописные и строчные буквы.
12. Файл **Reversed.xlsx** содержит столбец чисел. С помощью функции СМЕЩ запишите числа в обратном порядке: последнее первым, предпоследнее вторым и т. д.
13. В файле **Diagonal.xlsx** находится матрица чисел. Определите, как поместить диагональные элементы матрицы в один столбец.
14. В файле **Yeartodate.xlsx** находятся данные о ежемесячных продажах компании за период с января 2008 г по август 2013 г. Напишите формулу, вычисляющую совокупные продажи за год на данное число для заданного года и месяца. Например, если ввести 6 для месяца и 2010 для года, то по формуле должен быть вычислен общий объем продаж за январь — июнь 2010 г.
15. Покажите, как можно создать диаграмму ежемесячных продаж, на которой всегда отображаются продажи за последние шесть месяцев.
16. В файле **Transactiondata.xlsx** содержатся данные о торговых сделках подразделений A, B, C, D и E фармацевтической компании. С помощью функции СМЕЩ перенесите объемы продаж каждого подразделения в одну строку.

## ГЛАВА 23

# Функция ДВССЫЛ

### Обсуждаемые вопросы

- Мои формулы на листе часто содержат ссылки на ячейки, или диапазоны ячеек, или на то и другое. Вместо замены ссылок в формулах можно ли поместить эти ссылки в собственные ячейки и быстро заменять в них ссылки без изменения базовых формул?
- На каждом листе книги в ячейке D1 записываются ежемесячные продажи продукта. Можно ли написать и скопировать формулу для записи продаж продукта за каждый месяц на одном листе?
- Предположим, что я суммирую значения в диапазоне A5:A10 по формуле =СУММ(A5:A10). Если я вставляю пустую строку где-либо между строками 5 и 10, моя формула будет автоматически обновлена до =СУММ(A5:A11). Каким образом следует составить формулу, которая при вставке пустой строки между строками 5 и 10 по-прежнему суммировала бы стоимости в исходном диапазоне A5:A10?
- Как с помощью функции ДВССЫЛ обеспечить интерпретацию части формулы в качестве имени диапазона?
- Книга содержит продажи по каждому продукту компании в разных странах и регионах, и каждый континент отображен на отдельном листе. Как объединить эти данные на одном листе?
- Как составить список всех листов в книге?
- Моя книга содержит множество листов. Можно ли легко составить оглавление с гиперссылками так, чтобы была связь со всеми листами в книге?

Функция ДВССЫЛ (INDIRECT) — вероятно, одна из самых сложных для понимания функций Microsoft Excel. Однако если вы научитесь ею пользоваться, то сможете решать многие задачи, которые на первый взгляд кажутся неразрешимыми. В целом любая ссылка на ячейку в формуле с ДВССЫЛ обеспечивает немедленное вычисление ссылки на ячейку и приравнивание содержимого ячейки. Пример использования ДВССЫЛ есть в файле Indirectsimpleex.xlsx (рис. 23.1).

В ячейку C4 я ввел формулу =ДВССЫЛ(A4). Excel возвращает значение 6, поскольку ссылка на ячейку A4 незамедлительно заменяется текстовой строкой "B4", и формула вычисляется как =B4, что дает в результате значение 6. Аналогично, после ввода формулы =ДВССЫЛ(A5) в ячейку C5 будет возвращено значение в ячейке B5, равное 9.



	A	B	C
3		Value	Indirect Reference
4	B4	6	6
5	B5	9	9

**Рис. 23.1.** Простой пример с функцией ДВССЫЛ

## Ответы на вопросы

❓ Мои формулы на листе часто содержат ссылки на ячейки, или диапазоны ячеек, или на то и другое. Вместо замены ссылок в формулах можно ли поместить эти ссылки в собственные ячейки и быстро заменять в них ссылки без изменения базовых формул?

Данные для этого примера находятся в файле Sumindirect.xlsx (рис. 23.2). В диапазоне ячеек B4:H16 хранятся данные о ежемесячных продажах шести продуктов за 12 месяцев.

	B	C	D	E	F	G	H
1			Lower	Upper			
2			6	16			
3		C	D	E	F	G	H
4	Month	Prod 1	Prod 2	Prod 3	Prod 4	Prod 5	Prod 6
5	1	28	86	79	31	84	58
6	2	38	7	61	1	20	2
7	3	91	48	73	8	80	14
8	4	33	32	24	77	29	80
9	5	82	70	41	29	57	90
10	6	75	40	15	92	55	91
11	7	52	21	26	45	59	21
12	8	19	6	35	67	40	81
13	9	11	18	68	11	52	78
14	10	90	30	52	32	30	1
15	11	47	86	46	0	38	55
16	12	69	71	75	65	53	52
17							
18	Total	607	429	516	427	513	565
19							
20		=SUM(INDIRECT(C\$3&D\$2&"."&C\$3&E\$2))					

**Рис. 23.2.** Функция ДВССЫЛ позволяет изменить в формулах ссылки на ячейки без изменения самих формул

Я вычисляю общий объем продаж каждого продукта за месяцы 2–12. Самый простой способ это сделать — скопировать формулу =СУММ(C6:C16) из C18 в D18:H18. Предположим, однако, что вам нужно изменить месяцы, по которым подводится итог. Например, вы хотите узнать объем продаж за месяцы 3–12. Можно было бы

заменить формулу в C18 на =СУММ(C7:C16) и затем скопировать ее в D18:H18, но этот подход проблематичен, поскольку формула из C18 должна быть скопирована в D18:H18, и без просмотра формулы невозможно будет определить, какие строки суммировались.

Функция ДВССЫЛ позволяет реализовать другой подход. Начальную и конечную строки суммирования я указал в ячейках D2 и E2. Затем, благодаря функции ДВССЫЛ, мне останется только изменять ссылки на начальную и конечную строки в ячейках D2 и E2. Суммы после обновления будут включать требуемые строки. Кроме того, значения в ячейках D2 и E2 непосредственно указывают на суммируемые строки (месяцы). Необходимо только скопировать формулу =СУММ(ДВССЫЛ(C\$3&\$D\$2&"&\$C\$3&\$E\$2)) из C18 в D18:H18.

Если вы хотите увидеть, как Excel вычисляет ссылки в функции ДВССЫЛ, воспользуйтесь следующим приемом. Выделите часть формулы (например, C\$3) и нажмите F9. Отобразится значение, вычисленное по выделенной части формулы. Например, C\$3 дает в результате значение C. Обязательно нажмите Esc для отмены изменения. Каждой ссылке на ячейку в части формулы с функцией ДВССЫЛ приравнивается содержимое ячейки. C\$3 вычисляется как C, \$D\$2 вычисляется как 6, и \$E\$2 вычисляется как 16. С учетом амперсанда (&), включенного в качестве символа конкатенации, в результате будет получена формула =СУММ(C6:C16) — как раз то, что вам нужно. Формула в ячейке C18 возвращает значение 38 + 91 + ... + 69 = 607. В ячейке D18 формула также приобретает требуемый вид =СУММ(D6:D16). Далее, если необходимо суммировать продажи за месяцы 4–6, просто введите 8 в ячейку D2 и 10 в ячейку E2, и по формуле в ячейке C18 будет получен результат 33 + 82 + 75 = 190. (Об использовании амперсанда при объединении значений см. в главе 6 «Текстовые функции и инструмент Мгновенное заполнение».)

**❓ На каждом листе книги в ячейке D1 записываются ежемесячные продажи продукта. Можно ли написать и скопировать формулу для записи продаж продукта за каждый месяц на одном листе?**

В файле Indirectmultisheet.xlsx (рис. 23.3) находятся семь листов. На каждом листе в ячейке D1 хранятся данные о продажах продукта за конкретный месяц. Пусть Лист1 содержит данные за первый месяц, Лист2 — за второй месяц и т. д. Например, объем продаж за первый месяц — 1, за второй месяц — 4 и т. д.

Предположим теперь, что вам нужно составить список продаж для всех месяцев на одном листе. Было бы утомительно вносить продажи в список для первого месяца по формуле =Лист1!D1, для второго месяца по формуле =Лист2!D1 и т. д. до седьмого месяца и формулы =Лист7!D1. При наличии данных за 100 месяцев этот подход требует много времени и чреват ошибками. Гораздо изящнее записать продажи за месяц 1 в ячейку E10 листа Лист1 с помощью формулы =ДВССЫЛ(\$C\$10&D10&"!D1"). В Excel \$C\$10 обрабатывается как "Лист", D10 как "1" и "!D1" как текстовая строка "!D1". Вся формула обрабатывается как =Лист1!D1

и дает в результате объем продаж за месяц 1 из ячейки D1 листа Лист1. Копирование этой формулы в ячейки E11:E16 переносит записи из ячеек D1 на листах Лист2–Лист7. При копировании формулы из E10 в E11 ссылка на D10 заменяется на D11, и ячейка E11 возвращает значение в ячейке Лист2!D1. И если вы поместите курсор в ячейку, скажем, E12 и выберете **Вычислить формулу**, то увидите, как Excel наращивает формулу до Лист3!D1.

	C	D	E
1		1	
2			
3			
4			
5			
6			
7			
8			
9		Sheet#	Cell D1 entry
10	Sheet	1	1
11		2	4
12		3	0
13		4	12
14		5	15
15		6	3
16		7	4

**Рис. 23.3.** Список ежемесячных продаж продукта (за месяцы 1–7), составленный с помощью функции ДВССЫЛ

❓ Предположим, что я суммирую значения в диапазоне A5:A10 по формуле **=СУММ(A5:A10)**. Если я вставлю пустую строку где-либо между строками 5 и 10, моя формула будет автоматически обновлена до **=СУММ(A5:A11)**. Каким образом следует составить формулу, которая при вставке пустой строки между строками 5 и 10 по-прежнему суммировала бы стоимости в исходном диапазоне A5:A10?

В файле Indirectinsertrow.xlsx на листе SUM(A5A10), представленном на рис. 23.4, показано несколько способов суммирования чисел в диапазоне A5:A10. (Обратите внимание, что лист Row\_inserted может показаться похожим на этот листно на деле сильно отличается.) В ячейку A12 я ввел традиционную формулу **=СУММ(A5:A10)**, дающую в результате  $6 + 7 + 8 + 9 + 1 + 2 = 33$ .

Аналогично по формуле **=СУММ(\$A\$5:\$A\$10)** в ячейке E9 вычисляется результат 33. Как вы скоро увидите, если вставить пустую строку между строками 5 и 10, обе формулы попытаются суммировать значения в диапазоне A5:A11.

С функцией ДВССЫЛ у вас есть минимум два способа суммировать значения в диапазоне A5:A10. В ячейку F9 я ввел формулу **=СУММ(ДВССЫЛ("A5:A10"))**. Поскольку

в Excel ДВССЫЛ("A5:A10") обрабатывается как текстовая строка "A5:A10", при вставке пустой строки записи по-прежнему будут суммироваться в диапазоне A5:A10.

	A	B	C	D	E	F	G
1							
2							
3			Begin	End			
4			5	10			
5	6		Indirect Way				
6	7		33				
7	8						
8	9				Absolute reference	Indirect Reference	Insert a blank row and see what happens?
9	1				33	33	
10	2						
11	Old way				=SUM(SAS5:SAS10)	=SUM(INDIRECT("A5:A10"))	

**Рис. 23.4.** Несколько способов суммирования значений в диапазоне ячеек A5:A10

Другой способ суммирования значений в диапазоне A5:A10 с функцией ДВССЫЛ — это формула =СУММ(ДВССЫЛ("A"&C4&":A"&D4)), введенная в ячейку C6. Ссылка на C4 обрабатывается как 5, а ссылка на D4 — как 10. Таким образом, формула принимает вид =СУММ(A5:A10). Вставка пустой строки между строками 5 и 10 не оказывает влияния на вычисления по формуле, так как ссылка на C4 по-прежнему обрабатывается как 5, а ссылка на D4 — как 10. На рис. 23.5 показаны суммы, вычисленные по четырем формулам после вставки пустой строки под строкой 7. Эти данные можно найти в файле Indirectinsertrow.xlsx на листе Row\_inserted.

	A	B	C	D	E	F
1						
2						
3			Begin	End		
4			5	10		
5	6		Indirect Way			
6	7		31			
7	8					
8						
9	9				Absolute reference	Indirect Reference
10	1				33	31
11	2					
12	Old way				=SUM(SAS5:SAS11)	=SUM(INDIRECT("A5:A10"))
13	33					

**Рис. 23.5.** Результат вычислений по формулам суммирования после вставки пустой строки в исходный диапазон

Как видите, классические формулы суммирования без функции **ДВССЫЛ** изменились для суммирования значений в диапазоне **A5:A11**, так что по ним по-прежнему вычисляется значение **33**. Две формулы суммирования с функцией **ДВССЫЛ** продолжают складывать значения в диапазоне **A5:A10**, и поэтому значение **2** (теперь в ячейке **A11**) больше не участвует в вычислении суммы. Формулы суммирования с функцией **ДВССЫЛ** дают в результате значение **31**.

### ❓ Как с помощью функции **ДВССЫЛ** обеспечить интерпретацию части формулы в качестве имени диапазона?

Допустим, вы создали на листе несколько имен диапазонов в соответствии с квартальными продажами продукта. (См. рис. 23.6 и файл **Indirectrange.xlsx**.) Например, диапазон **D4:E6** (с именем **Квартал1**) содержит фиктивные продажи различных продуктов Microsoft в первом квартале.

	D	E	F	G	H	I	J
3	<b>Quarter1</b>						
4	<b>Office</b>	63					
5	<b>Windows</b>	66					
6	<b>Xbox</b>	70					
7							
8	<b>Quarter2</b>						
9	<b>Office</b>	93					
10	<b>Windows</b>	90					
11	<b>Xbox</b>	99					
12							
13	<b>Quarter3</b>						
14	<b>Office</b>	77					
15	<b>Windows</b>	58					
16	<b>Xbox</b>	60			<b>Office</b>	<b>Windows</b>	<b>Xbox</b>
17				<b>Quarter1</b>	63	66	70
18	<b>Quarter4</b>			<b>Quarter2</b>	93	90	99
19	<b>Office</b>	97		<b>Quarter3</b>	77	58	60
20	<b>Windows</b>	56		<b>Quarter4</b>	97	56	95

**Рис. 23.6.** Создание ссылки на имя диапазона в формуле при помощи функции **ДВССЫЛ**

Было бы здорово написать легко копируемую формулу и получить на листе объем продаж для каждого продукта в каждом квартале в одном прямоугольном диапазоне, как показано в диапазоне **H17:J20**. Казалось бы, для этого следует ввести формулу **=ВПР(H\$16;\$G17;2;ЛОЖЬ)** в ячейку **H17** и затем скопировать ее в диапазон **H17:J20**. К сожалению, в Excel **\$G17** не распознается как ссылка на имя **Квартал1**, а обрабатывается как текстовая строка **"Квартал1"**, и при вычислении формулы произойдет ошибка **#Н/Д**. Для исправления этой ошибки просто введите формулу **=ВПР(H\$16; ДВССЫЛ(\$G17);2;ЛОЖЬ)** в ячейку **H17** и затем скопируйте ее в диапазон **H17:J20**. Этот прием работает отлично! **ДВССЫЛ(\$G17)** вычисляется как **Квартал1**.

и распознается как имя диапазона. Теперь нетрудно записать данные по продажам для всех продуктов во всех четырех кварталах.

**❓ Книга содержит продажи по каждому продукту компании в разных странах и регионах, и каждый континент отображен на отдельном листе. Как объединить эти данные на одном листе?**

В ячейках E7:E9 файла Indirectconsolidate.xlsx содержатся объемы продаж автомобилей, грузовиков и самолетов по разным континентам. Для каждого континента выделен собственный лист. Как суммировать эти данные на одном листе?

На листе Summary я создал сводку, представленную на рис. 23.7.

	C	D	E	F	G	H	I
5			=LA!E7				
6			LA	NA	Asia	Africa	Europe
7	E7	Cars	100	200	500	100	80
8	E8	Trucks	150	150	400	50	120
9	E9	Planes	200	100	200	25	100

**Рис. 23.7.** Создание сводки по продажам продукции с помощью функции ДВССЫЛ

Я скопировал формулу =ДВССЫЛ(E\$6&"!"&\$C7) из ячейки E7 в диапазон E7:I9. По этой формуле создается формула =LA!E7, вычисляющая продажи автомобилей в Латинской Америке. Копирование формулы позволяет вычислить продажи каждого типа продукции для каждого континента.

На листе Another Summary (рис. 23.8) я создал такую же таблицу несколькими иным способом. Я автоматизировал создание адресов ячеек для извлечения продаж автомобилей, грузовиков и самолетов, скопировав формулу =АДРЕС(СТРОКА();СТОЛБЕЦ()+2) из C7 в C8:C9. При этом в адресе ячейки создается столбец E и строки 7, 8 и 9. Функция СТРОКА(), введенная в ячейку, возвращает номер строки этой ячейки, а функция СТОЛБЕЦ() возвращает номер столбца текущей ячейки (в данном случае 3). При вводе в качестве третьего аргумента функции АДРЕС выражение СТОЛБЕЦ()+2 возвращает имя столбца для ячейки на два столбца правее текущей ячейки (в данном случае столбец E). Затем я скопировал формулу =ДВССЫЛ(E\$6&"!"&\$C7) из ячейки E7 в диапазон E7:I9.

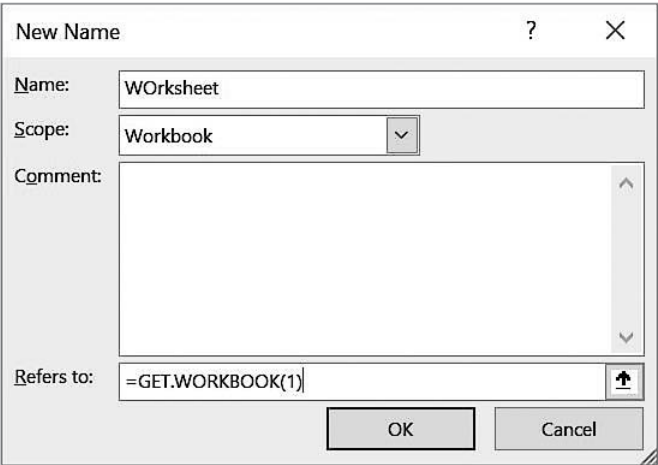
	C	D	E	F	G	H	I
5			=LA!E7				
6			LA	NA	Asia	Africa	Europe
7	\$E\$7	Cars	100	200	500	100	80
8	\$E\$8	Trucks	150	150	400	50	120
9	\$E\$9	Planes	200	100	200	25	100

**Рис. 23.8.** Создание сводки по всей книге с использованием копирования адресов ячеек

**? Как составить список всех листов в книге?**

В предыдущем примере было нетрудно перечислить имена всех листов. Теперь допустим, у нас 100 листов или больше. Для раскрытия полного потенциала функции ДВССЫЛ по извлечению данных из множества листов необходим эффективный метод составления списка имен листов в книге. Для перечисления имен всех листов в книге надо выполнить следующие действия. (См. книгу Worksheetnames.xlsm и рис. 23.9 и 23.10.)

- На вкладке Формулы (Formulas) в группе Определенные имена (Defined Names) выберите Присвоить имя (Define Name); в диалоговом окне Создание имени (New Name) создайте диапазон имен и назовите его РабочийЛист с помощью формулы =ПОЛУЧИТЬ.РАБОЧУЮ.КНИГУ(1) (рис. 23.9). ПОЛУЧИТЬ.РАБОЧУЮ.КНИГУ(1) — это макрос (макросы рассматриваются в главе 92, «Запись макросов»), поэтому у файла должно быть расширение xlsm.



**Рис. 23.9.** Создание имени диапазона для извлечения имен листов

	H	I	J
11			
12	How many worksheets?		
13	3	=SHEETS()	
14	{=worksheet}	{=worksheet}	{=worksheet}
15	[Worksheetnames.xlsm]Bob	[Worksheetnames.xlsm]Allen	[Worksheetnames.xlsm]Jill
16	Bob	Allen	Jill
17			
18			
19	=RIGHT(H15.LEN(H15)-FIND("I",H15,1))		

**Рис. 23.10.** Имена листов в строке 16 и число листов, вычисленное по формуле =ЛИСТЫ()

- Выберите горизонтальный массив, содержащий столько строк, сколько листов в книге; в этом примере только три листа. Точное число листов в большой книге можно узнать с помощью функции, появившейся в Excel 2013, **ЛИСТЫ** (SHEETS). В ячейке **H13** формула **=ЛИСТЫ()** отображает количество листов в книге (в данном случае — **3**). Кстати, функция **ЛИСТ** (SHEET) возвращает номер текущего листа в книге. После ввода формулы **=ЛИСТ()** на листе **Боб** будет получено значение **1**, на листе **Аллен** — значение **2**, на листе **Джил** — значение **3**.
- В левом углу выделенного диапазона (в данном случае в ячейке **H15**) введите **=РабочийЛист** и нажмите комбинацию **Ctrl+Shift+Enter**. Имя диапазона **РабочийЛист** действует как формула обработки массива (о формулах обработки массивов см. главу 91 «Формулы массива и функции, возвращающие массив»), поэтому для работы формулы необходимо нажать комбинацию **Ctrl+Shift+Enter**. В диапазоне ячеек **H15:J15** можно увидеть имя книги, заключенное в квадратные скобки, за которым следуют имена листов по порядку.

Если требуется получить имя листа в отдельной ячейке, просто с помощью функций **ПРАВСИМВ** и **НАЙТИ** из главы 6 извлеките все символы справа от закрывающей квадратной скобки. Скопируйте формулу **=ПРАВСИМВ(I15;ДЛСТР(I15)-НАЙТИ("]",I15;1))** из ячейки **H16** в ячейки **I16:J16**. Теперь имена листов перечислены в диапазоне **H16:J16**.

### ❓ **Моя книга содержит множество листов. Можно ли легко составить оглавление с гиперссылками так, чтобы была связь со всеми листами в книге?**

Трудно быстро передвигаться по книге со множеством листов. Если мы знаем названия листов, то ничего не стоит создать шикарное оглавление с гиперссылками на каждый лист.

Прежде чем показать вам, как создаются оглавления, нужно объяснить, как функция **ДВССЫЛ** работает с листами, имена которых содержат пробел. Книга под названием **Indirect with spaces.xlsx** содержит два листа, в именах которых есть пробел: **Data Table** и **Goal Seek**. Нам нужно создать название листа в ячейке **D9** (рис. 23.11) и, может быть,

	B	C	D	E	F	G	H	I	J
1		5							
2									
3									
4	7								
5									
6			Data Table						
7			Goal Seek						
8			Sheet	Cell					
9			Goal Seek	C1		7	=INDIRECT("'"&D9&'"'"&"!b4")		
10						5	=INDIRECT("'"&D9&'"'"&"!"&E9)		
11					#REF!		=INDIRECT(D9&"!"&E9)		

**Рис. 23.11.** Применение функции **ДВССЫЛ** в случае, когда имена листов содержат пробелы



адрес ячейки в E9; нам также нужно, чтобы формула в F10 возвращала запись в комбинации ячеек на выбранном листе. В ячейке F11 пишем формулу =ДВССЫЛ(D9&"!"&E9). Оказывается, она не работает, потому что в именах листов есть пробелы. В ячейке F10 вводим теперь правильную формулу =ДВССЫЛ("'"&D9&"'"&"!"&E9). Апостроф, помещенный в двойные кавычки, приводит к тому, что формула начинает работать нормально, и в таком случае она возвращает запись (5) из ячейки C1 листа, названного Goal Seek. Формула =ДВССЫЛ("'"&D9&"'"&"!b4") в ячейке F9 всегда будет возвращать запись в ячейке B4 листа, выбранного в ячейке D9.

После рассмотрения функции ЯЧЕЙКА (CELL) мы можем показать, как скомбинировать функции ГИПЕРССЫЛКА, ЯЧЕЙКА и ДВССЫЛ, для того чтобы создать удобное оглавление для книги.

Функция ЯЧЕЙКА (CELL) (см. лист Cellfunction.xlsx и рис. 23.12) возвращает важную информацию о ячейке или диапазоне ячеек. Наиболее важные параметры функции ЯЧЕЙКА представлены на рис. 23.12. Более подробная информация — в «Справке Excel» по функции ЯЧЕЙКА, которую можно посмотреть в Мастере для этой функции.

	C	D
1		
2		
3		5
4		
5	\$C\$3	=CELL("address",C3)
6		3 =CELL("col",C3)
7		5 =CELL("contents",C3)
8	C:\Users\Owner\Documents\beckeroc5\beckeroc17 coolstuff\[Cellfunction.xlsx]Sh	=CELL("filename",C5)
9		3 =CELL("row",C3)
10		70 =CELL("width",C3)

**Рис. 23.12.** Применение функции ЯЧЕЙКА

В файле Hyperlinktoc.xlsx показано, как правильно создать оглавление для рабочей книги (рис. 23.13). Сначала мы применим функцию ПОЛУЧИТЬ.РАБОЧУЮ.КНИГУ(1) для получения имен листов (в нашем случае это будут имена Pivot Table, Goal Seek и Sheet1). Затем просто скопируем следующую формулу из F13 в F14:F15:

=ГИПЕРССЫЛКА("#"&ЯЧЕЙКА("address";ДВССЫЛ("'"&E13&"'"&"\$A\$1")));E13).

Знак «решетки» # позволяет нашей формуле распознать текущую книгу. Функция ЯЧЕЙКА с параметром адреса, примененная к той части формулы, где находится функция ДВССЫЛ, возвращает ячейку A1 с листа Pivot Table. Последняя ячейка E13 позволяет использовать имя каждого листа как читабельный текст для нашей гиперссылки.

	E	F	G	H	I	J
13	Pivot table	Pivot table	=HYPERLINK("#"&CELL("address",INDIRECT(""&E13&"!\$A\$1")),E13)			
14	Goal seek	Goal seek	=HYPERLINK("#"&CELL("address",INDIRECT(""&E14&"!\$A\$1")),E14)			
15	Sheet1	Sheet1	=HYPERLINK("#"&CELL("address",INDIRECT(""&E15&"!\$A\$1")),E15)			

Рис. 23.13. Создание оглавления

## Задания

1. Функция **АДРЕС** возвращает фактический адрес ячейки, соотнесенный со строкой и столбцом. Например, формула **=АДРЕС(3;4)** дает в результате **\$D\$3**. Какой результат будет получен при вводе формулы **=ДВССЫЛ(АДРЕС(3;4))**?
2. В книге **P23\_2.xlsx** находятся данные по продажам пяти продуктов в четырех регионах (**East**, **West**, **North** и **South**). Напишите формулу с функцией **ДВССЫЛ**, оптимально вычисляющую общий объем продаж для любой комбинации последовательно пронумерованных продуктов, такой как **Product\_1–Product\_3**, **Product\_2–Product\_5** и т. д.
3. В файле **P23\_3.xlsx** шесть листов. Лист **Sheet\_1** содержит продажи за месяц **Month\_1** для продуктов **Product\_1–Product\_4**. Эти объемы продаж всегда указаны в диапазоне **E5:H5**. С помощью функции **ДВССЫЛ** внесите в таблицу на отдельном листе продажи каждого продукта по месяцам.
4. Напишите формулу, суммирующую числа в диапазоне **G2:K2** даже после вставки одного или нескольких столбцов между столбцами **G** и **K**.
5. В файле **Marketbasketdata.xlsx** содержатся данные о продажах различных товаров. Для каждой строки единица (1) в столбцах **C–K** отмечает приобретенный товар, а ноль (0) — товар, который не был приобретен. В столбце **Day Week** единица (1) означает, что сделка имела место в понедельник, двойка (2) — во вторник и т. д. Для каждого товара в ячейках **K9:K14** вычислите процент сделок, в которых товар был приобретен. Кроме того, вычислите долю сделок, происходящих каждый день.
6. В файле **Verizonindirectdata.xlsx** содержатся данные об отработанных каждым сотрудником часах и о рейтинге сотрудников за январь — май. Создайте объединенный лист, на котором можно выбрать любого сотрудника и получить отработанные этим сотрудником часы за каждый месяц наряду с рейтингом за этот месяц.

# Условное форматирование

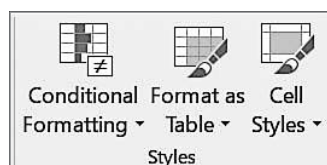
### Обсуждаемые вопросы

- Как можно визуально отобразить, связаны ли последние температурные данные с глобальным потеплением?
- Как работают правила выделения ячеек в условном форматировании?
- Как можно проверить и настроить созданные правила?
- Как работают гистограммы?
- Как работают цветовые шкалы?
- Как работают наборы значков?
- Как применить цветовой код к ежемесячной доходности акций для отображения каждого удачного месяца одним цветом, а каждого неудачного — другим?
- При поступлении данных о ежеквартальных доходах корпорации как можно выделить одним цветом кварталы, в которых доходы увеличились по сравнению с предыдущим кварталом, и другим цветом кварталы, в которых доходы снизились по сравнению с предыдущим кварталом?
- Как в заданном списке дат выделить цветом даты, приходящиеся на выходные дни?
- Наш тренер по баскетболу присвоил каждому игроку рейтинг от 1 до 10 на основе способностей игрока к игре в защите, нападении или в качестве центрального. Можно ли создать лист, визуально демонстрирующий способность каждого игрока играть на позиции, на которую его поставили?
- Для чего нужны флажки «Остановить, если истина» (Stop If True) в диалоговом окне Управление правилами (Manage Rules)?
- Как с помощью инструмента Формат по образцу (Format Painter) скопировать условный формат?

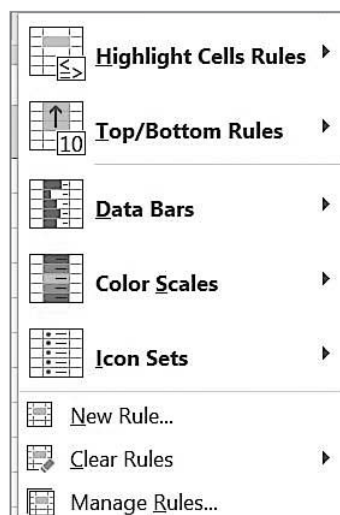
Условное форматирование позволяет отформатировать диапазон ячеек на основе содержимого этого диапазона. Если взять за пример оценки, полученные студентами на экзаменах, можно с помощью условного форматирования отобразить красным цветом имена тех студентов, которые в среднем набрали 90 баллов. Когда вы настраиваете форматирование для диапазона ячеек, Microsoft Excel 2019 проверяет каждую ячейку в диапазоне на предмет того, удовлетворяет ли

она заданным вами условиям (таким как «экзаменационный балл не меньше 90»). Ко всем удовлетворяющим условию ячейкам автоматически применяется выбранный формат. Если содержимое ячейки не удовлетворяет ни одному условию, ее формат не меняется. В Excel 2007 Условное форматирование было полностью пересмотрено и расширено. В Excel 2010 в Условное форматирование были внесены незначительные исправления. В данной главе я покажу, как применять все инструменты условного форматирования, имеющиеся в Microsoft Excel 2019.

Для просмотра параметров условного форматирования сначала выделите диапазон ячеек, который хотите отформатировать. Затем на вкладке Главная (Home) в группе Стили (Styles) нажмите на стрелку раскрывающегося списка Условное форматирование (Conditional Formatting), показанного на рис. 24.1, для открытия списка инструментов условного форматирования (рис. 24.2).




**Рис. 24.1.** Раскрывающийся список Условное форматирование



**Рис. 24.2.** Инструменты условного форматирования

Далее приведены краткие пояснения к каждому инструменту.

- **Правила выделения ячеек (Highlight Cells Rules)** — эти правила позволяют присваивать формат ячейкам, содержимое которых отвечает одному из следующих критериев:
  - находится в определенном диапазоне числовых значений;
  - совпадает с определенной текстовой строкой;
  - находится в определенном диапазоне дат (относительно текущей даты);
  - встречается в выделенном диапазоне несколько раз (или только один раз).

- **Правила отбора первых и последних значений (Top/Bottom Rules)** — эти правила позволяют присваивать формат в следующих случаях:
  - $N$  самых больших или самых маленьких значений в диапазоне;
  - $n\%$  наибольших или наименьших чисел в диапазоне;
  - числа выше или ниже среднего значения для всех чисел в диапазоне.
- **Гистограммы (Data Bars), Цветовые шкалы (Color Scales) и Наборы значков (Icon Sets)** — эти форматы предназначены для простой идентификации больших, малых или промежуточных значений в выбранном диапазоне. Более высокие столбцы гистограммы связаны с большими значениями. С помощью цветовых шкал можно, например, отображать меньшие значения красным цветом, а большие значения — синим цветом, с плавными переходами между ними по мере отображения значений в диапазоне от меньших к большим. Наборы значков позволяют применять для идентификации различных диапазонов значений до пяти символов. Например, можно отобразить стрелку , указывающую вверх для индикации больших значений, вправо для индикации промежуточных значений и вниз для индикации меньших значений.
- **Создать правило (New Rule)** — используется для создания собственных формул, определяющих, должна ли ячейка иметь определенный формат. Например, если значение в ячейке превышает значение в вышестоящей ячейке, можно окрасить ячейку в зеленый цвет. Если в ячейке содержится пятое по величине значение в столбце, можно окрасить ее в красный и т. д.
- **Удалить правила (Clear Rules)** — этот инструмент используется при удалении всего условного форматирования, созданного для выделенного диапазона или целого листа.
- **Управление правилами (Manage Rules)** — этот инструмент предназначен для просмотра, редактирования и удаления правил условного форматирования, а также создания новых правил или изменения порядка, в котором правила условного форматирования применяются в Excel.

## Ответы на вопросы

- ❓ **Как можно визуально отобразить, связаны ли последние температурные данные с глобальным потеплением?**

Этот вопрос предоставляет прекрасную возможность применить правила отбора первых и последних значений в условном форматировании. Файл `Globalwarming2018.xlsx` содержит данные (часть этих данных представлена на рис. 24.4) о годовых колебаниях температуры в мире за 1950–1960 гг. в градусах Цельсия. Если имеет место глобальное потепление, значит, температура в последние годы должна быть выше, чем в более ранние. Чтобы узнать, произошло ли потепление в последние годы, отметим красным цветом ячейки для 20 самых

теплых лет. Выделите диапазон **B4:B171**, содержащий значения температур. На вкладке Главная (Home) в группе Стили (Styles) в раскрывающемся списке Условное форматирование (Conditional Formatting) выберите Правила отбора первых и последних значений (Top/Bottom Rules). Выберите Первые 10 элементов (Top 10 Items) и заполните поля в диалоговом окне, как показано на рис. 24.3. Измените значение по умолчанию (10) на 20 и оставьте выделение по умолчанию (светло-красная заливка и темно-красный текст). Теперь 20 самых теплых лет отмечены красным. Обратите внимание, что в 20 самых теплых лет входят года с 2001-го по 2017-й. Если бы не было никакой закономерности в колебаниях температур, на этот период пришлось бы только два теплых года (примерно).



**Рис. 24.3.** Выделение 20 значений самых высоких температур красным цветом

#### ПРИМЕЧАНИЕ

В правой части окна на рис. 24.3 в раскрывающемся списке можно найти другие варианты, в том числе Пользовательский формат (Custom Format). При выборе этого варианта откроется диалоговое окно Формат ячеек (Format Cells), в котором вы можете создать пользовательский формат, который применится, когда вы нажмете ОК в обоих диалоговых окнах.

А теперь выберите диапазон **C4:C171**, вернитесь к правилам отбора первых и последних значений и выберите Первые 10% (Top 10%). Примите значения, указанные по умолчанию (10% самых высоких значений температур выделить красным цветом), и затем выделите Последние 10% (Bottom 10 percent) зеленым цветом. Наконец, в столбце D выделите зеленым цветом температуры выше средней, а красным — температуры ниже средней. Обратите внимание, что 15 лет после 2001 г. попадают в диапазон 10% самых высоких значений температур. Если бы никакой закономерности в колебаниях температуры не было, на этот период также пришлось бы примерно два теплых года. Также следует отметить, что во все годы после 1977 г. температура была выше средней! Если бы закономерность отсутствовала, вероятность превышения среднегодовой температуры составила бы около 50%, и в этом случае годы 1977–2017 показывают необоснованный подъем в течение 41 года! Условное форматирование является мощным инструментом, который позволяет продемонстрировать, что на Земле (по тем или иным причинам) в последнее время становится теплее.

	A	B	C	D
2	1950–Average			
3	Year	Top 20	Top and bottom 10%	Above and below average
4	1850	-0.49	-0.49	-0.49
5	1851	-0.35	-0.35	-0.35
6	1852	-0.34	-0.34	-0.34
7	1853	-0.38	-0.38	-0.38
8	1854	-0.31	-0.31	-0.31
9	1855	-0.28	-0.28	-0.28
10	1856	-0.46	-0.46	-0.46
11	1857	-0.57	-0.57	-0.57
12	1858	-0.40	-0.40	-0.40
13	1859	-0.37	-0.37	-0.37
14	1860	-0.48	-0.48	-0.48
149	1995	0.41	0.41	0.41
150	1996	0.31	0.31	0.31
151	1997	0.45	0.45	0.45
152	1998	0.60	0.60	0.60
153	1999	0.37	0.37	0.37
154	2000	0.38	0.38	0.38
155	2001	0.52	0.52	0.52
156	2002	0.60	0.60	0.60
157	2003	0.59	0.59	0.59
158	2004	0.49	0.49	0.49
159	2005	0.67	0.67	0.67
160	2006	0.62	0.62	0.62
161	2007	0.63	0.63	0.63
162	2008	0.49	0.49	0.49
163	2009	0.62	0.62	0.62
164	2010	0.70	0.70	0.70
165	2011	0.58	0.58	0.58
166	2012	0.60	0.60	0.60
167	2013	0.62	0.62	0.62
168	2014	0.68	0.68	0.68
169	2015	0.82	0.82	0.82
170	2016	0.96	0.96	0.96
171	2017	0.85	0.85	0.85

**Рис. 24.4.** Условное форматирование с реализацией правил отбора

### ? Как работают правила выделения ячеек в условном форматировании?

В файле `Highlightcells.xlsx` (рис. 24.5) продемонстрировано использование инструмента выделения ячеек. Допустим, вы хотите выделить красным цветом все повторяющиеся имена в диапазоне `C2:C11`. Просто выделите диапазон ячеек `C2:C11`, на вкладке Главная (Home) в группе Стили (Styles) откройте раскрывающийся список Условное форматирование (Conditional Formatting) и выберите Правила выделения ячеек (Highlight Cells Rules) и Повторяющиеся значения (Duplicate Values). Затем выберите светло-красная заливка и темно-красный текст (Light Red Fill With Dark Red Text). Нажмите ОК для применения правила, и все имена, встречающиеся больше одного раза (John и Josh), будут выделены красным цветом.

Допустим, теперь вам нужно выделить красным все ячейки в диапазоне `D2:D11` с текстом Eric. Выделите диапазон `D2:D11`, в списке Условное форматирование (Conditional Formatting) выберите Правила выделения ячеек (Highlight Cells Rules) и затем Текст содержит (Text That Contains). Введите в левое поле Eric и выберите справа светло-красная заливка и темно-красный текст (Light Red Fill With Dark Red Text). Как показано на рис. 24.5, выделены ячейки, содержащие текст Eric и Erica. (Erica, естественно, содержит текстовую строку Eric.) Опция Текст содержит нечувствительна к регистру символов, поэтому если набрать eric строчными буквами, результат поиска будет такой же.

	C	D	E	F
		text containing Eric red fill then yesterday red text green	last 7 days red	
1	duplicates red fill			
2	John	John	6/1/2006	
3	Eric	Eric	6/27/2018	today()-1
4	James	James	6/23/2018	today()-5
5	John	John	5/15/2007	
6	Erica	Erica	6/14/2006	
7	JR	JR	2/3/2003	
8	Adam	Adam	5/12/2006	
9	Josh	Josh	6/17/2005	
10	Babe	Babe	8/1/2006	
11	Josh	Josh	9/2/2005	

**Рис. 24.5.** Применение правил выделения ячеек

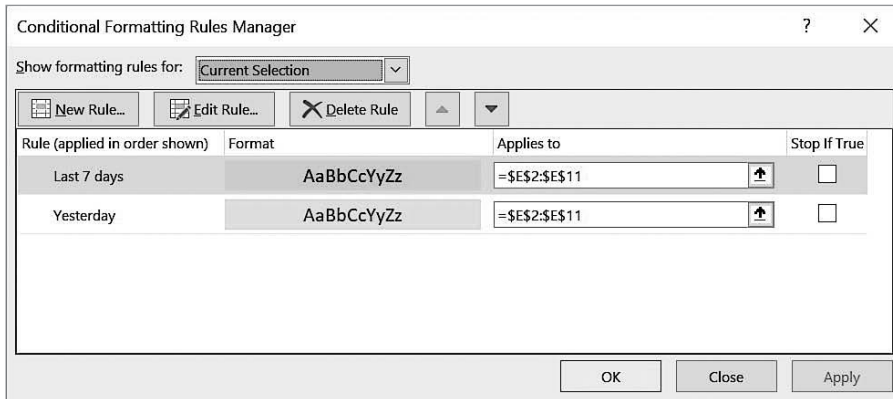
Пусть имеется список дат (такой как в диапазоне E2:E11), и требуется выделить зеленым все ячейки, содержащие вчерашнюю дату, а все ячейки с датами в пределах семи последних дней — красным (см. файл Highlightcells.xlsx, лист Right Way). Предположим, сегодня 28 июня 2018 г. (как на рис. 24.5). Обратите внимание, что ячейка E3 содержит формулу =СЕГОДНЯ()-1, поэтому в ней всегда отображается вчерашняя дата. Ячейка E4 содержит формулу =СЕГОДНЯ()-5.

Для начала выделим диапазон ячеек, который требуется отформатировать (E2:E11). Список Условное форматирование (Conditional Formatting), далее Правила выделения ячеек (Highlight Cells Rules), далее Дата (A Date Occurring), а затем в диалоговом окне выберите из раскрывающихся списков Вчера (Yesterday) и Зеленая заливка и темно-зеленый цвет (Green Fill With Dark Green Text) и нажмите ОК. Затем снова в списке Условное форматирование (Conditional Formatting) выберите Правила выделения ячеек (Highlight Cells Rules) и еще раз Дата (A Date Occurring). В диалоговом окне Дата (A Date Occurring) укажите За последние 7 дней (In The Last 7 Days) и светло-красная заливка и темно-красный текст (Light Red Fill With Dark Red Text). Правила форматирования, созданные позднее, имеют приоритет над правилами, созданными ранее (если только приоритет правил не изменен, как описано далее в главе). Это объясняет, почему вчерашняя дата выделена красным цветом, а не зеленым.

### ? Как можно проверить и настроить созданные правила?

После создания правил условного форматирования эти правила можно просмотреть с помощью инструмента Управление правилами (Manage Rules). Например, выделите даты в диапазоне E2:E11 и выберите в раскрывающемся списке Условное форматирование (Conditional Formatting) инструмент Управление правилами (Manage Rules); в диалоговом окне появятся правила (рис. 24.6). Как видите, правило форматирования За последние 7 дней (In The Last 7 Days) применяется до правила форматирования Вчера (Yesterday).

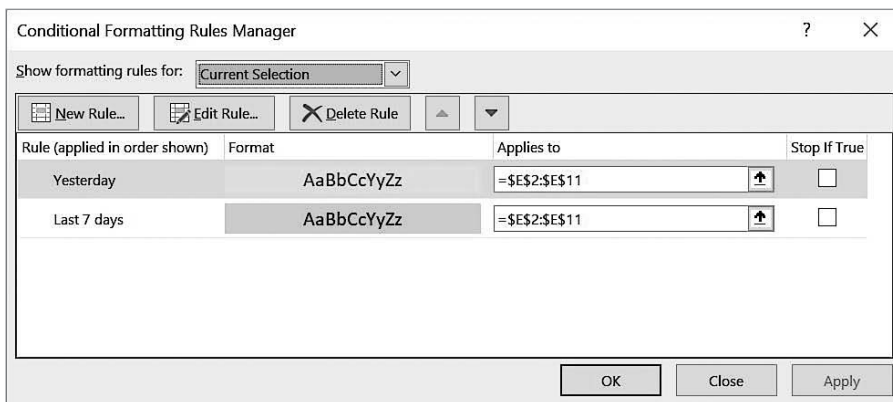




**Рис. 24.6.** Диалоговое окно Диспетчер правил условного форматирования

В диалоговом окне Диспетчер правил условного форматирования (Conditional Formatting Rules Manager) можно выполнить следующее:

- создать правило, нажав Создать правило (New Rule);
- отредактировать или изменить правило, нажав Изменить правило (Edit Rule);
- удалить правило, выбрав его и нажав Удалить правило (Delete Rule);
- изменить приоритет правила, выбрав правило и нажав кнопки со стрелками вверх и вниз.



**Рис. 24.7.** Теперь правило форматирования Вчера имеет более высокий приоритет, чем правило За последние 7 дней

Для иллюстрации возможностей диалогового окна Диспетчер правил условного форматирования (Conditional Formatting Rules Manager) я скопировал предыдущий лист в файле Highlightcells.xlsx (щелкните правой кнопкой мыши на имени листа

и выберите Переместить или скопировать (Move Or Copy), затем установите флажок Создать копию (Create A Copy)) и назвал лист **Right Way**. В диалоговом окне я выбрал правило **Вчера (Yesterday)** и щелкнул на кнопке со стрелкой вверх. Теперь правило **Вчера (Yesterday)** имеет более высокий приоритет, чем правило **За последние 7 дней (Last 7 Days)**, и ячейка **E3** будет выделяться зеленым цветом, а не красным. Соответствующее диалоговое окно **Диспетчер правил условного форматирования (Conditional Formatting Rules Manager)** представлено на рис. 24.7, а на рис. 24.8 показана выделенная зеленым цветом ячейка **E3** и выделенная красным цветом ячейка **E4**, что и требовалось.

	C	D	E	F
		text containing	yesterday	
1	duplicates red fill	Eric red fill red text	green then last 7 days red	
2	John	John	6/1/2006	
3	Eric	Eric	6/27/2018	today()-1
4	James	James	6/23/2018	today()-5
5	John	John	5/15/2007	
6	Erica	Erica	6/14/2006	
7	JR	JR	2/3/2003	
8	Adam	Adam	5/12/2006	
9	Josh	Josh	6/17/2005	
10	Babe	Babe	8/1/2006	
11	Josh	Josh	9/2/2005	

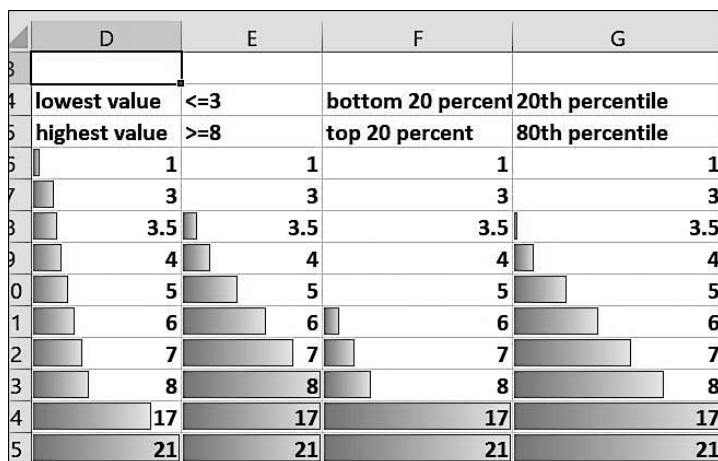
**Рис. 24.8.** После изменения приоритета правил ячейка со вчерашней датой выделена зеленым цветом

## ❓ Как работают гистограммы?

Когда у вас есть длинный список чисел, неплохо иметь визуальный индикатор больших и малых величин. Гистограммы, цветовые шкалы и наборы значков (все внедрено еще в Microsoft Excel 2007) являются превосходными инструментами для графического отображения различий. В файле **Scalesiconsdatabars.xlsx** показано использование этих инструментов.

На рис. 24.9 (см. файл **Data bars.xlsx**) дан пример использования гистограмм. Для своих примеров я начал с того, что применил гистограммы по умолчанию к данным в **D6:D15**. Я выделил данные в этом диапазоне и в раскрывающемся списке **Условное форматирование (Conditional Formatting)** выбрал инструмент **Гистограммы (Data Bars)**. В группе **Градиентная заливка (Gradient Fill)** я щелкнул на **Синяя гистограмма (Blue Data Bars)** для создания формата, представленного в столбце **D** на рис. 24.9. Более длинный горизонтальный синий столбец соответствует большому значению в ячейке. По умолчанию самый короткий синий столбец связан с наименьшим числом в выделенном диапазоне, а самый длинный — с наибольшим.

Как видите, в столбце D размер горизонтального столбца прямо пропорционален значению данных, то есть горизонтальный столбец для значения 8 в два раза длиннее, чем столбец для значения 4, и т. д.

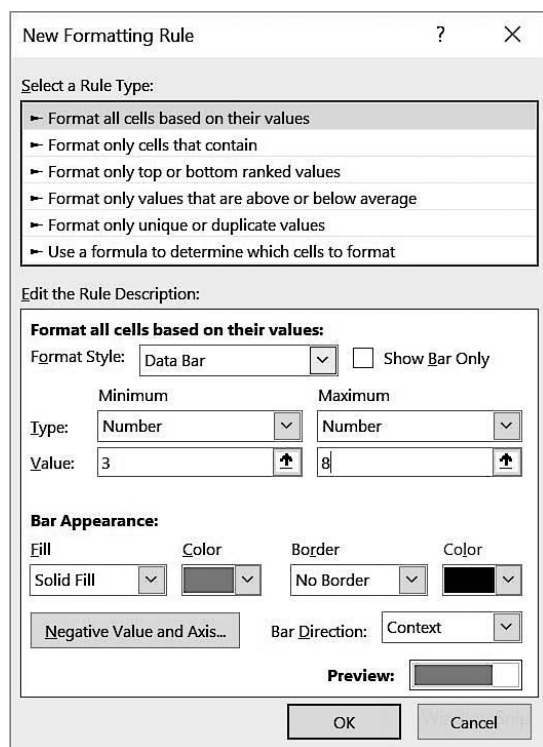


**Рис. 24.9.** Гистограмма позволяет графически отобразить численную разницу

Если после выбора инструмента Гистограммы (Data Bars) вы щелкнете на Другие правила (More Rules), откроется диалоговое окно Создание правила форматирования (New Formatting Rule), показанное на рис. 24.10. (Его можно также открыть, выбрав инструмент Управление правилами (Manage Rules) и нажав Изменить правило (Edit Rule) или дважды щелкнув по правилу.) В этом диалоговом окне можно поменять критерии привязки столбцов гистограммы к ячейкам. В диапазоне E6:E15 не связывайте столбцы гистограммы с числами не больше 3 и свяжите самые длинные столбцы с числами не меньше 8. Как показано на рис. 24.9, со всеми числами в столбце E, меньше или равными 3, столбцы гистограммы не связаны, а со всеми числами, больше или равными 8, связаны самые длинные горизонтальные столбцы. С числами от 3 до 8 связаны ранжированные горизонтальные столбцы. Обратите внимание, что в диалоговом окне Изменение правила форматирования (Edit Formatting Rule) можно установить флажок Показывать только столбец (Show Bar Only) для отображения в условно отформатированных ячейках цветного столбца, но не значения, хранящегося в ячейке.

Далее в столбце F в диапазоне F6:F15 не связывайте столбцы гистограммы с числами, входящими в нижние 20%, и свяжите самые длинные столбцы с числами из верхних 20%. Другими словами, со всеми числами не больше  $1 + 0,2 \times (21 - 1) = 5$  столбцы гистограммы не связаны, а со всеми числами не меньше  $1 + 0,8 \times (21 - 1) = 17$  связаны самые длинные столбцы. Как видно на рис. 24.9, в столбце F с первыми пятью числами столбцы гистограммы не связаны, а с числами 17 и 21 связаны самые длинные столбцы.

В диапазоне G6:G13 со всеми числами на уровне 20-го перцентиля (3) или ниже столбцы гистограммы не связаны, а со всеми числами на уровне 80-го перцентиля (17) или выше связаны самые длинные столбцы. В раскрывающемся списке **Направление столбца (Bar Direction)** можно выбрать, от какой границы ячейки, правой или левой, должен начинаться столбец гистограммы.



**Рис. 24.10.** Диалоговое окно настройки гистограммы

В Excel 2019 допускается сплошная заливка в столбцах гистограммы и ориентация столбцов гистограммы, связанных с отрицательными значениями, в направлении, противоположном столбцам, связанным с положительными значениями. В диалоговом окне **Изменение правила форматирования (Edit the Rule Description)** щелкните на **Отрицательные значения и ось (Negative Values and Axis)** и передвиньте ось из обычного положения в центр ячейки, а также укажите для отрицательных и положительных столбцов одно направление или противоположные направления (по умолчанию). В файле **Negativedatabars.xlsx** и на рис. 24.11 представлены гистограммы для набора данных с отрицательными значениями. Для двух последних столбцов на рис. 24.11 ось была установлена автоматически, благодаря чему осталось больше места для положительных значений, поскольку они больше по модулю, чем отрицательные значения.

	D	E	F	G	H	I
1	Left to Right Centered		Right to Left Centered			Automatic
2	Year	Sales	Year	Sales	Year	Sales
3	2000	-5	2000	-5	2000	-5
4	2001	56	2001	56	2001	500
5	2002	85	2002	85	2002	500
6	2003	-31	2003	-31	2003	-31
7	2004	-26	2004	-26	2004	-26
8	2005	-40	2005	-40	2005	-40
9	2006	85	2006	85	2006	200
10	2007	34	2007	34	2007	34
11	2008	50	2008	50	2008	50
12	2009	-46	2009	-46	2009	-46
13	2010	4	2010	4	2010	4
14	2011	47	2011	47	2011	47
15	2012	-5	2012	-5	2012	-5
16	2013	44	2013	44	2013	44

**Рис. 24.11.** Гистограммы для отрицательных значений со сплошной и градиентной заливкой

### ? Как работают цветовые шкалы?

Давайте используем цветовые шкалы для обобщения некоторых наборов данных. Как и в правилах выделения ячеек, в цветовой шкале для отображения численной разницы значений в ячейках используется заливка. В этом разделе я описываю пример с трехцветной шкалой. (См. файл *Colorscaleinvestment.xlsx* и рис. 24.12. Обратите внимание, что строки 19–80 скрыты; для их отображения выберите строки 18 и 81, щелкните правой кнопкой мыши и в контекстном меню выберите Показать (Unhide)).

Для своих примеров я выбрал ежегодные доходы от инвестиций в акции, казначейские векселя и казначейские облигации в ячейках B6:D93. Затем в Условном форматировании (Conditional Formatting) я выбрал Цветовые шкалы (Color Scales) и затем Другие правила (More Rules). Отобразилось диалоговое окно Изменение правила форматирования (Edit Formatting Rule), показанное на рис. 24.13.

Я выбрал красный для самого низкого дохода, зеленый для самого высокого дохода и желтый для среднего дохода. Удивительно, но Excel может незначительно изменять заливку каждой ячейки, основываясь на ее значении. В столбце В на рис. 24.12 ячейка с самым низким доходом залита красным. Отметим, что 1931 и 2008 гг. (как мы все знаем) были самыми неудачными для инвестиций в акции. По мере приближения дохода к 50-му процентилю цвет ячеек постепенно меняется на желтый. Затем по мере роста дохода от 50-го процентиля к самому высокому доходу цвет ячеек меняется от желтого к зеленому. Большая часть зеленых и красных ячеек относится к акциям, поскольку ежегодные доходы от акций более

	D	E	F	G
5	Year	S&P 500	3-month T.Bill	10-year T. Bond
6	1928	43.81%	3.08%	0.84%
7	1929	-8.30%	3.16%	4.20%
8	1930	-25.12%	4.55%	4.54%
9	1931	-43.84%	2.31%	-2.56%
10	1932	-8.64%	1.07%	8.79%
11	1933	49.98%	0.96%	1.86%
12	1934	-1.19%	0.32%	7.96%
13	1935	46.74%	0.18%	4.47%
14	1936	31.94%	0.17%	5.02%
15	1937	-35.34%	0.30%	1.38%
16	1938	29.28%	0.08%	4.21%
17	1939	-1.10%	0.04%	4.41%
18	1940	-10.67%	0.03%	5.40%
81	2003	28.36%	1.03%	0.38%
82	2004	10.74%	1.23%	4.49%
83	2005	4.83%	3.01%	2.87%
84	2006	15.61%	4.68%	1.96%
85	2007	5.48%	4.64%	10.21%
86	2008	-36.55%	1.59%	20.10%
87	2009	25.94%	0.14%	-11.12%
88	2010	14.82%	0.13%	8.46%
89	2011	2.10%	0.03%	16.04%
90	2012	15.89%	0.05%	2.97%
91	2013	32.15%	0.07%	-9.10%
92	2014	13.52%	0.05%	10.75%
93	2015	1.36%	0.21%	1.28%

Рис. 24.12. Трехцветная шкала

New Formatting Rule ? X

Select a Rule Type:

► Format all cells based on their values

► Format only cells that contain

► Format only top or bottom ranked values

► Format only values that are above or below average

► Format only unique or duplicate values

► Use a formula to determine which cells to format

Edit the Rule Description:

**Format all cells based on their values:**

Format Style: 3-Color Scale

Minimum

Midpoint

Maximum

Type: Lowest Value

Percentile

Highest Value

Value: (Lowest value)

50

(Highest value)

Color:

Preview:

OK Cancel

Рис. 24.13. Настройка трехцветной шкалы

изменчивы, чем доходы от казначейских векселей и облигаций. Эта изменчивость приводит к довольно частой смене низкого и высокого дохода от акций. Все ежегодные доходы от казначейских векселей и облигаций отмечены желтым цветом, так как малая изменчивость ежегодного дохода от этих инвестиций означает, что большую часть времени доходы имеют среднюю величину.

В файле **Scalesiconsdatabars.xlsx**, показанном на рис. 24.14, я создал несколько двухцветных шкал. Я выделил диапазон ячеек, выбрал **Условное форматирование (Conditional Formatting)** и **Цветовые шкалы (Color Scales)**. Далее можно выбрать цветовую комбинацию из заданного списка или создать собственную комбинацию, щелкнув на **Другие правила (More Rules)**.

Для своих примеров я выбрал двухцветную шкалу, показывающую меньшие значения белым цветом, а большие — синим.

- В диапазоне ячеек D19:D28 я выбрал для минимального значения белый цвет, а для максимального значения — синий. По мере роста чисел ячейки становятся темнее.
- В диапазоне E19:E28 я выбрал для значений, меньше или равных 3, белый цвет, а для значений, больше или равных 8, — синий. Для чисел от 3 до 8 по мере увеличения чисел заливка ячеек становится темнее.
- В диапазоне F19:F28 я выбрал для значений из нижних 20% белый цвет, а для значений из верхних 20% — синий. Для чисел из средних 60% по мере увеличения чисел заливка ячеек становится темнее.

	D	E	F
16	<b>Color Scales</b>		
17	<b>lowest value</b>	<b>&lt;=3</b>	<b>20th percentile</b>
18	<b>highest value</b>	<b>&gt;=8</b>	<b>80th percentile</b>
19	1	1	1
20	2	2	1
21	3	3	2
22	4	4	2
23	5	5	4
24	6	6	5
25	7	7	5
26	8	8	7
27	9	9	9
28	10	10	10

**Рис. 24.14.** Двухцветные шкалы

### ❓ Как работают наборы значков?

Численную разницу можно также отобразить с помощью наборов значков. (См. рис. 24.15 и файл **Scalesiconsdatabars.xlsx**.) Набор значков включает от трех до пяти символов. Для связи значков со значениями в ячейках необходимо определить



критерии. Например, можно указать стрелку вниз для меньших чисел, стрелку вверх для больших чисел и горизонтальную стрелку для промежуточных значений. В диапазоне E32:F41 показаны два набора значков. Для каждого столбца можно использовать, например, красную стрелку вниз, желтую горизонтальную стрелку и зеленую стрелку вверх.

	D	E	F
30	down arrow	<=3	20th percentile
31	up arrow	>=8	80th percentile
32		↓	1 ↓
33		↓	2 ↓
34		↓	3 ⇒
35		⇒	4 ⇒
36		⇒	5 ⇒
37		⇒	6 ⇒
38		⇒	7 ⇒
39		↑	8 ⇒
40		↑	9 ↑
41		↑	10 ↑

**Рис. 24.15.** Применение набора значков

Вот как я привязал значки к диапазону численных значений:

1. Выбрав диапазон E32:E41, я щелкнул на Условное форматирование (Conditional Formatting) и Наборы значков (Icon Sets). Выбрал Другие правила (More Rules) и набор из трех цветных стрелок в списке Стиль значка (Icon Styles). В столбце E я хочу отобразить числа меньше 4 стрелкой вниз, числа от 4 до 7 горизонтальной стрелкой и числа, больше или равные 8, стрелкой вверх. Для этого я установил параметры в диалоговом окне Создание правила форматирования (New Formatting Rule), как показано на рис. 24.16.
2. Аналогичным образом я создал правило форматирования для диапазона F31:F42, в котором стрелки вверх соответствуют ячейкам, содержащим числа на уровне 80-го перцентиля и выше ( $\geq 8$ ), а стрелки вниз — ячейкам, содержащим числа ниже 20-го перцентиля ( $\leq 1$ ). В диалоговом окне Создание правила форматирования (New Formatting Rule) я выбрал тот же стиль значков (стрелки), в поле Тип оставил значение по умолчанию Процент и изменил верхнее Значение (Value) на 80 и нижнее на 20. Соответствующие параметры форматирования показаны на рис. 24.17.

Необязательные параметры включают Обратный порядок значков (Reverse Icon Order), который связывает значки слева с меньшими числами, а значки справа с большими числами, и флажок Показать только значок (Show Icon Only), скрывающий содержимое ячейки. В Excel 2019 можно скрыть подмножество значков. Также можно настроить набор значков. В файле Historicalinvest.xlsx на листе Customize Icons для обобщения доходов по инвестициям я использовал три значка. Верхняя треть доходов обозначена серой стрелкой вверх, нижняя треть — красной стрелкой вниз. Для сред-



ней трети доходов горизонтальную стрелку можно скрыть. На рис. 24.18 показаны полученные в результате значки, а на рис. 24.19 — диалоговое окно, открытое через Управление правилами (Manage Rules), с правилом форматирования. Я настроил набор значков из трех серых стрелок, заменив серую стрелку вниз красной стрелкой.

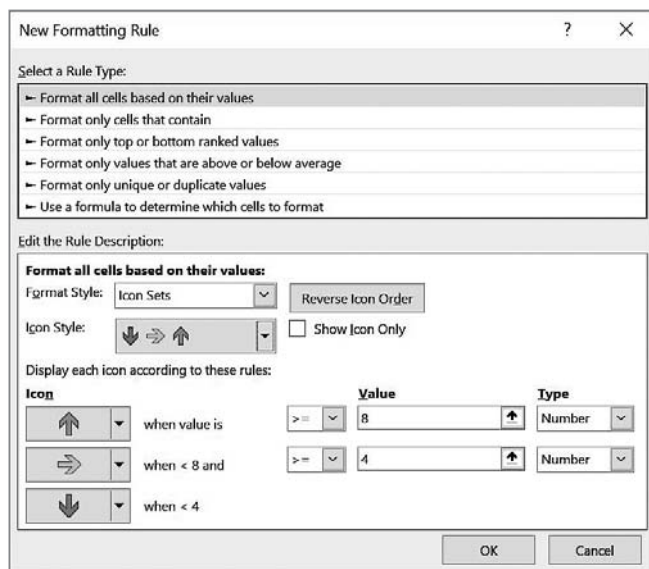


Рис. 24.16. Установление соответствия значков численным значениям

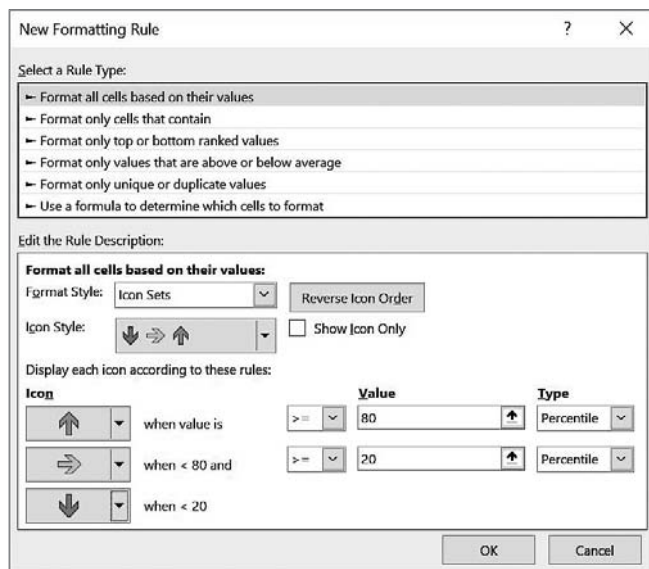


Рис. 24.17. Установление соответствия значков процентильным значениям

	A	B	C	D
6		Annual Returns on Investments in		
7	Year	Stocks	T.Bills	T.Bonds
8	1928	43.81%	3.08%	0.84%
9	1929	-8.30%	3.16%	4.20%
10	1930	-25.12%	4.55%	4.54%
11	1931	-43.84%	2.31%	-2.56%
12	1932	-8.64%	1.07%	8.79%
13	1933	49.98%	0.96%	1.86%
14	1934	-1.19%	0.30%	7.96%
15	1935	46.74%	0.23%	4.47%
80	2000	-9.03%	5.37%	16.66%
81	2001	-11.85%	5.73%	5.57%
82	2002	-22.10%	17.84%	3.83%
83	2003	28.68%	1.45%	1.65%
84	2004	10.88%	8.51%	1.02%
85	2005	4.91%	7.81%	1.20%
86	2006	15.79%	1.19%	2.98%
87	2007	5.49%	9.88%	4.66%
88	2008	-37.00%	25.87%	1.60%
89	2009	26.46%	-14.90%	0.10%

Рис. 24.18. Измененный набор значков со скрытым значком

New Formatting Rule?×

Select a Rule Type:

Format all cells based on their values

Format only cells that contain

Format only top or bottom ranked values

Format only values that are above or below average

Format only unique or duplicate values

Use a formula to determine which cells to format

Edit the Rule Description:

Format all cells based on their values:

Format Style: Icon Sets

Reverse Icon Order

Icon Style: Custom

Show Icon Only

Display each icon according to these rules:

Icon

Value

Type

↑

when value is

>=

67

↑

Percent

No Cell Icon

when < 67 and

>=

33

↑

Percent


↓

when < 33

OK

Cancel

Рис. 24.19. Диалоговое окно для изменения и скрытия значков

 Как применить цветовой код к ежемесячной доходности акций для отображения каждого удачного месяца одним цветом, а каждого неудачного — другим?

В файле *Sandp.xlsx* (рис. 24.20) содержатся ежемесячные данные о ценах и доходах по фондовому индексу Standard&Poor's 500. Допустим, вы хотите выделить зеленым цветом каждый месяц, в котором индекс S&P поднимался более чем на 3%, и красным цветом каждый месяц, в котором он падал более чем на 3%.

	A	B	C
6		S&P	
7		Comp.	
8	Date	P	Change
9	1871.01	4.44	
10	1871.02	4.5	0.0135
11	1871.03	4.61	0.0244
12	1871.04	4.74	0.0282
13	1871.05	4.86	0.0253
14	1871.06	4.82	-0.008
15	1871.07	4.73	-0.019
16	1871.08	4.79	0.0127
17	1871.09	4.84	0.0104
18	1871.1	4.59	-0.052
19	1871.11	4.64	0.0109
20	1871.12	4.74	0.0216
21	1872.01	4.86	0.0253
22	1872.02	4.88	0.0041
23	1872.03	5.04	0.0328
24	1872.04	5.18	0.0278

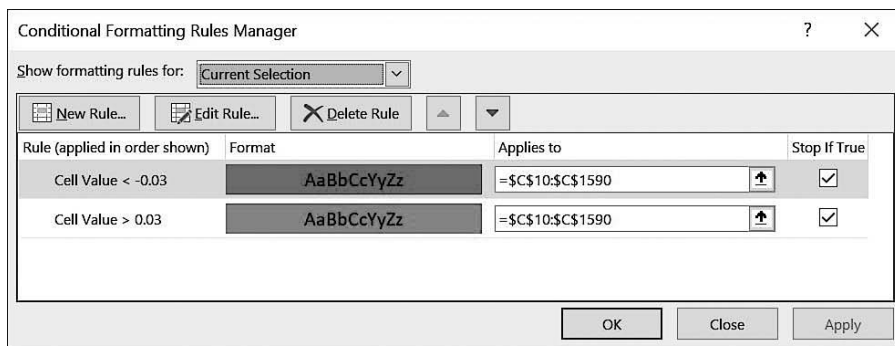
**Рис. 24.20.** Условное форматирование для выделения доходов в фондовом индексе S&P

Начните с ячейки C10 (первый месяц, содержащий доход по индексу S&P) и выберите все ежемесячные доходы, нажав **Ctrl+Shift+↓**. Затем в раскрывающемся списке **Условное форматирование (Conditional Formatting)** выберите **Выделить ячейки (Highlight Cells)** и **Больше чем (Greater Than)**. Введите данные в диалоговом окне, как показано на рис. 24.21.

**Рис. 24.21.** Применение специального форматирования к доходам по S&P больше 3%

Во втором списке в области измените описание правила и выберите значение **больше**, а в правом поле установите значение **0,03**. Обратите внимание, что в диа-

логовом окне **Формат ячеек (Format Cells)** списки шрифтов и размеров недоступны, так что при выборе форматирования изменить эти атрибуты невозможно. На вкладке **Заливка (Fill)** можно указать цвет ячеек, а на вкладке **Границы (Borders)** создать границы для ячеек, удовлетворяющих критериям форматирования. Для возврата в диалоговое окно **Диспетчер правил условного форматирования (Conditional Formatting Rules Manager)** нажмите **ОК** в диалоговых окнах **Формат ячеек (Format Cells)** и **Создание правила форматирования (New Formatting Rule)**. Теперь можно снова нажать **Создать правило (New Rule)** и аналогичным образом установить заливку красным цветом для всех ячеек, содержащих значения меньше чем  $-0,03$  (рис. 24.22).



**Рис. 24.22.** Отображение доходности акций меньше  $-3\%$  красным цветом и больше  $3\%$  — зеленым цветом

После нажатия **ОК** в диалоговом окне **Диспетчер правил условного форматирования (Conditional Formatting Rules Manager)** все доходы по S&P больше  $3\%$  (см. ячейку C23, например) отобразятся зеленым цветом, а все доходы меньше чем  $-3\%$  (см. ячейку C18) — красным. Для ячеек, значения которых не соответствуют ни одному из этих условий, сохраняется исходное форматирование.

Далее приведены некоторые советы по условному форматированию.

- Для удаления условного форматирования (или любого формата), примененного к диапазону ячеек, просто выделите этот диапазон ячеек, в раскрывающемся списке **Условное форматирование (Conditional Formatting)** выберите **Удалить правила (Clear Rules)** и затем **Удалить правила из выделенных ячеек (Clear rules from selected cells)**.
- Для выделения всех ячеек на листе, к которым применено **Условное форматирование**, нажмите **F5** для открытия диалогового окна **Переход (Go To)**. В нижнем левом углу этого окна нажмите **Выделить (Special)**, установите переключатель в положение **Условные форматы (Conditional Formats)** и щелкните на **ОК**.
- Отредактировать правило условного форматирования можно, нажав **Условное форматирование**, затем **Управление правилами (Manage Rules)** и затем дважды щелкнув по правилу или нажав кнопку **Изменить правило (Edit Rule)**.

- Удалить правило условного форматирования можно, нажав **Условное форматирование**, затем **Управление правилами (Manage Rules)** и затем дважды щелкнув по правилу или нажав **Удалить правило (Delete Rule)**.

Обратите внимание, что после того как оба правила будут определены, правило форматирования красным цветом окажется в списке первым (поскольку оно было создано позже правила форматирования зеленым цветом). В диалоговом окне **Диспетчер правил условного форматирования (Conditional Formatting Rules Manager)** правила указаны в порядке приоритета. В этом примере не важно, какое правило указано первым, поскольку ни одна ячейка не удовлетворяет критериям обоих правил. При возникновении конфликта правил более высокий приоритет имеет правило, стоящее в списке первым. Для изменения порядка применения правил выделите правило и нажмите стрелку вверх для повышения его приоритета или стрелку вниз для понижения приоритета.

**❓ При поступлении данных о ежеквартальных доходах корпорации как можно выделить одним цветом кварталы, в которых доходы увеличились по сравнению с предыдущим кварталом, и другим цветом кварталы, в которых доходы снизились по сравнению с предыдущим кварталом?**

В файле **Amazon.xlsx** содержатся данные о ежеквартальных доходах (в млн долл.) для Amazon.com с 1995 по 2017 гг. Доходы с 1995 по 2002 гг. показаны на рис. 24.23. Мы хотим выделить зеленым цветом кварталы, доходы в которых выросли по сравнению с предыдущим кварталом, и красным цветом кварталы, доходы в которых снизились по сравнению с предыдущим кварталом.

В диалоговом окне **Диспетчер правил условного форматирования (Conditional Formatting Rules Manager)** тип правила **Использовать формулу для определения форматизируемых ячеек (Use a formula to determine which cells to format)** позволяет указать формулу, определяющую условия, которые Excel проверяет, прежде чем применить форматирование к ячейке. Я использую эту возможность в данном ответе, но сначала давайте посмотрим, как в Excel работают некоторые логические функции (см. файл **Logicalexamples.xlsx**).

Что произойдет, если в ячейку **B4** ввести формулу **=B3<2**? Если в ячейке **B3** находится число меньше 2, то в ячейке **B4** Excel возвратит значение **ИСТИНА**; в противном случае Excel возвратит значение **ЛОЖЬ**. Другие подобные примеры, в том числе комбинации логических функций **И**, **ИЛИ** и **НЕ** в формулах, можно найти в файле **Logicalexamples.xlsx** (рис. 24.24).

- В ячейке **B6** по формуле **=ИЛИ(B3<3;C3>5)** возвращается значение **ИСТИНА**, если одно из условий (**B3<3** или **C3>5**) истинно. Поскольку значение в **C3** больше 5, возвращено значение **ИСТИНА**.
- В ячейке **B7** по формуле **=И(B3=3;C3>5)** возвращается значение **ИСТИНА**, если **B3=3** и **C3>5**. Поскольку значение в **B3** не равно 3, возвращено значение **ЛОЖЬ**. Однако в ячейке **B8** по формуле **=И(B3>3;C3>5)** возвращено значение **ИСТИНА**, поскольку оба условия (**B3>3** и **C3>5**) истинны.

- В ячейке B9 по формуле `=НЕ(В3<2)` возвращено значение **ИСТИНА**, поскольку `В3<2` возвращает значение **ЛОЖЬ**, а неложное значение — это истинное значение.

	C	D	E
3			
4	<b>FiscalYear</b>	<b>FiscalQuarter</b>	<b>Revenue</b>
5	1995	4	0.511
6	1996	1	0.875
7	1996	2	2.230
8	1996	3	4.173
9	1996	4	8.468
10	1997	1	16.005
11	1997	2	27.855
12	1997	3	37.887
13	1997	4	66.040
14	1998	1	87.361
15	1998	2	115.982
16	1998	3	153.649
17	1998	4	252.893
18	1999	1	293.643
19	1999	2	314.376
20	1999	3	355.778
21	1999	4	676.042
22	2000	1	573.889
23	2000	2	577.876
24	2000	3	637.858
25	2000	4	972.360
26	2001	1	700.356
27	2001	2	667.625
28	2001	3	639.281
29	2001	4	1115.171
30	2002	1	847.422

**Рис. 24.23.** Рост продаж выделен зеленым цветом, падение продаж — красным

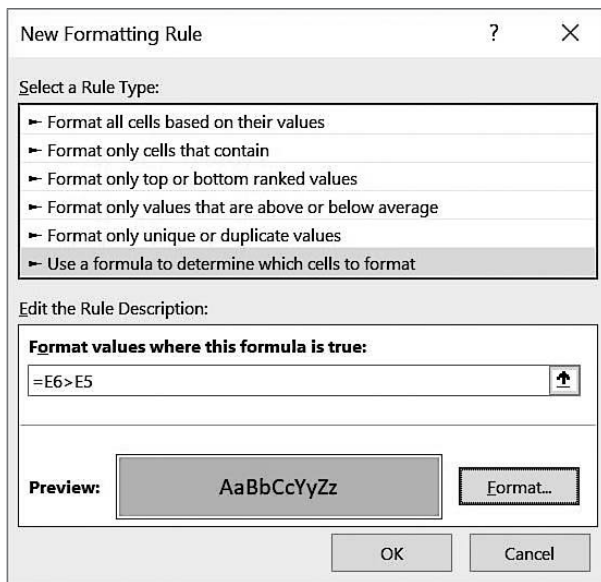
	A	B	C
2		<b>Logical formulas</b>	
3		4	6
4	<code>B3&lt;2</code>	FALSE	
5	<code>B3&gt;3</code>	TRUE	
6	<code>OR(B3&lt;3,C3&gt;5)</code>	TRUE	
7	<code>AND(B3=3,C3&gt;5)</code>	FALSE	
8	<code>AND(B3&gt;3,C3&gt;5)</code>	TRUE	
9	<code>NOT(B3&lt;2)</code>	TRUE	

**Рис. 24.24.** Логические функции

Теперь посмотрим, как параметр **Использовать формулу для определения форматируемых ячеек (Use a formula...)** позволяет создать условный формат в диапазоне ячеек. Начнем с выделения диапазона ячеек, к которому хотим применить условный формат. В раскрывающемся списке **Условное форматирование (Conditional Formatting)** выберите **Управление правилами (Manage Rules)** для открытия диалогового окна **Диспетчер правил условного форматирования (Conditional Formatting Rules Manager)**. Нажмите **Создать правило (New Rule)** и выберите в последней строке **Использовать формулу для определения форматируемых ячеек (Use a formula...)**. Я просто называю это *вариант «формула»*. Далее вводится формула (которая должна начинаться со знака равенства), значение которой истинно тогда и только тогда, когда ячейке в левом верхнем углу должен быть присвоен выбранный формат. Эта логическая формула может быть скопирована как обычная формула в оставшуюся часть выделенного диапазона, так что для правильного применения формулы к ячейкам диапазона необходимо внимательно отнестись к использованию знаков доллара (\$). Нажмите

Формат (Format) и введите желаемое форматирование. Нажмите **ОК** в диалоговых окнах **Формат ячеек (Format Cells)** и **Создание правила форматирования (New Formatting Rule)**. После нажатия **ОК** в диалоговом окне **Диспетчер правил условного форматирования (Conditional Formatting Rules Manager)** формула и форматирование скопируются во весь диапазон ячеек. Форматирование применяется к любой ячейке выделенного диапазона, которая удовлетворяет условию, определенному в формуле.

Теперь вернемся к файлу **Amazon.xlsx** и отметим зеленым цветом кварталы, в которых доходы выросли. Выделите диапазон **E6:E93** (для сравнения с цифрой доходов в ячейке **E5** нет предшествующего квартала) и создайте правило для выделения зеленым цветом каждой ячейки, значение в которой больше значения в ячейке над ней. Такое правило показано на рис. 24.25.



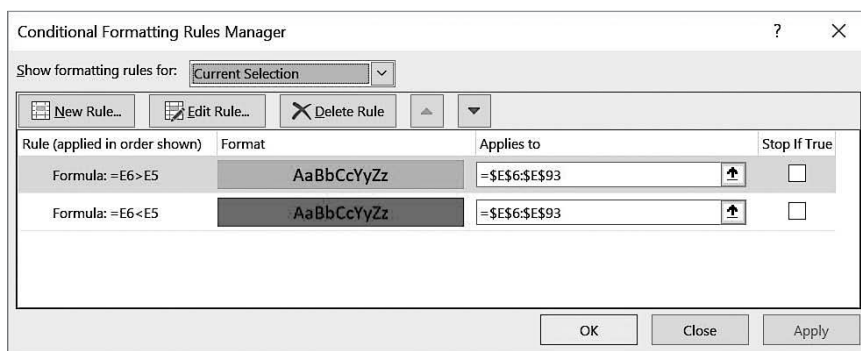
**Рис. 24.25.** Настройки условного форматирования для выделения зеленым цветом ячеек с выросшими квартальными доходами

При вводе формулы **=E6>E5** путем указания соответствующей ячейки не забудьте удалить знаки доллара ( **\$** ) из формулы в диалоговом окне **Диспетчер правил условного форматирования (Conditional Formatting Rules Manager)**, или формула не будет правильной при копировании. Вероятно, самый простой способ вставить или удалить знаки доллара — использовать клавишу **F4**. Например, при выделении ссылки на ячейку **A3** нажатие **F4** вставит или удалит знаки доллара в следующем порядке: **A3**, **\$A\$3**, **A\$3**, **\$A3**. Таким образом, если начать с **\$A\$3**, нажатие **F4** изменит ссылку на **A\$3**. Формула в этом примере гарантирует, что ячейка **E6** будет выделена зеленым цветом тогда и только тогда, когда продажи в соответствующем квартале превысят продажи в предыдущем квартале. После нажатия **ОК** все кварталы



с выросшим доходом будут отмечены зеленым цветом. В ячейке E7, например, формула, скопированная обычным способом, превращается в `=E7>E6`.

Для добавления условия к форматированию ячеек, в которых доходы снизились, выделите еще раз диапазон E6:E93, откройте диалоговое окно Диспетчер правил условного форматирования (Conditional Formatting Rules Manager), нажмите Создать правило (New Rule) и затем выберите Использовать формулу для определения форматлируемых ячеек (Use a formula...). Введите формулу `=E6<E5` и нажмите Формат (Format). На вкладке Заливка (Fill) укажите для заливки красный цвет и нажмите ОК дважды. Откроется диалоговое окно Диспетчер правил условного форматирования (Conditional Formatting Rules Manager) — рис. 24.26.



**Рис. 24.26.** Условия для отображения ячеек с выросшими доходами зеленым цветом и со снизившимися доходами — красным

Точно так же формулы можно использовать при условном форматировании с помощью гистограмм, цветовых шкал и наборов значков. Для этого при определении критериев для гистограмм, цветовых шкал и наборов значков выберите тип правила Использовать формулу для определения форматлируемых ячеек (Use a formula...).

### ? Как в заданном списке дат выделить цветом даты, приходящиеся на выходные дни?

В файле Weekendformatting.xlsx (рис. 24.27) содержится несколько дат. Я хочу выделить красным все субботы и все воскресенья. Для этого я сначала скопировал формулу `=ДЕНЬНЕД(C6;2)` из D6 в D7:D69. При выборе для аргумента тип значения 2 функция ДЕНЬНЕД (WEEKDAY) возвращает значение 1 для понедельников, 2 для вторников и т. д. Таким образом, функция возвращает значение 6 для суббот и 7 для воскресений.

Затем я выбрал диапазон D6:D69, в раскрывающемся списке Условное форматирование (Conditional Formatting) щелкнул Управление правилами (Manage Rules), Создать правило (New Rule) и тип правила Использовать формулу для определения форматлируемых ячеек (Use a formula...). Я заполнил диалоговое окно, как показано на рис. 24.28.



	C	D	E	F
3		Monday = 1 Tuesday = 2 ,etc.		
4				
5	Date	Weekday		
6	2/8/2003	6		
7	1/2/2007	2		
8	1/2/2005	7		
9	10/25/2005	2		
10	10/10/2004	7		
11	10/13/2006	5		
12	9/26/2006	2		
13	9/25/2006	1		
14	11/1/2005	2		
15	11/29/2006	3		
16	2/16/2005	3		
17	7/27/2007	5		
18	3/24/2004	3		
19	10/6/2008	1		
20	4/11/2007	3		
21	2/3/2004	2		
22	1/22/2009	4		
23	10/29/2006	7		
24	6/9/2005	4		
25	8/16/2008	6		
26	12/13/2006	3		
27	3/5/2007	1		
28	8/8/2008	5		
29	9/1/2007	6		

**Рис. 24.27.** Применение функции ДЕНЬНЕД для выделения красным цветом суббот и воскресений

New Formatting Rule
?
X

Select a Rule Type:

- Format all cells based on their values
- Format only cells that contain
- Format only top or bottom ranked values
- Format only values that are above or below average
- Format only unique or duplicate values
- Use a formula to determine which cells to format

Edit the Rule Description:

Format values where this formula is true:
=OR(D6=6,D6=7)

Preview:
AaBbCcYyZz
Format...

OK
Cancel

**Рис. 24.28.** Настройки в диалоговом окне Создание правила форматирования для выделения красным цветом суббот и воскресений

После нажатия ОК день недели отобразился шрифтом красного цвета во всех ячейках, в которых функция ДЕНЬНЕД возвратила значение 6 (для суббот) или 7 (для воскресений). Формула =ИЛИ(D6=6;D6=7) подразумевает активацию красного цвета для шрифта при записи в ячейку 6 или 7. Конечно, я мог также выбрать тип правила Форматировать только ячейки, которые содержат (Format only cells that contain) и формулу >=6 или >5.

**?** Наш тренер по баскетболу присвоил каждому игроку рейтинг от 1 до 10 на основе способностей игрока к игре в защите, нападении или в качестве центрального. Можно ли создать лист, визуально демонстрирующий способность каждого игрока играть на позиции, на которую его поставили?

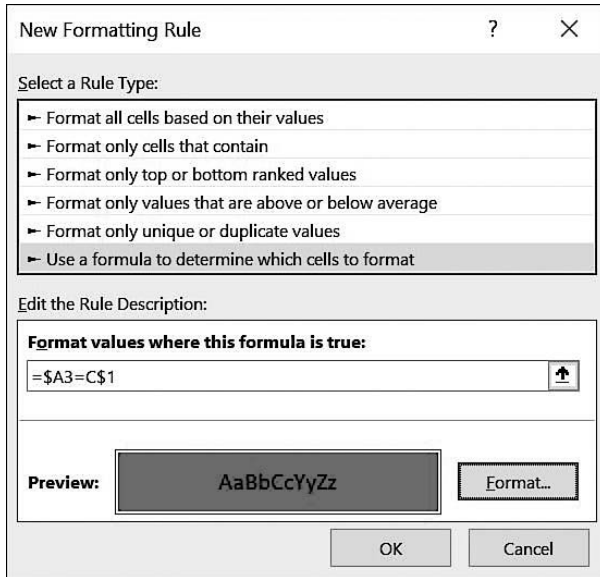
В файле Basketball.xlsx (рис. 24.29) содержатся данные о рейтингах 20 игроков по позициям и позиции каждого игрока (1 — защитник, 2 — нападающий, 3 — центрровой). Цель состоит в выделении красным цветом ячеек, содержащих рейтинги игроков для тех игровых позиций, на которые они поставлены.

	A	B	C	D	E	F
1			1	2	3	
2	Position played	Player	Guard rating	Forward rating	Center rating	
3	1	1	1	9	2	1=Guard
4	1	2	4	3	9	2=Forward
5	2	3	7	3	7	3=Center
6	2	4	9	8	8	
7	2	5	5	8	9	
8	3	6	2	7	2	
9	3	7	7	6	6	
10	3	8	4	4	3	
11	3	9	3	8	10	
12	3	10	6	1	4	
13	2	11	6	7	5	
14	2	12	2	6	5	
15	2	13	8	6	9	
16	1	14	1	1	3	
17	1	15	3	6	8	
18	2	16	4	10	1	
19	2	17	8	5	1	
20	2	18	1	7	7	
21	3	19	9	2	7	
22	3	20	10	3	10	

**Рис. 24.29.** Рейтинги способности игроков играть на определенных позициях

Начнем с выделения диапазона ячеек C3:E22, содержащего рейтинги игроков. В раскрывающемся списке Условное форматирование (Conditional Formatting) щелкнем на Управление правилами (Manage Rules). Нажмите Создать правило (New Rule) и выберите Использовать формулу для определения форматируемых ячеек (Use a formula...). Заполните диалоговое окно, как показано на рис. 24.30.

Формула =A3=C\$1 сравнивает позиции игроков с заголовками столбцов 1(2 или 3) в строке 1. Если для позиции игрока установлено значение 1 (защитник), ячей-



**Рис. 24.30.** Настройки в диалоговом окне Создание правила форматирования для выделения рейтингов игроков красным цветом

ка с его рейтингом в столбце С (рейтингом в защите) выделяется красным. Аналогично если для позиции игрока установлено значение 2, ячейка с его рейтингом в нападении в столбце D выделяется красным. Наконец, если для позиции игрока установлено значение 3, ячейка с его рейтингом при игре на месте центрального в столбце Е выделяется красным. Учтите, если не поставить знак доллара перед А и 1 в формуле, формула будет скопирована неправильно.

Хочу также отметить, что в Excel 2019 допускается создание условных форматов с формулами, включающими ссылки на данные на других листах.

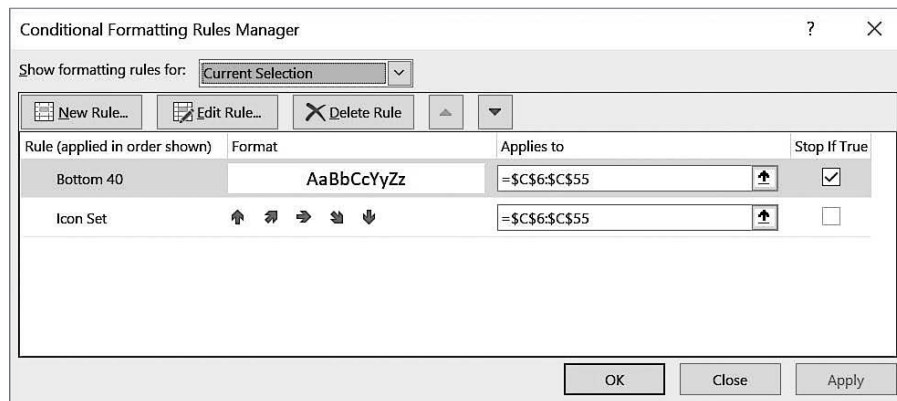
**❓ Для чего нужны флажки «Остановить, если истина» (Stop If True) в диалоговом окне «Управление правилами» (Manage Rules)?**

Предположим, что для правила установлен флажок Остановить, если истина (Stop If True). Если ячейка удовлетворяет этому правилу, все правила с более низким приоритетом игнорируются. Используем файл Income.xlsx для иллюстрации работы с установленным флажком Остановить, если истина (Stop If True). В этом файле содержатся данные о средних показателях доходов для всех штатов США между 1984 и 2010 г. Пусть (как показано на рис. 24.31) требуется отметить 10 штатов с самыми высокими показателями доходов в 2010 г. стрелками вверх, а другие штаты не отмечать значками. Ключ к решению задачи — создать первое правило, в котором отсутствует форматирование для 40 самых низких показателей доходов, и установить флажок Остановить, если истина (Stop If True). Затем создать требуемый набор значков.

	A	C
3		2010
4	State	Median income
5		
6	Alabama	40,976
7	Alaska	58,198
8	Arizona	47,279
9	Arkansas	38,571
10	California	54,459
11	Colorado	60,442
12	Connecticut	66,452
13	Delaware	55,269
14	Florida	44,243
15	Georgia	44,108
16	Hawaii	58,507
17	Idaho	47,014
18	Illinois	50,761
19	Indiana	46,322
20	Iowa	49,177
21	Kansas	46,229
22	Kentucky	41,236
23	Louisiana	39,443
24	Maine	48,133
25	Maryland	64,025
26	Massachusetts	61,333
27	Michigan	46,441
28	Minnesota	52,554
29	Mississippi	37,985
30	Missouri	46,184
31	Montana	41,467
32	Nebraska	52,728
33	Nevada	51,525
34	New Hampshire	66,707

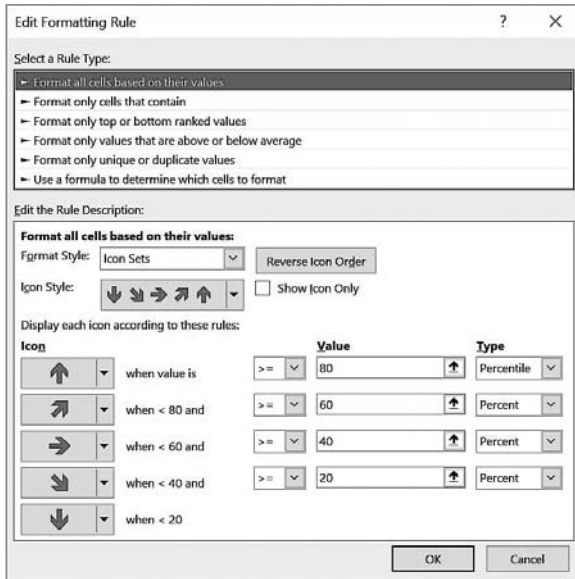
**Рис. 24.31.** Десять штатов с самыми высокими показателями доходов отмечены стрелками вверх

Параметры, которые скрывают стрелки для 40 штатов с самыми низкими показателями доходов в 2010 г., показаны на рис. 24.32.



**Рис. 24.32.** Параметры, гарантирующие, что только 10 самых высоких показателей доходов будут отмечены стрелками вверх

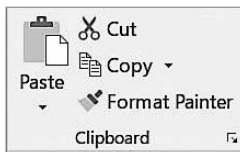
Параметры для второго правила показаны на рис. 24.33. Обратите внимание, что 20% штатов это  $0,2 \times 50 = 10$  штатов. Остальные параметры не имеют значения.



**Рис. 24.33.** Параметры, гарантирующие, что 10 штатов с самыми высокими показателями доходов будут отмечены стрелками вверх

### ❓ Как с помощью инструмента Формат по образцу скопировать у словный формат?

Значок Формат по образцу (Format Painter) в виде кисти (рис. 24.34) позволяет применить формат (включая условные форматы) из любой ячейки или группы ячеек в любой другой группе ячеек. Выделите ячейку или группу ячеек с форматом, который требуется скопировать, и щелкните на Формат по образцу (Format Painter). С помощью кисти выберите ячейки, в которые требуется скопировать формат.



**Рис. 24.34.** Значок Формат по образцу

Если требуется скопировать формат в диапазон, включающий несмежные ячейки, дважды щелкните по значку Формат по образцу (Format Painter) и выберите, как

обычно, все ячейки, в которые требуется скопировать формат . Для отключения инструмента щелкните по значку **Формат по образцу (Format Painter)** еще раз.

## Задания

1. На основе данных файла **Sandp.xlsx** примените **Условное форматирование** в следующих ситуациях:
  - используйте полужирный шрифт для каждого месяца, в котором фондовый индекс S&P рос, и подчеркивание для каждого месяца, в котором он снижался;
  - выделите зеленым цветом каждый месяц, в котором индекс S&P изменился максимум на 2%;
  - выделите наибольший индекс S&P красным цветом, а наименьший — фиолетовым.
2. На основе данных файла **Toysrusformat.xlsx** выделите доходы во всех четвертых кварталах синим цветом, а во всех первых кварталах — красным.
3. В файле **Test.xlsx** содержатся данные об экзаменационных баллах студентов. Десять лучших студентов получают оценку А, следующие 20 студентов — оценку В, прочие — оценку С. Выделите баллы, соответствующие оценке А, красным цветом, оценке В — зеленым цветом и оценке С — синим цветом. Подсказка: функция **НАИБОЛЬШИЙ(D4:D63;10)** возвращает десятый по величине экзаменационный балл.
4. В файле **Weekendformatting.xlsx** выделите все субботы и воскресенья красным цветом. Выделите синим цветом все дни, которые приходятся на первые 10 дней месяца. Какой цвет будет иметь преимущество, если оба цвета применяются к одной ячейке?
5. Все работники финансового отдела Microsoft были распределены по четырем группам. Руководитель каждой группы оценил своих работников по шкале 0–10 баллов, а каждый работник оценил удовлетворенность от работы с каждой из четырех групп. (См. файл **Satissuper.xlsx**.) Для каждого работника в соответствии с группой, в которую он распределен, выделите оценку руководителя и рейтинг удовлетворенности сотрудника работой в этой группе.
6. В файле **Varianceanalysis.xlsx** содержатся ежемесячные прогнозы прибыли и ежемесячные фактические продажи. Отклонение оборота за месяц составляет:

$$\frac{\text{фактические\_продажи} - \text{прогнозируемые\_продажи}}{\text{прогнозируемые\_продажи}}$$

Выделите красным цветом все месяцы с благоприятным отклонением не менее 20% и зеленым цветом все месяцы с неблагоприятным отклонением более 20%.

7. На основе данных для примера с разработкой препаратов (файл `Offsetcost.xlsx` из главы 22) отобразите на листе все затраты на первом этапе красным цветом, все затраты на втором этапе — зеленым цветом и все затраты на третьем этапе — фиолетовым цветом.
8. Файл `Names.xlsx` содержит список имен. Выделите все повторы зеленым цветом, а все имена, содержащие символы `Ja`, — красным цветом.
9. В файле `Duedates.xlsx` находятся данные о сроках оплаты различных счетов. Выделите красным цветом все счета, подлежащие оплате в конце следующего месяца.
10. В файле `Historicalinvest.xlsx` создайте Условное форматирование с тремя значками и отметьте ячейки с верхними 10% доходов стрелкой вверх, нижними 10% доходов — стрелкой вниз и промежуточными 80% доходов — горизонтальной стрелкой.
11. Файл `Nbasalaries.xlsx` содержит зарплаты игроков НБА в миллионах долларов. Создайте гистограммы, обобщающие эти данные. Зарплате менее 1 млн долл. должен соответствовать самый короткий столбец гистограммы, а зарплате более 15 млн долл. — самый длинный.
12. Создайте трехцветную шкалу для обобщения данных о зарплате игроков НБА. Измените цвет для нижних 10% всех зарплат на зеленый, а для верхних 10% — на красный.
13. Обобщите данные о зарплатах игроков НБА с помощью пяти значков. Создайте точки перелома для значков при 3, 6, 9 и 12 млн долл.
14. Создайте правило, при котором значки для игроков с зарплатой от 7 до 8 млн долл. не отображаются. Подсказка: используйте параметр **Остановить, если истина** (`Stop If True`).
15. Файл `Fractiondefective.xlsx` содержит данные о проценте дефектных единиц продукции, производимых в день. Количество считается приемлемым, если в день было произведено не более 2% единиц продукции с дефектами. Отметьте все приемлемые дни зеленым флажком. На другом листе отметьте все приемлемые дни красным флажком. Подсказка: установите флажок **Остановить, если истина** (`Stop If True`) для гарантии, что значок не появится в какой-либо ячейке, содержащей число больше 2.
16. Файл `Globalwarming2011.xlsx` содержит средние мировые температуры за 1881–2011 гг. Обобщите данные о глобальном потеплении с помощью трехцветной шкалы. Отметьте более низкие температуры синим цветом, промежуточные температуры желтым цветом и более высокие температуры красным цветом.
17. Вы откладываете деньги на высшее образование для своего ребенка. Вы кладете на счет одинаковую сумму в конце каждого года. Ваша цель — накопить \$100 000. Определите, каким должен быть ежегодный вклад при ежегодных доходах по инвестициям от 4 до 15% и сроке инвестирования от 5 до 15 лет. Допустим, вы можете откладывать \$10 000 в год. С помощью условного фор-

матирования выделите минимальный срок накопления \$100 000 для каждой годовой ставки доходности.

18. Файл **Amazon.xlsx** содержит ежеквартальные доходы **Amazon.com**. Примените условное форматирование для выделения объемов продаж в разных кварталах различными цветами.
19. С помощью условного форматирования раскрасьте ячейки **A1:H8** в шахматном порядке, чередуя черный и белый цвета. Подсказка: функция **СТРОКА()** возвращает номер строки ячейки, а функция **СТОЛБЕЦ()** — номер столбца.
20. Когда студенты сдают Вивиан экзамен по бухгалтеру, они должны ввести в систему Scantron свои адреса электронной почты. К сожалению, они часто ошибаются. У Вивиан есть список фактических адресов электронной почты студентов группы. Она хочет выделить желтым цветом введенные студентами неправильные адреса электронной почты. Тогда она сможет их легко исправить. Необходимые данные находятся в файле **Scantrons.xlsx**. Помогите Вивиан!
21. Файл **Top5.xlsx** содержит данные о доходах 50 компаний. Для каждого года отметьте 5 компаний с самыми большими доходами. Подсказка: функция **НАИБОЛЬШИЙ(B1:B50;5)** возвращает пятое по величине число в диапазоне **B1:B50**.
22. Файл **Threetimes.xlsx** содержит список имен. Выделите красным цветом каждое имя, которое встречается не менее трех раз.
23. Файл **GNP.xlsx** содержит данные о валовом национальном продукте США. Воспользуйтесь тем, что в гистограммах есть возможность обработки отрицательных значений, и обобщите ежеквартальные темпы роста ВВП.
24. На основе данных из файла **Globalwarming2011temp.xlsx** (из папки **Templates**) отметьте годы (а не температуры), в которые температура превышала средние значения.
25. В файле **Accountsums.xlsx** (рис. 24.35) представлены еженедельные зарплаты в тысячах долларов (перечисленные в хронологическом порядке), выплаченные нескольким высококлассным консуль тантам. Например, Бритни (Britney) за первую неделю получила \$91 000, а за последнюю — \$5700.
  - Используйте тип правила **Использовать формулу для определения форматированных ячеек (Use a formula...)** для выделения имени каждого консуль танта желтым цветом.
  - Используйте тип правила **Использовать формулу для определения форматированных ячеек (Use a formula...)** для выделения зарплаты каждого консуль танта за первую неделю оранжевым цветом.
26. Файл **Shading.xlsx** содержит 27 квартальных объемов продаж.
  - Выделите ячейки поочередно зеленым и желтым цветом.
  - Выделите по четыре ячейки поочередно зеленым и желтым цветом.
27. Файл **Problem27data.xlsx** содержит имена нескольких человек. Нужно вводить имя в ячейку **E2**. Используйте условное форматирование, чтобы при вводе имеющегося в файле имени оно выделялось.



	B	C
4	Britney	
5		86
6		52
7		79
8		76
9		82
10		79
11		82
12		96
13		94
14		81
15		90
16		94
17		51
18		83
19		51
20	Lindsay	
21		80
22		87
23		75
24		71
25		93

**Рис. 24.35.** Данные для задания 25

28. В файле **Problem28data.xlsx** содержится информация об избирателях по 20 избирательным округам. В ячейку **I2** введите номер избирательного округа. Ваша задача состоит в том, чтобы задать форматирование, которое выделит все строки данных, относящиеся к округу, указанному в ячейке **I2**.
29. В файле **Problem29data.xlsx** столбцы **E–I** содержат следующую информацию об акциях: начальная цена, владение акцией на начало дня, куплена, продана, владение акцией на конец дня. Используя условное форматирование, выделите все строки с данными в дни, когда акция была продана.
30. Файл **Problem30data.xlsx** содержит статистику по квотербекам НФЛ за 2014 г. В ячейке **J2** введите число ярдов  $x$ . Используя условное форматирование, выделите всю информацию по всем квотербекам, имеющим значение  $YDS \geq x$ .
31. Используя данные в файле с именем *Amazon.xlsx*, найдите кварталы, в которых продажи превышали продажи в течение каждого из последних двух кварталов.
32. В файле **Problem32data.xlsx** содержится информация о 2423 транзакциях. Предполагается, что возможно мошенничество, если выполняется одно из следующих условий:
  - сумма чека составляет не менее 500 долларов, и никто не авторизовал чек;
  - сумма чека составляет не менее 500 долларов, и тот же человек подтвердил и обналичил чек.

Выделите все подозрительные транзакции.

33. Файл **Problem33data.xlsx** содержит информацию о 2000 компаниях. В **K3** вводится страна, а в **L3** — отрасль. Затем в электронной таблице должна быть выделена вся информация для компаний в выбранной стране и отрасли.
34. Файл **Problem34data.xlsx** содержит адреса, на которые были отправлены чеки. Если адрес включает в себя **Р.О. Вох**, то, скорее всего, транзакция была мошеннической. Используйте условное форматирование, чтобы выделить все адреса, содержащие **Р.О. Вох**. Подсказка: функция **ПОИСК** не чувствительна к регистру, а функция **ЕОШИБКА** выдает **Истина** тогда и только тогда, когда формула выдает ошибку.
35. Файл **Problem35data.xlsx** содержит списки счетов, которые были обналичены. Иногда номер чека пропускали. Используйте условное форматирование, чтобы выделить все чеки, где предыдущий номер чека не был обналичен. Например, чек 15 должен быть выделен, потому что чек 14 не был обналичен.
36. В файле **Задача 36** содержатся следующие данные для возмещений: номер счета, сумма и номер поставщика. Выделите все транзакции, где сумма и номер счета-фактуры происходят более одного раза.

## ГЛАВА 25

# Сортировка в Excel

### Обсуждаемые вопросы

- Как можно отсортировать данные торговых сделок сначала по продавцам, потом по продуктам, по количеству проданных единиц товара и в хронологическом порядке от самых старых до самых новых?
- Мне всегда хотелось отсортировать данные на основе цвета ячеек или цвета шрифта. Возможно ли это в Excel 2016?
- Мне нравятся большие наборы значков, описанные в главе 24. Можно ли отсортировать данные на основе значков в ячейке?
- На листе имеется столбец, в котором указан месяц для каждой продажи. При сортировке по этому столбцу я получаю первым или первый месяц по алфавиту, или последний. Как можно отсортировать данные по этому столбцу в хронологическом порядке, то есть сначала получить сделки за январь, затем за февраль и т. д.?
- Можно ли отсортировать данные без диалогового окна Сортировка (Sort)?
- Мне часто требуется напечатать список городов, в которых моя компания имеет служебные подразделения. Имеется ли возможность создать свой список, а потом при вводе только первого города из списка и перетаскивании курсора вниз автоматически заполнить ячейки остальными городами из этого списка?

Почти каждому пользователю Microsoft Excel приходилось хотя бы раз сортировать столбцы данных по алфавиту или по числовым значениям. Вот несколько примеров того, насколько эффективно работает сортировка в Excel 2019.

## Ответы на вопросы

**?** Как можно отсортировать данные торговых сделок сначала по продавцам, потом по продуктам, по количеству проданных единиц товара и в хронологическом порядке от самых старых до самых новых?

JAC — небольшая компания, продающая косметику В файле Makeupsorttemp.xlsx из папки Templates на листе Makeup (рис. 25.1) содержатся следующие данные о продажах:

- номер сделки;
- имя продавца;
- дата сделки;

- проданный продукт;
- количество;
- полученная сумма в долларах;
- место проведения сделки.

	E	F	G	H	I	J	K
4	1	Betsy	4/1/2004	lip gloss	45	\$ 137.20	south
5	2	Hallagan	3/10/2004	foundation	50	\$ 152.01	midwest
6	3	Ashley	2/25/2005	lipstick	9	\$ 28.72	midwest
7	4	Hallagan	5/22/2006	lip gloss	55	\$ 167.08	west
8	5	Zaret	6/17/2004	lip gloss	43	\$ 130.60	midwest
9	6	Colleen	11/27/2005	eye liner	58	\$ 175.99	midwest
10	7	Cristina	3/21/2004	eye liner	8	\$ 25.80	midwest
11	8	Colleen	12/17/2006	lip gloss	72	\$ 217.84	midwest
12	9	Ashley	7/5/2006	eye liner	75	\$ 226.64	south
13	10	Betsy	8/7/2006	lip gloss	24	\$ 73.50	east
14	11	Ashley	11/29/2004	mascara	43	\$ 130.84	east
15	12	Ashley	11/18/2004	lip gloss	23	\$ 71.03	west
16	13	Emilee	8/31/2005	lip gloss	49	\$ 149.59	west
17	14	Hallagan	1/1/2005	eye liner	18	\$ 56.47	south
18	15	Zaret	9/20/2006	foundation	-8	\$ (21.99)	east
19	16	Emilee	4/12/2004	mascara	45	\$ 137.39	east
20	17	Colleen	4/30/2006	mascara	66	\$ 199.65	south
21	18	Jen	8/31/2005	lip gloss	88	\$ 265.19	midwest
22	19	Jen	10/27/2004	eye liner	78	\$ 236.15	south
23	20	Zaret	11/27/2005	lip gloss	57	\$ 173.12	midwest
24	21	Zaret	6/2/2006	mascara	12	\$ 38.08	west

**Рис. 25.1.** Данные сделок до сортировки

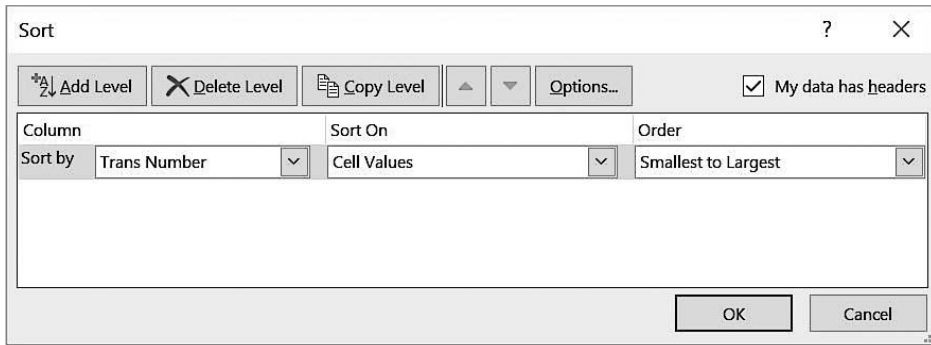
Я хочу отсортировать данные с соблюдением следующих условий.

- Сделки должны быть перечислены по продавцам в алфавитном порядке. Требуется отсортировать их в обычном порядке, от А до Z, то есть сначала все сделки Ashley (Эшли), а все сделки Zaret (Зарет) в конце списка.
- Сделки каждого продавца сортируются по продуктам в алфавитном порядке. Таким образом, сделки Эшли (Ashley) с подводкой для глаз (eye liner) должны следовать перед сделками с основной под макияж (foundation) и т. д.
- Для каждого продавца и продукта сделки перечисляются по количеству проданных единиц продукции (в порядке убывания).
- Если у продавца две или более сделок с одним и тем же продуктом при одинаковом количестве единиц, сделки указываются в хронологическом порядке.

До версии Excel 2007 было трудно выполнять сортировку более чем по трем критериям. Теперь Excel позволяет применять до 64 критериев в одной сортировке. Для сортировки данных по сделкам я должен сначала выделить данные (диапазон ячеек E3:K1894). Вот два простых способа:

- установите курсор в левом верхнем углу данных (ячейка E3), нажмите **Ctrl+Shift+→** и затем **Ctrl+Shift+↓**;
- установите курсор в любом месте требуемого диапазона и нажмите **Ctrl+Shift+\***.

Затем на вкладке **Данные (Data)** в группе **Сортировка и фильтр (Sort & Filter)** я щелкнул на **Сортировка (Sort)** для открытия диалогового окна **Сортировка (Sort)**, показанного на рис. 25.2.

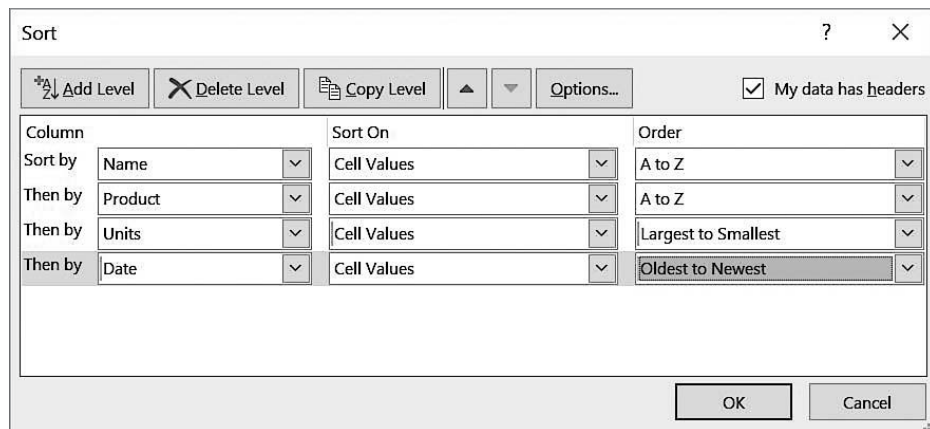


**Рис. 25.2.** Не заполненное данными диалоговое окно Сортировка

Поскольку в строке 3 находятся заголовки столбцов данных, я установил флажок **Мои данные содержат заголовки (My data has headers)**. Теперь я выберу следующие четыре критерия в указанном порядке:

1. Сортировать по столбцу **Имя** значения (это означает содержимое ячеек) в порядке от А до Я. В списке **Столбец (Column)** вы выбираете **Имя (Name)**. На верхнем уровне сортировки в строке **Сортировать по (Sort By)** в списке **Сортировка (Sort On)** оставляете значение по умолчанию **Значения (Values)**. В списке **Порядок (Order)** стоит по умолчанию **От А до Я (A To Z)**, и для этой сортировки вы оставляете этот порядок.
2. Сортировать по столбцу **Продукт** значения в порядке от А до Я. Нажимаете пункт меню **Добавить уровень (Add Level)**, под уровнем **Сортировать по (Sort By)** добавляется уровень **Затем по (Then By)**. В списке **Столбец** вы выбираете **Продукт (Product)**, в списке **Сортировка** оставляете значение **Значения (Values)** и в списке **Порядок** — **От А до Я (A To Z)**.
3. Сортировать по столбцу **Количество** значения в порядке по убыванию. Нажимаете пункт меню **Добавить уровень (Add Level)**, на новом уровне выбираете **Количество (Units)** в списке **Столбец**, в списке **Сортировка** оставляете значение **Значения** и выбираете **По убыванию (Largest To Smallest)** в списке **Порядок**.
4. Сортировать по столбцу **Дата** значения в хронологическом порядке, от старых к новым. Добавляете уровень, выбираете **Дата (Date)** в списке **Столбец** и оставляете значение по умолчанию **От старых к новым (Oldest To Newest)** в списке **Порядок**.

Теперь диалоговое окно выглядит так, как на рис. 25.3.



**Рис. 25.3.** Диалоговое окно Сортировка для сортировки данных в примере со сделками

Нажимаю ОК, чтобы применить заданные параметры сортировки. Результат сортировки показан на рис. 25.4.

	E	F	G	H	I	J	K
3	Trans Num	Name	Date	Product	Units	Dollars	Location
4	785	Ashley	4/10/2005	eye liner	92	\$ 278.34	east
5	1879	Ashley	8/18/2006	eye liner	90	\$ 271.85	midwest
6	1685	Ashley	11/5/2005	eye liner	88	\$ 265.96	south
7	1737	Ashley	3/28/2006	eye liner	88	\$ 265.53	east
8	1579	Ashley	6/6/2004	eye liner	87	\$ 262.85	east
45	1858	Ashley	7/7/2005	foundation	95	\$ 286.83	south
46	1491	Ashley	7/18/2005	foundation	93	\$ 281.34	south
47	551	Ashley	4/19/2006	foundation	93	\$ 281.17	midwest
201	1290	Betsy	7/7/2005	eye liner	95	\$ 286.61	east
202	1777	Betsy	12/19/2005	eye liner	95	\$ 286.41	midwest
203	855	Betsy	1/23/2005	eye liner	94	\$ 284.42	south
204	1609	Betsy	4/10/2005	eye liner	94	\$ 283.46	east
205	1483	Betsy	7/29/2005	eye liner	94	\$ 284.31	midwest
206	735	Betsy	5/2/2005	eye liner	91	\$ 274.96	west
207	1509	Betsy	4/30/2006	eye liner	83	\$ 251.40	midwest
208	91	Betsy	8/18/2006	eye liner	83	\$ 251.18	south
209	872	Betsy	9/20/2006	eye liner	79	\$ 239.06	south
210	487	Betsy	1/23/2005	eye liner	75	\$ 227.10	west

**Рис. 25.4.** Отсортированные данные по сделкам

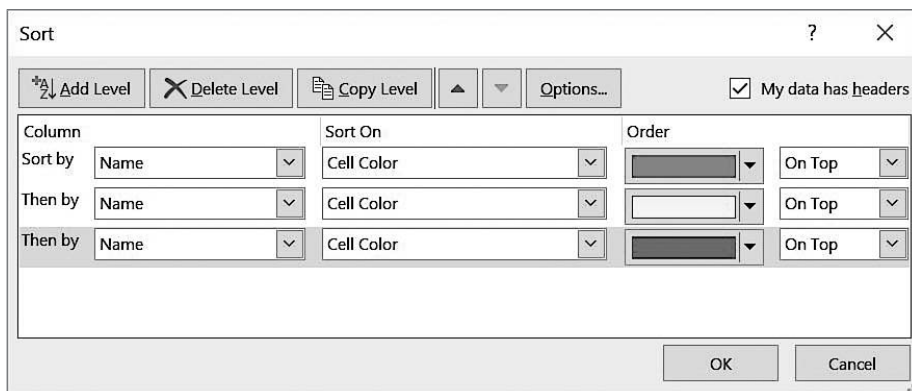
Обратите внимание, что первыми перечислены сделки Эшли с подводкой для глаз, затем с основой для макияжа и т. д. Сделки с подводкой для глаз перечислены по

убыванию количества проданных единиц товара. Если количество проданных единиц товара совпадает (см. строки 6 и 7), сделки указаны в хронологическом порядке.

В диалоговом окне **Сортировка (Sort)** можно добавить критерии сортировки ( **Добавить уровень (Add Level)**), удалить критерии ( **Удалить уровень (Delete Level)**), скопировать настройки, определяющие уровень сортировки ( **Копировать уровень (Copy Level)**), и указать, есть ли в выделенных данных заголовки. Нажав **Параметры (Options)**, можно установить флажок для учета регистра при сортировке или даже отсортировать данные, находящиеся в разных столбцах (вместо более привычной сортировки по строкам).

**❓ Мне всегда хотелось отсортировать данные на основе цвета ячеек или цвета шрифта. Возможно ли это в Excel 2019?**

В Excel 2019 сортировка на основе цвета ячеек или цвета шрифта выполняется элементарно. Рассмотрим в файле **Makeupsorttemp.xlsx** лист **Makeup**. Некоторые имена в столбце **F** выделены разными цветами. Например, ячейка **F31** с именем **Cici** выделена красным, а ячейка **F23** с именем **Colleen** — желтым. Допустим, вы хотите перенести в верхнюю часть списка имена в ячейках с зеленой заливкой, за ними расположить ячейки с желтой заливкой и затем с красной. Остальные строки должны располагаться под ними. Для сортировки столбца **Имя** по цвету выделите диапазон, который требуется отсортировать (**E3:K1894**), на вкладке **Данные (Data)** выберите инструмент **Сортировка (Sort)** и нажмите **Добавить уровень (Add Level)**. Выберите столбец **Имя**, затем в раскрывающемся списке **Сортировка (Sort On)** выберите **Цвет ячейки (Cell Color)**. (Выберите **Цвет шрифта (Font Color)**, если необходимо отсортировать ячейки по цвету шрифта.) Для первого уровня из раскрывающегося списка **Порядок (Order)** выберите зеленый цвет и оставьте упорядочивание по умолчанию **Сверху (On Top)**. Затем добавьте еще два уровня; выберите для второго уровня желтый цвет и для третьего уровня красный цвет. Окончательный вид диалогового окна показан на рис. 25.5. Результат сортировки представлен на листе **Colors** в файле **Makeupsort.xlsx** в папке (**Practice Files**) и на рис. 25.6.



**Рис. 25.5.** Диалоговое окно **Сортировка** с настройками для сортировки по цвету ячеек



	E	F	G	H	I	J	K
3	Trans Number	Name	Date	Product	Units	Dollars	Location
4	105	Cristina	9/13/2004	lipstick	51	\$ 155.30	midwest
5	165	Hallagan	12/19/2005	foundatio	25	\$ 76.99	east
6	86	Jen	8/9/2005	eye liner	-2	\$ (4.24)	east
7	23	Colleen	2/1/2006	mascara	25	\$ 77.31	midwest
8	14	Hallagan	1/1/2005	eye liner	18	\$ 56.47	south
9	33	Cici	6/17/2004	mascara	41	\$ 125.27	west
10	785	Ashley	4/10/2005	eye liner	92	\$ 278.34	east
11	1879	Ashley	8/18/2006	eye liner	90	\$ 271.85	midwest
12	1685	Ashley	11/5/2005	eye liner	88	\$ 265.96	south
13	1737	Ashley	3/28/2006	eye liner	88	\$ 265.53	east

**Рис. 25.6.** Результат сортировки по цвету ячеек

**?** Мне нравятся большие наборы значков, описанные в главе 24. Можно ли отсортировать данные на основе значков в ячейке?

Для сортировки по значкам в диалоговом окне Сортировка (Sort) в раскрывающемся списке Сортировка (Sort On) выберите Значок ячейки (Cell Icon). В раскрывающемся списке Порядок (Order) выберите для первого уровня значок, отмечающий ячейки, которые требуется поместить в верхнюю часть списка, и добавьте новый уровень для каждого значка, по которому вы хотите отсортировать данные.

**?** На листе имеется столбец, в котором указан месяц для каждой продажи. При сортировке по этому столбцу я получаю первым или первый месяц по алфавиту, или последний. Как можно отсортировать данные по этому столбцу в хронологическом порядке, то есть сначала получить сделки за январь, затем за февраль и т. д.?

В файле Makeupsorttemp.xlsx на листе Dates находится список месяцев (рис. 25.7). Я хотел бы отсортировать месяцы в хронологическом порядке, начиная с января. Начнем с выделения диапазона D6:D15 и укажем сортировку столбца D по значению. Для порядка сортировки я выбрал Настраиваемый список (Custom List), а затем — янв, фев, мар, апр, .... (jan, feb, mar,...). Обратите внимание, что данные можно также отсортировать по дням недели. Я нажал ОК для подтверждения выбора и вернулся в диалоговое окно Сортировка. Окончательный вид диалогового окна Сортировка (Sort) представлен на рис. 25.8, а результаты сортировки — на рис. 25.9.

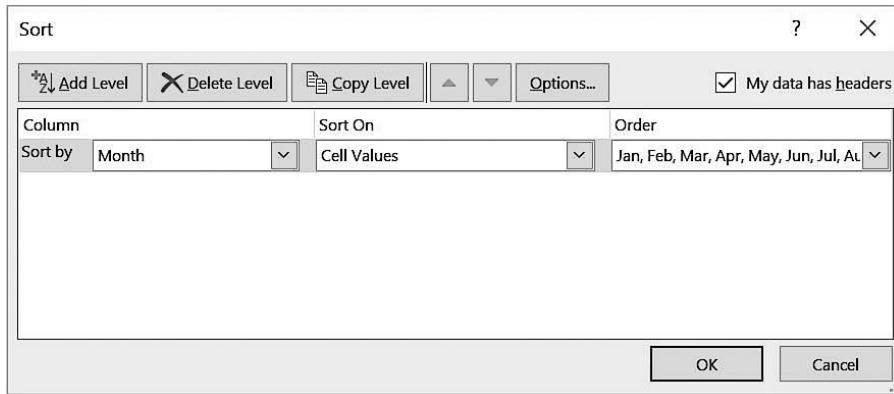
Обратите внимание, что в диалоговом окне Списки (Custom Lists) можно создать собственный список порядка сортировки. Выберите Новый список (New List); в разделе Элементы списка (List Entries) введите элементы в порядке, требуемом для сортировки, и нажмите Добавить (Add). Новый список будет включен в меню

	D
5	Month
6	Feb
7	Feb
8	Jan
9	Aug
10	Mar
11	Apr
12	May
13	Jan
14	Feb
15	Feb

**Рис. 25.7.** Месяцы, которые необходимо отсортировать



выбора. Например, если ввести в разделе **Элементы списка (List Entries)** Jack, John и Alan (на отдельных строках или через запятую), все записи с именем Jack будут в списке первыми, за ними будут следовать записи с именем John и завершат этот список записи с именем Alan.



**Рис. 25.8.** Диалоговое окно Сортировка, настроенное на сортировку по месяцам

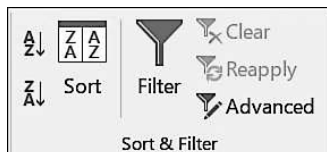
	D
5	Month
6	Jan
7	Jan
8	Feb
9	Feb
10	Feb
11	Feb
12	Mar
13	Apr
14	May
15	Aug

**Рис. 25.9.** Месяцы, отсортированные в хронологическом порядке

### ❓ Можно ли отсортировать данные без диалогового окна Сортировка (Sort)?

Иногда удобнее сортировать данные без диалогового окна **Сортировка (Sort)**. Предположим, что нам снова нужно отсортировать данные сделок в файле **Makeupsorttemp.xlsx** на листе **Makeup** сначала по продавцам, затем по продуктам, по количеству проданных единиц товара и, наконец, в хронологическом порядке от старых к новым. Для начала выделите наименее важный столбец для сортировки, которым является столбец с датами (**G3:G1894**). Затем на вкладке **Данные (Data)** в группе **Сортировка и фильтр (Sort&Filter)** нажмите кнопку **Сортировка от А до Я (Sort A To Z)**, показанную

на рис. 25.10, и в появившемся диалоговом окне, оставив переключатель в положении автоматически расширить выделенный диапазон (Expand The Selection), кнопку Сортировка (Sort) для сортировки всех столбцов. Кнопка Сортировка от А до Я (Sort A To Z) предназначена для сортировки текстовых данных от А до Я, а также для сортировки числовых данных по возрастанию и сортировки дат, так что в верхней части списка окажутся меньшие числа или более старые даты.



**Рис. 25.10.** Значок Сортировка для сортировки без диалогового окна Сортировка

Кнопка Сортировка от Я до А (Sort Z To A) предназначена для сортировки текстовых данных от Я до А, а также для сортировки числовых данных по убыванию и сортировки дат, так что в верхней части списка окажутся самые большие числа и самые новые даты.

Далее отсортируйте второй наименее важный столбец (**Количество**) от Я до А, поскольку более крупные сделки должны находиться в верхней части списка. Затем отсортируйте столбец **Продукт** от А до Я и, наконец, столбец с именами продавцов от А до Я. Эта последовательность действий даст те же результаты, которые показаны ранее на рис. 25.4.

**?** Мне часто требуется напечатать список городов, в которых моя компания имеет служебные подразделения. Имеется ли возможность создать свой список, а потом при вводе только первого города из списка и перетаскивании курсора вниз автоматически заполнить ячейки остальными городами из этого списка?

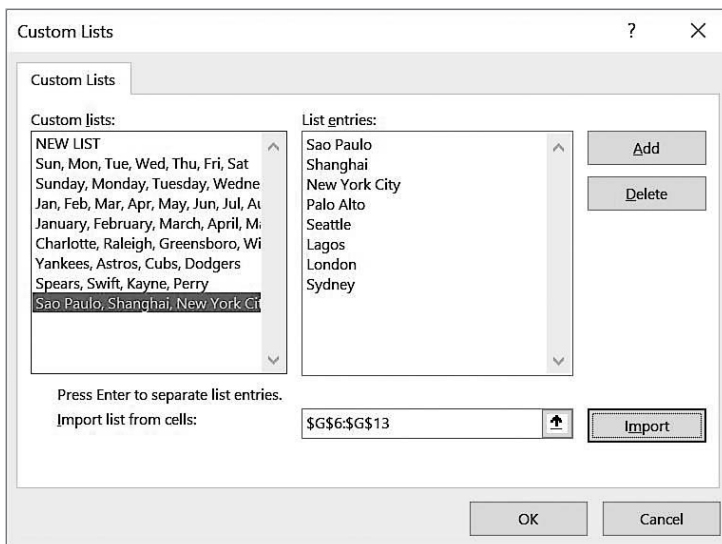
Предположим, ваша компания имеет подразделения в городах, приведенных на рис. 25.11 (см. файл Listtemp.xlsx в папке Templates).

	G
6	Sao Paulo
7	Shanghai
8	New York City
9	Palo Alto
10	Seattle
11	Lagos
12	London
13	Sydney

**Рис. 25.11.** Перечень городов, из которого требуется создать свой список

Для создания списка сделайте следующее:

- Нажмите вкладку **Файл (File)**, выберите **Параметры (Options)** в нижней части левой панели, в левой части панели диалогового окна **Параметры Excel (Excel Options)** нажмите **Дополнительно (Advanced)**, прокрутите экран вниз и в нижней части раздела **Общие (General)** нажмите кнопку **Изменить списки...** (**Edit Custom List**).
- Откроется диалоговое окно **Списки (Custom Lists)**, приведенное на рис. 25.12. В поле **Импорт списка из ячеек (Import List From Cells)** задайте диапазон (**G6:G13**), содержащий список, и нажмите кнопку **Импорт (Import)**. Вы увидите, что ваш список городов добавился в набор встроенных списков. Нажмите **OK**, чтобы закрыть окно **Списки (Custom Lists)**, и еще раз **OK**, чтобы закрыть окно **Параметры Excel (Excel Options)**.



**Рис. 25.12.** Создан пользовательский список городов

Теперь, если вы введете Sao Paulo в любой ячейке, наведете курсор на квадратик в нижнем правом углу этой ячейки и протащите курсор вниз, вы увидите названия городов, перечисленные в том порядке, который был задан вами в списке (Sao Paulo, Shanghai и т. д.). Если вы протащите курсор вниз достаточно сильно, список начнет повторяться. Конечно, если вам требуется создать две копии списка, введите Sao Paulo в двух ячейках и затем протащите вниз эти две ячейки.

## Задания

1. В файле **Makeupsort.xlsx** отсортируйте данные сделок в алфавитном порядке по месту проведения сделки, затем по типу продукта, имени, дате сделки и количеству проданных единиц товара.

2. Файл **Sortday.xlsx** содержит данные об отработанных часах в разные дни недели. Отсортируйте данные по дням недели, то есть за данными для понедельника должны следовать данные для вторника и т. д.
3. Файл **Sorticons.xlsx** содержит данные о годовых доходах по инвестициям, где зеленой стрелкой вверх отмечены удачные годы, желтой горизонтальной стрелкой — относительно удачные годы и красной стрелкой вниз — неудачные годы. Отсортируйте данные в столбце **Stocks** по значкам: сначала стрелки вверх, затем горизонтальные стрелки, и за ними стрелки вниз.
4. В файле **Makeupsortfont.xlsx** находятся данные о косметической продукции, в которых некоторые даты записаны синим, красным или коричневым шрифтом. Отсортируйте данные следующим образом: в верхней части списка даты, записанные коричневым шрифтом, за ними даты, записанные красным шрифтом, и затем даты, записанные синим шрифтом.

## ГЛАВА 26

# Таблицы

### Обсуждаемые вопросы

- Я указал на листе количество проданных единиц товара и общую выручку для каждого продавца, и теперь легко могу вычислить среднюю цену единицы товара для каждого продавца. Как создать формат, который при вводе новых данных копировался бы автоматически? Кроме того, существует ли простой способ автоматически скопировать формулы при добавлении новых данных?
- Я ввел в книгу цены на природный газ за несколько лет и создал линейную диаграмму, отображающую ежемесячное изменение цен. Можно ли добиться, чтобы при добавлении новой информации о ценах диаграмма обновлялась автоматически?
- Для каждой торговой сделки указаны продавец, дата, продукт, место продажи и объем продажи. Можно ли подсчитать, например, общий объем продаж помады на Востоке для Джен или Коллин?
- Как срезы (добавлены в Microsoft Excel 2013) позволяют выполнить срез данных в таблице Excel?
- Как сослаться на фрагмент таблицы в других частях листа?
- Применяются ли условные форматы автоматически к новым данным, добавляемым в таблицу?

Работая в Excel, мы часто добавляем в книгу новые данные, а затем вручную обновляем формулы, форматы и диаграммы. Скучища! Благодаря новым возможностям обработки таблиц она нам больше не грозит.

## Ответы на вопросы

**?** Я указал на листе количество проданных единиц товара и общую выручку для каждого продавца, и теперь легко могу вычислить среднюю цену единицы товара для каждого продавца. Как создать формат, который при вводе новых данных копировался бы автоматически? Кроме того, существует ли простой способ автоматически скопировать формулы при добавлении новых данных?

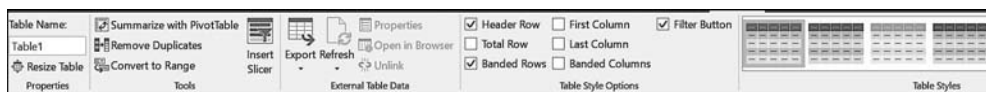
Файл Tableexampletemp.xlsx в папке Templates (рис. 26.1) содержит данные о количестве проданных единиц товара и выручке для каждого из шести продавцов. Известно, что новые данные будут добавлены начиная со строки 12, и в столбце H

потребуется вычислить среднюю цену единицы товара (выручка/количество) для каждого продавца. Кроме того, вы хотите создать хороший формат для отображения данных и формулу расчета средней цены, автоматически копируемую при добавлении новых данных.

	E	F	G
5	<b>Name</b>	<b>Units</b>	<b>Revenue</b>
6	<b>John</b>	<b>814</b>	<b>39886</b>
7	<b>Adam</b>	<b>594</b>	<b>26136</b>
8	<b>Dixie</b>	<b>528</b>	<b>13200</b>
9	<b>Tad</b>	<b>806</b>	<b>20956</b>
10	<b>Erica</b>	<b>826</b>	<b>27258</b>
11	<b>Gabrielle</b>	<b>779</b>	<b>28044</b>

**Рис. 26.1.** Данные для создания таблицы

Создание таблицы приведет к автоматическому обновлению формулы и форматирования при добавлении новых данных. Начните с выделения текущего диапазона и заголовков столбцов (E5:G11). Затем на вкладке **Вставка (Insert)** выберите **Таблица (Table)** или нажмите **Ctrl+T**. Убедитесь, что флажок **Таблица с заголовками (My table has headers)** выставлен, и нажмите **OK**. Таблица (E5:G11) прекрасно отформатирована. Это форматирование автоматически распространяется на новые данные, вводимые в таблицу. При работе с таблицей на вкладке **Конструктор (Design)**, показанной на рис. 26.2, доступно множество стилей и параметров. Эта вкладка отображается, только когда активная ячейка находится внутри таблицы. На вкладке можно выбрать стиль форматирования, который будет применяться к текущей таблице данных и к любым новым данным, добавляемым в нее.



**Рис. 26.2.** Конструктор таблиц

Обратите внимание, что заголовки таблицы снабжены раскрывающимися списками (рис. 26.3). В этих списках перечислены функциональные возможности для сортировки и фильтрации таблицы (фильтрация рассмотрена далее в этой главе).

Ячейки в выделенной таблице (за исключением заголовков) получают по умолчанию имя **Таблица1**. Переименуем ячейки в **Продажи** на вкладке **Конструктор (Design)** в группе **Свойства (Properties)**. Если затем перейти на вкладку **Формулы (Formulas)** и открыть **Диспетчер имен (Name Manager)**, можно увидеть, что диапазон E6:G12 теперь называется **Продажи**. Дополнительную привлекательность концепции диапазонов (и таблицы) придает автоматическое расширение диапазона при

	E	F	G
5	Name ▾	Units ▾	Revenue ▾
6	John	814	39886
7	Adam	594	26136
8	Dixie	528	13200
9	Tad	806	20956
10	Erica	826	27258
11	Gabrielle	779	28044

**Рис. 26.3.** Таблица для форматирования с фильтрами

добавлении новых строк в нижнюю часть таблицы и новых столбцов в ее правую часть. В главе 22 динамический диапазон мы создавали с помощью функции **СМЕЩ** (**OFFSET**), но новые возможности работы с таблицей превосходят эту функцию по простоте использования.

Допустим, в ячейке **D15** вы хотите вычислить общую выручку. Начните вводить формулу **=СУММ(П.** Автоматически появятся варианты завершения ввода, в том числе вариант с таблицей **Продажи**. Для автозавершения формулы дважды щелкните на имени таблицы. Автозавершение также можно активировать, выбрав из списка таблицу **Продажи** (или спустившись к ее названию с помощью **<↓>**) и нажав **Tab**. Затем после **=СУММ(Продажи** введите открывающую квадратную скобку (**[**), функция автозавершения формул предложит добавить заголовки столбцов из таблицы **Продажи**. Завершите написание формулы **=СУММ(Продажи[Выручка])**, и в ячейке появится результат вычисления общей выручки в размере \$155 480 (рис. 26.4). (Пример выбора в окне автозавершения элемента, начинающегося со знака «#», рассмотрен далее в этой главе при описании структуры таблицы.)

	B	C	D	E	F	G
5				Name ▾	Units ▾	Revenue ▾
6				John	814	39886
7				Adam	594	26136
8				Dixie	528	13200
9				Tad	806	20956
10				Erica	826	27258
11				Gabrielle	779	28044
12						
13						
14						
15	revenue	155480				
16		=SUM(Table1[Revenue])				

**Рис. 26.4.** Общая выручка для исходных данных

При добавлении новых строк данные из этих строк автоматически включаются в формулу. Для примера добавим новые данные в строку 12: Аманда продала 400 единиц товара на \$5000. Как показано на рис. 26.5, общая выручка увеличилась на \$5000 и составила \$160 480.

	B	C	D	E	F	G
5				Name	Units	Revenue
6				John	814	39886
7				Adam	594	26136
8				Dixie	528	13200
9				Tad	806	20956
10				Erica	826	27258
11				Gabrielle	779	28044
12				Amanda	400	5000
13						
14						
15	revenue	160480				
16		=SUM(Table1[Revenue])				

**Рис. 26.5.** Общая выручка с учетом новых данных в 12-й строке таблицы

Форматирование также было расширено до строки 12, и формула вычисления общей выручки была изменена для включения данных Аманды. Даже если данные вводятся в среднюю, а не нижнюю часть таблицы, все обновляется как надо.

Пусть теперь в столбце H необходимо вычислить среднюю цену единицы товара, вырученную каждым продавцом. Введите заголовок столбца **Цена единицы продукции** в ячейку H5. В H6 введите =Продажи[. С помощью функции автозавершения формул выберите **Выручка** и введите завершающую скобку (]). Формула примет вид =Продажи[Выручка]. Введите слеш (/), с помощью функции автозавершения формул закончите формулу: =Продажи[Выручка]/Продажи[Количество]. Произойдет удивительная вещь. Формула будет скопирована автоматически до ячейки H12 в нижней части таблицы (рис. 26.6). В пределах таблицы в любой ячейке столбца H формула имеет вид =[Выручка]/[Количество]. Безусловно, такую формулу намного проще понять, чем формулу =G6/F6. Эта формула интерпретируется как деление того, что находится в текущей строке столбца **Выручка**, на то, что находится в текущей строке столбца **Количество**.

При выборе какой-либо ячейки в пределах таблицы на ленте появляется контекстуальная вкладка **Конструктор (Table Tools)**, на которой доступны следующие возможности.

- В поле **Имя таблицы (Table Name)** можно переименовать таблицу, например, изменить имя **Таблица1** (по умолчанию) на **Продажи**. Находится в группе **Свойства (Properties)**.



	B	C	D	E	F	G	H
5				Name ▾	Units ▾	Revenue ▾	Unit Price ▾
6				John	814	39886	49
7				Adam	594	26136	44
8				Dixie	528	13200	25
9				Tad	806	20956	26
10				Erica	826	27258	33
11				Gabrielle	779	28044	36
12				Amanda	400	5000	12.5
13							
14							
15	revenue	160480					
16		=SUM(Table1[Revenue])					

**Рис. 26.6.** Автоматическое копирование формулы расчета цены в таблице

- Инструмент **Размер таблицы (Resize Table)** позволяет добавить или удалить строки и/или столбцы в указанной таблице. Находится в группе **Свойства (Properties)**.
- Инструмент **Преобразовать в диапазон (Convert to Range)** преобразует таблицу в обычные ячейки и удаляет табличную структуру. Находится в группе **Сервис (Tools)**.
- Инструмент **Удалить дубликаты (Remove Duplicates)** удаляет строки, содержащие дубликаты. Например, если в диалоговом окне **Удалить дубликаты (Remove Duplicates)** указать столбец **Имя**, то никакое имя не появится в столбце больше одного раза. Если выбрать и столбец **Имя**, и столбец **Количество**, то в таблице будут отсутствовать строки с одинаковыми значениями в столбцах **Имя** и **Количество**. Находится в группе **Сервис (Tools)**.
- Если выставлен флажок **Строка заголовка (Header Row)**, то на листе отображается содержимое строки заголовка. Если флажок снят, то строка заголовка пуста. Находится в группе **Параметры стилей таблиц (Table Style Options)**.
- Флажок **Строка итогов (Total Row)** мы рассмотрим далее в этой главе.
- Если выставлен флажок **Первый столбец (First Column)**, к первому столбцу таблицы применяется специальный формат.
- Если выставлен флажок **Последний столбец (Last Column)**, то специальный формат применяется к последнему столбцу таблицы.
- Если выставлен флажок **Чередующиеся строки (Banded Rows)**, то к четным строкам таблицы применяется формат, отличный от формата для нечетных строк.
- Если выставлен флажок **Чередующиеся столбцы (Banded Columns)**, то к нечетным столбцам таблицы применяется формат, отличный от формата для четных столбцов.

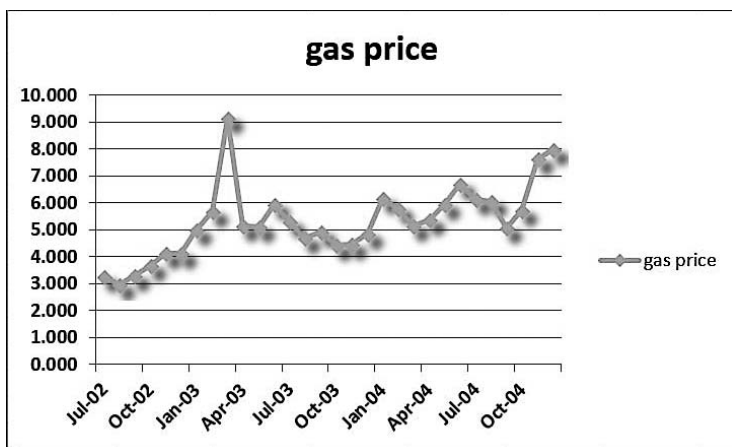
- В группе **Стили таблиц (Table Styles)** можно выбрать любой из представленных форматов. Если таблица расширяется или сжимается, формат применяется в соответствии с этими изменениями.

❓ **Я ввел в книгу цены на природный газ за несколько лет и создал подходящую линейную диаграмму, отображающую ежемесячное изменение цен. Можно ли добиться, чтобы при добавлении новой информации о ценах диаграмма обновлялась автоматически?**

В файле *Gasprices507.xlsx* на листе *Original* находятся данные о ценах на природный газ с июля 2002 г. по декабрь 2004 г. (рис. 26.7). Как я писал выше, вы можете выбрать диапазон **B4:C34** (содержащий месяцы и цены) и нажать **Ctrl+T** для создания таблицы из этого диапазона. Затем для просмотра данных можно создать линейный график. Для этого на вкладке **Вставка (Insert)** в группе **Диаграммы (Charts)** выберите **График (Line)** и затем четвертый тип линейного графика, **График с маркерами**. Уже созданный линейный график показан на рис. 26.8.

	B	C
4	month	gas price
5	Jul-02	3.278
6	Aug-02	2.976
7	Sep-02	3.288
8	Oct-02	3.686
9	Nov-02	4.126
10	Dec-02	4.140
11	Jan-03	4.988
12	Feb-03	5.660
13	Mar-03	9.133
14	Apr-03	5.146
15	May-03	5.123
16	Jun-03	5.945
17	Jul-03	5.291
18	Aug-03	4.693
19	Sep-03	4.927
20	Oct-03	4.430
21	Nov-03	4.459
22	Dec-03	4.860
23	Jan-04	6.150
24	Feb-04	5.775
25	Mar-04	5.150
26	Apr-04	5.365
27	May-04	5.935
28	Jun-04	6.680
29	Jul-04	6.141
30	Aug-04	6.048
31	Sep-04	5.082
32	Oct-04	5.723
33	Nov-04	7.626
34	Dec-04	7.976

**Рис. 26.7.** Цены на газ в 2002–2004 гг.

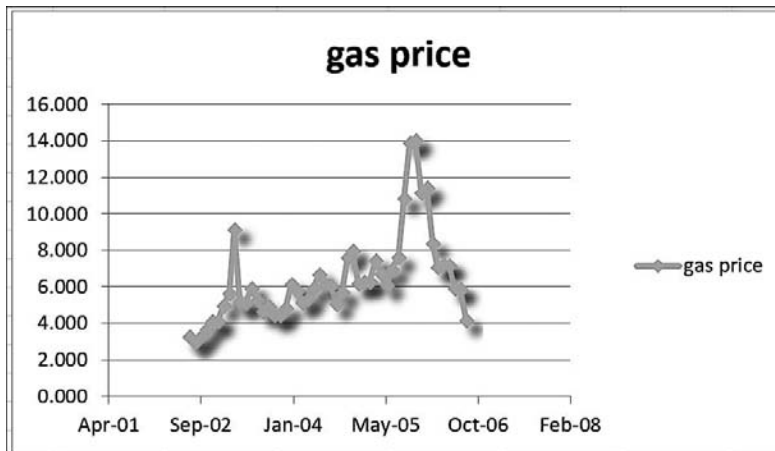


**Рис. 26.8.** Линейный график цен на газ в 2002–2004 гг.

А теперь скопируйте этот лист. Для этого щелкните правой кнопкой мыши на имени листа, затем выберите команду **Переместить или скопировать (Move or copy sheet)** и выставите флажок **Создать копию (Make a copy)**. Добавьте цены на газ по июлю 2006 г. включительно. Теперь данные занимают ячейки до строки 54. При-

своему новому листу имя **Newdata**. Обратите внимание, что при добавлении новых данных линейный график на этом листе обновляется автоматически (рис. 26.9).

Этот пример показывает, что для основанного на таблице графика новые данные включаются в график автоматически.



**Рис. 26.9.** Линейный график цен на газ в 2002–2006 гг.

**?** Для каждой торговой сделки указаны продавец, дата, продукт, место продажи и объем продажи. Можно ли подсчитать, например, общий объем продаж помады на Востоке для Джен или Коллин?

Файл **Tablemakeuptemp.xlsx** содержит данные о торговых сделках (рис. 26.10). Для каждой сделки указана следующая информация: номер сделки, имя, дата, продукт, место, сумма в долларах и количество проданных единиц товара. Если сгруппировать эти данные в виде таблицы, можно добавить строку итогов для столбцов **Количество** и **Сумма** и затем с помощью фильтров создать строку итогов для любого подмножества сделок. Сначала установите курсор на какую-либо ячейку диапазона и нажмите **Ctrl+T** для создания таблицы. Обратите внимание, что при просмотре таблицы с помощью прокрутки строка заголовков остается видимой. Выделите какую-либо ячейку таблицы и на вкладке **Конструктор (Design)** в группе **Параметры стилей таблиц (Table Style Options)** выставите флажок **Строка итогов (Total Row)**. По умолчанию общее количество строк таблицы появится в ячейке **K1895**, которую можно удалить. Выделите ячейку **I1895**, нажмите справа от ячейки на стрелку раскрывающегося списка и выберите **Сумма (Sum)**. Пропишите то же самое для ячейки **J1895**. Так будут итоговые показатели для всех записей в столбцах **Количество** и **Сумма**. Таким образом, на текущий момент общая выручка составляет \$239 912,67 и всего продано 78 707 единиц товара. (См. рис. 26.11 и файл **Tablemakeuptotals.xlsx**.)

	E	F	G	H	I	J	K
3	Trans Number	Name	Date	Product	Units	Dollars	Location
4		1 Betsy	38078	lip gloss	45	137.2045583	south
5		2 Hallagan	38056	foundation	50	152.0073031	midwest
6		3 Ashley	38408	lipstick	9	28.71948312	midwest
7		4 Hallagan	38859	lip gloss	55	167.0753225	west
8		5 Zaret	38155	lip gloss	43	130.6028724	midwest
9		6 Colleen	38683	eye liner	58	175.9909741	midwest
10		7 Cristina	38067	eye liner	8	25.80069218	midwest
11		8 Colleen	39068	lip gloss	72	217.8396539	midwest
12		9 Ashley	38903	eye liner	75	226.6423269	south
13		10 Betsy	38936	lip gloss	24	73.50234217	east
14		11 Ashley	38320	mascara	43	130.8353684	east
15		12 Ashley	38309	lip gloss	23	71.03436769	west
16		13 Emilee	38595	lip gloss	49	149.5927969	west
17		14 Hallagan	38353	eye liner	18	56.47199923	south
18		15 Zaret	38980	foundation	-8	-21.99304472	east
19		16 Emilee	38089	mascara	45	137.3903759	east
20		17 Colleen	38837	mascara	66	199.6543347	south
21		18 Jen	38595	lip gloss	88	265.1875515	midwest
22		19 Jen	38287	eye liner	78	236.1469779	south

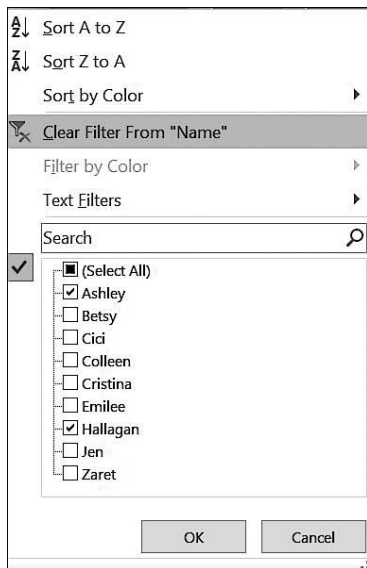
Рис. 26.10. Таблица с данными продаж косметики

	Trans Number	Name	Date	Product	Units	Dollars	Location
1879	1885	Ashley	38848	lip gloss	12	37.83771126	west
1880	1886	Cristina	38573	mascara	89	269.1475428	south
1881	1887	Zaret	38463	lip gloss	61	185.3147883	midwest
1882	1888	Colleen	38551	eye liner	24	73.81115186	west
1883	1889	Emilee	39046	eye liner	76	229.9178081	west
1884	1890	Cici	38518	foundation	16	49.7539874	east
1885	1891	Betsy	38452	foundation	39	119.1888319	east
1886	1892	Cici	38771	mascara	92	278.4349111	west
1887	1893	Cici	38199	foundation	20	61.92385747	midwest
1888	1894	Colleen	38122	lip gloss	60	181.8703479	east
1889	1895	Emilee	38683	eye liner	15	47.16102233	east
1890	1896	Ashley	38397	foundation	36	109.8425992	east
1891	1897	Colleen	38661	lip gloss	46	140.4088994	west
1892	1898	Zaret	38001	lipstick	72	217.8358862	west
1893	1899	Hallagan	39024	eye liner	28	85.65682953	south
1894	1900	Cristina	38881	eye liner	54	164.4873342	midwest
1895	Total				78707	239912.6741	1891

Рис. 26.11. Общая выручка и количество проданных единиц товара

Для получения итоговых показателей по сделкам только с помадой на Востоке для Эшли (Ashley) и Халлаган (Hallagan) нажмите стрелку раскрывающегося списка в ячейке F3 (справа от заголовка Имя). Снимите флажок Выделить все (Select All), а затем выставите флажки только для продавцов Ashley и Hallagan (рис. 26.12) и нажмите ОК. Затем нажмите стрелку в ячейке с заголовком Продукт и выставите флажок только для помады (lipstick) и, наконец, нажмите стрелку в ячейке

с заголовком **Место** и выставите флажок для Востока ( **East**). Теперь перед вами только данные, отобранные по критериям фильтрации (см. рис. 26.13 и лист **Slicers** в файле **Tablemakeuptotals.xlsx** в папке **Program Files**). Как видите, эти ребята продали 564 штук помады на Востоке на общую сумму \$1716,56. Инструмент фильтрации таблицы позволяет получить итоговые результаты для любого подмножества строк на листе **Excel**. При необходимости можно также скопировать строки, связанные с продажей помады Коллин или Джен на Востоке, и вставить их в каком-либо другом месте листа.



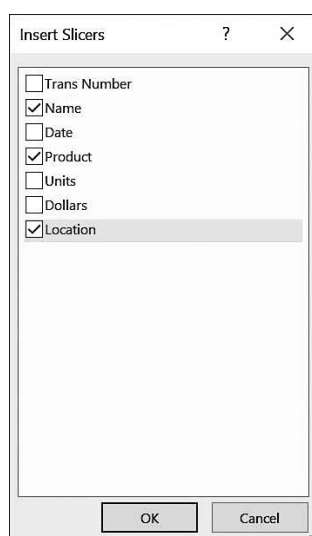
**Рис. 26.12.** Фильтр по именам в таблице

	E	F	G	H	I	J	K
3	Trans Number	Name	Date	Product	Units	Dollars	Location
282	288	Ashley	38606	lipstick	-8	-21.9125458	east
565	571	Hallagan	38925	lipstick	60	182.2924607	east
670	676	Hallagan	38023	lipstick	8	25.63945706	east
702	708	Ashley	38375	lipstick	71	215.1433462	east
968	974	Ashley	39002	lipstick	72	218.0571538	east
1142	1148	Ashley	38727	lipstick	42	127.8685805	east
1173	1179	Ashley	38452	lipstick	70	211.6914545	east
1186	1192	Hallagan	38023	lipstick	40	122.5516179	east
1282	1288	Ashley	38320	lipstick	84	254.1193721	east
1332	1338	Ashley	38969	lipstick	50	152.3101916	east
1423	1429	Hallagan	38441	lipstick	24	73.61606446	east
1469	1475	Ashley	38705	lipstick	51	155.1847283	east
1895	Total				564	1716.561881	12

**Рис. 26.13.** Отфильтрованные подмножества для единиц товара и выручки

### ? Как срезы (добавлены в Microsoft Excel 2013) позволяют выполнить срез данных в таблице Excel?

В Excel 2010 срезы были введены для упрощения фильтрации сводных таблиц (о сводных таблицах и срезах см. в главе 43). Начиная с Excel 2013 срезы используются для фильтрации обычных таблиц. Преимущество срезов в том, что они позволяют легко просмотреть фильтры и общие вычисления. Создадим фильтры для примера с косметикой (файл `Tablemakeuptotals.xlsx`). Установите курсор внутрь таблицы и на вкладке **Вставка (Insert)** в группе **Фильтры (Filters)** выберите **Срез (Slicers)**. Для создания срезов по имени, продукту и месту продажи выставите соответствующие флажки (рис. 26.14).



**Рис. 26.14.** Создание срезов по имени, продукту и месту продажи

Срезы представлены на рис. 26.15. Работая со срезом, вы можете выбрать фильтры, используя **Shift** для выбора смежных элементов или **Ctrl** для выбора несмежных элементов.

В срезах на рис. 26.15 данные отфильтрованы по всем продажам помады на Востоке для Эшли (Ashley) и Халлаган (Hallagan). Обратите внимание, что строка итогов и видимые строки идентичны строкам на рис. 26.13. Щелчок по значку воронки в правом верхнем углу среза удаляет фильтр по соответствующему столбцу таблицы.

Размер среза можно изменить, выбрав его и перетащив его края или углы с нажатой клавишей **Ctrl**. После выбора среза на ленте появится вкладка **Параметры (Slicer tools options)** для настройки среза. На этой вкладке можно изменить многие свойства среза, включая стиль среза, имя, размер и количество столбцов в срезе.



	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1														
2														
3	Trans Number	Name	Date	Product	Units	Dollars	Location							
282	288	Ashley	38606	lipstick	-8	-21.9125458	east							
265	571	Hallagan	38925	lipstick	60	182.2924607	east							
570	676	Hallagan	38023	lipstick	8	25.63945706	east							
702	708	Ashley	38375	lipstick	71	215.1433462	east							
968	974	Ashley	39002	lipstick	72	218.0571538	east							
142	1148	Ashley	38727	lipstick	42	127.8685805	east							
173	1179	Ashley	38452	lipstick	70	211.6914545	east							
186	1192	Hallagan	38023	lipstick	40	122.5516179	east							
282	1288	Ashley	38320	lipstick	84	254.1193721	east							
332	1338	Ashley	38969	lipstick	50	152.3101916	east							
423	1429	Hallagan	38441	lipstick	24	73.61606446	east							
469	1475	Ashley	38705	lipstick	51	155.1847283	east							
895	Total				564	1716.561881	12							
896														
897														
898														
899														

Рис. 26.15. Фильтрация таблицы с помощью срезов

Вычисление итоговых значений с применением срезов — не самая очевидная задача. Рассмотрим файл `Sumwithslicerstemp.xlsx` (в папке `Templates`). В нем представлены годовые продажи в США и за границей. Подмножество данных (месяцы с марта по сентябрь скрыты) приведены на рис. 26.16. Задача состоит в том, чтобы с помощью срезов выделить некоторое подмножество лет, выбрать внутренние и международные продажи и вычислить общий объем продаж для этого подмножества данных. Сначала вычислим сумму всех продаж в таблице, введя в ячейку Q15 формулу `=СУММ(Table1[[January]:[December]])`. Получим общий доход, равный 14 977 000 долларов. Затем зададим срезы для лет и типа продаж (годы 2010–2011 и международные продажи, INT).

	A	B	C	D	M	N
15	Year	Type	January	February	November	December
16	2010	US	\$177,000	\$107,000	\$190,000	\$113,000
17	2010	INT	\$74,000	\$71,000	\$51,000	\$70,000
18	2011	US	\$149,000	\$128,000	\$172,000	\$183,000
19	2011	INT	\$63,000	\$75,000	\$61,000	\$67,000
20	2012	US	\$184,000	\$103,000	\$126,000	\$127,000
21	2012	INT	\$67,000	\$58,000	\$52,000	\$64,000
22	2013	US	\$109,000	\$173,000	\$122,000	\$110,000
23	2013	INT	\$74,000	\$51,000	\$75,000	\$70,000
24	2014	US	\$112,000	\$136,000	\$123,000	\$149,000
25	2014	INT	\$51,000	\$74,000	\$58,000	\$65,000
26	2015	US	\$113,000	\$181,000	\$165,000	\$185,000
27	2015	INT	\$59,000	\$57,000	\$67,000	\$65,000

Рис. 26.16. Суммирование в таблице со срезами

Как можно увидеть на рис. 26.17, мы снова получим суммарный доход 14 977 000 долларов. Но это неверно! Проблема состоит в том, что Excel не

	A	B	C	D	M	N	P	Q
1								
2							=SUM(Table1[[January]:[December]])	\$14,977,000.00
3		SUMSLICERS.XLSX						
4							\$3,509,000.00	
5							=AGGREGATE(9,5,Table1[[January]:[December]])	
6								
7								
8								
9								
10								
11								
12								
13								
14								
15	Year	Type	January	February	November	December		
16	2010	US	\$177,000	\$107,000	\$190,000	\$113,000		
18	2011	US	\$149,000	\$128,000	\$172,000	\$183,000		
28								
29								

**Рис. 26.17.** Применение срезов к суммарному доходу

игнорирует скрытые строки. Если ввести в ячейку P4 формулу =АГРЕГАТ(9,5,Table1[[January]:[December]]), то, если снова выделить тип продаж INT и годы 2010–2011 с помощью срезов, мы получим правильное значение объема продаж, 1 506 000 долларов. Первый аргумент, 9, функции АГРЕГАТ указывает, что требуется вычислить сумму, а второй аргумент, 5, сообщает Excel, что требуется игнорировать скрытые выбором срезов строки.

### ❓ Как сослаться на фрагмент таблицы в других частях листа?

В файле Tablestructure.xlsx приведено множество примеров ссылок на фрагменты таблицы при работе за пределами табличного диапазона. Такие ссылки часто называют *структурированными ссылками* (рис. 26.18). После ввода в формуле имени таблицы, за которым следует открывающая квадратная скобка (⌈), функция автозавершения формул предоставляет для выбора имени столбцов и определенные части таблицы.

- **Имя таблицы** — все ячейки в таблице, за исключением заголовков и строк итогов.
- **#Все** — все ячейки в таблице, включая строку заголовков и строку итогов (если таковая имеется).
- **#Данные** — все ячейки в таблице, за исключением строки заголовков и строки итогов.
- **#Заголовки** — только строка заголовков.
- **#Итоги** — только строка итогов. Если строка итогов отсутствует, возвращается пустой диапазон ячеек.
- **@** или **Эта строка** — все элементы в текущей строке таблицы. Начиная с Excel 2010 символ @ заменил обозначение **Эта строка**.

Ссылка на столбец включает все ячейки в столбце таблицы, кроме заголовка и записи в строке итогов (если таковая имеется).



	A	B	C	D	E	F	G	H
4				Product	Q1	Q2	Q3	Q4
5				food1	37	42	24	32
6				mag1	20	23	24	41
7				drug1	47	34	41	28
8		158		drug2	41	28	49	40
9				food2	44	22	46	50
10	=SUM(Table1[@])			drug1	39	25	38	29
11				mag2	26	35	31	30
12				food4	48	49	50	50
13				mag3	65	34	35	43
14				Total	367	292	338	343
15				55 =COUNTA(Table1[#All])				
16				45 =COUNTA(Table1)				
17				45 =COUNTA(Table1[#Data])				
18				5 =COUNTA(Table1[#Headers])				
19				367 =SUM(Table1[Q1])				
20				1340 =SUM(Table1[#Totals])				
21				997 =SUM(Table1[#Data],[Q1]:[Q3])				

**Рис. 26.18.** Структурированные ссылки на таблицу

Вот несколько примеров использования определенных частей таблицы в формулах (лист *Original* в файле *Tablestructure.xlsx*):

- в B8 по формуле =СУММ(Таблица1[@]) суммируются элементы в строке 8 (41 + 28 + 49 + 40);
- в C15 формула =СЧЁТЗ(Таблица1[#Все]) дает в результате 55, поскольку в таблице 55 элементов;
- в C16 формула =СЧЁТЗ(Таблица1) дает в результате 45, поскольку строки заголовков и итогов не учитываются;
- в C17 формула =СЧЁТЗ(Таблица1[#Данные]) дает в результате 45, поскольку это ссылка на диапазон ячеек D5:H13;
- в C18 формула =СЧЁТЗ(Таблица1[#Заголовки]) дает в результате 5, поскольку это ссылка только на строку заголовков (D4:H4);
- в C19 формула =СУММ(Таблица1[Q1]) дает в результате 367, поскольку по этой формуле суммируются элементы в диапазоне E5:E13;
- в C20 по формуле =СУММ(Таблица1[#Итоги]) суммируются все элементы в строке итогов, и в результате получается 1340, что является суммой всех элементов таблицы;
- в C21 по формуле =СУММ(Таблица1[#Данные];[Q1]:[Q3]) суммируются все элементы данных в столбцах Q1:Q3 включительно (ячейки E5:G13). Таким образом, имена столбцов, разделенные двоеточием, включают все введенные значения между этими столбцами и элементы данных в этих столбцах.

Безусловно, при добавлении новых данных в таблицу все эти формулы обновляются автоматически.

### ❓ Применяются ли условные форматы автоматически к новым данным, добавляемым в таблицу?

Да, условные форматы автоматически охватывают новые данные (рис. 26.19). Как мы увидим, сложнее обеспечить, чтобы условное форматирование, созданное с помощью правила **Использовать формулу для определения форматируемых ячеек (Use A Formula)**, автоматически изменялось при добавлении в таблицу новых данных.

Для примера в файле **Tablestructure.xlsx** на листе **Original** я установил условный формат для выделения трех самых больших значений объема продаж **Q1** в столбце **E**. Записи в строках 7, 12 и 13 выделены красным цветом. На листе **Add Biggersale** в ячейку **E14** добавлено значение **90**. Оказалось, что это самое большое значение в столбце, которое незамедлительно было выделено красным цветом. Ячейка **E7** не выделена красным, поскольку значение в ней больше не входит в три самых больших значения в столбце **E**.

	D	E	F	G	H
4	Product	Q1	Q2	Q3	Q4
5	food1	37	42	24	32
6	mag1	20	23	24	41
7	drug1	47	34	41	28
8	drug2	41	28	49	40
9	food2	44	22	46	50
10	drug1	39	25	38	29
11	mag2	26	35	31	30
12	food4	48	49	50	50
13	mag3	65	34	35	43
14	drug3	90	45	34	23

**Рис. 26.19.** Автоматическое распространение условного форматирования на новые данные в таблице

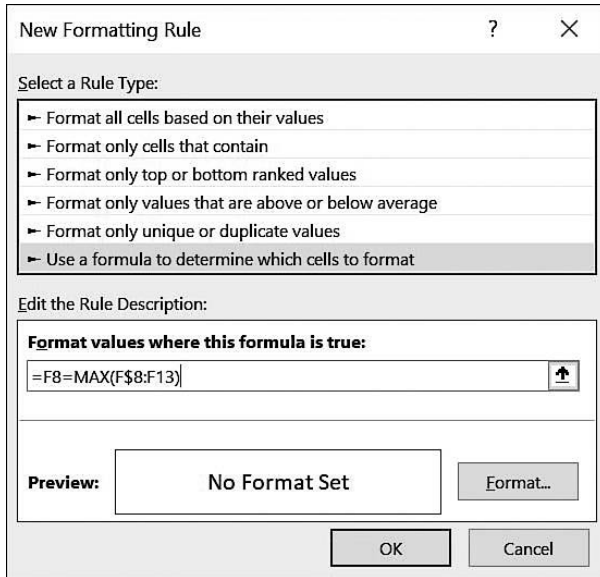
	F	G	H
7	Wash	Ore	Idaho
8	22	33	44
9	33	55	23
10	52	2	12
11	75	43	98
12	98	54	112
13	135	77	143

**Рис. 26.20.** Выделение наибольшего числа в каждом столбце

Теперь предположим, что нам требуется выделить максимальное число в каждом столбце в файле **Formattablesfinal.xlsx**, представленном на рис. 26.20.

Мы можем легко создать такое форматирование с помощью правила **Использовать формулу для определения форматируемых ячеек (Use A Formula)**, но если требуется, чтобы этот формат применялся к новым строкам, добавленным в таблицу, нам нужно заполнить диалоговое окно **Создание правила форматирования (New Formatting Rule)**, как показано на рис. 26.21. (На вкладке **Главная (Home)** в группе **Стили (Styles)** выберите **Условное форматирование (Conditional Formatting)** и затем **Создать правило (New Rule)**).

При выборе значений для форматирования введите формулу **=F8=МАКС(F\$8:F13)**. Главное в ней — не поставить знак доллара перед числом 13. Если ввести **F\$13**, Excel заблокирует форматирование так, что оно не сможет распространиться за пределы строки 13. Введите новую строку, содержащую большие числа, чтобы проверить, как автоматически меняется форматирование с учетом новых данных.



**Рис. 26.21.** Такие настройки гарантируют, что при добавлении новых данных Excel выделит максимальное число в каждом столбце

## Задания

1. Файл `Singers.xlsx` содержит список песен, исполненных различными певцами, а также продолжительность звучания каждой песни. Вычислите общее количество песен в исполнении Эминема и среднюю продолжительность звучания его песен. При добавлении новых данных все формулы должны обновляться автоматически.
2. В файле `Tableexample.xlsx` включите в лист ранг для каждого продавца согласно общей выручке и количеству проданных единиц товара. При добавлении новых данных ранги должны обновляться автоматически. Советую воспользоваться функцией **РАНГ**. Ее синтаксис в данном случае: **РАНГ(число;массив;0)**. Эта функция возвращает ранг числа в массиве. У самого большого числа ранг равен 1.
3. Файл `Lookupdata.xlsx` содержит коды продуктов и цены. Сделайте так, чтобы при введении кода продукта возвращалась цена продукта. При добавлении новых продуктов формула по-прежнему должна выдавать правильный ответ.
4. Файл `Productlookup.xlsx` содержит данные о продаже товаров в любой день недели. Напишите формулу, по которой рассчитывается объем продаж любого продукта в любой день недели. При добавлении в данные новых продуктов в формуле должны быть учтены объемы продаж для этих продуктов.

5. В файле **Tablepie.xlsx** хранится информация о продажах различных продуктов в небольшом магазине. Для подведения итогов по этим данным необходимо создать круговую диаграмму. При добавлении новых категорий продуктов круговая диаграмма должна автоматически включать новые данные.
6. В файле **Tablexnpvdata.xlsx** указаны денежные потоки малого бизнеса. Создайте формулу для вычисления (по состоянию на 5 января 2007 г. ) чистой приведенной стоимости (ЧПС) всех денежных потоков. Допустим, годовая ставка дисконтирования составляет 10%. При вводе новые потоки должны учитываться в формуле автоматически.
7. Файл **Nikedata.xlsx** содержит данные о ежеквартальных доходах от продаж компании Nike. Постройте график доходов Nike, автоматически включающий новые данные о доходах.
8. Основываясь на файле **Tablemakeuptemp.xlsx**, узнайте общее количество проданных единиц товара и выручку для помады и блеска для губ, проданных Джен (Jen) и Эшли (Ashley) на Юге и Востоке.
9. В файле **Closest.xlsx** находится список сотрудников и их зарплат. Необходимо по любому заданному числу определить сотрудника, размер зарплаты которого ближе всего к этому числу. Цель должна быть достигнута даже в случае добавления новых или удаления из списка существующих имен.
10. В файле **Adagency.xlsx** содержатся данные о зарплатах и возрасте сотрудников рекламного агентства. Вычислите среднюю зарплату и средний возраст сотрудников. При найме на работу и увольнении сотрудников вычисления должны обновляться автоматически.
11. Файл **Problem11data.xlsx** содержит статистику по нападающим НФЛ. Имя нападающего нужно вводить в ячейку I2. Задайте формат, в соответствии с которым будет выделяться вся строка с данными по нападающему, имя которого введено в ячейку I2. Ваше форматирование должно автоматически учитывать появление новых строк с данными в листе.
12. Файл **Problem 12data.xlsx** содержит в одной ячейке название штата и население. Используйте формулы, чтобы извлечь население в новый столбец. Если новая строка добавлена, то формулы должны автоматически работать с новыми данными. Предполагается, что ни одно название штата не состоит из более чем двух слов.
13. Файл **Problem13data.xlsx** содержит данные о продажах 2000 компаний. Напишите формулы для подсчета количества банковских фирм и расчета доли банковского дохода, полученного банками, составляющего не менее 20 миллиардов долларов. Ответы должны автоматически обновляться, если добавляются новые строки данных.

# Счетчики, полосы прокрутки, переключатели, флажки, группы и поля со списками

### Обсуждаемые вопросы

- Мне необходимо запустить анализ чувствительности с множеством ключевых исходных данных, таких как объем продаж за первый год, годовой рост продаж, цена реализации в первый год и себестоимость единицы товара. Каким образом можно быстро изменить эти исходные данные и сразу отследить влияние такого изменения, например, на расчет ЧПС?
- Как создать простой флажок, с помощью которого можно было бы включать и отключать условное форматирование?
- Как создать лист, который позволит персоналу, занятому в цепочке поставок, одним нажатием кнопки назначить высокую, низкую или среднюю цену продукта?
- Как предоставить пользователю простой способ ввода дня недели без необходимости набирать какой-либо текст?

Пользователи Microsoft Excel 2019 могут добавлять с помощью пользовательских форм на лист множество удобных элементов управления. В этой главе я покажу, как просто можно применять счетчики, полосы прокрутки, флажки, переключатели, группы и поля со списками. Для доступа к пользовательским формам Excel откройте на ленте вкладку Разработчик (Developer) и в группе Элементы управления (Controls) выберите Вставить (Insert). Не путайте элементы управления формы с элементами управления ActiveX, которые обычно используются в языке программирования Microsoft Visual Basic for Applications (VBA).

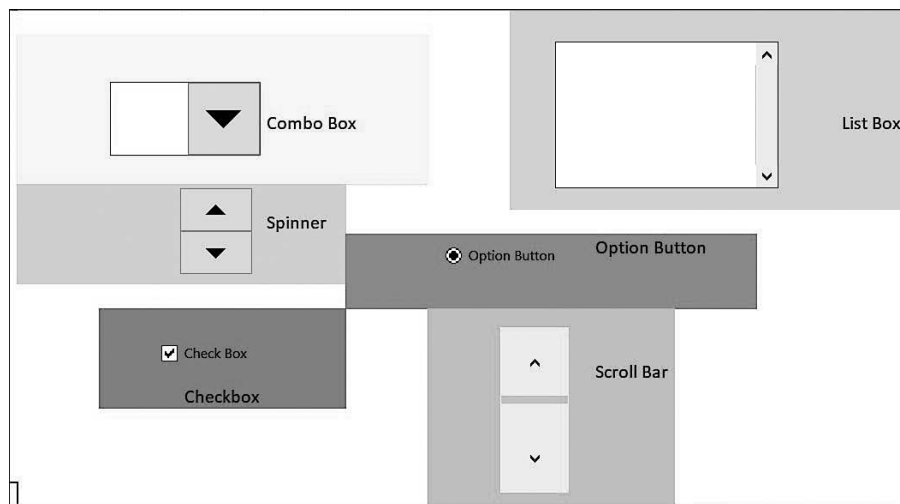
### ПРИМЕЧАНИЕ

---

Для отображения вкладки Разработчик (Developer) откройте вкладку Файл (File) и выберите команду Параметры (Options). В левой части окна выберите раздел Настройка ленты (Customize Ribbon) и затем в списке Основные вкладки (Main Tabs) установите флажок Разработчик (Developer).

---

Пользовательские формы, которые я описываю в этой главе, представлены на рис. 27.1. (См. также файл Controls.xlsx в папке Practice Files к этой главе.)



**Рис. 27.1.** Элементы управления пользовательских форм Excel

## Ответы на вопросы

**?** Мне необходимо запустить анализ чувствительности с множеством ключевых исходных данных, таких как объем продаж за первый год, годовой рост продаж, цена реализации в первый год и себестоимость единицы товара. Каким образом можно быстро изменить эти исходные данные и сразу отследить влияние такого изменения, например, на расчет ЧПС?

Как мы обсудили в главе 19, для изменения группы исходных ячеек в целях просмотра соответствующих изменений в конечных результатах служит Диспетчер сценариев (Scenario Manager) (вкладка Данные (Data)), группа Работа с данными (Forecast), пункт меню Анализ "что, если" (What-If Analysis)). К сожалению, Диспетчер сценариев (Scenario Manager) требует вводить каждый сценарий по отдельности, что затрудняет создание достаточно большого количества сценариев. Предположим, например, что для модели ЧПС четырьмя ключевыми входными параметрами являются объем продаж за первый год, рост продаж, цена реализации в первый год и себестоимость в первый год. (См. файл NPVspinners.xlsx.) Необходимо проследить, как изменяется ЧПС при изменении входных параметров в пределах, указанных ниже:

Исходные данные	Нижнее значение	Верхнее значение
Объем продаж в первый год	\$5000	\$30 000
Годовой рост продаж, %	0	50
Цена реализации в первый год	\$6	\$20
Себестоимость в первый год	\$2	\$15

Создание сценариев, в которых исходные значения в ячейках варьируют в заданных пределах, при помощи **Диспетчера сценариев** займет много времени. Однако с помощью **счетчиков** можно быстро создать массу сценариев, в которых каждый входной параметр изменяется от нижнего до верхнего значения.

Счетчик представляет собой элемент управления, связанный с конкретной ячейкой. Если нажать на счетчике стрелку вверх или вниз, значение в связанной ячейке изменится. Далее я покажу, как изменяется результат вычислений (например, по формуле расчета ЧПС автомобиля) в ответ на изменение исходных данных. Нажимая на стрелки в ячейках D2:D5, посмотрите, как счетчики меняют значения.

А теперь о том, как создать счетчики, позволяющие изменять объем продаж за первый год (Year1sales), годовой рост продаж (Sales growth), цену реализации в первый год (Year1price) и себестоимость в первый год (Year1cost) в требуемых интервалах. Лист **Original Model** (см. файл NPVspinnerstemp.xlsx в папке **Templates**) представлен на рис. 27.2.

	A	B	C	D	E	F
1		taxrate	0.4			
2		Year1sales	10000			
3		Sales growth	0.48		48	
4		Year1price	\$ 9.00			
5		Year1cost	\$ 6.00			
6		intrate	0.15			
7		costgrowth	0.05			
8		pricegrowth	0.03			
9	Year	1	2	3	4	5
10	Unit Sales	10000	14800	21904	32417.92	47978.5216
11	unit price	\$ 9.00	\$ 9.27	\$ 9.55	\$ 9.83	\$ 10.13
12	unit cost	\$ 6.00	\$ 6.30	\$ 6.62	\$ 6.95	\$ 7.29
13	Revenues	\$ 90,000.00	\$ 137,196.00	\$ 209,141.58	\$ 318,815.43	\$ 486,002.24
14	Costs	\$ 60,000.00	\$ 93,240.00	\$ 144,894.96	\$ 225,166.77	\$ 349,909.16
15	Before Tax Profits	\$ 30,000.00	\$ 43,956.00	\$ 64,246.62	\$ 93,648.66	\$ 136,093.08
16	Tax	\$ 12,000.00	\$ 17,582.40	\$ 25,698.65	\$ 37,459.46	\$ 54,437.23
17	Aftertax Profits	\$ 18,000.00	\$ 26,373.60	\$ 38,547.97	\$ 56,189.20	\$ 81,655.85
18						
19	NPV	\$133,664.07				

**Рис. 27.2.** Лист **Original Model**, без счетчиков

Для создания счетчиков выделите строки, в которые требуется вставить значки счетчиков (в этом примере выбраны строки 2–5), и затем увеличьте высоту строк, щелкнув правой кнопкой мыши и выбрав **Высота строки (Row Height)**. Для размещения стрелок счетчика обычно достаточно высоты, равной 27. Иначе нарисуйте значок и, удерживая **Alt**, измените его размеры, чтобы он помещался в ячейке.

Откройте палитру пользовательских форм, нажав **Вставить (Insert)** в группе **Элементы управления (Controls)** на вкладке **Разработчик (Developer)** (вспомните, что эта вкладка подключается на странице **Настройка ленты (Custom Ribbon)** в диалоговом окне **Параметры (Options)**). Выберите значок счетчика в **Элементах управления формы** (см. рис. 27.1). Курсор примет форму знака «+». С помощью мыши вы можете поместить значок счетчика в требуемом месте и придать ему нужную форму. Сформируйте значок счетчика в ячейке **D2**. Для изменения формы элемента управления или его перемещения щелкните на нем, удерживая нажатой клавишу **Ctrl**. Когда курсор принимает форму четырехсторонней стрелки, вы можете перетянуть элемент в другое место. Когда курсор принимает форму двусторонней стрелки, вы можете изменить размеры элемента.

Мы поместили значок счетчика в ячейку **D2**. Этот счетчик можно использовать для изменения значения объема продаж за первый год. Щелкните правой кнопкой мыши на счетчике и выберите команду **Копировать (Copy)** в контекстном меню. Щелкните правой кнопкой мыши в ячейке **D3** и выберите команду **Вставить (Paste)**. Также вставьте счетчик в ячейки **D4** и **D5**. Теперь у вас есть четыре счетчика, как показано на рис. 27.3.

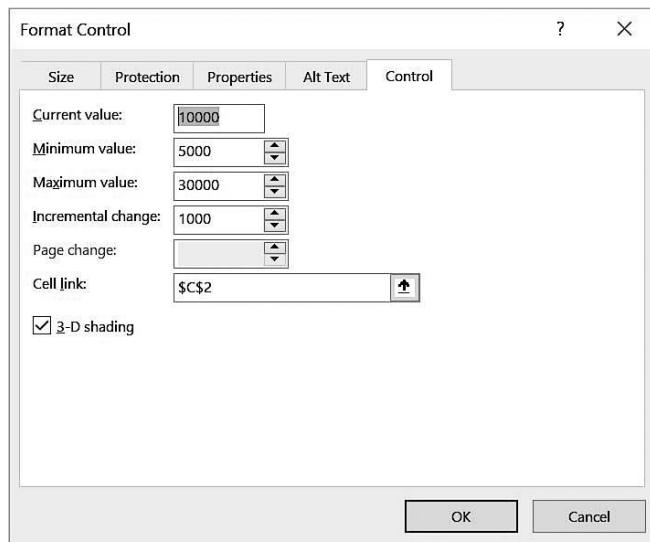
	B	C	D
1	taxrate	0.4	
2	Year1sales	10000	
3	Sales growth	0.46	
4	Year1price	\$ 9.00	
5	Year1cost	\$ 6.00	

**Рис. 27.3.** Размещение счетчиков в ячейках листа

Теперь требуется связать каждый из счетчиков с ячейкой с исходными данными. Для связывания счетчика из ячейки **D2** с ячейкой **C2** щелкните правой кнопкой мыши на счетчике в **D2** и выберите **Формат объекта... (Format Control)**. Заполните поля в диалоговом окне **Формат элемента управления (Format Control)**, как показано на рис. 27.4: например, 9000 для текущего значения, 5000 для минимального значения, 30 000 для максимального значения, 1000 для шага изменения и **\$C\$2** для связи с ячейкой. Оставьте выбранным **Объемное затенение** и нажмите **ОК**.

Параметр **Текущее значение** не так уж важен. Остальные параметры сообщают Excel, что этот счетчик связан со значениями в **C2 (Year1sales)**, что каждый щелчок по стрелке вверх будет увеличивать значение в **C2** на 1000, а каждый щелчок по стрелке вниз будет уменьшать значение в **C2** на 1000. Когда значение в **C2** достигнет 30 000, последующие щелчки по стрелке вверх увеличивать его больше не будут; когда значение в **C2** достигнет 5000, последующие щелчки по стрелке вниз уменьшать его больше не будут.





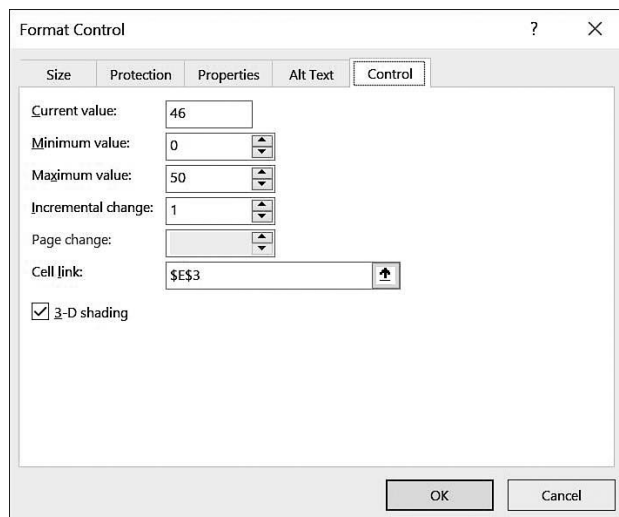
**Рис. 27.4.** Связывание счетчика с объемом продаж за первый год в диалоговом окне Формат элемента управления

Затем, используя диалоговое окно **Формат элемента управления**, свяжите счетчик в D4 с ценой реализации в первый год (ячейка C4). (Счетчик в ячейке D3 мы рассмотрим последним.) Для текущего значения возьмем 9. Минимальное значение — 6, максимальное значение — 20, шаг изменения — 1. Счетчик связан с ячейкой C4. Щелчки по стрелкам счетчика в ячейке D4 изменяют значение цены реализации в первый год от \$6 до \$20 с шагом в \$1.

Для связывания счетчика в ячейке D5 с себестоимостью в первый год (ячейка C5) возьмем 6 для текущего значения, 2 для минимального значения, 15 для максимального значения и 1 для шага изменения. Свяжем счетчик с ячейкой C5. Щелчки по стрелкам счетчика в ячейке D5 изменяют себестоимость в первый год от \$2 до \$15 с шагом в \$1.

Связывание счетчика в ячейке D3 со значением роста продаж — несколько более сложная задача. Необходимо, чтобы счетчик изменял значения в интервале от 0 до 50%. Проблема в том, что минимальный шаг изменения для счетчика — 1. Поэтому нужно связать этот счетчик с фиктивным значением в ячейке E3 и поместить в ячейку C3 формулу E3/100. Теперь ячейка E3 изменяется от 1 до 50, а рост продаж изменяется от 1 до 50%. Рисунок 27.5 показывает связывание счетчика с ячейкой E3: текущее значение 48, минимальное значение 0, максимальное значение 50, шаг изменения 1 и связь с ячейкой E3. Помните, что рост продаж в ячейке C3 — это просто число из E3, деленное на 100.

Если щелкнуть на элементе управления с нажатой клавишей **Ctrl**, можно легко изменить его размеры с помощью меток-манипуляторов.



**Рис. 27.5.** Связывание счетчика в ячейке D3 с ячейкой E3 в диалоговом окне Формат элемента управления

Щелкая на стрелке одного из счетчиков, вы можете наблюдать, как изменение отдельного входного значения — при фиксированных остальных входных значениях на листе — изменяет ЧПС автомобиля. Чтобы просмотреть влияние изменений, выберите ячейку F9, нажмите **Закрепить области (Freeze Panes)** на вкладке **Вид (View)** в группе **Окно (Window)** и затем выберите пункт **Закрепить области (Freeze Panes)**. Эта команда фиксирует данные выше строки 9 и слева от столбца F. Теперь вы можете с помощью полос прокрутки расположить окно, как показано на рис. 27.6.

	A	B	C	D	E	F
1		taxrate	0.4			
2		Year1sales	10000			
3		Sales growth	0.48		48	
4		Year1price	\$ 9.00			
5		Year1cost	\$ 6.00			
6		intrate	0.15			
7		costgrowth	0.05			
8		pricegrowth	0.03			
19	NPV	\$133,664.07				

**Рис. 27.6.** Фиксация областей для просмотра результатов вычислений в других частях листа

Нажимая стрелки счетчика значений роста продаж при зафиксированных остальных входных значениях, мы видим, что увеличение роста продаж на 1% увеличивает ЧПС примерно на \$2000. (Для возвращения листа в прежнее состояние выберите **Закрепить области (Freeze Panes)** на вкладке **Вид (View)**, а затем выберите пункт **Снять закрепление областей (Unfreeze Panes)**.)

Элемент «прокрутка» очень схож с элементом «счетчик». Главное отличие состоит в том, что перемещая курсор по серой области в середине полосы прокрутки, вы можете вызвать плавное изменение значения в связанной ячейке.

Выбрав в контекстном меню строку **Формат объекта (Format Control)** и изменив в диалоговом окне **Формат элемента управления** значение в поле **Шаг изменения по страницам (Page Change)**, вы можете управлять скоростью изменения значения в связанной ячейке.

### ❓ Как создать простой флажок, с помощью которого можно было бы включать и отключать условное форматирование?

Флажок представляет собой форму управления, которая записывает в ячейку значение **ИСТИНА**, когда флажок включен, и **ЛОЖЬ**, когда флажок снят. Флажки можно применять для создания тумблеров, включающих или отключающих определенные функции. В качестве примера возьмем флажок, включающий и отключающий условное форматирование.

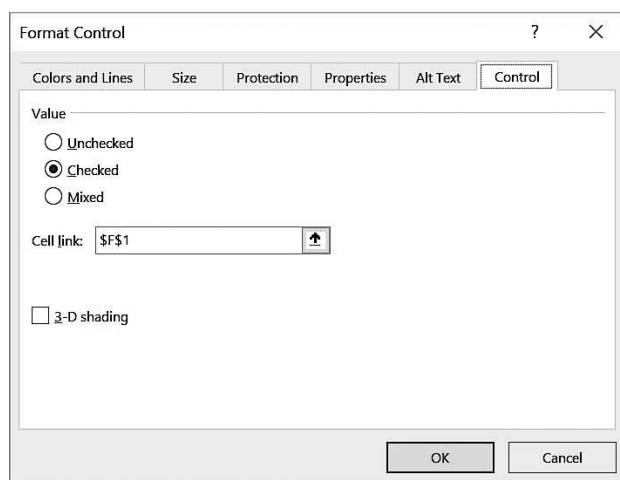
Пусть на листе содержатся данные о ежемесячных продажах и требуется выделить пять самых крупных продаж зеленым цветом, а пять самых мелких продаж красным цветом (откройте файл **Checkbox.xlsx** и щелкните на флажке, чтобы увидеть, как меняется форматирование). Введите в **G4** формулу **=НАИБОЛЬШИЙ(Продажи;5)**, вычисляющую пятое по величине значение объема продаж. В ячейке **H4** вычислите пятое наименьшее значение объема продаж по формуле **=НАИМЕНЬШИЙ(Продажи;5)**, как показано на рис. 27.7.

Затем создайте флажок и свяжите его с ячейкой **F1** для записи в нее значений **ИСТИНА** или **ЛОЖЬ**. Откройте файл **Checkboxtemp.xlsx** из папки **Templates**. На вкладке **Разработчик (Developer)** в группе **Элементы управления (Controls)** выберите **Вставить (Insert)** и затем из палитры **Элементы управления формы (Forms Controls)** выберите флажок. Вид курсора изменится на перекрестье. Перетащите флажок в ячейку **G9** и замените в нем текст на **ВКЛ ИЛИ ОТКЛ ФОРМАТИРОВАНИЕ** (для этого щелкните правой кнопкой мыши на флажке и в контекстном меню выберите **Изменить текст**). Вам может потребоваться увеличить размер элемента, чтобы в нем помещался весь текст. Для этого щелкните правой кнопкой мыши на флажке, в контекстном меню выберите **Формат объекта (Format Control)** и в диалоговом окне введите данные, как показано на рис. 27.8: установите значение **установлен (Checked)**, свяжите флажок с ячейкой **F1** и нажмите **ОК**.

Теперь в ячейке **F1** каждый раз при установке флажка будет появляться значение **ИСТИНА**, а при снятии флажка — значение **ЛОЖЬ**.

	D	E	F	G	H	I
1			TRUE			
2						
3	Sales			5th largest	5th smallest	
4	1010			1050	584	
5	619					
6	524					
7	1114					
8	619					
9	1097			<input checked="" type="checkbox"/> TURN FORMATTING ON OR OFF		
10	627					
11	578					
12	947					
13	1020					
14	1046					
15	678					
16	510					
17	674					
18	756					
19	665					
20	609					
21	556					

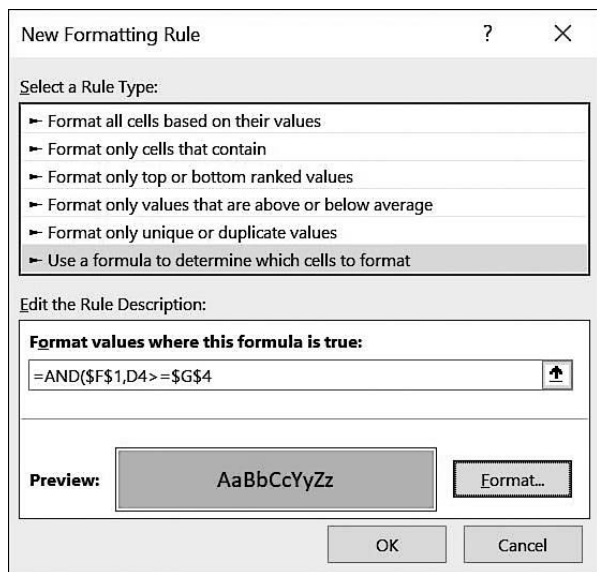
**Рис. 27.7.** Флажок для включения и отключения функции условного форматирования



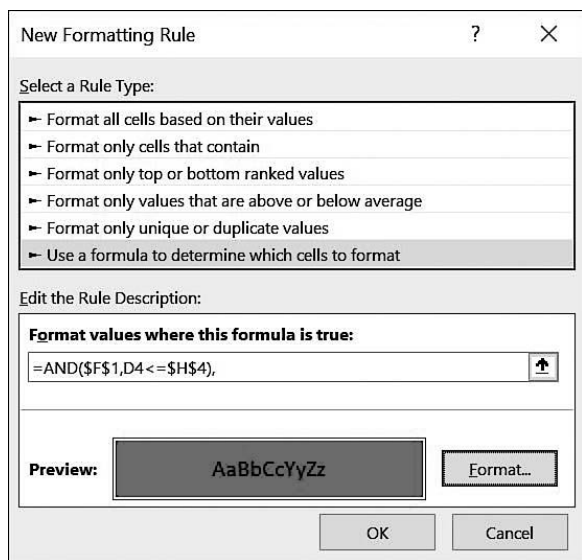
**Рис. 27.8.** Диалоговое окно Форматирование объекта для флажка

Выделите диапазон ячеек D4:D29, на вкладке Главная (Home) в группе Стили (Styles) откройте список Условное форматирование (Conditional Formatting) и выберите Создать правило (New Rule). Выберите правило Использовать формулу для определения форматируемых ячеек (Use a formula...) и введите формулу, показанную на рис. 27.9, `=И($F$1,D4>=$G$4)`, для выделения зеленым цветом пяти самых крупных продаж (выберите Формат (Format), далее вкладку Заливка (Fill), затем зеленый цвет и затем ОК). Затем введите формулу, показанную на рис. 27.10, `=И($F$1,D4<=$H$4)`, для выделения красным цветом пяти самых мелких продаж. Обратите внимание, что

часть формулы И(\$F\$1) гарантирует применение форматирования в случае, если в ячейке F1 записано значение **ИСТИНА**. Находится ли в ячейке F1 значение **ИСТИНА** или **ЛОЖЬ**, определяется флажком, так что если флажок не установлен, ячейки не будут выделяться зеленым или красным цветом.



**Рис. 27.9.** Формат для выделения зеленым цветом пяти наибольших значений



**Рис. 27.10.** Формат для выделения красным цветом пяти наименьших значений

Если требуется скрыть значения **ИСТИНА** и **ЛОЖЬ**, можно выделить ячейку F1 и изменить цвет шрифта на белый.

**?** Как создать лист, который позволит персоналу, занятому в цепочке поставок, одним нажатием кнопки назначить высокую, низкую или среднюю цену продукта?

Пусть требуется назначить одну из трех цен продукта: высокую, среднюю или низкую. Эти цены указаны в диапазоне B7:B9 в файле Optionbuttons.xlsx (пощелкайте переключателями). Если бы пользователь ввел значение высокой, средней или низкой цены, ее легко можно было бы отобразить с помощью таблицы поиска. Однако лучшим решением было бы, если бы пользователь с помощью переключателя просто выбирал уровень цены: высокая, средняя или низкая, а сама цена была бы вычислена по формуле автоматически (рис. 27.10).

Откройте файл Optionbuttonstemp.xlsx в папке Templates. Сначала выберите в палитре Элементы управления формы (Forms Controls) элемент управления группа (расположен в левом нижнем углу палитры) и поместите его на лист в ячейки A14:C21. Щелкните правой кнопкой мыши на элементе, выберите Изменить текст (Edit Text) и замените в нем текст на Выберите цену.

	A	B	C	D	E	F
4						3
5			2			
6					Price level	Price
7	High	\$ 8.00			Low	\$ 3.00
8	Medium	\$ 6.00				
9	Low	\$ 3.00				
10						
11						
12						
13						
14	Select price					
15	<input type="radio"/> High price					
16						
17	<input type="radio"/> Medium price					
18						
19	<input checked="" type="radio"/> Low price					
20						
21						

**Рис. 27.11.** Переключатели для выбора цены продукта

Затем перетащите в этот групповой блок переключатели для каждого варианта выбора. Для трех уровней цен потребуется поместить три переключателя. На вкладке Разработчик (Developer) в группе Элементы управления (Controls) выберите Вставить (Insert), выберите элемент Переключатель (см. рис. 27.1) и перетащите

курсор в ячейку B16. Щелкните правой кнопкой мыши на любом из переключателей (дайте переключателям имена **Высокая цена**, **Средняя цена** и **Низкая цена**) и с помощью команды **Формат объекта (Format Control)** в контекстном меню свяжите его с какой-либо ячейкой, предположим E4. Все переключатели в группе окажутся связанными с одной и той же ячейкой. В ячейке E4 при выборе первого переключателя устанавливается значение 1, при выборе второго переключателя — значение 2 и при выборе третьего переключателя — значение 3.

В ячейку E7 введите формулу **=ИНДЕКС(A7:A9;E4;1)**, определяющую описание цены в соответствии с положением, в которое установлены переключатели. Цена, соответствующая установленному положению переключателей, вычисляется в ячейке F7 по формуле **=ВПР(E7;A7:B9;2;ЛОЖЬ)**.

### ❓ Как предоставить пользователю простой способ ввода дня недели без необходимости набирать какой-либо текст?

В файле **Combobox.xlsx** показано, как создать поле со списком или список (рис. 27.12), в котором пользователь легко может выбрать требуемый элемент из списка доступных элементов. (Раскрывающиеся списки также можно создать при проверке данных; см. главу 40.) Здесь целью является расчет количества часов, отработанных сотрудником в указанный день. Отработанные часы для каждого дня хранятся в ячейках G9:G15. Для выбора записи из списка можно создать элемент управления поле со списком или список. Если поле со списком или список связаны с какой-либо ячейкой через **Формат объекта (Format Control)**, то при выборе первой записи в списке в эту ячейку записывается значение 1; при выборе второй записи в списке в ячейку записывается значение 2 и т. д. Для достижения поставленной в задаче цели на вкладке **Разработчик (Developer)** в группе **Элементы**

	A	B	C	D	E	F	G
1						=INDEX(F9:F15,A8,1)	=VLOOKUP(F3,\$F\$9:\$G\$15,2,FALSE)
2						day	
3			Combo Box		combo box	Tuesday	7
4					list box	Monday	6
5							
6			Tuesday				
7							
8		2					hrs worked
9						Monday	6
10						Tuesday	7
11						Wednesday	8
12						Thursday	9
13						Friday	5
14		1				Saturday	4
15						Sunday	5
16							
17							
18							
19							

**Рис. 27.12.** Поле со списком и список для выбора дня недели

управления (Controls) выберите Вставить (Insert) и затем из палитры Элементы управления формы (Forms Controls) выберите и перетащите поле со списком (см. значок на рис. 27.1) в ячейку C5, а затем список в ячейку D16. Щелкните правой кнопкой мыши на поле со списком и выберите команду Формат объекта (Format Control). Выделите диапазон исходных данных F9:F15 (содержащий дни недели) и свяжите его с ячейкой A8. Затем щелкните правой кнопкой мыши на списке и выберите Формат объекта (Format Control). Выделите диапазон исходных данных F9:F15 и свяжите его с ячейкой A13. После этих шагов при выборе, например, вторника в поле со списком и понедельника из списка, в ячейке A8 появится значение 2, а в ячейке A13 — значение 1.

В ячейке F3 формула =ИНДЕКС(F9:F15;A8;1) вычисляет день недели, соответствующий выбору в поле со списком. В ячейке G3 формула =ВПР(F3;\$F\$9:\$G\$15;2;ЛОЖЬ) определяет количество часов, отработанных в день, указанный в поле со списком. Аналогично формулы в ячейках F4 и G4 определяют день недели и отработанные часы в соответствии с выбором в списке.

## Задания

1. Добавьте в решение примера с ЧПС автомобиля счетчик, позволяющий варьировать ставку налога от 30 до 50%.
2. Добавьте в решение примера с ЧПС автомобиля счетчик, позволяющий варьировать процентную ставку от 5 до 20%.
3. В диалоговом окне Формат элемента управления (Format Control) допускается минимальное значение 0. Несмотря на это ограничение, можно ли найти способ изменять с помощью счетчика рост продаж в интервале между –10 и 20%?
4. Для примера с лимонадом из главы 17 создайте счетчики, позволяющие изменять исходные данные в следующих пределах:

Исходные данные	Нижнее значение	Верхнее значение
Цена	\$2,00	\$5,00
Себестоимость единицы товара	\$0,20	\$1,00
Постоянные затраты	\$20 000,00	\$100 000,00

5. Для примера с ипотекой из главы 17 создайте счетчики, позволяющие изменять исходные данные в следующих пределах:

Исходные данные	Нижнее значение	Верхнее значение
Сумма кредита	\$270 000	\$800 000
Количество месяцев	120	360
Годовая процентная ставка	4%	10%



6. Для примера с выходными днями в главе 24 создайте флажок, позволяющий включать и отключать **Условное форматирование**.
7. В финансовых формулах, описанных в главе 10, последнему аргументу присваивалось значение 1 для денежных потоков в конце месяца и 0 для денежных потоков в начале месяца. В Excel значение **ИСТИНА** эквивалентно 1, а значение **ЛОЖЬ** эквивалентно 0. Создайте лист, на котором пользователь для расчета ежемесячного платежа может ввести количество месяцев кредита, основную сумму кредита и годовую процентную ставку. Затем создайте флажок для выбора платежей в начале или конце месяца.
8. На основе данных к заданию 15 в главе 22 создайте счетчик для построения графика продаж за последние несколько месяцев (от 3 до 8 месяцев).
9. В файле **Suppliers.xlsx** содержатся данные о ценах за единицу товара для различных поставщиков и количестве закупленной продукции. Создайте список, с помощью которого можно выбрать поставщика и получить цену за единицу товара и количество закупленной продукции для этого поставщика.

## ГЛАВА 28

# Революция в аналитике

### Обсуждаемые вопросы

- Что такое аналитика?
- Что такое прогнозная аналитика?
- Что такое предписывающая аналитика?
- Почему значение аналитики возрастает?
- Насколько важна аналитика для вашей организации?
- Что необходимо знать для проведения анализа?
- Какие трудности могут возникнуть при осуществлении анализа?
- Какие тенденции будут влиять на развитие аналитики?

## Ответы на вопросы

### ❓ Что такое аналитика?

Аналитика — это практика применения данных для принятия решения, оптимально отвечающего целям организации. Типичными целями являются максимизация прибыли, минимизация затрат, снижение финансовых рисков, повышение качества медицинского обслуживания и качества образования, снижение уровня преступности и другие задачи. Безусловно, применение данных для принятия решения — не новое явление. Например, греческий философ Фалес использовал опционы (см. главу 79) для увеличения прибыли и снижения риска, сдавая в аренду оливковые прессы. Термин «аналитика» стал популярен после публикации в 2007 г. книги Томаса Дэвенпорта «Аналитика как конкурентное преимущество» (Thomas Davenport. *Competing on Analytics* Boston. — Harvard Business School Press, 2007)<sup>1</sup>.

### ❓ Что такое прогнозная аналитика?

Прогнозная аналитика — это просто использование данных для точных предсказаний относительно интересующих количественных величин. Вот несколько примеров.

---

<sup>1</sup> Дэвенпорт Д., Харрис Дж. Аналитика как конкурентное преимущество. — М.: BestBusinessBooks, 2010.

- Предсказание результатов президентских выборов (см. главу 60).
- Прогнозирование результатов игр НФЛ (см. главу 35).
- Определение влияния выкладки продукта в магазине на его продажи (см. главу 62).
- Определение влияния вида лечения на выживаемость онкологических больных.
- Определение влияния комплекса маркетинга (рекламы, снижения цен, выкладки товара) на ежедневные продажи клейкой лентой 3M Scotch. Здесь трудность в том, что на продажах сказывается сезонность, и модель прогнозирования должна это учитывать.
- Каковы шансы каждой команды выиграть турнир Национальной студенческой спортивной ассоциации? (См. главу 80.)
- Ряд тем прогнозной аналитики рассмотрен в главах 35, 55–67 и 74.

### **❓ Что такое предписывающая аналитика?**

Предписывающая аналитика предполагает использование математических моделей, как правило, основанных на данных, для принятия решений, доводящих до максимума или до минимума соответствующую количественную цель. Далее приведено несколько примеров таких целей.

- Какая цена продукта максимизирует прибыль? (См. главы 88–90.)
- Какой товарный ассортимент максимизирует ежемесячную прибыль компании Eli Lilly? (См. главу 29.)
- Как спланировать трудовую деятельность сотрудников для минимизации эксплуатационных расходов? (См. главу 30.)
- Как компания может минимизировать затраты на логистику продукции? (См. главу 32.)
- Как организация может выбрать проекты, максимально обеспечивающие достижение корпоративных целей при условии ограниченных ресурсов? Например, как с учетом ограниченного капитала и количества программистов Microsoft может максимизировать прибыльность новых продуктов? Как местный школьный совет может использовать свои ограниченные средства для максимального улучшения оценок выпускников средней школы? Эти вопросы обсуждаются в главе 33.
- Сколько необходимо класть на счет каждый год для получения суммы, достаточной для выхода на пенсию? (См. главу 34.)
- Как букмекеры могут установить пороги безубыточности, гарантирующие получение прибыли? (См. главу 35.)
- Где должны быть расположены склады для минимизации перевозок? (См. главу 36.)
- Как распределить работу между сотрудниками, учитывая удовлетворенность работой и начальников, и сотрудников? (См. главу 37.)

- В какой последовательности работник почтовой службы должен доставлять пакеты для минимизации времени доставки всех пакетов? (См. главу 38.)
- Какое количество почтовых открыток, напечатанных ко Дню святого Валентина, позволит компании Hallmark максимизировать ожидаемую прибыль? (См. главу 77.)
- Какую цену в заявке на проект строительства должна указать компания для получения максимальной ожидаемой прибыли? (См. главу 78.)
- Какое распределение активов минимизирует риск, необходимый для получения желаемого ожидаемого дохода? (См. главу 79.)

### **❓ Почему значение аналитики возрастает?**

В последнее время аналитика становится все более важной по нескольким причинам. Во-первых, в свободном доступе появилось гораздо больше интересных данных. Например, компания Factual пытается собрать все представляющие интерес данные; она владеет базой данных, содержащей местоположения всех автозаправочных станций США и информацию о питательной ценности всех товаров, продаваемых в розничных сетях. Новые источники важных данных возникают со все возрастающей быстротой. Например, все арены НБА оборудованы камерами, отслеживающими местоположение каждого игрока и мяча в каждую секунду игры.

Кроме того, компьютеры и программы все время становятся более быстродействующими, что облегчает обработку больших массивов данных. Например, надстройка PowerPivot (см. главу 45) упрощает создание отчетов по бизнес-аналитике на основе сотен миллионов строк данных. Такое программное обеспечение, как SAS и HADOOP, созданное не на основе Microsoft Excel, позволяет относительно легко обрабатывать и анализировать огромные массивы данных.

### **❓ Насколько важна аналитика для вашей организации?**

Значимость аналитики для организации зависит от преимуществ, получаемых от аналитического подхода, в противовес затратам на его реализацию. Например, перевешивает ли выгода от оценивания преподавателей аналитическими методами затраты на обработку данных?

Свет на важность аналитики может пролить применение аналитики в профессиональном спорте. В бейсболе (денежном виде спорта, как вы знаете) для оценки бейсболистов есть множество доступных данных. Эти данные достаточно легко поддаются анализу, так что почти в каждой крупной команде есть аналитический отдел. Раньше такие отделы были только у нескольких команд НБА. В 2013 г. на конференции по спортивной аналитике 2013 Sloan Conference on Sports Analytics не была представлена только одна команда («Лейкерс», и мы знаем, как они выступили). Однако в командах НФЛ аналитических отделов пока немного, возможно, потому, что многие важные проблемы в футбольной аналитике не поддаются решению. Например, в бейсболе и баскетболе нетрудно выявить справедливую ценность игрока. В американском футболе сделать это гораздо сложнее. Как мож-

но оценить блокировку великолепного паса при нехватке статистических данных для оценки неадекватных действий судьи на линии?

### **❓ Что необходимо знать для проведения анализа?**

Во многих университетах существуют программы для студентов и аспирантов, а также онлайн-программы по сертификации аналитиков. В 2013 г. по аналогии с экзаменом на получение сертификата дипломированного бухгалтера-ревизора (Certified Public Accountant, CPA) институт INFORMS (Institute for Operations Research and Management Science) предложил ввести экзамен для сертификации специалистов по аналитике. Как минимум, специалист по аналитике должен обладать знаниями в следующих областях:

- Уметь обрабатывать данные в любой форме. Сюда входит обработка не количественных данных, таких как твиты в Twitter.
- Знать, как собрать вместе разрозненные наборы данных для аналитического исследования. Например, ресторану быстрого питания требуется снизить текучесть кадров. Для этого нужны сведения о каждом работнике (образование, возраст, результаты тестов и другие статистические данные), а также о его успехах. Тогда можно разработать модель, предсказывающую текучесть кадров на основе того, что было известно при приеме сотрудника на работу. К сожалению, во многих компаниях сведения о нанятых сотрудниках и сведения об их работе находятся в отдельных базах данных, не пересекающихся друг с другом.
- Иметь навыки программирования на таких языках, как Java или Python.
- Уметь анализировать и обрабатывать данные в Excel на уровне опытного пользователя.
- Уметь анализировать большие наборы данных с помощью статистических пакетов, таких как R, SPSS или SAS.
- Знать статистику, в том числе все формы прогнозного анализа и алгоритмы классификации. Например, как можно причислить пациента к группе высокого риска возникновения инфаркта?
- Уметь моделировать неопределенные ситуации, такие как будущая стоимость пенсионного портфеля по плану 401k или очередь в приемном покое больницы.
- Использовать оптимизацию для поиска наилучшего способа сделать что-либо. Например, какое распределение площадей под различные товары максимизирует прибыль универсама?

### **❓ Какие трудности могут возникнуть при осуществлении анализа?**

При осуществлении анализа могут возникнуть некоторые проблемы. Вот несколько примеров.

- Выбор метрики для оценки успеха часто является предметом споров. Например, пока не появился гуру саберметрики Билл Джеймс, все думали, что левых игроков необходимо оценивать на основе перемещения по полю, то есть

по доле отскочивших возле полевого игрока мячей, которые были успешно отыграны. Очевидно, что такая метрика не учитывает тот факт, что медленный шорт-стоп (игрок между второй и третьей базой) позволяет множеству мячей уйти в дальнюю часть поля и превратиться в хит. Билл Джеймс ввел ранжирующий фактор, который гораздо лучше подходит для оценки перемещения игрока по полю. Для шорт-стопа, например, ранжирующий фактор равен количеству выводов в аут и передач из расчета на иннинг для этого игрока, разделенному на среднее количество выводов в аут и передач из расчета на иннинг для всех шорт-стопов. Ранжирующий фактор, превышающий 1, показывает, что игрок добирается до тех мячей, которые средний шорт-стоп упускает. Оззи Смит является одним из самых титулованных шорт-стопов. Его коэффициент перемещения по полю за карьеру, 0,966, не особо впечатляет, но в лучшие годы ранжирующий фактор был у него просто великолепным. Кроме того, Оззи играл на искусственном покрытии, что делает этот показатель еще более поразительным. Мораль: не стесняйтесь сомневаться в метриках, которые, по вашему мнению, не отражают реальной производительности вашей организации.

- Как упоминалось ранее, часто источники данных, требуемые для аналитического проекта, не связаны друг с другом. Например, компании-разработчику программного обеспечения требовался прогноз будущих доходов. Известно, что будущие доходы зависят от пробных версий, отправленных потенциальным клиентам, и звонков потенциальным клиентам. К сожалению, данные о продажах и данные о клиентах находились в разных, никак не связанных базах данных, поэтому проект пострадал от длительной, совершенно излишней задержки.
- При проведении прогнозного анализа аналитики часто не знают, какие данные необходимы для принятия правильного решения. Я предлагаю работать в обратном направлении. После определения того, что требуется спрогнозировать, подумайте о переменных, которые помогут предсказать интересующую переменную. Вот несколько примеров ситуаций, в которых для завершения аналитического проекта потребовалось создать новый набор данных.
  - Известный спортивный аналитик Джефф Сагарин и автор этой книги решили разработать для игроков НБА метод оценки способностей к игре в защите. Поскольку подсчет очков в НБА включает мало статистики по защите (перехваты и блоки), потребовался новый набор данных. Мы рассудили, что хороший защитник поможет своей команде потерять меньше очков, когда находится в игре, а его отсутствие на площадке обернется для команды потерей большего количества очков. Мы создали набор данных, включивший минуты, проведенные игроками на площадке, и количество проигранных за это время очков. На основе подобных данных мы могли бы определить вклад в защиту таких великих защитников, как Рон Артест, Кевин Гарнетт и Марк Газоль.
  - Компания United Health решила снизить расходы на медицинское обслуживание. Компания постановила, что ей необходимо знать, лечили ли

сотрудников от болезней, которые могут вызвать большие расходы на медицинское обслуживание в будущем. Большие расходы на медицинское обслуживание в будущем (и ужасные проблемы) вызывает, например, диабет. Компания выделила своим сотрудникам по \$450 для обследования на диабет и других подобных болезней. Удивительно, но 30% сотрудников являлись диабетиками или были предрасположены к диабету, но не знали об этом. Лечение таких пациентов позволило компании United Health избежать повышения расходов на медицинское обслуживание и улучшить качество жизни своих сотрудников в будущем.

### **❓ Какие тенденции будут влиять на развитие аналитики?**

Более быстрые компьютеры и улучшенные алгоритмы для анализа данных и выполнения оптимизаций фактически гарантируют, что аналитика станет более важной в будущем. В будущем улучшения в визуализации облегчат начинающему аналитику обнаружение паттернов в сложных данных. В главе 54 вы узнаете множество интересных приемов создания диаграмм, позволяющих лучше понять данные.

Вероятно, самой важной тенденцией в аналитике станет доступность множества аналитических приложений в любом месте в любое время на мобильных устройствах. Например, в 2011 г. в Сан-Рамоне, Калифорния, было разработано приложение для смартфона, в котором зарегистрированы люди, умеющие делать искусственное дыхание и способные при необходимости оказать помощь тем, кто в ней нуждается. Если при поступлении вызова 911 скорая помощь не может приехать вовремя, всем зарегистрированным лицам отправляется текстовое сообщение с местоположением пострадавшего человека. Таким образом, больной человек получит помощь гораздо быстрее.

Еще одной тенденцией в области аналитики станет увеличение количества прогнозов в реальном времени, приводящих к действиям в реальном времени. Например, если прогноз покажет, что в целевом магазине пепси-кола закончится к 11 утра, то на основе аналитического приложения с данными о местоположении всех грузовиков компании Pepsi водителю соответствующего грузовика будет отправлено сообщение о необходимости пополнить запасы пепси-колы в этом магазине к 11 утра.

## ГЛАВА 29

# Введение в оптимизацию с надстройкой Поиск решения

Во многих ситуациях требуется найти оптимальный способ сделать что-либо. Скажем, вам необходимо найти на листе значения в определенных ячейках, оптимизирующие (увеличивающие до максимума или сводящие к минимуму) определенную цель. Надстройка Поиск решения (Solver) поможет вам найти ответы на следующие вопросы.

- Как крупная фармацевтическая компания может определить ежемесячный ассортимент продукции, производимый на заводе в Индианаполисе, который максимизирует прибыль компании?
- Если компания Microsoft производит приставки Xbox в трех месторасположениях, каким образом минимизировать затраты на удовлетворение спроса на них?
- Какая цена на приставки Xbox и игры обеспечит компании Microsoft максимальную прибыль от продаж Xbox?
- Компания Microsoft намеревается предпринять 20 стратегических инициатив, с которыми в течение следующих пяти лет будут связаны денежные средства и квалифицированные программисты. Имеющихся ресурсов для всех 20 проектов недостаточно. Какие инициативы ей следует предпочесть?
- Как букмекеры определяют оптимальный набор рейтингов при поиске правильной разницы в счете для команд НФЛ?
- Как распределить пенсионный портфель между акциями высокотехнологичных компаний, недооцененными акциями, облигациями, наличными деньгами и золотом?

Оптимизационная модель состоит из трех частей: целевая ячейка, изменяемые ячейки и ограничения. Целевая ячейка представляет собой цель. Как правило, величину в целевой ячейке необходимо либо минимизировать, либо максимизировать. В ранее приведенном вопросе об ассортименте продукции фармацевтической компании предполагается, что директору завода необходимо каждый месяц максимизировать прибыльность завода. Ячейка, в которой вычисляется прибыльность, и будет целевой ячейкой. Целевые ячейки для каждой из ситуаций, описанных в начале главы, указаны в табл. 29.1. Имейте в виду, что в некоторых ситуациях возможно наличие нескольких целевых ячеек. Например, у компании Microsoft может быть и вторичная цель — максимизировать долю рынка приставок Xbox.



**Таблица 29.1.** Список целевых ячеек

<b>Модель</b>	<b>Максимизировать или минимизировать</b>	<b>Целевая ячейка</b>
Ассортимент продукции фармацевтической компании	Максимизировать	Ежемесячная прибыль
Поставка приставок Xbox	Минимизировать	Расходы на реализацию
Цены на приставки Xbox	Максимизировать	Прибыль от приставок Xbox и игр
Проектные инициативы Microsoft	Максимизировать	Чистая приведенная стоимость, внесенная выбранными проектами
Рейтинги НФЛ	Минимизировать	Разница между очками по рейтингу и фактически набранными очками в игре
Пенсионный портфель	Минимизировать	Фактор риска для портфеля

Изменяемые ячейки — это ячейки на листе, значения в которых можно изменить для оптимизации значения в целевой ячейке. В примере с фармацевтической компанией директор завода может регулировать количество каждого продукта, производимого за месяц. В этой модели изменяемыми ячейками являются ячейки, в которых записываются эти количества. Для моделей, описанных в начале главы, соответствующие изменяемые ячейки определены в табл. 29.2, а соответствующие ограничения указаны в табл. 29.3.

**Таблица 29.2.** Список изменяемых ячеек

<b>Модель</b>	<b>Изменяемые ячейки</b>
Ассортимент продукции фармацевтической компании	Количество каждого продукта, произведенного за месяц
Поставка приставок Xbox	Количества, произведенные на каждом заводе за каждый месяц, поставляемые каждому клиенту
Цены на приставки Xbox	Цена на приставки и игры
Проектные инициативы Microsoft	Выбираемые проекты
Рейтинги НФЛ	Рейтинги команд
Пенсионный портфель	Доли денег, вложенные в каждый класс активов

Удобнее всего изучать надстройку Поиск решения (Solver) на подробных примерах. В последующих главах я опишу способы применения надстройки Поиск решения (Solver) ко всем описанным в данной главе ситуациям, а также к другим важным бизнес-задачам.

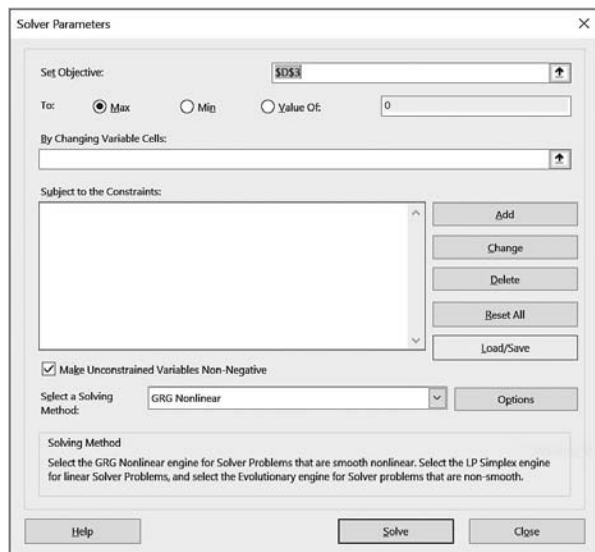
Таблица 29.3. Список ограничений

Модель	Ограничения
Ассортимент продукции фармацевтической компании	Для ассортимента должны использоваться только доступные ресурсы. Производить столько продукта, сколько можно продать
Поставка приставок Xbox	Количество отгружаемых единиц за каждый месяц не должно превышать мощность завода. Убедитесь, что каждый клиент получает требуемое количество приставок Xbox
Цены на приставки Xbox	Цены не могут намного отличаться от цен конкурентов
Проектные инициативы Microsoft	Для выбранных проектов можно использовать только доступное количество денег и опытных программистов
Рейтинги НФЛ	Нет
Пенсионный портфель	Инвестировать деньги куда-либо (например, в наличные деньги). Ожидаемый доход должен быть не менее 10% от инвестиций

Для активации надстройки Поиск решения (Solver) на вкладке Файл (File) выберите команду Параметры (Options), затем среди параметров Excel выберите раздел Надстройки (Add-Ins). В нижней части диалогового окна в раскрывающемся списке Управление (Manage) выберите Надстройки Excel (Excel Add-Ins) и нажмите Перейти (Go). В диалоговом окне Надстройки (Add-Ins) включите флажок Поиск решения (Solver Add-In) и нажмите ОК. После активации надстройку Поиск решения (Solver) можно выбрать и запустить на вкладке Данные (Data) в группе Анализ (Analysis). На рис. 29.1 показано диалоговое окно Параметры поиска решения (Solver Parameters). В следующих нескольких главах вы увидите, как использовать это диалоговое окно для настройки целевой ячейки, изменяемых ячеек и ограничений.

Обратите внимание на раскрывающийся список Выберите метод решения (Select A Solving Method). В этом списке для решения оптимизационной задачи необходимо выбрать соответствующий метод решения.

- Линейные оптимизационные задачи решаются симплекс-методом. Как вы увидите в главах 30–34, линейная оптимизационная задача — это задача, в которой целевая ячейка и ограничения создаются путем добавления условий в формуле (изменяемая\_ячейка)\*(константа).
- Нелинейный метод обобщенного понижающего градиента (ОПГ) используется для решения оптимизационных задач, в которых значения в целевой ячейке или значения для некоторых ограничений не являются линейными и вычисляются с использованием общих математических операций, таких как умножение, деление или возведение в степень значений в изменяемых ячейках, либо с применением экспоненциальных или тригонометрических функций и т. п. Метод ОПГ включает эффективный параметр Несколько начальных то-



**Рис. 29.1.** Диалоговое окно Параметры поиска решения

чек (Multistart), позволяющий решить множество задач, которые в предыдущих версиях Excel решались неправильно. Нелинейный метод ОПГ подробно рассмотрен в главах 35–36.

- Эволюционный поиск решения используется в случае, когда целевая ячейка или ограничения содержат негладкие функции со ссылками на изменяемые ячейки. Негладкая функция — это функция с резко меняющимся градиентом (тангенсом угла наклона касательной, проведенной в исследуемой точке). Например, при  $x = 0$  градиент функции  $|x|$  резко меняется с  $-1$  на  $+1$ . Если целевая ячейка или ограничения содержат функции ЕСЛИ (IF), СУММЕСЛИ (SUMIF), СЧЁТЕСЛИ (COUNTIF), СУММЕСЛИМН (SUMIFS), СЧЁТЕСЛИМН (COUNTIFS), СРЗНАЧЕСЛИ (AVERAGEIF), СРЗНАЧЕСЛИМН (AVERAGEIFS), ABS, МАКС (MAX) или МИН (MIN) со ссылками на изменяемые ячейки, то это негладкие функции, и, вероятно, для решения такой оптимизационной задачи более всего подойдет эволюционный поиск. Эволюционный поиск решения рассматривается в главах 37–38.

Что делает надстройка Поиск решения (Solver) после того, как вы задали целевую ячейку, изменяемые ячейки и ограничения? Для ответа на этот вопрос вам необходимо освоить некоторую терминологию. Любой перечень изменяемых ячеек, удовлетворяющих ограничениям модели, является допустимым решением. Например, допустимым решением является любой ассортимент продуктов, удовлетворяющий следующим трем условиям:

- для его производства не требуется больше сырья и рабочей силы, чем доступно;
- каждого продукта не производится больше, чем требуется;
- для любого продукта не производится отрицательное количество продукта.

По сути, надстройка Поиск решения (Solver) выполняет поиск всех допустимых решений и отбирает решение с оптимальным значением в целевой ячейке (с наибольшим значением для оптимизации по максимуму и с наименьшим значением для оптимизации по минимуму). Такое решение называется *оптимальным решением*. Как вы увидите в главе 30, некоторые модели поиска не имеют оптимального решения, а некоторые модели имеют единственное решение. Остальные модели поиска решения имеют несколько (фактически бесконечное число) оптимальных решений. В следующей главе мы начнем изучение методов поиска решения с примера про ассортимент продукции фармацевтической компании.

## Задания

Для каждой описанной ситуации определите целевую ячейку, изменяемые ячейки и ограничения.

1. Я взял \$100 000 в ипотеку на 15 лет. Годовая процентная ставка составляет 8%. Как определить ежемесячные выплаты по ипотеке?
2. Как автоконцерн должен распределить свой бюджет на рекламу между различными рекламными форматами?
3. Как соблюсти рациональный баланс при перевозке учащихся в более отдаленные школы?
4. Где должна быть расположена единственная в городе больница?
5. Как фармацевтическая компания должна распределить усилия своего торгового персонала между товарами?
6. Фармацевтическая компания выделила 2 млрд долл. на покупку биотехнологических компаний. Какие именно компании следует купить?
7. Ставка налога, взимаемого с фармацевтической компании, зависит от страны/региона, где производится продукт. Как фармацевтическая компания может определить, где следует производить тот или иной препарат?

## ГЛАВА 30

# Поиск решения при определении оптимального ассортимента продукции

### Обсуждаемые вопросы

- Как определить ежемесячный ассортимент продукции, максимизирующий прибыльность?
- Всегда ли модель поиска решения имеет решение?
- Что означает выдаваемый моделью поиска решения ответ «Значения ячейки целевой функции не сходятся»?

## Ответы на вопросы

### ❓ Как определить ежемесячный ассортимент продукции, максимизирующий прибыльность?

Компаниям часто приходится определять, какое количество продуктов нужно произвести каждый месяц. В своей простейшей форме задача определения ассортимента продукции сводится к определению количества каждого продукта, производимого за месяц, для максимизации прибыли. Ассортимент продукции должен соответствовать следующим ограничениям.

- Для производства ассортимента продукции могут использоваться только доступные ресурсы.
- Поскольку спрос на каждый продукт ограничен, не следует производить продуктов за месяц больше, чем диктует спрос. Избыточное производство — это напрасный расход ресурсов (например, производство препарата с коротким сроком хранения).

Теперь приступим к решению задачи выбора ассортимента продукции. Допустимое решение этой задачи см. в файле **Prodmix.xlsx** на листе **Feasible Solution** (рис. 30.1). Пробные значения количеств каждого препарата введены в строке 2.

Допустим, вы работаете на фармацевтическую компанию, которая производит на своем заводе шесть препаратов. Для производства каждого препарата требуются

	B	C	D	E	F	G	H	I
1								
2		<b>Pounds made</b>	<b>150</b>	<b>160</b>	<b>170</b>	<b>180</b>	<b>190</b>	<b>200</b>
3	<b>Available</b>	<b>Product</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
4	<b>4500</b>	<b>Labor</b>	<b>6</b>	<b>5</b>	<b>4</b>	<b>3</b>	<b>2.5</b>	<b>1.5</b>
5	<b>1600</b>	<b>Raw Material</b>	<b>3.2</b>	<b>2.6</b>	<b>1.5</b>	<b>0.8</b>	<b>0.7</b>	<b>0.3</b>
6		<b>Unit price</b>	<b>\$ 12.50</b>	<b>\$ 11.00</b>	<b>\$ 9.00</b>	<b>\$ 7.00</b>	<b>\$ 6.00</b>	<b>\$ 3.00</b>
7		<b>Variable cost</b>	<b>\$ 6.50</b>	<b>\$ 5.70</b>	<b>\$ 3.60</b>	<b>\$ 2.80</b>	<b>\$ 2.20</b>	<b>\$ 1.20</b>
8		<b>Demand</b>	<b>960</b>	<b>928</b>	<b>1041</b>	<b>977</b>	<b>1084</b>	<b>1055</b>
9		<b>Unit profit cont.</b>	<b>\$ 6.00</b>	<b>\$ 5.30</b>	<b>\$ 5.40</b>	<b>\$ 4.20</b>	<b>\$ 3.80</b>	<b>\$ 1.80</b>
10								
11								
12		<b>Profit</b>	<b>\$ 4,504.00</b>					
13					<b>Available</b>			
14		<b>Labor Used</b>	<b>3695</b>	<b>&lt;=</b>	<b>4500</b>			
15		<b>Raw Material Used</b>	<b>1488</b>	<b>&lt;=</b>	<b>1600</b>			

**Рис. 30.1.** Допустимое решение для задачи выбора ассортимента продукции

рабочая сила и сырье. На рис. 30.1 для производства фунта каждого продукта в строке 4 указано требуемое количество рабочего времени в часах, а в строке 5 — необходимое количество фунтов сырья. Например, для производства фунта препарата 1 требуется 6 часов рабочего времени и 3,2 фунта сырья. Цена за фунт для каждого препарата указана в строке 6, себестоимость производства фунта препарата — в строке 7, а отчисление прибыли с каждого фунта — в строке 9. Например, препарат 2 продается по \$11 за фунт, себестоимость производства фунта препарата 2 равна \$5,70, и прибыль с каждого фунта составляет \$5,30. Ежемесячная потребность в каждом препарате указана в строке 8. Например, потребность в препарате 3 составляет 1,041 фунта. В текущем месяце в распоряжении компании имеется 4500 часов рабочего времени и 1600 фунтов сырья. Каким образом компания может максимизировать прибыль?

Если бы вы не знали про существование инструмента Поиск решения (Solver), то приступили бы к решению задачи с такого бока: создали лист для отслеживания прибыли и использования ресурсов, связанных с ассортиментом продукции. Затем для оптимизации получаемой прибыли вы методом проб и ошибок меняли бы ассортимент продукции в рамках доступного рабочего времени, объемов сырья и спроса на любой производимый препарат. В этом процессе на стадии проб и ошибок можно использовать Поиск решения (Solver). По существу Поиск решения (Solver) представляет собой метод оптимизации, безусловно выполняющий поиск методом проб и ошибок.

Ключом к решению задачи с выбором ассортимента продукции является эффективный расчет используемых ресурсов и прибыли, связанных с любым заданным ассортиментом продукции. Важным инструментом в этом расчете является функция СУММПРОИЗВ (SUMPRODUCT). Эта функция перемножает соответствующие значения в диапазонах ячеек и возвращает их сумму. Все диапазоны ячеек в аргументах функции СУММПРОИЗВ должны иметь одинаковую размерность, то есть функцию СУММПРОИЗВ можно использовать для двух строк или двух столбцов, но не для столбца и строки.

В качестве примера использования функции СУММПРОИЗВ вычислим применение ресурсов. Используемое рабочее время вычисляется по формуле:

(рабочее\_время\_для\_фунта\_препарата 1)\*(количество\_фунтов\_препарата 1) + (рабочее\_время\_для\_фунта\_препарата 2)\*(количество\_фунтов\_препарата 2) + ... + (рабочее\_время\_для\_фунта\_препарата 6)\*(количество\_фунтов\_препарата 6).

Вы можете рассчитать рабочее время как  $=D2*D4+E2*E4+F2*F4+G2*G4+H2*H4+I2*I4$ . Аналогично объем сырья вычисляется как  $=D2*D5+E2*E5+F2*F5+G2*G5+H2*H5+I2*I5$ . Однако ввод таких формул даже для шести продуктов отнимает много времени. Только представьте, сколько времени это займет в компании, производящей на своем заводе, например, 50 продуктов. Гораздо проще вычислить рабочее время и объемы сырья, скопировав формулу  $=СУММПРОИЗВ(\$D\$2:\$I\$2,D4:I4)$  из D14 в D15. Эта формула так же вычисляет выражение  $D2*D4+E2*E4+F2*F4+G2*G4+H2*H4+I2*I4$  (используемое количество рабочего времени), как и предыдущая, но ее проще ввести. Укажите знак \$ в диапазоне D2:I2 для использования ассортимента продукции из строки 2 при копировании формулы. В ячейке D15 по такой же формуле вычисляется используемый объем сырья.

Аналогичным образом прибыль высчитывается по формуле:

(прибыль\_с\_фунта\_препарата 1)\*(количество\_препарата 1) + (прибыль\_с\_фунта\_препарата 2)\*(количество\_препарата 2) + ... + (прибыль\_с\_фунта\_препарата 6)\*(количество\_препарата 6).

Прибыль вычисляется в ячейке D12 по простой формуле

$=СУММПРОИЗВ(D9:I9,\$D\$2:\$I\$2)$ .

Теперь вы можете определить три компонента модели поиска решения для выбора ассортимента продукции:

- **Целевая ячейка** — цель состоит в максимизации прибыли, вычисляемой в D12.
- **Изменяемые ячейки** — объем каждого произведенного продукта в фунтах указан в диапазоне D2:I2.
- **Ограничения** — имеются следующие ограничения:
  - не использовать больше рабочего времени и сырья, чем доступно. Это означает, что значения в ячейках D14:D15 (используемые ресурсы) должны быть меньше или равны значениям в ячейках F14:F15 (доступные ресурсы);
  - не производить больше препарата, чем требуется. Это означает, что значения в ячейках D2:I2 (количество произведенного препарата в фунтах) должны быть меньше или равны значениям потребности в каждом препарате (указанным в ячейках D8:I8);
  - невозможно произвести отрицательное количество любого препарата.

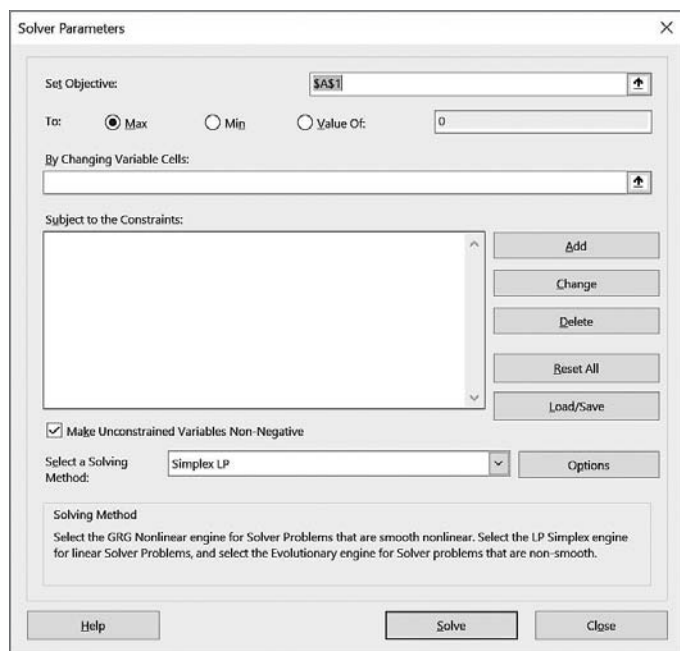
Я покажу вам, как ввести целевую ячейку, изменяемые ячейки и ограничения в инструменте Поиск решения (Solver). После этого для поиска максимально при-

большого ассортимента выпускаемой продукции останется только нажать **Найти решение (Solve)**.

### ПРИМЕЧАНИЕ

Как я говорил в главе 29, инструмент Поиск решения (Solver) можно активировать так: вкладка **Файл | Параметры | Надстройки (FILE | Options | Add-Ins)**. Затем в списке **Управление (Manage)** выберите **Надстройки Excel (Excel Add-ins)** и нажмите **Перейти (Go)**. Включите флажок **Поиск решения (Solver Add-In)** и нажмите **OK**.

Для начала откройте вкладку **Данные (Data)** и в группе **Анализ (Analysis)** выберите инструмент **Поиск решения (Solver)**. Откроется диалоговое окно **Параметры поиска решения (Solver Parameters)**, показанное на рис. 30.2.

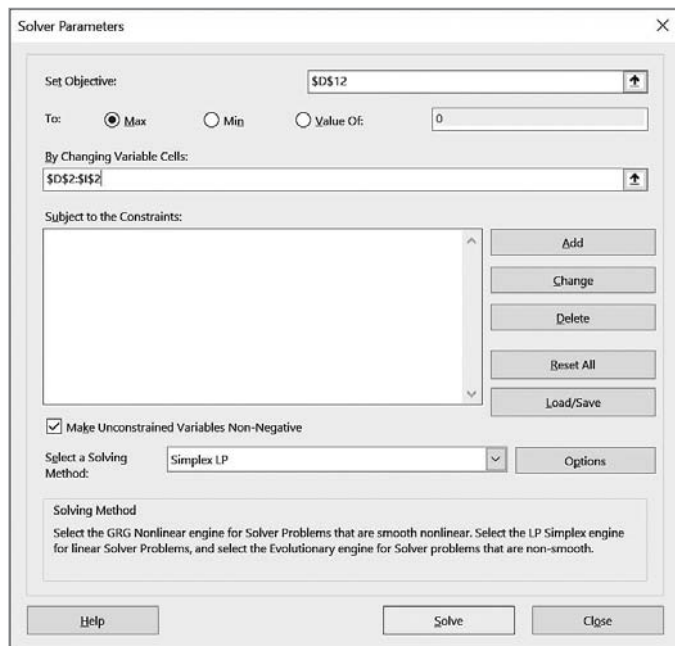


**Рис. 30.2.** Диалоговое окно Параметры поиска решения

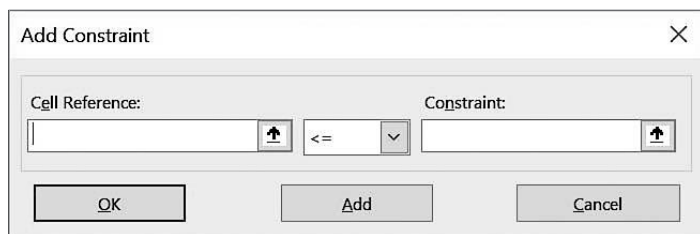
В поле **Оптимизировать целевую функцию (Set Objective)** выберите ячейку со значением прибыли (D12). В поле **Изменяя ячейки переменных (By Changing Variable Cells)** выберите диапазон **D2:I2**, содержащий количество каждого произведенного продукта в фунтах. Диалоговое окно должно выглядеть так, как на рис. 30.3.

Теперь все готово для добавления в модель ограничений. Нажмите **Добавить (Add)**. Откроется диалоговое окно **Добавление ограничения (Add Constraint)**, показанное на рис. 30.4.





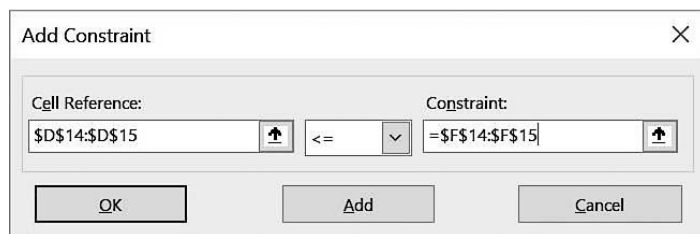
**Рис. 30.3.** Диалоговое окно Параметры поиска решения, в котором определены целевая ячейка и изменяемые ячейки



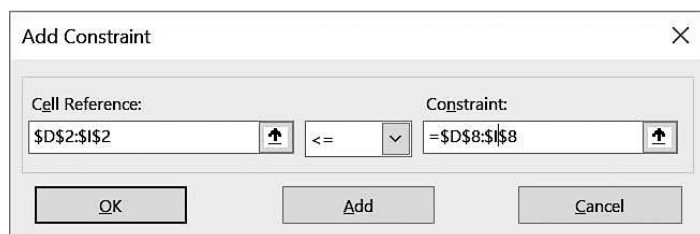
**Рис. 30.4.** Диалоговое окно Добавление ограничения

Для добавления ограничений по ресурсам в поле **Ссылка на ячейки (Cell Reference)** выберите диапазон **D14:D15**. Затем выберите из среднего списка **<=**, затем в поле **Ограничение (Constraint)** выберите диапазон **F14:F15**. Диалоговое окно **Добавление ограничения (Add Constraint)** должно выглядеть так, как на рис. 30.5.

Вы гарантировали, что при подборе различных значений для изменяемых ячеек будут приниматься во внимание только комбинации, удовлетворяющие как условию **D14<=F14** (используемое рабочее время не превышает доступное рабочее время), так и условию **D15<=F15** (объем используемого сырья не превышает доступный объем сырья). Нажмите **Добавить (Add)** и заполните диалоговое окно как на рис. 30.6.



**Рис. 30.5.** Диалоговое окно Добавление ограничения с заданными ограничениями по ресурсам

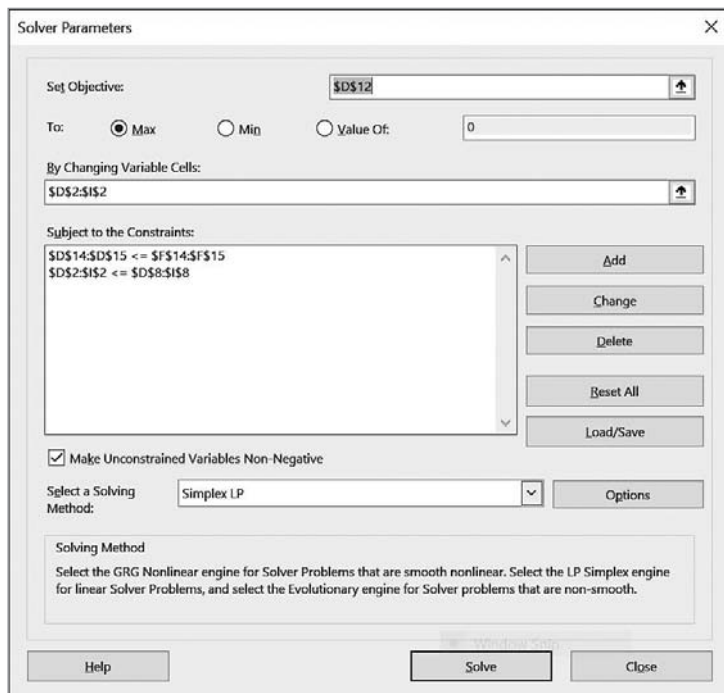


**Рис. 30.6.** Диалоговое окно Добавление ограничения с заданным ограничением по потребности в препаратах

Добавление этих ограничений гарантирует, что при подборе различных значений для изменяемых ячеек будут учитываться только комбинации, удовлетворяющие следующим условиям:

- $D2 \leq D8$  (количество произведенного препарата 1 меньше или равно потребности в препарате 1);
- $E2 \leq E8$  (количество произведенного препарата 2 меньше или равно потребности в препарате 2);
- $F2 \leq F8$  (количество произведенного препарата 3 меньше или равно потребности в препарате 3);
- $G2 \leq G8$  (количество произведенного препарата 4 меньше или равно потребности в препарате 4);
- $H2 \leq H8$  (количество произведенного препарата 5 меньше или равно потребности в препарате 5);
- $I2 \leq I8$  (количество произведенного препарата 6 меньше или равно потребности в препарате 6).

В диалоговом окне **Добавление ограничения (Add Constraint)** нажмите **OK**. Включение флажка **Сделать переменные без ограничений неотрицательными (Make Unconstrained Variables Non-Negative)** гарантирует, что значения во всех изменяемых ячейках будут больше или равны 0. Диалоговое окно **Параметры поиска решения** должно выглядеть так, как на рис. 30.7.



**Рис. 30.7.** Окончательный вид диалогового окна Параметры поиска решения для задачи выбора ассортимента продукции

Затем выберите в списке Выберите метод решения (Select A Solving Method) вариант Поиск решения лин. задач симплекс-методом (Simplex LP). Его необходимо выбрать потому, что задача выбора ассортимента продукции представляет собой особый тип задачи поиска решения, называемый *линейной моделью*. Модель поиска решения является линейной при следующих условиях:

- значение в целевой ячейке вычисляется путем суммирования членов в виде (изменяемая\_ячейка)\*(константа);
- каждое ограничение удовлетворяет требованию линейной модели. Это означает, что каждое ограничение вычисляется путем суммирования членов в виде (изменяемая\_ячейка)\*(константа) и сравнения сумм с константой.

Почему задача поиска решения для выбора ассортимента является линейной? Значение в целевой ячейке (прибыль) вычисляется следующим образом:

(прибыль\_с\_фунта\_препарата 1)\*(количество\_препарата 1) + (прибыль\_с\_фунта\_препарата 2)\*(количество\_препарата 2) + ... + (прибыль\_с\_фунта\_препарата 6)\*(количество\_препарата 6).

Это соответствует образцу, в котором значение в целевой ячейке получается путем суммирования членов в виде (изменяемая\_ячейка)\*(константа).

Ограничение по рабочему времени вычисляется путем сравнения значения выражения

$$(\text{рабочее\_время\_для\_фунта\_препарата } 1) * (\text{количество\_фунтов\_препарата } 1) + (\text{рабочее\_время\_для\_фунта\_препарата } 2) * (\text{количество\_фунтов\_препарата } 2) + \dots + (\text{рабочее\_время\_для\_фунта\_препарата } 6) * (\text{количество\_фунтов\_препарата } 6)$$

с доступным количеством рабочего времени.

Следовательно, ограничение по рабочему времени вычисляется путем сложения членов в виде  $(\text{изменяемая\_ячейка}) * (\text{константа})$  и последующего сравнения суммы с константой. И ограничение по рабочему времени, и ограничение по сырью удовлетворяют требованиям линейной модели.

Ограничения по потребности в препарате принимают следующий вид:

$$\begin{aligned} (\text{количество произведенного препарата } 1) &\leq (\text{потребность в препарате } 1) \\ (\text{количество произведенного препарата } 2) &\leq (\text{потребность в препарате } 2) \\ &\dots \\ (\text{количество произведенного препарата } 6) &\leq (\text{потребность в препарате } 6) \end{aligned}$$

Каждое ограничение по потребности в препарате также удовлетворяет требованиям линейной модели, поскольку оно вычисляется путем сложения членов в виде  $(\text{изменяемая\_ячейка}) * (\text{константа})$  и сравнения суммы с константой.

Если модель ассортимента продукции линейная, нам беспокоиться не о чем.

- Если модель для поиска решения линейная и вы выбрали симплекс-метод, для модели гарантированно будет найдено оптимальное решение. Если модель нелинейная, оптимальное решение может быть (а может и не быть) найдено.
- Если модель для поиска решения линейная и вы выбрали симплекс-метод, поиск оптимального решения для модели будет выполняться по эффективному алгоритму (симплекс-методом). Если модель для поиска решения линейная, но вы выбрали не симплекс-метод, поиск решения будет выполнен по малоэффективному алгоритму, методом обобщенного понижающего градиента (Generalized Reduced Gradient, GRG2), и поиск оптимального решения будет сопряжен с трудностями.

Когда вы нажмете **Найти решение (Solve)**, для модели выбора ассортимента продукции будет найдено оптимальное решение (если оно существует). Как я утверждал в главе 29, оптимальным решением по множеству всех допустимых решений для модели выбора ассортимента продукции является набор значений в изменяемых ячейках (количество произведенных препаратов в фунтах), максимизирующих прибыль. Допустимое решение представляет собой набор значений в изменяемых ячейках, удовлетворяющих всем ограничениям. Значения в изменяемых ячейках, показанные на рис. 30.8, являются допустимым решением, поскольку все объемы производства неотрицательны, объем производства не превышает спрос на товар и используемые ресурсы не превышают доступные ресурсы. См. лист **Feasible Solution** в файле **Prodmix.xlsx**.

	B	C	D	E	F	G	H	I
1								
2		Pounds made	150	160	170	180	190	200
3	Available	Product	1	2	3	4	5	6
4	4500	Labor	6	5	4	3	2.5	1.5
5	1600	Raw Material	3.2	2.6	1.5	0.8	0.7	0.3
6		Unit price	\$ 12.50	\$ 11.00	\$ 9.00	\$ 7.00	\$ 6.00	\$ 3.00
7		Variable cost	\$ 6.50	\$ 5.70	\$ 3.60	\$ 2.80	\$ 2.20	\$ 1.20
8		Demand	960	928	1041	977	1084	1055
9		Unit profit cont.	\$ 6.00	\$ 5.30	\$ 5.40	\$ 4.20	\$ 3.80	\$ 1.80
10								
11								
12		Profit	\$ 4,504.00					
13					Available			
14		Labor Used	3695	<=	4500			
15		Raw Material Used	1488	<=	1600			
16								

**Рис. 30.8.** Допустимое решение задачи выбора ассортимента продукции соответствует ограничениям

Измененные значения ячеек, показанные на рис. 30.9 (и на листе **Infeasible Solutions**), являются недопустимым решением по следующим причинам:

- препарата 5 производится больше, чем требуется;
- используется больше рабочего времени, чем доступно;
- используется больше сырья, чем доступно.

	B	C	D	E	F	G	H	I
1								
2		Pounds made	300	0	0	0	1085	1000
3	Available	Product	1	2	3	4	5	6
4	4500	Labor	6	5	4	3	2.5	1.5
5	1600	Raw Material	3.2	2.6	1.5	0.8	0.7	0.3
6		Unit price	\$ 12.50	\$ 11.00	\$ 9.00	\$ 7.00	\$ 6.00	\$ 3.00
7		Variable cost	\$ 6.50	\$ 5.70	\$ 3.60	\$ 2.80	\$ 2.20	\$ 1.20
8		Demand	960	928	1041	977	1084	1055
9		Unit profit cont.	\$ 6.00	\$ 5.30	\$ 5.40	\$ 4.20	\$ 3.80	\$ 1.80
10								
11								
12		Profit	\$ 7,723.00					
13					Available			
14		Labor Used	6012.5	<=	4500			
15		Raw Material Used	2019.5	<=	1600			

**Рис. 30.9.** Недопустимое решение задачи из-за несоответствия заданным ограничениям

Когда вы нажмете **Найти решение (Solve)**, для модели будет найдено оптимальное решение, показанное на рис. 30.10. Выберите **Сохранить найденное решение (Keep Solver Solution)** для сохранения значений оптимального решения на листе.

Фармацевтическая компания может максимизировать свою ежемесячную прибыль до \$6625,20, произведя 596,67 фунта препарата 4 и 1084 фунта препарата 5. Невозможно определить, существует ли другой способ максимизировать прибыль

	B	C	D	E	F	G	H	I
1								
2		Pounds made	0	0	0	596.6667	1084	0
3	Available	Product	1	2	3	4	5	6
4	4500	Labor	6	5	4	3	2.5	1.5
5	1600	Raw Material	3.2	2.6	1.5	0.8	0.7	0.3
6		Unit price	\$ 12.50	\$ 11.00	\$ 9.00	\$ 7.00	\$ 6.00	\$ 3.00
7		Variable cost	\$ 6.50	\$ 5.70	\$ 3.60	\$ 2.80	\$ 2.20	\$ 1.20
8		Demand	960	928	1041	977	1084	1055
9		Unit profit cont.	\$ 6.00	\$ 5.30	\$ 5.40	\$ 4.20	\$ 3.80	\$ 1.80
10								
11								
12		Profit	\$ 6,625.20					
13					Available			
14		Labor Used	4500	<=	4500			
15		Raw Material Used	1236.1333	<=	1600			

**Рис. 30.10.** Оптимальное решение задачи выбора ассортимента продукции

в размере \$6625,20. Единственное, что можно утверждать — в этом месяце невозможно получить прибыль свыше \$6625,20, при ограничениях на ресурсы и данном спросе на препараты.

### ❓ Всегда ли модель поиска решения имеет решение?

Допустим, удовлетворение спроса на каждый продукт является обязательным условием. (См. в файле Prodmix.xlsx лист No Feasible Solution.) Тогда вам нужно изменить ограничения на спрос с D2:I2<=D8:I8 на D2:I2>=D8:I8. Откройте инструмент Поиск решения (Solver), выберите ограничение D2:I2<=D8:I8 и нажмите Изменить (Change). Появится диалоговое окно Изменение ограничения (Change Constraint), показанное на рис. 30.11.

Change Constraint

Cell Reference:

Constraint:

**Рис. 30.11.** Диалоговое окно Изменение ограничения

Выберите >= и щелкните на ОК. Теперь при поиске решения гарантированно будут учитываются только те изменения ячеек, которые удовлетворяют спрос на все препараты. Если вы щелкнете на Найти решение (Solve), появится сообщение «В ходе поиска не удалось найти допустимого решения» (Solver could not find a feasible solution). Это не значит, что в модели есть ошибка, просто ограничения на ресурсы

не позволяют удовлетворить спрос на все продукты. Таким образом, для удовлетворения спроса на все продукты необходимо добавить рабочее время или сырье или то и другое.

**? Что означает выдаваемый моделью поиска решения ответ «Значения ячейки целевой функции не сходятся»?**

Давайте посмотрим, что случится, если вы разрешите неограниченный спрос на каждый продукт и производство отрицательных количеств каждого препарата (см. в файле Prodmix.xlsx лист Set Values Do Not Converge). Для поиска оптимального решения в такой ситуации откройте инструмент Поиск решения (Solver) и в диалоговом окне Параметры поиска решения (Solver Parameters) снимите флажок Сделать переменные без ограничений неотрицательными (Make Unconstrained Variables Non-Negative). Далее выберите ограничение на спрос  $D2:I2 \leq D8:I8$  и нажмите Удалить (Delete). Затем щелкните Найти решение (Solve). Появится сообщение «Значения ячейки целевой функции не сходятся» (Set Cell Values Do Not Converge). Оно означает, что если значение целевой ячейки должно быть максимальным (как в данном примере), то существуют допустимые решения со сколь угодно большими значениями в целевой ячейке. Если значение в целевой ячейке должно быть минимальным, сообщение «Значения ячейки целевой функции не сходятся» означает, что существуют допустимые решения со сколь угодно малыми значениями в целевой ячейке. В этой ситуации при допущении негативных количеств препаратов создаются ресурсы, которые могут быть использованы для производства сколь угодно больших объемов других препаратов. С учетом неограниченного спроса на препараты можно получить неограниченную прибыль. В реальной ситуации бесконечно большую денежную сумму получить невозможно. Таким образом, появление сообщения «Значения ячейки целевой функции не сходятся» означает наличие в модели ошибки.

## Задания

1. Пусть фармацевтическая компания может приобрести до 500 часов рабочего времени на \$1 за час дороже, чем нынешние затраты на рабочую силу. Каким образом это максимизирует прибыль?
2. На заводе-изготовителе микросхем четыре техника (А, В, С и D) производят три продукта (продукт 1, 2 и 3). В этом месяце производитель чипов может продать 80 единиц продукта 1, 50 единиц продукта 2 и не больше 50 единиц продукта 3. Техник А может производить только продукт 1 и 3. Техник В производит только продукт 1 и 2. Техник С может производить только продукт 3. Техник D производит только продукт 2. С каждой произведенной единицы продукции завод получает следующую прибыль: продукт 1 — \$6, продукт 2 — \$7 и продукт 3 — \$10. Время (в часах), необходимое каждому технику для производства продукта, указано в таблице ниже.

Продукт	Техник А	Техник В	Техник С	Техник D
1	2	2,5	Не производит	Не производит
2	Не производит	3	Не производит	3,5
3	3	Не производит	4	Не производит

Каждый техник может работать 120 часов в месяц. Каким образом производитель чипов может максимизировать ежемесячную прибыль? Допустим, возможно производить дробное количество единиц продукции.

- Завод производит мыши, клавиатуры и джойстики для видеоигр. В нижеследующей таблице приведена прибыль с единицы продукции, количество рабочего времени, затрачиваемое на единицу продукции, ежемесячный спрос на продукты и количество машинного времени, затрачиваемое на единицу продукции.

	Мышь	Клавиатура	Джойстик
Прибыль/единица	\$8	\$11	\$9
Рабочее время/единица	0,2 часа	0,3 часа	0,24 часа
Машинное время/единица	0,04 часа	0,055 часа	0,04 часа
Ежемесячный спрос	15 000	29 000	11 000

Каждый месяц доступно в общей сложности 13 000 часов рабочего времени и 3000 часов машинного времени. Каким образом завод может максимизировать ежемесячную прибыль?

- Решите наш пример с препаратами при условии, что минимальный спрос на каждый препарат составляет 200 единиц.
- Джейсон изготавливает бриллиантовые браслеты, ожерелья и серьги. Он собирается работать максимум 160 часов в месяц. У него имеется 800 унций алмазов. В нижеследующей таблице указана прибыль, рабочее время в часах и количество алмазов в унциях для каждого продукта. Как Джейсон может максимизировать свою прибыль при неограниченном спросе на каждый продукт?

Продукт	Прибыль/единица	Рабочее время/единица	Унции/единица
Браслет	\$300	0,35	1,2
Ожерелье	\$200	0,15	0,75
Серьги	\$100	0,05	0,5



6. В примере с ассортиментом продукции предположим, что каждый раз при продаже более 400 фунтов любого продукта необходимо давать скидку \$1 за каждый фунт, проданный сверх этого количества. Изменит ли это решение задачи?
7. При условии неограниченного спроса и неограниченных ресурсов найдите решения для нашей задачи с ассортиментом продуктов, которое максимизирует прибыль до одного миллиарда долларов.
8. Shoeco производит три типа обуви. Спрос на обувь не ограничен, и доступны 40 часов в неделю машинного времени и рабочего времени. В нижеследующей таблице указана прибыль от продажи одной пары обуви и количество затраченного машинного и рабочего времени. При условии, что количество пар обуви, производимое каждый месяц, это целое число, как Shoeco может максимизировать еженедельную прибыль?

	<b>Обувь 1</b>	<b>Обувь 2</b>	<b>Обувь 3</b>
Прибыль	\$40	\$25	\$30
Рабочее время, минуты	40	20	30
Машинное время, минуты	45	25	26

9. Woodco производит столы и стулья. В нижеследующей таблице указана прибыль от продажи одной единицы товара, количество использованной древесины (в квадратных футах) и количество часов, требуемых опытному плотнику для производства одной единицы товара.

	<b>Стол</b>	<b>Стул</b>
Прибыль	\$250	\$150
Количество исполь- зованной древе- сины	22	18
Количество часов работы плотника	14	8

Производится целое число столов и стульев, и спрос неограничен. Как мы можем максимизировать еженедельную прибыль?

## ГЛАВА 31

# Поиск решения при планировании расписания работы сотрудников

### Обсуждаемый вопрос

- Каким спланировать рабочее расписание сотрудников для удовлетворения потребности в рабочей силе?

Многие организации (например банки, рестораны, почтовые службы) имеют различную потребность в рабочей силе в разные дни недели, поэтому им необходим эффективный метод для составления расписания своих работников. Для решения такой задачи вы можете применить инструмент Поиск решения.

## Ответ на вопрос

- ❓ Каким спланировать рабочее расписание сотрудников для удовлетворения потребности в рабочей силе?

В Bank 24 чеки обрабатываются семь дней в неделю. В файле Bank24.xlsx в строке 14 указано количество работников, необходимое для ежедневной обработки чеков (рис. 31.1). Например, для выполнения работы во вторник необходимы 13 работников, в среду — 15 работников и т. д. Все служащие банка работают пять дней подряд. Какое минимальное количество сотрудников покрывает потребность банка в рабочей силе?

	A	B	C	D	E	F	G	H	I	J
2	Total				There are multiple solutions all of which use 20 workers.					
3	20		Working?							
4	Number starting	Day worker starts	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	
5	1	Monday	1	1	1	1	1	0	0	
6	3	Tuesday	0	1	1	1	1	1	0	
7	0	Wednesday	0	0	1	1	1	1	1	
8	4	Thursday	1	0	0	1	1	1	1	
9	1	Friday	1	1	0	0	1	1	1	
10	2	Saturday	1	1	1	0	0	1	1	
11	9	Sunday	1	1	1	1	0	0	1	
12		Number working	17	16	15	17	9	10	16	
13			>=	>=	>=	>=	>=	>=	>=	
14		Number needed	17	13	15	17	9	9	12	

Рис. 31.1. Данные для задачи планирования работы служащих банка

Начните с определения целевой ячейки, изменяемых ячеек и ограничений для модели решения.

- **Целевая ячейка** — минимизация общего количества служащих банка.
- **Изменяемые ячейки** — количество служащих банка, которые начинают работу (в первый из пяти подряд идущих дней) в конкретный день недели. Значение в каждой изменяемой ячейке должно быть неотрицательным целым числом.
- **Ограничения** — для каждого дня недели количество работающих служащих должно быть больше или равно требуемому количеству служащих.

Для создания модели решения этой задачи вам нужно отслеживать количество служащих, работающих каждый день. Я начал с ввода пробных значений для количества служащих, начинающих свою пятидневку в каждый день недели, в диапазон A5:A11. Например, в ячейку A5 я ввел значение 1, показывающее, что в понедельник к работе приступает один служащий, который проработает с понедельника по пятницу. Ежедневно необходимое количество работников я указал в диапазоне C14:I14.

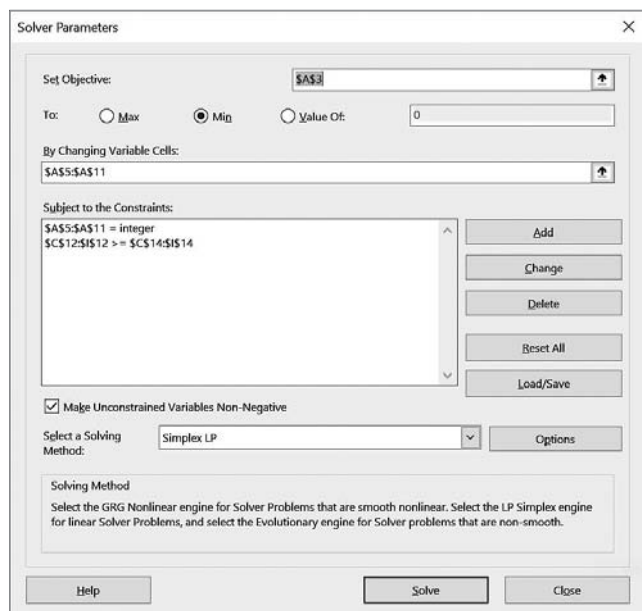
Для отслеживания количества служащих банка, работающих каждый день, в каждую ячейку диапазона C5:I11 я ввел 1 или 0. Значение 1 указывает, что служащие, приступившие к работе в день, связанный с этой строкой, работают в день, связанный с этим столбцом. Например, 1 в ячейке G5 показывает, что служащие, которые начали работать в понедельник, работают в пятницу; 0 в ячейке H5 показывает, что служащие, которые начали работать в понедельник, не работают в субботу.

Количество служащих, работающих каждый день, я вычислил по формуле =СУММПРОИЗВ(\$A\$5:\$A\$11;C5:C11), скопировав ее из C12 в D12:I12. Например, в ячейке C12 эта формула эквивалентна формуле =A5+A8+A9+A10+A11, вычисляющей (число приступивших к работе в понедельник) + (число приступивших к работе в четверг) + (число приступивших к работе в пятницу) + (число приступивших к работе в субботу) + (число приступивших к работе в воскресенье). Это общее количество сотрудников, работающих в понедельник.

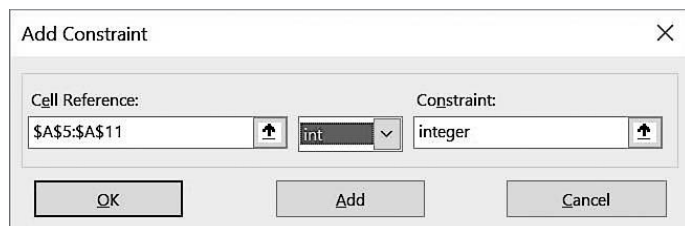
После ввода формулы =СУММ(A5:A11) в ячейку A3 для вычисления общего количества служащих я могу ввести данные для модели в окне **Параметры поиска решения (Solver Parameters)**, как показано на рис. 31.2.

В целевой ячейке (A3) я хочу минимизировать общее количество служащих. Ограничение C12:I12>=C14:I14 гарантирует, что количество работающих каждый день служащих не меньше количества необходимых каждый день служащих. Ограничение A5:A11=целое гарантирует, что количество приступающих каждый день к работе служащих является целым числом. Для добавления этого ограничения я выбрал в диалоговом окне **Параметры поиска решения (Solver Parameters)** **Добавить (Add)** и заполнил диалоговое окно **Добавление ограничения (Add Constraint)**, как показано на рис. 31.3.

Обратите внимание, что эта модель является линейной, поскольку значение в целевой ячейке создается путем суммирования значений в изменяемых ячейках,



**Рис. 31.2.** Диалоговое окно Параметры поиска решения для решения задачи составления рабочего расписания



**Рис. 31.3.** Это ограничение определяет количество приступающих каждый день к работе служащих как целое число

а ограничение создается путем сравнения результата суммирования значений в изменяемых ячейках, умноженных на константу (или 1 или 0), с требуемым количеством служащих. Поэтому я выбрал симплекс-метод для решения задачи. Кроме того, поскольку количество приступающих к работе служащих в каждый из дней не может быть отрицательным, я включил флажок **Сделать переменные без ограничений неотрицательными** (Make Unconstrained Variables Non-Negative). После нажатия **Найти решение** (Solve) появится оптимальное решение, показанное ранее на рис. 31.1.

В общей сложности для такой работы потребуется 20 служащих. Один служащий приступает к работе в понедельник, три — во вторник, четыре — в четверг, один — в пятницу, два — в субботу и девять — в воскресенье. Обратите внимание, что на

самом деле эта модель имеет несколько оптимальных решений для 20 служащих. При повторном запуске поиска решения, возможно, будет найдено одно из альтернативных решений.

## Задания

1. В Bank 24 работает 22 служащих и нужно составить расписание с максимальным количеством выходных дней для служащих в субботу и воскресенье. Как это сделать?
2. Служащие Bank 24 получают \$150 в день за работу в первые пять дней и \$350 за переработку. Каким образом банк должен составить расписание?
3. В нижеследующей таблице указано количество телефонных операторов, необходимых авиакомпании в течение рабочего дня. Каждый оператор работает в одну из шестичасовых смен: 00:00–06:00, 06:00–12:00, 12:00–18:00, 18:00–24:00. Каково минимальное требуемое количество операторов?

Время	Требуемое количество операторов
00:00—04:00	12
04:00—08:00	16
08:00—12:00	22
12:00—16:00	30
16:00—20:00	31
20:00—24:00	22

4. На рис. 31.4 показано количество людей в различных демографических группах, которые смотрят различные телепередачи, и стоимость (в тысячах долларов) размещения 30-секундной рекламы в каждой передаче. Например, размещение 30-секундной рекламы в сериале «Друзья» стоит \$160 000. Этот сериал смотрят 6 млн мужчин в возрасте 18–35 лет, 3 млн мужчин в возрасте

	C	D	E	F	G	H	I	J
4	000s	needed	60	60	28	60	60	28
5	Cost	Show	M 18-35	M 36-55	M >55	W 18-35	W 36-55	W >55
6	180	Modern Family	6	3	1	9	4	2
7	100	MNF	6	5	3	1	1	1
8	80	Simpsons	5	2	0	4	2	0
9	9	Sports Center	0.5	0.5	0.3	0.1	0.1	0
10	13	MTV	0.7	0.2	0	0.9	0.1	0
11	16	GLOW	0.1	0.1	0	0.6	1.3	0.4
12	8	CNN	0.1	0.2	0.3	0.1	0.2	0.3
13	85	Suits	1	2	4	1	3	4

Рис. 31.4. Данные к заданию 4

36–55 лет, 1 млн мужчин старше 55 лет, 9 млн женщин в возрасте 18–35 лет, 4 млн женщин в возрасте 36–55 лет и 2 млн женщин старше 55 лет. Данные также включают количество людей в каждой группе (в млн), которым рекламодатель хотел бы показать рекламу. Например, рекламодателю хотелось бы показать свою рекламу по крайней мере 60 млн мужчин в возрасте 18–35 лет. Каков наиболее дешевый способ достижения этих целей?

5. Кредитный союз Pine Valley Credit Union пытается составить расписание для кассиров банка. Кредитный союз работает с 8:00 до 18:00, а количество кассиров, необходимое каждый час для работы, указано в нижеследующей таблице. Полный рабочий день кассира длится с 8:00 до 17:00 (с обеденным перерывом с 12:00 до 13:00) или с 9:00 до 18:00 (с обеденным перерывом с 13:00 до 14:00).

Сокращенный рабочий день кассира длится с 10:00 до 14:00. Кассиры, работающие полный рабочий день, получают \$300 в день, а кассиры, работающие сокращенный день, — \$60. На сокращенный рабочий день можно нанять не более четырех кассиров. Каким образом кредитный союз может минимизировать свои ежедневные затраты на оплату труда кассиров?

Требуемое количество кассиров	Время
4	8–9
8	9–10
6	10–11
4	11–12
9	12–13
8	13–14
5	14–15
4	15–16
4	16–17
5	17–18

6. Кредитный союз Houston Credit Union работает по рабочим дням с 9:00 до 17:00. Кассиры, работающие полный день (без обеденного перерыва), получают \$200 в день. Кассиры, работающие сокращенный рабочий день с 11:00 до 13:00, получают \$80 в день. Количество кассиров, необходимое каждый час для работы, указано на рис. 31.5. Создайте линейную модель решения для определения, как кредитному союзу минимизировать ежедневные затраты на оплату труда.

	C	D
11	Time	Needed
12	9-10 AM	12
13	10-11 Am	7
14	11 AM- Noon	13
15	Noon-1 PM	15
16	1 PM-2 PM	11
17	2 PM- 3 PM	10
18	3 PM-4 PM	6
19	4 PM-5PM	7

**Рис. 31.5.** Данные для задания 6

## ГЛАВА 32

# Поиск решения для задач транспортировки и распределения

### Обсуждаемый вопрос

- Как фармацевтическая компания может определить, где следует производить препараты и откуда их лучше отгружать заказчикам?

Многие компании производят свои продукты в различных местах (часто называемых *точками поставки*) и отправляют свою продукцию заказчикам (в так называемые *точки спроса*). Возникает естественный вопрос: «Как дешево произвести и доставить продукцию заказчикам, удовлетворив при этом потребность в спросе?» Такой тип задач называется транспортной задачей. Транспортная задача может быть представлена в виде линейной модели для инструмента Поиск решения (Solver) следующим образом:

- **Целевая ячейка** — минимизация общей стоимости производства продукции и транспортных расходов.
- **Изменяемые ячейки** — количество произведенной продукции в каждой точке поставки, отправляемое в каждую точку спроса.
- **Ограничения** — количество, отправляемое из каждой точки поставки, не может превышать производственные мощности соответствующего завода. Спрос должен быть удовлетворен в каждой точке спроса. Кроме того, значение в каждой изменяемой ячейке не может быть отрицательным.

## Ответ на вопрос

- ❓ **Как фармацевтическая компания может определить, где следует производить препараты и откуда их лучше отгружать заказчикам?**

Решение этой задачи приведено в файле **Transport.xlsx**. Допустим, компания производит некий препарат на своих заводах в Лос-Анджелесе, Атланте и Нью-Йорке. Каждый месяц завод в Лос-Анджелесе может произвести 10 000 фунтов препарата. В Атланте может быть произведено 12 000 фунтов и в Нью-Йорке — 14 000 фунтов. Каждый месяц компания должна отправлять препарат (количества в фунтах указаны в ячейках **B2:E2**) в четыре региона США: на Восток, Средний Запад, Юг и Запад (рис. 32.1). Например, на Восток каждый месяц должно быть поставлено

не менее 13 000 фунтов препарата. Стоимость производства фунта препарата на каждом заводе и доставки его в каждый регион страны указана в ячейках B4:E6. Например, производство одного фунта препарата в Лос-Анджелесе и отправка его на Средний Запад стоит \$3,50. Каков наиболее дешевый способ поставки требуемого количества препарата в каждый регион?

	A	B	C	D	E	F	G	H
1								
2	DEMAND	9000	6000	6000	13000			
3		EAST	MIDWEST	SOUTH	WEST	CAPACITY		
4	LA	\$ 5.00	\$ 3.50	\$ 4.20	\$ 2.20	10000		
5	ATLANTA	\$ 3.20	\$ 2.60	\$ 1.80	\$ 4.80	12000		
6	NEW YORK CITY	\$ 2.50	\$ 3.10	\$ 3.30	\$ 5.40	14000		
7								
8	Shipments							
9		EAST	MIDWEST	SOUTH	WEST	Sent		Capacity
10	LA	0	0	0	10000	10000	<=	10000
11	ATLANTA	0	3000	6000	3000	12000	<=	12000
12	NEW YORK CITY	9000	3000	0	0	12000	<=	14000
13	Received	9000	6000	6000	13000			
14		>=	>=	>=	>=			
15	Demand	9000	6000	6000	13000			
16								
17								
18	Total Cost	\$ 86,800.00						

Рис. 32.1. Данные для транспортной задачи

Для выражения целевой ячейки необходимо отследить общие транспортные расходы. После ввода в диапазон B10:E12 пробных значений для отгрузки из каждой точки поставки в каждый регион общие транспортные расходы могут быть вычислены как

(количество\_отправленное\_из\_Лос-Анджелеса\_на\_Восток)\*(стоимость\_фунта\_препарата\_отправленного\_из\_Лос-Анджелеса\_на\_Восток)+(количество\_отправленное\_из\_Лос-Анджелеса\_на\_Средний\_Запад)\*(стоимость\_фунта\_препарата\_отправленного\_из\_Лос-Анджелеса\_на\_Средний\_Запад)+(количество\_отправленное\_из\_Лос-Анджелеса\_на\_Юг)\*(стоимость\_фунта\_препарата\_отправленного\_из\_Лос-Анджелеса\_на\_Юг)+(количество\_отправленное\_из\_Лос-Анджелеса\_на\_Запад)\*(стоимость\_фунта\_препарата\_отправленного\_из\_Лос-Анджелеса\_на\_Запад)+...+(количество\_отправленное\_из\_Нью-Йорка\_на\_Запад)\*(стоимость\_фунта\_препарата\_отправленного\_из\_Нью-Йорка\_на\_Запад).

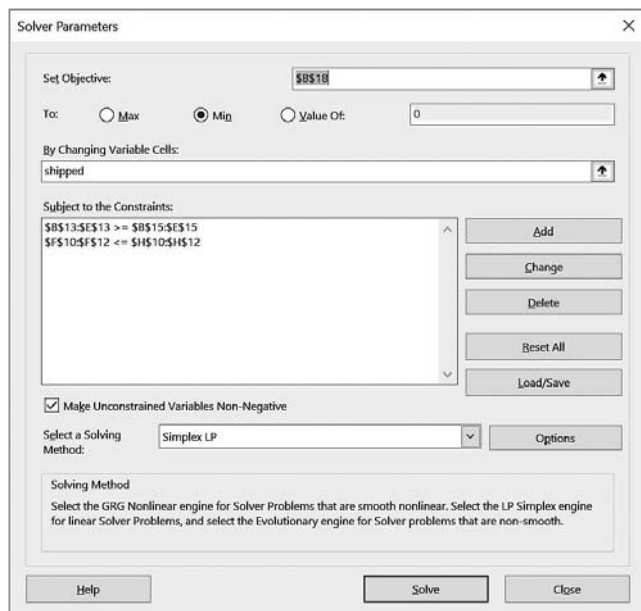
Функция СУММПРОИЗВ (SUMPRODUCT) перемножит соответствующие элементы двух отдельных прямоугольных диапазонов (при условии, что прямоугольники имеют одинаковый размер) и суммирует произведения. Я присвоил диапазону B4:E6 имя Costs, а диапазону изменяемых ячеек ( B10:E12) — имя Shipped. Таким образом, общие транспортные и производственные расходы можно вычислить в ячейке B18 по формуле =СУММПРОИЗВ(Costs;Shipped).



Для выражения ограничений модели я сперва вычислил общее количество препарата, отправленное из каждой точки поставки. По формуле `=СУММ(B10:E10)` в ячейке F10 я вычислил общее количество в фунтах, отправленное из Лос-Анджелеса (на Восток) + (на Средний Запад) + (на Юг) + (на Запад). Будучи скопирована в F11:F12, эта формула вычисляет общее количество препарата, отправленное из Атланты и Нью-Йорка. Позже я добавлю ограничение (называемое *ограничением по поставке*), гарантирующее, что количество препарата, отправленное из каждого места, не превышает производственные мощности завода.

Далее я вычислил общее количество полученного препарата в каждой точке спроса. Сначала я ввел формулу `=СУММ(B10:B12)` в ячейку B13. Эта формула вычисляет общее количество препарата в фунтах, полученное на Востоке (из Лос-Анджелеса) + (из Атланты) + (из Нью-Йорка). Скопировав ее из B13 в C13:E13, я получил фунты препарата, полученные на Среднем Западе, Юге и Востоке. Позже я добавлю ограничение (называемое *ограничением по спросу*), гарантирующее получение в каждом регионе требуемого количества препарата.

Теперь я открываю диалоговое окно **Параметры поиска решения (Solver Parameters)** (вкладка **Данные (Data)**, группа **Анализ (Analysis)**, инструмент **Поиск решения (Solver)**) и заполняю его, как показано на рис. 32.2.



**Рис. 32.2.** Параметры поиска решения транспортной задачи

Мне нужно минимизировать общие транспортные расходы (вычисленные в ячейке B18). Изменяемые ячейки содержат количество препарата в фунтах, отправ-

ленное с каждого завода в каждый регион. (Эти количества указаны в диапазоне с именем **Shipped**, состоящем из ячеек **B10:E12**.) Ограничение **F10:F12<=H10:H12** (ограничение по поставке) гарантирует, что количество препарата, отправленного с каждого завода, не превысит производственную мощность этого завода. Ограничение **B13:E13>=B15:E15** (ограничение по спросу) гарантирует получение каждым регионом требуемого количества препарата.

Эта модель является линейной моделью поиска решения, так как значение в целевой ячейке определяется путем суммирования членов в виде **(изменяемая\_ячейка)\*(константа)**, и оба ограничения, как по поставке, так и по спросу, созданы путем сравнения суммы значений в изменяемых ячейках с константой. Так как модель линейная, я выбрал симплекс-метод для решения. Очевидно, что поставки должны быть неотрицательными, поэтому я включил флажок **Сделать переменные без ограничений неотрицательными** (**Make Unconstrained Variables Non-Negative**).

Теперь в окне **Параметры поиска решения** (**Solver Parameters**) я нажал **Найти решение** (**Solve**) и получил оптимальное решение, представленное ранее на рис. 32.1. Минимальные расходы на удовлетворение спроса заказчиков составляют \$86 800. Такой минимум может быть достигнут, если компания использует следующий план производства и транспортировки:

- отправка 10 000 фунтов препарата из Лос-Анджелеса на Запад;
- отправка 3000 фунтов препарата из Атланты на Запад и отправка такого же количества препарата из Атланты на Средний Запад. Отправка 6000 фунтов препарата из Атланты на Юг;
- отправка 9000 фунтов препарата из Нью-Йорка на Восток и отправка 3000 фунтов из Нью-Йорка на Средний Запад.

## Задания

1. В нижеследующей таблице приведены расстояния между городами Бостон, Чикаго, Даллас, Лос-Анджелес и Майами. Каждый город потребляет 40 000 кВт·ч (киловатт-часов) электроэнергии, а Чикаго, Даллас и Майами способны производить 70 000 кВт·ч. Предположим, что передача 1000 кВт·ч на 100 миль обходится в \$4. Откуда следует передавать энергию для минимизации расходов на удовлетворение спроса в каждом городе?

	Бостон	Чикаго	Даллас	Лос-Анджелес	Майами
Чикаго	983	0	1205	2112	1390
Даллас	1815	1205	0	801	1332
Майами	1539	1390	1332	2757	0

2. Решите задачу, описанную в данной главе, при условии, что спрос на Западе возрос на 13 000 фунтов.

3. Ваша компания производит и продает препараты в нескольких местах. Решение о том, где производить товары для каждого места продажи, может оказать большое влияние на прибыльность. Модель решения этой задачи аналогична модели, рассмотренной в данной главе. Определите, где должны быть произведены препараты, при условии, что:

- препараты производятся в шести местах и продаются в шести регионах;
- ставка налога и переменные затраты на производство зависят от места производства. Например, стоимость любой единицы продукции, произведенной в местоположении 3, равна \$6; прибыль с этой продукции облагается налогом в 20%;
- розничная цена каждого препарата зависит от места продажи. Например, каждый продукт в местоположении 2 продается за \$40;

Место производства	1	2	3	4	5	6
Розничная цена	\$45	\$40	\$38	\$36	\$39	\$34
Ставка налога	31%	40%	20%	40%	35%	18%
Переменные затраты на производство	\$8	\$7	\$6	\$9	\$7	\$7

- каждый из шести заводов может производить до 6 млн единиц продукции в год;
- ежегодный спрос на продукцию (в млн единиц) в каждом местоположении указан в нижеследующей таблице:

Место продажи	1	2	3	4	5	6
Спрос	1	2	3	4	5	6

- расходы на транспортировку единицы продукции зависят от завода, на котором производится продукт, и места продажи продукта, как видно из следующей таблицы:

	Место продажи 1	Место продажи 2	Место продажи 3	Место продажи 4	Место продажи 5	Место продажи 6
Завод 1	\$3	\$4	\$5	\$6	\$7	\$8
Завод 2	\$5	\$2	\$6	\$9	\$10	\$11
Завод 3	\$4	\$3	\$1	\$6	\$8	\$6
Завод 4	\$5	\$5	\$7	\$2	\$5	\$5
Завод 5	6	\$9	\$6	\$5	\$3	\$7
Завод 6	\$7	\$7	\$8	\$9	\$10	\$4

Например, расходы на транспортировку единицы продукции, произведенной на заводе 1 и проданной в местоположении 3, составляют \$5. Как можно максимизировать прибыль после уплаты налогов при ограниченной производственной мощности?

4. Каждый день в северной, центральной и южной части Калифорнии используется по 100 млрд галлонов воды. Пусть в северной и центральной частях Калифорнии доступно по 120 млрд галлонов воды, в то время как в южной части Калифорнии доступно 40 млрд галлонов воды. Стоимость транспортировки 1 млрд галлонов воды между этими тремя регионами указана в нижеследующей таблице.

	Северная	Центральная	Южная
Северная	\$5000	\$7000	\$10 000
Центральная	\$7000	\$5000	\$6000
Южная	\$10 000	\$6000	\$5000

Спрос на воду не может быть удовлетворен полностью, поэтому предположим, что каждый миллиард галлонов воды неудовлетворенного спроса влечет потери от дефицита, указанные в нижеследующей таблице.

	Северная	Центральная	Южная
Потери от дефицита/млрд галлонов	\$6000	\$5500	\$9000

Как следует распределить воду в Калифорнии, чтобы минимизировать сумму транспортных расходов и потерь от дефицита?

5. Магазин техники имеет на складах в Хьюстоне, Далласе и Сан-Антонио 50 газонокосилок. В текущем месяце будет продано 35 газонокосилок в Хьюстоне, Далласе, Сан-Антонио и Остине. Стоимость транспортировки газонокосилок с каждого склада в каждый город указана ниже. Определите, как удовлетворить спрос при минимальных затратах (рис. 32.3).

	D	E	F	G	H
16		Houston	Dallas	San Antonio	Austin
17	Houston	\$1.00	\$7.00	\$5.00	\$6.00
18	Dallas	\$7.00	\$1.50	\$10.00	\$9.00
19	San Antonio	\$8.00	\$10.00	\$2.00	\$4.00

**Рис. 32.3.** Данные к заданию 5

6. Как использовать значок Автосумма для решения заданий транспортировки?

## ГЛАВА 33

# Поиск решения для бюджетирования капиталовложений

Обсуждаемый вопрос

- Как с помощью инструмента Поиск решения компания может определить, какой из проектов ей следует принять к реализации?

Каждый год такие компании, как Eli Lilly , определяют, какие препараты разрабатывать; такие как Microsoft — какое программное обеспечение разрабатывать; такие как Procter&Gamble — какие потребительские товары разрабатывать. Принять решение поможет инструмент Поиск решения от Microsoft Excel 2019.

## Ответ на вопрос

❓ **Как с помощью инструмента Поиск решения компания может определить, какой из проектов ей следует принять к реализации?**

Большинство корпораций предпочитает реализовывать проекты, которые при условии ограниченных ресурсов (обычно это капитал и рабочая сила) имеют наибольшую чистую приведенную стоимость (ЧПС). Скажем, компания — разработчик программного обеспечения пытается определить, какие из 20 проектов ей следует осуществить. ЧПС (в млн долларов) каждого проекта, а также затраты (в млн долларов) и количество программистов, необходимое в течение следующих трех лет, представлены в файле *Capbudget.xlsx* на листе **Basic Model** (рис. 33.1). Например, проект 2 принесет в результате \$908 млн. Для его реализации требуется вложить \$151 млн в год 1, \$269 млн в год 2 и \$248 млн в год 3. Для работы над проектом 2 необходимо задействовать 139 программистов в год 1, 86 программистов в год 2 и 83 программиста в год 3. В ячейках **E4:G4** указаны денежные средства (в млн долл.), доступные в течение каждого года из следующих трех лет, а в ячейках **H4:J4** показано количество доступных программистов. Например, в год 1 доступно не более \$2,5 млрд и 900 программистов.

Для каждого проекта компания должна решить, имеет ли смысл его реализовать. Предположим, что компания не может реализовать часть программного проекта.

Например, при выделении половины необходимых ресурсов компания получает неработающую программу, которая принесет \$0 дохода.

Прием, применяемый в ситуациях моделирования, в которых необходимо либо что-то делать, либо что-то не делать, состоит в использовании *бинарных изменяемых ячеек*. Значение в бинарной изменяемой ячейке всегда только 0 или 1. Если значение в соответствующей проекту бинарной изменяемой ячейке равно 1, проект принимается. Если значение в такой ячейке равно 0, проект отклоняется. Используемый диапазон бинарных изменяемых ячеек при поиске решения вы можете задать, добавив ограничение — указав в диалоговом окне **Добавление ограничения (Add Constraint)** ссылку на требуемые изменяемые ячейки и выбрав из списка **бин (Bin)**.

	A	B	C	D	E	F	G	H	I	J
1		Total NPV								
2		9293		Used	2460	2684	2742	876	895	702
3					<=	<=	<=	<=	<=	<=
4				Available	2500	2800	2900	900	900	900
5	Do IT?		NPV		Cost Year 1	Cost Year 2	Cost Year 3	Labor Year 1	Labor Year 2	Labor Year 3
6	0 Project 1	928			398	180	368	111	108	123
7	1 Project 2	908			151	269	248	139	86	83
8	1 Project 3	801			129	189	308	56	61	23
9	0 Project 4	543			275	218	220	54	70	59
10	0 Project 5	944			291	252	228	123	141	70
11	1 Project 6	848			80	283	285	119	84	37
12	1 Project 7	545			203	220	77	54	44	42
13	1 Project 8	808			150	113	143	67	101	43
14	1 Project 9	638			282	141	160	37	55	64
15	1 Project 10	841			214	254	355	130	72	62
16	0 Project 11	664			224	271	130	51	79	58
17	0 Project 12	546			225	150	33	35	107	63
18	0 Project 13	699			101	218	272	43	90	71
19	1 Project 14	599			255	202	70	3	75	83
20	1 Project 15	903			228	351	240	60	93	80
21	1 Project 16	859			303	173	431	60	90	41
22	0 Project 17	748			133	427	220	59	40	39
23	0 Project 18	668			197	98	214	95	96	74
24	1 Project 19	888			313	278	291	66	75	74
25	1 Project 20	655			152	211	134	85	59	70

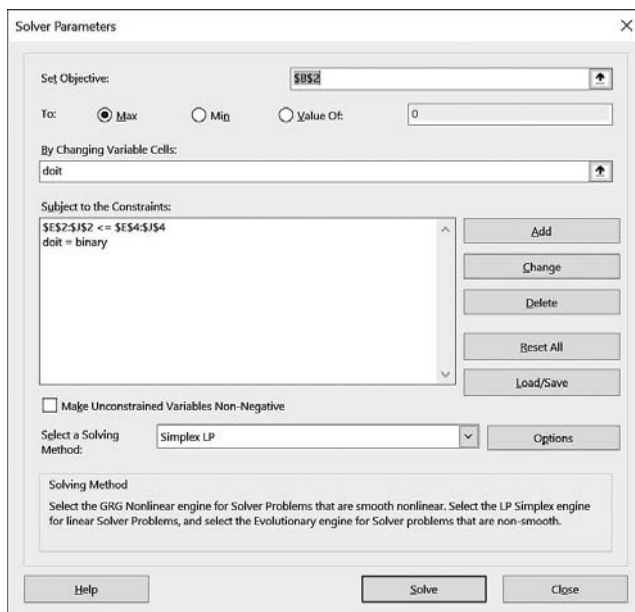
**Рис. 33.1.** Данные для определения предпочтительных проектов с помощью инструмента Поиск решения

Теперь мы готовы перейти к решению задачи отбора проектов. Как и в любой модели поиска решения, нужно начать с определения целевой ячейки, изменяемых ячеек и ограничений.

- **Целевая ячейка** — максимальная ЧПС выбранных проектов.
- **Изменяемые ячейки** — подбор значения 0 или 1 в изменяемой ячейке, соответствующей каждому проекту. Я выбрал диапазон A6:A25 (с именем *doit*). Например, 1 в ячейке A6 покажет, что проект 1 одобрен, а 0 в ячейке A6 покажет, что проект 1 отклонен.
- **Ограничения** — для каждого года  $t$  ( $t = 1, 2, 3$ ) используемые денежные средства должны быть меньше или равны доступным средствам, а количество задействованных программистов не должно превышать количество доступных программистов.

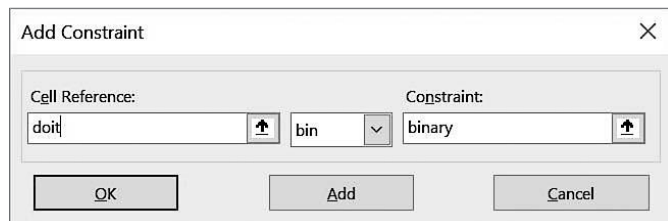
Как видите, на листе должна вычисляться ЧПС, ежегодные затраты и необходимое количество программистов для любых проектов. В ячейке **B2** я использовал формулу `=СУММПРОИЗВ(doit;NPV)` для вычисления ЧПС выбранных проектов. (Имя `NPV` относится к диапазону `C6:C25`.) Для каждого проекта со значением **1** в столбце **A** эта формула выбирает ЧПС проекта, а для каждого проекта со значением **0** в столбце **A** эта формула ЧПС не учитывает. Таким образом, я могу вычислить ЧПС всех проектов. Задача является линейной, поскольку значение в целевой ячейке рассчитывается путем суммирования членов в виде `(изменяемая_ячейка)*(константа)`. Ежегодные затраты и необходимое количество программистов я вычислил аналогичным образом, скопировав формулу `=СУММПРОИЗВ(doit;E6:E25)` из **E2** в **F2:J2**.

Далее я заполнил диалоговое окно **Параметры поиска решения (Solver Parameters)**, как показано на рис. 33.2.



**Рис. 33.2.** Диалоговое окно **Параметры поиска решения** для модели выбора проектов

Цель состоит в получении максимальной ЧПС выбранных проектов (ячейка **B2**). Изменяемые ячейки (диапазон с именем `doit`) являются бинарными ячейками для каждого проекта. Ограничение `E2:J2<=E4:J4` гарантирует, что каждый год затраты и рабочая сила не превышают доступных значений. Для добавления ограничения, по которому изменяемые ячейки становятся бинарными, в диалоговом окне **Параметры поиска решения (Solver Parameters)** я щелкнул на **Добавить (Add)** и затем в диалоговом окне **Добавление ограничения (Add Constraint)** указал ссылку на эти ячейки и выбрал из списка в центре значение **бин (Bin)**. Диалоговое окно **Добавление ограничения (Add Constraint)** показано на рис. 33.3.



**Рис. 33.3.** Выбор Bin в диалоговом окне Добавление ограничения для бинарных изменяемых ячеек (со значениями 0 или 1)

Эта модель является линейной, поскольку значение в целевой ячейке рассчитывается как сумма членов в виде (изменяемая\_ячейка)\*(константа), а ограничения на использование ресурсов вычисляются путем сравнения суммы членов в виде (изменяемая\_ячейка)\*(константа) с константой. Следовательно, я выбрал для решения симплекс-метод.

Теперь, когда диалоговое окно Параметры поиска решения (Solver Parameters) заполнено, я щелкаю на **Найти решение (Solve)** и получаю результаты, показанные ранее на рис. 33.1. Компания может получить максимальную ЧПС в размере \$9293 млн (\$9,293 млрд), выбрав для реализации проекты 2, 3, 6–10, 14–16, 19 и 20.

## Обработка других ограничений

Иногда модели отбора проектов включают и другие ограничения. Например, предположим, что если вы выбрали проект 3, то должны выбрать и проект 4. Поскольку текущее оптимальное решение не включает проект 4, для новой задачи оно не будет оптимальным. Для решения новой задачи просто добавьте следующее ограничение: значение в бинарной изменяемой ячейке для проекта 3 меньше или равно значению в бинарной изменяемой ячейке для проекта 4.

Этот пример (рис. 33.4) вы найдете в файле **Capbudget.xlsx** на листе **If 3 then 4**. Ячейка L9 ссылается на бинарное значение, связанное с проектом 3, а ячейка L12 — на бинарное значение, связанное с проектом 4. После добавления ограничения  $L9 \leq L12$  при выборе проекта 3 значение в L9 равно 1, и это ограничение принудительно устанавливает значение 1 для L12 (бинарное значение проекта 4). Это ограничение не оказывает влияние на бинарное значение в изменяемой ячейке, соответствующей проекту 4, если проект 3 не выбран. Если проект 3 не выбран, значение в L9 равно 0, и бинарное значение для проекта 4 может быть равно 0 или 1, что и нам и нужно. На рис. 33.4 показано новое оптимальное решение.

Теперь предположим, что вы можете осуществить только четыре проекта из проектов 1–10. (См. лист **At Most 4 Of P1-P10**, показанный на рис. 33.5.) В ячейке L8 по формуле  $=\text{СУММ}(A6:A15)$  вычисляется сумма бинарных значений, связанных с проектами 1–10. Добавьте ограничение  $L8 \leq L10$ , гарантирующее выбор не более четырех проектов из первых десяти. Новое оптимальное решение представлено на рис. 33.5. ЧПС упала до \$9,014 млрд.



	A	B	C	D	E	F	G	H	I	J	K	L
1		Total NPV										
2		9157		Used	2444	2760	2837	866	895	659		
3					<=	<=	<=	<=	<=	<=		
4				Available	2500	2800	2900	900	900	900		
5	Do IT?		NPV		Cost Year 1	Cost Year 2	Cost Year 3	Labor Year 1	Labor Year 2	Labor Year 3		
6	0	Project 1	928		398	180	368	111	108	123		
7	1	Project 2	908		151	269	248	139	86	83		
8	1	Project 3	801		129	189	308	56	61	23	Proj 3	1
9	1	Project 4	543		275	218	220	54	70	59		
10	0	Project 5	944		291	252	228	123	141	70	<=	1
11	1	Project 6	848		80	283	285	119	84	37	Proj 4	1
12	1	Project 7	545		203	220	77	54	44	42		
13	1	Project 8	808		150	113	143	67	101	43		
14	1	Project 9	638		282	141	160	37	55	64		
15	0	Project 10	841		214	254	355	130	72	62		
16	0	Project 11	664		224	271	130	51	79	58		
17	0	Project 12	546		225	150	33	35	107	63		
18	0	Project 13	699		101	218	272	43	90	71		
19	0	Project 14	599		255	202	70	3	75	83		
20	1	Project 15	903		228	351	240	60	93	80		
21	1	Project 16	859		303	173	431	60	90	41		
22	1	Project 17	748		133	427	220	59	40	39		
23	1	Project 18	668		197	98	214	95	96	74		
24	1	Project 19	888		313	278	291	66	75	74		
25	0	Project 20	655		152	211	134	85	59	70		

**Рис. 33.4.** Новое оптимальное решение для ситуации обязательного выбора проекта 4 после выбора проекта 3

	A	B	C	D	E	F	G	H	I	J	K	L
1		Total NPV										
2		9014		Used	2378	2734	2755	778	896	702		
3					<=	<=	<=	<=	<=	<=		
4				Available	2500	2800	2900	900	900	900		
5	Do IT?		NPV		Cost Year 1	Cost Year 2	Cost Year 3	Labor Year 1	Labor Year 2	Labor Year 3		
6	0	Project 1	928		398	180	368	111	108	123		
7	0	Project 2	908		151	269	248	139	86	83	At most 4 of Projects 1-10	4
8	1	Project 3	801		129	189	308	56	61	23		
9	0	Project 4	543		275	218	220	54	70	59	<=	4
10	0	Project 5	944		291	252	228	123	141	70		
11	0	Project 6	848		80	283	285	119	84	37		
12	1	Project 7	545		203	220	77	54	44	42		
13	1	Project 8	808		150	113	143	67	101	43		
14	0	Project 9	638		282	141	160	37	55	64		
15	1	Project 10	841		214	254	355	130	72	62		
21	1	Project 16	859		303	173	431	60	90	41		
22	1	Project 17	748		133	427	220	59	40	39		
23	1	Project 18	668		197	98	214	95	96	74		
24	1	Project 19	888		313	278	291	66	75	74		
25	1	Project 20	655		152	211	134	85	59	70		

**Рис. 33.5.** Оптимальное решение для ситуации выбора не более четырех проектов из проектов 1–10

## Решение задач бинарного и целочисленного программирования

Линейные модели, в которых все или несколько изменяемых ячеек должны содержать бинарные или целые значения, обычно сложнее решать, чем линейные модели, в которых все изменяемые ячейки могут содержать дроби. Поэтому аналитики для задач бинарного или целочисленного программирования часто удовлетворяются решением, близким к оптимальному. Если поиск решения для модели слишком затянулся, попробуйте в **Параметры поиска решения (Solver Options)** щелкнуть **Параметры (Options)**, открыть диалоговое окно и откорректировать значение в поле **Целочисленная оптимальность (Integer Optimality)** — рис. 33.6. Например, значение

0,5 показывает, что поиск решения будет остановлен, как только будет найдено первое допустимое решение со значением в целевой ячейке, находящимся в пределах 0,5% от *теоретического оптимального значения*. (Теоретическое оптимальное значение в целевой ячейке — это оптимальное значение, найденное при отсутствии бинарных и целочисленных ограничений.) Часто вам придется делать выбор между поиском решения в пределах 10% от оптимального в течение 10 минут и поиском оптимального решения, для которого требуется 2 недели машинного времени! По умолчанию целочисленная оптимальность равна 5%; это означает, что поиск решения прекращается, когда найдено значение целевой ячейки в пределах 5% от теоретического оптимального значения в целевой ячейке. Когда мы решали пример с проектами разработки программного обеспечения в первый раз, для целочисленной оптимальности было установлено значение 5%, и для целевой ячейки было найдено оптимальное значение 9269. После изменения значения целочисленной оптимальности на 0,50% для целевой ячейки было получено большее значение (9293).

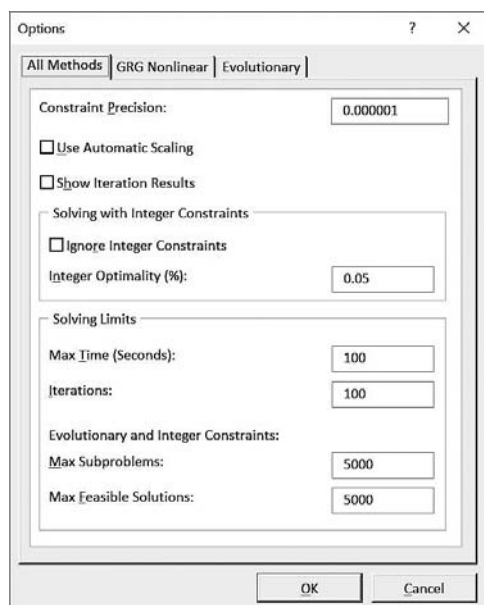


Рис. 33.6. Корректировка параметра Целочисленная оптимальность

## Задания

1. Компания рассматривает девять проектов. В таблице ниже указаны ЧПС каждого проекта и денежные средства, необходимые для его реализации в следующие два года. (Все значения указаны в миллионах.) Например, проект 1 добавляет ЧПС в размере \$14 млн и требует затрат в год 1 на сумму \$12 млн

и в год 2 на сумму \$3 млн. В год 1 для реализации проектов доступно \$50 млн, в год 2 — \$20 млн.

	ЧПС	Затраты, год 1	Затраты, год 2
Проект 1	14	12	3
Проект 2	17	54	7
Проект 3	17	6	6
Проект 4	15	6	2
Проект 5	40	32	35
Проект 6	12	6	6
Проект 7	14	48	4
Проект 8	10	36	3
Проект 9	12	18	3

- Как максимизировать ЧПС при условии, что одобрить часть проекта невозможно (можно принять только весь проект целиком или отклонить его)?
- Предположим, что проект 5 должен быть одобрен в случае, если одобрен проект 4. Каким образом максимизировать ЧПС?

2. Издательство пытается определить, какие из 36 книг должны быть изданы в этом году. Данные по каждой книге приведены в файле **Pressdata.xlsx**:

- прогнозируемые доходы и затраты на разработку (в тысячах долларов);
- количество страниц в каждой книге;
- ориентированность книги на разработчиков ПО (указана значением в столбце E).

В этом году издательство может выпустить книги общим объемом до 8500 страниц, и по крайней мере четыре книги должны быть предназначены для разработчиков ПО. Каким образом издательству максимизировать прибыль?

3. В уравнении **SEND + MORE = MONEY** каждая буква представляет цифру от 0 до 9. Какая цифра связана с каждой буквой?

4. Джил пытается составить свое расписание на следующий семестр. Семестр состоит из двух частей по семь недель каждая. Джил должна пройти четыре курса в каждой из половин семестра. В каждом семестре по пять временных интервалов. Естественно, Джил не может пройти один и тот же курс дважды. Для каждого курса и соответствующего временного интервала Джил указала определенную ценность. Эти данные содержатся в файле **Classdata.xlsx**. Например, курсу 1 в течение временного интервала 5 в первой половине семестра

присвоена ценность 5. Какие курсы должна пройти Джил в течение каждой половины семестра для максимального увеличения ценности пройденных за семестр курсов?

5. Используйте **Условное форматирование** для выделения желтым цветом каждой строки, соответствующей выбранному проекту.
6. Используйте **Поиск решения (Solver)** для определения минимального числа монет для суммы 92 цента.
7. Нефтяная компания рассматривает 12 проектов на бурение скважин. ЧПС и средства, необходимые на выполнение каждого из этих проектов, представлены в файле **Problem7.xlsx**. На реализацию проектов компания может выделить \$50 млн. Какие проекты следует одобрить?
8. Я собираюсь в поход и складываю в рюкзак наиболее полезные вещи. Вес в фунтах и степень полезности каждого предмета представлены в файле **Problem8data.xlsx**. Рюкзак вмещает 26 фунтов. Необходимо взять как минимум один напиток (воду или Gatorade) и как минимум один белковый продукт (сыр или вяленую говядину). Как собрать рюкзак таким образом, чтобы получилась максимальная полезность всех вошедших в него предметов?
9. Вы производите дорогие ожерелья для особ королевских кровей. У вас достаточно работников для производства трех различных видов ожерелий. Розничная цена каждого ожерелья, доступные драгоценные камни и их количество, необходимое для изготовления каждого ожерелья, указаны в файле **Problem9data.xlsx**. Производство каких ожерелий максимизирует ваш доход?

## ГЛАВА 34

# Поиск решения при финансовом планировании

### Обсуждаемые вопросы

- Можно ли с помощью инструмента Поиск решения проверить точность функции ПЛТ или определить ипотечные платежи с плавающей процентной ставкой?
- Можно ли с помощью инструмента Поиск решения определить, сколько денег нужно отложить до выхода на пенсию?

Инструмент Excel 2019 Поиск решения эффективен при анализе задач финансового планирования. Во многих задачах подобного типа такие суммы, как невыплаченный остаток по кредиту или денежная сумма, которую нужно скопить к выходу на пенсию, меняются со временем. Возьмем для примера ситуацию с кредитом. Поскольку невыплаченный остаток по кредиту снижается только за счет непроцентной части ежемесячного платежа, справедливо следующее уравнение (уравнение 1):

(невыплаченный остаток по кредиту на конец периода  $t$ ) = (невыплаченный остаток по кредиту на начало периода  $t$ ) – [(платеж за месяц  $t$ ) – (выплаченные проценты за месяц  $t$ )].

Теперь предположим, что вы откладываете средства, на которые собираетесь жить, когда выйдете на пенсию. Вплоть до выхода на пенсию вы вносите денег на свой пенсионный счет в начале каждого периода (допустим, что период равен году), и в течение года ваш пенсионный фонд инвестирует деньги и получает определенный доход от инвестиций. После выхода на пенсию вы снимаете деньги со своего счета в начале каждого года, и ваш пенсионный фонд все еще получает доход от инвестиций. Следующее уравнение (уравнение 2) описывает отношение между вкладами, снятием денег и доходом от инвестиций:

(пенсионные накопления на конец года  $(t + 1)$ ) = (пенсионные накопления на конец года  $t$  + пенсионный вклад в начале года  $(t + 1)$  – снятие денег в начале года  $(t + 1)$ ) \*  $(1 + \text{инвестиционный доход в \% за год } (t + 1))$ .

Применив к подобным базовым взаимосвязям инструмент Поиск решения (Solver), вы сможете ответить на мириады интересующих вас вопросов, связанных с финансовым планированием.

## Ответы на вопросы

❓ Можно ли с помощью инструмента Поиск решения проверить точность функции ПЛТ или определить ипотечные платежи с плавающей процентной ставкой?

Напомню, что в главе 10 «Еще несколько финансовых функций Excel» ежемесячный платеж (при условии платежей в конце месяца) по 10-месячному кредиту в размере \$8000 при годовой ставке 10% составил \$1037,03. Можем ли мы использовать Поиск решения (Solver) для вычисления ежемесячного платежа? Ответ вы найдете в файле Finmathsolver.xlsx на листе PMT By Solver (рис. 34.1).

	A	B	C	D	E
1			rate	0.00667	
2					
3		From PMT function	\$1,037.03		
4	Month	Beginning Balance	Payment	Interest Owed	Ending Balance
5	1	\$ 10,000.00	\$ 1,037.03	\$ 66.67	\$ 9,029.63
6	2	\$ 9,029.63	\$ 1,037.03	\$ 60.20	\$ 8,052.80
7	3	\$ 8,052.80	\$ 1,037.03	\$ 53.69	\$ 7,069.45
8	4	\$ 7,069.45	\$ 1,037.03	\$ 47.13	\$ 6,079.55
9	5	\$ 6,079.55	\$ 1,037.03	\$ 40.53	\$ 5,083.05
10	6	\$ 5,083.05	\$ 1,037.03	\$ 33.89	\$ 4,079.90
11	7	\$ 4,079.90	\$ 1,037.03	\$ 27.20	\$ 3,070.07
12	8	\$ 3,070.07	\$ 1,037.03	\$ 20.47	\$ 2,053.51
13	9	\$ 2,053.51	\$ 1,037.03	\$ 13.69	\$ 1,030.16
14	10	\$ 1,030.16	\$ 1,037.03	\$ 6.87	\$ 0.00

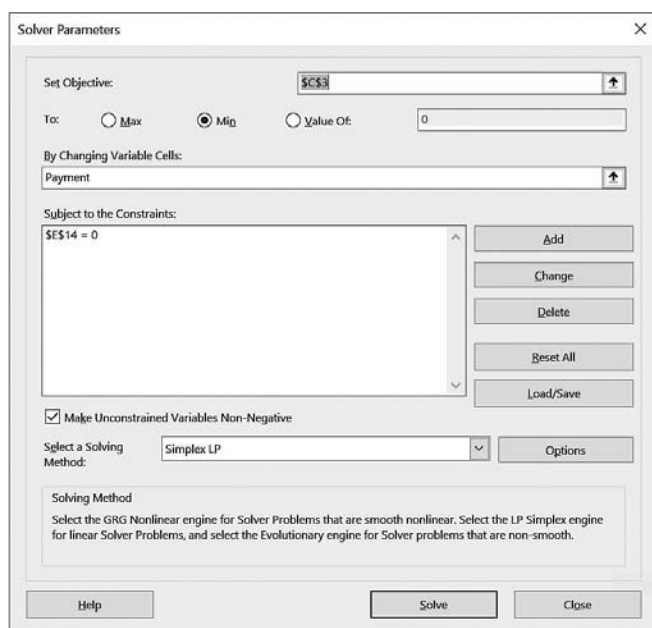
**Рис. 34.1.** Модель поиска решения для вычисления ежемесячного платежа по кредиту

Ключом к этой модели является применение уравнения 1 (приведенного ранее в данной главе) для отслеживания остатка на начало месяца. В целевой ячейке должен находиться минимальный размер ежемесячного платежа. Изменяемая ячейка — это ячейка с ежемесячным платежом. Единственным ограничением является нулевой остаток на конец месяца 10.

Я ввел остаток на начало месяца в ячейку B5. В ячейку C5 (с именем Payment) я ввел пробную величину ежемесячного платежа и затем скопировал его в диапазон C6:C14. Поскольку я подразумеваю, что платежи вносятся в конце каждого месяца, проценты отражаются на остатке на начало каждого месяца. Ежемесячная процентная ставка вычисляется в ячейке D1 (с именем rate) путем деления ежегодной ставки 0,08 на 12. Выплачиваемые каждый месяц проценты вычисляются путем копирования формулы =rate\*B5 из D5 в D6:D14. Проценты вычисляются как  $0,006666 \times (\text{остаток на начало месяца})$ . Для вычисления остатка на конец каждого месяца (согласно уравнению 1) я скопировал формулу =(B5-(Payment-D5)) из E5 в E6:E14. Поскольку остаток на начало месяца ( $t + 1$ ) равен остатку на конец месяца  $t$ , остаток на начало каждого месяца вычисляется копированием формулы =E5 из B6 в B7:B14.

Теперь все готово для использования Поиска решения (Solver) с целью вычисления ежемесячного платежа. Диалоговое окно Параметры поиска решения (Parameters) показано на рис. 34.2. (Чтобы открыть диалоговое окно Параметры поиска решения, нажмите Поиск решения (Solver) в группе Анализ (Analysis) на вкладке Данные (Data).)

Цель — минимизировать ежемесячный платеж (ячейка C5). Обратите внимание, что изменяемая ячейка является в то же время целевой ячейкой. Единственное ограничение — нулевой остаток на конец месяца 10. Добавление этого ограничения обеспечивает полную выплату кредита. После того как я выбрал симплекс-метод и неотрицательные переменные, Поиск решения вычислил \$1037,03, что соответствует сумме, полученной при помощи функции ПЛТ.



**Рис. 34.2.** Диалоговое окно Параметры поиска решения для вычисления ежемесячного платежа

Эта модель является линейной, поскольку значение в целевой ячейке равно значению в изменяемой ячейке, и ограничение создается путем суммирования значений, кратных значениям в изменяемой ячейке.

Следует отметить, что если модели поиска решения включают очень большие или очень малые числа, то Поиск решения иногда обрабатывает линейные модели как нелинейные. Во избежание этой проблемы в диалоговом окне Параметры (Options) рекомендую включить флажок Использовать автоматическое масштабирование (Use Automatic Scaling). После этого модель будет распознаваться как линейная.

**?** Можно ли с помощью инструмента Поиск решения определить, сколько денег нужно отложить до выхода на пенсию?

Сумму, которую необходимо отложить до выхода на пенсию, можно узнать с помощью уравнения 2 (приведенного ранее в данной главе). Вот пример.

В начале этого и каждого из последующих 39 лет я планирую откладывать некоторую сумму в свой пенсионный фонд. Каждый год я планирую увеличивать свой взнос на \$500. После выхода на пенсию через 40 лет я планирую снимать \$100 000 в год в течение 20 лет (в начале каждого года). Я сделал следующие предположения о доходности своего пенсионного инвестиционного портфеля:

- в первые 20 лет инвестиции будут приносить 10% в год;
- в остальные годы инвестиции будут приносить 5% в год.

Внесение и снятие денег осуществляется в начале года. Исходя из этих предположений, какова наименьшая сумма, которую необходимо внести в этом году и которой будет достаточно для обеспечения снятия требуемых сумм после выхода на пенсию?

Решение этой задачи приведено в файле `Finmathsolver.xlsx` на листе `Retire` и на рис. 34.3. Обратите внимание, что я скрыл много строк в модели.

	A	B	C	D	E	F
5	Year	Initial balance	Contribution	Return	Withdrawal	Ending Balance
6	1	\$0.00	\$ 1,387.87	10%	\$0.00	\$1,526.65
7	2	\$1,526.65	\$ 1,887.87	10%	\$0.00	\$3,755.98
8	3	\$3,755.98	\$ 2,387.87	10%	\$0.00	\$6,758.23
44	39	\$1,146,596.10	\$ 20,387.87	5%	\$0.00	\$1,225,333.17
45	40	\$1,225,333.17	\$ 20,887.87	5%	\$0.00	\$1,308,532.09
46	41	\$1,308,532.09		5%	\$100,000.00	\$1,268,958.69
47	42	\$1,268,958.69		5%	\$100,000.00	\$1,227,406.62
62	57	\$372,324.80		5%	\$100,000.00	\$285,941.04
63	58	\$285,941.04		5%	\$100,000.00	\$195,238.10
64	59	\$195,238.10		5%	\$100,000.00	\$100,000.00
65	60	\$100,000.00		5%	\$100,000.00	\$0.00

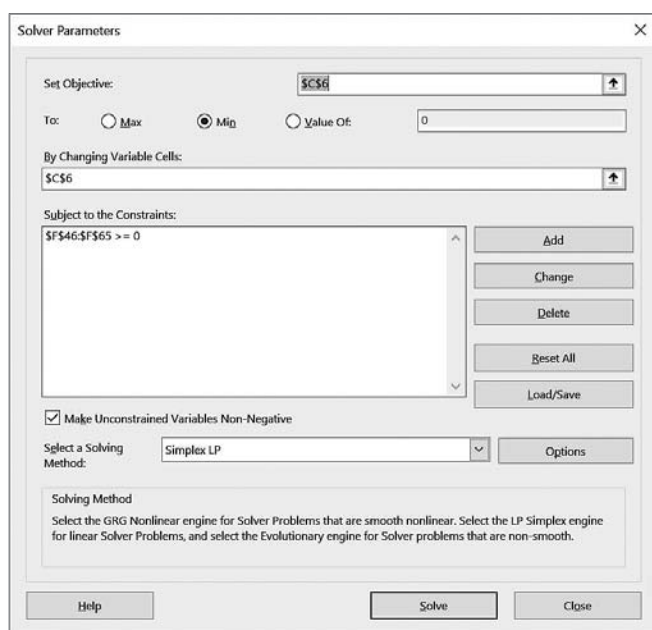
**Рис. 34.3.** Данные планирования пенсии для анализа с помощью инструмента Поиск решения

На этом листе отслеживается баланс на пенсионном счете на протяжении следующих 60 лет. Каждый год инвестиции приносят доход в соответствии с указанной процентной ставкой. Начнем с ввода пробного значения для взноса в год 1 в ячейку C6. Копирование формулы  $=C6+500$  из C7 в C8:C45 обеспечивает увеличение взноса на \$500 в каждый год со 2-го по 40-й. Я ввел в столбец D предполагаемый процент дохода от инвестиций для каждого года из 60. В ячейки E46:E65 я ввел ежегодно снимаемую со счета сумму (\$100 000) в годы с 41-го по 60-й. Копирование формулы  $=(B6+C6-E6)*(1+D6)$  из F6 в F7:F65 использует уравнение 2, чтобы вычислить для



каждого года баланс на пенсионном счете на конец периода. Для вычисления баланса на начало каждого года со 2-го по 60-й я скопировал формулу =F6 из B7 в B8:B65. Разумеется, начальный баланс в год 1 равен 0. Обратите внимание, что значение 6,8704E-07 в ячейке F65 приблизительно равно 0 с учетом погрешности округления.

Диалоговое окно **Параметры поиска решения (Solver Parameters)** для этой модели показано на рис. 34.4. Я хочу минимизировать взнос в год 1 (ячейка C6). Изменяемой ячейкой также является ячейка со значением взноса в год 1 (ячейка C6). Я гарантирую, что никогда не буду страдать от нехватки денег на пенсии, путем добавления ограничения F46:F65 >= 0, чтобы баланс на конец периода для каждого года с 41-го по 60-й был неотрицательным.



**Рис. 34.4.** Диалоговое окно Параметры поиска решения для задачи планирования пенсии

В диалоговом окне **Параметры поиска решения (Solver Parameters)** после выбора симплекс-метода и включения флажка **Сделать переменные без ограничений неотрицательными (Make Unconstrained Variables Non-Negative)** я щелкнул на **Найти решение (Solve)** и узнал, что взнос в год 1 должен составить \$1387,87.

Эта модель является линейной, поскольку целевая ячейка та же, что и изменяемая ячейка, а ограничение создается путем добавления значений, кратных значениям в изменяемых ячейках. Обратите внимание, что поскольку доход от инвестиций меняется, для этой задачи невозможно найти решение просто с помощью финансовых функций Excel. Инструмент **Поиск решения (Solver)** предоставляет общую

схему решения, позволяющую анализировать задачи финансового планирования, в которых ставки по ипотечным кредитам или доходность инвестиций не являются постоянными величинами.

## Задания

1. Я беру в кредит \$15 000 на покупку нового автомобиля. Я должен вносить платеж 60 месяцев подряд, в последних числах месяца. Годовая ставка по кредиту составляет 10%. Агент по продаже автомобилей — мой друг и он позволит мне вносить ежемесячные платежи за месяцы с 1-го по 30-й вдвое меньшие, чем платежи за месяцы с 31-го по 60-й. Каковы платежи в каждом месяце?
2. Решите задачу планирования пенсии при условии, что снимаете деньги в конце каждого года, а вносите деньги — в начале каждого года.
3. Решите задачу с ипотечным кредитом при условии, что платежи вносятся в начале каждого месяца.
4. Допустим, что в задаче планирования пенсии ваша зарплата в год 1 составляет \$40 000 и возрастает на 5% каждый год до выхода на пенсию. Каждый год, пока вы работаете, вы собираетесь откладывать один и тот же процент от зарплаты. Какой процент от зарплаты необходимо откладывать?
5. В задаче с ипотечным кредитом требуется увеличивать ежемесячный платеж на \$50 каждый месяц. Каким должен быть ежемесячный платеж?
6. Вы собираетесь взять на 20 лет ипотечный кредит размером \$300 000 с платежами в конце каждого месяца. Годовая процентная ставка составляет 6%. Через 20 лет вам нужно будет выполнить завершающий остаточный платеж в размере \$40 000. Поскольку вы ожидаете, что ваши доходы будут расти, вы собираетесь структурировать кредит таким образом, что в начале каждого года ваш ежемесячный платеж будет расти на 2%. Вычислите сумму ежемесячного платежа для каждого года.
7. Мать Блэра откладывает деньги на его высшее образование. В таблице ниже представлены суммы платежей, которые должны быть выполнены в указанное время.

Через 4 года	Через 5 лет	Через 6 лет	Через 7 лет
\$24 000	\$26 000	\$28 000	\$30 000

Доступны следующие инвестиции.

- Сегодня, через год, через два года, через три года и через четыре года она может инвестировать деньги на год и получить 6% дохода.
- Сегодня, через два года и через четыре года она может инвестировать деньги на два года и получить 14% дохода.

- Через три года она может инвестировать деньги на три года и получить 18% дохода.
- Сегодня она может инвестировать деньги на семь лет и получить 65% дохода.

Какова минимальная сумма, которую мать Блэра должна вложить в его образование сегодня для гарантированной оплаты счетов колледжа?

8. У меня \$10 000 на одной кредитной карте с годовой процентной ставкой 18% и \$5000 на другой карте с годовой процентной ставкой 12%. Проценты за месяц рассчитываются на основе остатка на начало месяца. Я могу выплачивать \$2000 в месяц, и минимальный ежемесячный платеж по каждой карте составляет 10% от неоплаченного остатка на начало месяца. Моя цель — расплатиться по обеим картам в течение двух лет. Какова минимальная сумма выплаты процентов?
9. Вы взяли в кредит \$50 000 и должны производить платежи в конце каждого из следующих 24 месяцев. Годовая ставка по кредиту составляет 10%. В 7-й, 13-й и 19-й месяцы ежемесячный платеж увеличивается на \$500. Сколько вы будете должны платить каждый месяц?

## ГЛАВА 35

# Поиск решения при оценке спортивных команд

Обсуждаемый вопрос

- Как можно с помощью Excel распределить очки для команд НФЛ?

Многие из нас смотрят баскетбол, американский футбол, хоккей или бейсбол. Букмекеры считают разницу в счете для матчей для всех этих видов спорта, а также для других видов. Например, наилучшим предположением букмекеров был выигрыш командой «Каролина Пэнтерс» Суперкубка в 2016 г. с разницей в 5 очков. Однако выиграла «Денвер Бронкос». В этой главе мы узнаем, насколько точно Поиск решения (Solver) может оценить относительную силу команд Национальной футбольной лиги (НФЛ).

С помощью простой модели поиска решения вы можете сгенерировать обоснованную разницу очков в играх на основе результатов сезона 2017 г. (включая все игры, кроме Суперкубка 2018 г.). Решение приведено в файле `Nfl2017.xlsx` (рис. 35.1). В качестве исходных данных используется счет в каждой игре команд НФЛ в сезоне 2017 г. Изменяемые ячейки для модели поиска решения — это рейтинги для каждой команды и величина преимущества на домашнем поле. Например, если у команды «Индианаполис Колтс» (Indianapolis Colts) рейтинг +5, а у «Нью-Йорк Джетс» (New York Jets) рейтинг +7, то считается, что вторая команда на два очка превосходит первую.

Что касается преимущества на домашнем поле, в течение многих лет футбольные команды колледжей и профессиональные футбольные команды, а также профессиональные баскетбольные команды, как правило, выигрывают на своем поле в среднем три очка (а баскетбольные команды колледжей, как правило, выигрывают в среднем пять очков). В данной модели определим преимущество на домашнем поле как изменяемую ячейку, и пусть Поиск решения (Solver) вычислит это значение. Можно определить результат игры в НФЛ как количество очков, на которое команда хозяев опережает команду гостей, и предсказать результат каждой игры с помощью следующего уравнения (уравнение 1):

(прогнозные очки, на которые команда хозяев опережает команду гостей) = (преимущество на домашнем поле) + (рейтинг команды хозяев) – (рейтинг команды гостей).

Например, если преимущество на домашнем поле составляет три очка, и команда «Колтс» принимает команду «Джетс», «Колтс» будет фаворитом с разницей

в одно очко ( $3 + 5 - 7$ ). Если «Джетс» принимает «Колтс», то «Джетс» будет фаворитом с разницей в пять очков ( $3 + 7 - 5$ ).

Какая целевая ячейка даст надежные рейтинги? Цель — найти множество значений для рейтингов команд и преимущества на домашнем поле, которое наилучшим образом предсказывает результаты всех игр. То есть предсказание для игры должно быть как можно ближе к результату игры. Таким образом, требуется минимизировать сумму разностей (фактический результат) – (прогнозируемый результат) по всем играм. Однако проблема в том, что ошибки предсказаний с разными знаками нейтрализуют друг друга. Например, если предсказана ошибочная разница для команды хозяев поля в плюс 50 очков в одной игре и в минус 50 очков в другой игре, то в целевой ячейке будет значение 0, показывающее идеальную точность, хотя по факту ошибка составляет 50 очков за игру. Эту проблему можно решить, минимизируя суммы по всем играм по формуле  $[(\text{фактический результат}) - (\text{прогнозируемый результат})]^2$ . Теперь ошибки с разными знаками не будут нейтрализовать друг друга.

## Ответ на вопрос

### ❓ Как можно с помощью Excel распределить очки для команд НФЛ?

Теперь посмотрим, как определить точные рейтинги команд НФЛ по результату игр регулярного сезона 2017 г. Данные к этой задаче находятся в файле NFL2017.xlsx (рис. 35.1).

	A	B	C	D	E	F	G	H	I
1		-8.8E-12			=home+VLOOKUP(C5,lookup,2,FALSE)-VLOOKUP(D5,lookup,2,FALSE)				
2	Team	Rating	home	2.454243577					SSE
3	Arizona Cardinals	-3.79							37836.51366
4	Atlanta Falcons	5.56	Home	Away	Home Points	Away Points	Home Margin	Prediction	Squared Error
5	Baltimore Ravens	3.16	New England Patriots	Kansas City Chiefs	27	42	-15	8.77	565.03
6	Buffalo Bills	-3.51	Washington Redskins	Philadelphia Eagles	17	30	-13	-9.62	11.44
7	Carolina Panthers	4.64	Detroit Lions	Arizona Cardinals	35	23	12	8.77	10.41
8	Chicago Bears	-1.41	Houston Texans	Jacksonville Jaguars	7	29	-22	-10.52	131.85
9	Cincinnati Bengals	-5.20	Tennessee Titans	Oakland Raiders	16	26	-10	3.56	184.00
10	Cleveland Browns	-11.19	Chicago Bears	Atlanta Falcons	17	23	-6	-4.52	2.20
11	Dallas Cowboys	1.79	Buffalo Bills	New York Jets	21	12	9	3.66	28.51
12	Denver Broncos	-6.55	Cleveland Browns	Pittsburgh Steelers	18	21	-3	-13.30	106.14
13	Detroit Lions	2.52	Cincinnati Bengals	Baltimore Ravens	0	20	-20	-5.91	198.61
14	Green Bay Packers	-2.09	Los Angeles Rams	Indianapolis Colts	46	9	37	20.70	265.66
15	Houston Texans	-6.61	San Francisco 49ers	Carolina Panthers	3	23	-20	-5.10	222.14
16	Indianapolis Colts	-10.25	Green Bay Packers	Seattle Seahawks	17	9	8	-1.51	90.53
17	Jacksonville Jaguars	6.36	Dallas Cowboys	New York Giants	19	3	16	11.68	18.66
18	Kansas City Chiefs	3.03	Minnesota Vikings	New Orleans Saints	29	19	10	1.17	77.93
19	Los Angeles Chargers	3.82	Denver Broncos	Los Angeles Chargers	24	21	3	-7.91	119.10
20	Los Angeles Rams	8.00	Cincinnati Bengals	Houston Texans	9	13	-4	3.86	61.79
21	Miami Dolphins	-6.02	Baltimore Ravens	Cleveland Browns	24	10	14	16.81	7.88
22	Minnesota Vikings	7.66	Jacksonville Jaguars	Tennessee Titans	16	37	-21	12.25	1105.40
23	New England Patriots	9.35	Kansas City Chiefs	Philadelphia Eagles	27	20	7	-5.33	151.99
24	New Orleans Saints	8.94	Indianapolis Colts	Arizona Cardinals	13	16	-3	-4.00	0.99
25	New York Giants	-7.44	Carolina Panthers	Buffalo Bills	9	3	6	10.60	21.15
26	New York Jets	-4.72	New Orleans Saints	New England Patriots	20	36	-16	2.05	325.74
27	Oakland Raiders	-4.54	Pittsburgh Steelers	Minnesota Vikings	26	9	17	-0.64	311.25
28	Philadelphia Eagles	10.81	Tampa Bay Buccaneers	Chicago Bears	29	7	22	2.73	371.31

**Рис. 35.1.** Данные для определения рейтингов команд НФЛ с помощью инструмента Поиск решения

Сначала поместим пробное значение преимущества на домашнем поле в ячейку D2 и дадим этой ячейке имя **home**. Кроме того, в ячейки B3:B34 введем пробные значения (подойдут любые числа) для рейтингов всех команд.

Начиная со строки 6 в столбцах C и D перечислены команды хозяев и команды гостей для каждой игры. Например, первую игру (в строке 6) команда Канзас-Сити играет в Новой Англии. В столбце E содержатся очки, набранные командой хозяев, а в столбце F — очки, набранные командой гостей. Как видно из рисунка, команда Patriots одержала победу над командой Chiefs со счетом 42:27. Теперь можно вычислить результат каждой игры (разницу в очках, с которой команда хозяев победила команду гостей) по формуле  $=E5-F5$  в ячейке G5. Установив указатель в правой нижней части этой ячейки и дважды щелкнув левой кнопкой мыши, можно скопировать формулу вниз до последней игры регулярного сезона, то есть до строки 270. (Кстати, проще всего выделить все данные, нажав **Ctrl+Shift+↓**. Действие, вызываемое этой комбинацией клавиш, распространяется до последней строки, заполненной данными, — строки 270 в данном случае.)

В столбце H создайте прогноз для каждой игры согласно уравнению 1. Прогноз для первой игры вычисляется в ячейке H5 следующим образом:

$=- \text{home} + \text{ВПР}(C5;\text{lookup};2;\text{ЛОЖЬ}) - \text{ВПР}(D5;\text{lookup};2;\text{ЛОЖЬ});$

По этой формуле прогноз для первой игры вычисляется путем сложения преимущества на домашнем поле с рейтингом команды хозяев и последующего вычитания рейтинга команды гостей. Член  $\text{ВПР}(C5;\text{lookup};2;\text{ЛОЖЬ})$  выполняет поиск рейтинга команды хозяев, а член  $\text{ВПР}(D5;\text{lookup};2;\text{ЛОЖЬ})$  — рейтинга команды гостей. (Об использовании функций поиска см. в главе 3 «Функции поиска».) В столбце I вычисляется квадратичная ошибка (фактический счет — прогнозируемый счет)<sup>2</sup> для каждой игры. Квадратичная ошибка для первой игры вычисляется в ячейке I6 по формуле  $=(G6-H6)^2$ . Выделите диапазон ячеек G5:I5, дважды щелкните кнопкой мыши и скопируйте формулу до строки 270. Затем вычислите цель в ячейке I3 путем суммирования всех квадратичных ошибок по формуле  $=\text{СУММ}(I5:I270)$ . Формулу для большого столбца чисел можно создать следующим образом: введите часть формулы  $=\text{СУММ}()$  и выделите первую ячейку в диапазоне, значения в котором требуется сложить. Затем нажмите **Ctrl+Shift+↓** для ввода диапазона от выделенной ячейки до нижней строки в столбце и добавьте закрывающую скобку.

Удобно принимать средний рейтинг команд равным 0. Команда с положительным рейтингом играет лучше средней команды, а команда с отрицательным рейтингом — хуже. Вычислите средний рейтинг команд в ячейке B1 по формуле  $=\text{СРЗНАЧ}(B3:B34)$ .

Теперь заполним диалоговое окно **Параметры поиска решения (Solver Parameters)**, как показано на рис. 35.2.

Сведите к минимуму сумму квадратичных ошибок прогнозирования по всем играм (вычисляемую в ячейке I3) путем изменения рейтингов команд (перечисленных в ячейках B3:B34, с именем диапазона **rating**) и преимущества в очках на

Solver Parameters

Set Objective:

To: ☐ Max ☒ Min ☐ Value Of:

By Changing Variable Cells:

Subject to the Constraints:

☐ Make Unconstrained Variables Non-Negative

Select a Solving Method:

Solving Method

Select the GRG Nonlinear engine for Solver Problems that are smooth nonlinear. Select the LP Simplex engine for linear Solver Problems, and select the Evolutionary engine for Solver problems that are non-smooth.

**Рис. 35.2.** Диалоговое окно Параметры поиска решения для определения рейтинга команд НФЛ

	J	K	L
4	Rank	Team	Rating
5	1	Philadelphia Eagles	10.81
6	2	New England Patriots	9.35
7	3	New Orleans Saints	8.94
8	4	Los Angeles Rams	8.00
9	5	Minnesota Vikings	7.66
10	6	Jacksonville Jaguars	6.36
11	7	Atlanta Falcons	5.56
12	8	Carolina Panthers	4.64
13	9	Pittsburgh Steelers	4.56
14	10	Los Angeles Chargers	3.82
15	11	Baltimore Ravens	3.16
16	12	Kansas City Chiefs	3.03
17	13	Detroit Lions	2.52
18	14	Seattle Seahawks	1.88
19	15	Dallas Cowboys	1.79
20	16	Tampa Bay Buccaneers	-1.14
21	17	Washington Redskins	-1.26
22	18	Chicago Bears	-1.41
23	19	Green Bay Packers	-2.09
24	20	San Francisco 49ers	-2.91

**Рис. 35.3.** Топ-20 команд НФЛ в сезоне 2017 г.

домашнем поле (ячейка B1). Ограничение  $B3=0$  гарантирует, что средний рейтинг команд равен 0. Обратите внимание, что мы не ставим флажок в диалоговом окне **Неотрицательные изменяющиеся ячейки**, потому что у некоторых команд должен быть отрицательный рейтинг, чтобы их позиции были ниже среднего.

На рис. 35.1 отображены рейтинги команд и значение преимущества на домашнем поле, вычисленные при помощи **Поиск решения**. Мы можем видеть, что преимущество команды хозяев над командой гостей составляет 2,45 очка. На рис. 35.3 представлен список 20 команд с самым высоким рейтингом (попадающих в список идущих на Суперкубок). Помните, что рейтинги в ячейках  $B3:B32$  вычисляются при помощи инструмента **Поиск решения (Solver)**. В файле шаблона можно начать с любых чисел в этих ячейках, и инструмент **Поиск решения (Solver)** все равно справится с вычислением «лучших» рейтингов.

При проведении Суперкубка 2018 года букмекеры ставили вперед команду «Новая Англия» с преимуществом в 7 очков. Наша модель предрекла победу команде Eagles, которая привела многострадальных фанатов Eagles в эйфорию (смотрите об этом *Silver Linings Playbook!*).

## Почему эта модель не является линейной моделью поиска решения?

Эта модель не является линейной, поскольку в целевой ячейке суммируются члены в виде (рейтинг команды хозяев + преимущество на домашнем поле – рейтинг команды гостей)<sup>2</sup>. Вспомните: в линейной модели поиска решения целевая ячейка должна быть создана путем сложения членов вида (изменяемая ячейка)\*(константа). В данном случае соотношение другое, поэтому модель не является линейной. Однако **Поиск решения (Solver)** получает правильный ответ для любой модели определения спортивного рейтинга, в которой в целевой ячейке минимизируется сумма квадратичных ошибок. Обратите внимание, что для решения выбран метод ОПГ (GRG), так как используется нелинейная модель, которая не включает нематематические или негладкие функции, такие как функция ЕСЛИ.

## Задания

1. Файлы Nfl0x.xlsx ( $x = 1, 2, 3, 4$ ) содержат результаты игр каждого регулярно-го сезона НФЛ 2001–2004 гг. Определите рейтинги команд в каждом сезоне. Какая команда в каждом сезоне по прогнозу должна получить Суперкубок?
2. Для сезона 2004 г. разработайте метод прогнозирования фактического счета в каждой игре. Подсказка: присвойте каждой команде рейтинг в нападении и рейтинг в защите. У какой команды было самое эффективное нападение? У какой команды была самая эффективная защита?
3. Справедливо ли утверждение «Команда НФЛ могла бы проиграть каждую игру и иметь рейтинг выше среднего»?



4. Файл `Nba01_02.xlsx` содержит результаты всех игр команд НБА в сезоне 2001–2002 гг. Определите рейтинги команд.
5. Файл `Nba02_03.xlsx` содержит результаты всех игр команд НБА в сезоне 2002–2003 гг. Определите рейтинги команд.
6. Файл `Worldball.xlsx` содержит все результаты чемпионата мира по баскетболу 2006 г. Определите рейтинги команд. Какие три команды были лучшими?
7. Этот метод определения рейтингов команд отлично работает для американского футбола и баскетбола. Какие проблемы возникают при применении этих методов в хоккее или бейсболе?
8. Файл `NFL2012data.xlsx` содержит результаты всех игр команд НФЛ в регулярном сезоне 2012 г. Определите рейтинги команд. Несмотря на результат 10:6 для команды «Колтс», ее рейтинг значительно ниже среднего. Чем объясняется это противоречие?

## ГЛАВА 36

# Расположение складов по методу ОПГ с несколькими начальными точками и согласно эволюционному поиску решения

Обсуждаемые вопросы

- В каком месте в США транспортная интернет-компания должна расположить единственный склад для минимизации суммарной протяженности транспортировки пакетов?
- В каком месте в США транспортная интернет-компания должна расположить два склада для минимизации суммарной протяженности транспортировки пакетов?

Начиная с Excel 2010 инструмент Поиск решения (Solver) непрерывно пополнялся новыми возможностями. В данной главе (а также главах 37 и 38) я поясню, как с помощью этих алгоритмов можно решить многие важные оптимизационные задачи.

## Метод ОПГ с несколькими начальными точками и эволюционный поиск решения

Как я указывал в главе 29 «Введение в оптимизацию с надстройкой Поиск Решения» Поиск решения (Solver) Microsoft Excel 2019 решает оптимизационные задачи тремя методами: линейные задачи симплекс-методом, гладкие нелинейные задачи методом ОПГ и негладкие задачи эволюционным методом. В следующих разделах мы подробно рассмотрим решение оптимизационных задач с помощью двух последних методов.

### Как Поиск решения применяется к линейным задачам?

Как я указывал в главах 30–34, модель поиска решения является линейной, если все ссылки на изменяемые ячейки в целевой ячейке и ограничения создаются путем суммирования членов вида (изменяемая\_ячейка)\*(константа). Для линейных

моделей всегда следует выбирать симплекс-метод, разработанный для эффективного поиска решений в линейных моделях. Поиск решения (Solver) в Excel 2019 может обрабатывать задачи, имеющие до 200 изменяемых ячеек и до 100 ограничений. На веб-сайте Solver.com доступны версии инструмента Поиск решения (Solver), способные обрабатывать более масштабные задачи.

## Как метод ОПГ применяется к нелинейным оптимизационным моделям?

Если целевая ячейка и/или ограничение содержат ссылки на изменяемые ячейки не в виде (изменяемая\_ячейка)\*(константа), значит, модель является нелинейной. Если  $x$  и  $y$  — изменяемые ячейки, то к нелинейной модели приводят следующие ссылки в целевой ячейке и/или каком-либо ограничении:

- $x^2$ ;
- $xy$ ;
- $\sin x$ ;
- $e^x$ ;
- $xe^{2y}$ .

Если нелинейные формулы включают обычные математические операторы (как в предыдущих примерах), примененный к модели надлежащим образом метод ОПГ должен быстро приводить к оптимальному решению. Для иллюстрации работы метода ОПГ предположим, что необходимо найти максимум функции —  $x^2$ . Графическое представление этой функции дано на рис. 36.1.

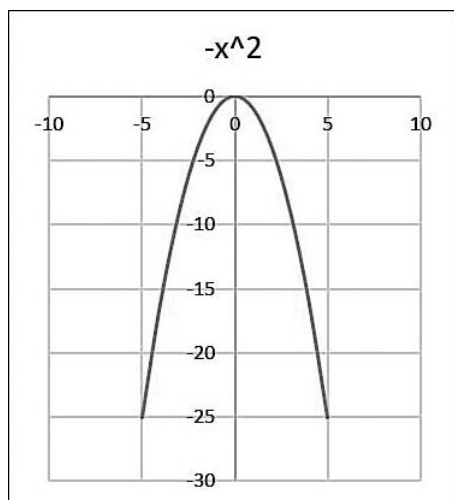
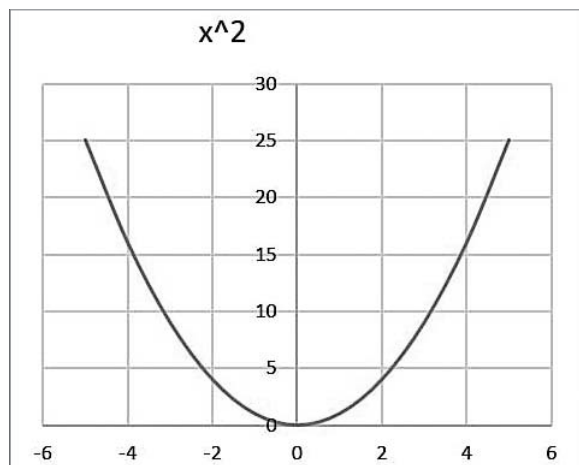


Рис. 36.1. Поиск максимума функции по методу ОПГ

Как видите, эта функция имеет максимум при  $x = 0$ . Отметим, что в точке  $x = 0$  градиент функции равен 0. При решении задачи по этому методу выполняется поиск точки, в которой градиент функции равен 0. Аналогично, если требуется найти минимум функции  $y = x^2$ , то задача по методу ОПГ решается путем определения точки, где градиент равен 0 (при  $x = 0$ ) — рис. 36.2.



**Рис. 36.2.** Поиск минимума функции по методу ОПГ

К сожалению, для многих функций невозможно найти максимум по точке, в которой градиент функции равен 0. Например, вы хотите найти максимум функции, представленной на рис. 36.3, для  $x$  от 0 до 12.



**Рис. 36.3.** Поиск максимума функции с несколькими максимумами

Как видите, у этой функции несколько максимумов. Если начать со значения  $x$ , близкого к 8, будет получено правильное решение задачи ( $x$  несколько меньше, чем 8). Если начать со значения, близкого к другому максимуму, скажем, к  $x = 2$ , будет найдено значение  $x$  несколько больше, чем 2, которое не является правильным. Поскольку в большинстве задач (особенно в тех, где изменяемых ячеек несколько) хорошая начальная точка неизвестна, похоже, что именно это и является основным препятствием. К счастью, в Excel 2019 есть параметр **Несколько начальных точек (Multistart)**.

Для установки этого параметра выберите на вкладке **Данные (Data)** в группе **Анализ (Analysis)** инструмент **Поиск решения (Solver)**. (Если вы еще не установили инструмент **Поиск решения**, обратитесь к главе 29.) В окне **Параметры поиска решения (Options)**, справа от **Выберите метод решения/Select A Solving Method**) нажмите кнопку **Параметры (Options)** и в открывшемся окне выберите вкладку **Поиск решения нелинейных задач методом ОПГ (GRG Nonlinear)**. На этой вкладке включите флажок **Использовать несколько начальных точек (Use Multistart)**. Тогда оптимальное решение будет выбрано из нескольких решений, полученных для нескольких начальных точек. Как правило, такой подход решает проблему наличия нескольких экстремумов.

Кстати, поиск решения можно остановить, нажав **Esc**. Также помните, что метод ОПГ с несколькими начальными точками работает эффективнее при выборе разумных верхних и нижних границ значений в изменяемых ячейках. (Например, не следует указывать, что значение в изменяемой ячейке  $< 100$  миллионов.)

Применение метода ОПГ также сопряжено с проблемами, если в целевой ячейке и/или в ограничениях используются негладкие функции: **МАКС (MAX)**, **МИН (MIN)**, **ABS**, **ЕСЛИ (IF)**, **СУММЕСЛИ (SUMIF)**, **СЧЁТЕСЛИ (COUNTIF)**, **СУММЕСЛИМН (SUMIFS)**, **СЧЁТЕСЛИМН (COUNTIFS)** и др., связанные с изменяемыми ячейками. Эти функции содержат точки, в которых не существует однозначно определенный градиент, поскольку в них происходит резкое изменение градиента. Предположим, что в оптимизационной задаче требуется смоделировать значение европейского колл-опциона с ценой исполнения \$40. Этот опцион позволяет купить акцию за \$40. Если цена акции равна  $s$  на дату истечения срока действия опциона, то стоимость колл-опциона может быть вычислена по формуле  $=\text{МАКС}(0; s-40)$  или представленной графически на рис. 36.4. Очевидно, что при  $s = 40$  функция стоимости опциона не имеет градиента, поэтому метод ОПГ не приведет к правильному результату.

Как показано на рис. 36.5, модели поиска решения, включающие функцию модуля (напомним, что модуль числа — это лишь расстояние от числа до точки 0), не имеют градиента для  $x = 0$ . В Excel модуль  $x$  возвращает функция **ABS(x)**.

Оптимизационные задачи, в которых целевая ячейка и/или любое ограничение не имеют градиента для какого-либо из значений в изменяемых ячейках, называются *негладкими* оптимизационными задачами. При решении задач такого типа метод ОПГ наталкивается на трудности. Поэтому я рекомендую выбрать **Эволюционный поиск решения (Solver Evolutionary)**. В нелинейных моделях поиска решения допускается использовать 100 изменяемых ячеек и 100 ограничений.

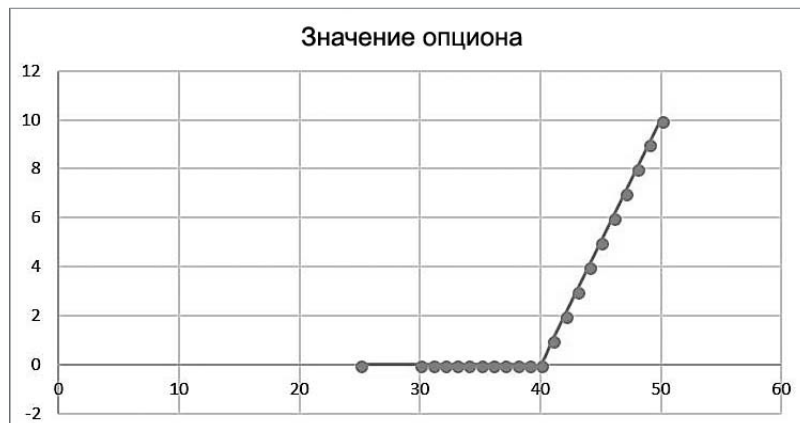


Рис. 36.4. Функция стоимости опциона не имеет градиента при цене \$40 за акцию

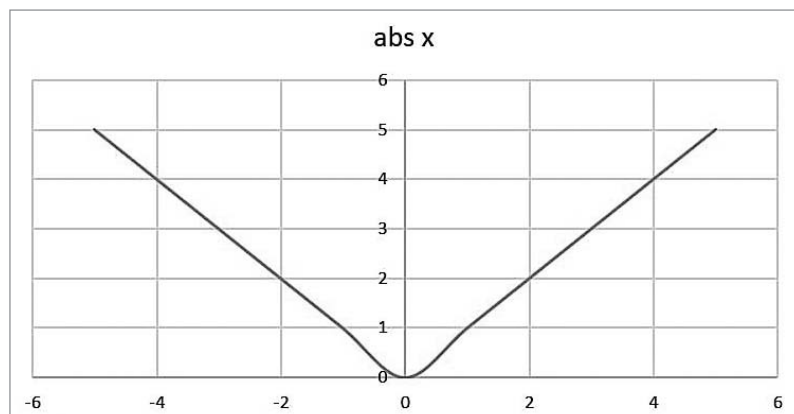


Рис. 36.5. Функция  $|x|$  не имеет градиента для  $x = 0$

## Как метод эволюционного поиска решения применяется к негладким оптимизационным задачам?

Эволюционный поиск решения в Excel 2019 основан на генетических алгоритмах, разработанных Джоном Холландом, профессором информатики в Мичиганском университете. Для применения метода эволюционного поиска решения сначала берется от 50 до 100 точек в области допустимых решений задачи (т. е. множество точек, удовлетворяющих ограничениям). Это множество точек называется *популяцией*. Затем для каждой точки вычисляется значение в целевой ячейке. Согласно принципу «выживает наиболее приспособленный» из теории эволюции точки в популяции следует заменять таким образом, чтобы повышалась вероятность размещения будущих членов популяции возле предыдущих членов популяции с подходящим значением в целевой ячейке. Поскольку этот подход основан

на значениях в целевой ячейке, а не на градиентах, наличие нескольких экстремумов не вызывает проблем. Кроме того, функции, не имеющие градиента (так называемые негладкие функции), также становятся меньшей проблемой. Метод эволюционного поиска решения (подобно методу ОПГ с несколькими начальными точками) работает эффективнее при выборе разумных верхних и нижних границ значений в изменяемых ячейках. Выбрав **Эволюционный поиск решения (Evolutionary Solver)** в раскрывающемся списке **Выберите метод решения (Select A Solving Method)** в диалоговом окне **Параметры поиска решения (Solver Parameters)**, щелкните на **Параметры (Options)**, откройте вкладку **Эволюционный поиск решения (Evolutionary)** и замените скорость изменения на 0,5. Включите флажок **Обязательные границы для переменных (Required Bounds On Variables)** и увеличьте максимальное время без улучшения до 3600 секунд. Увеличение скорости изменения снижает вероятность того, что **Поиск решения** застопорится возле неподходящего решения. Увеличение максимального времени без улучшения до 3600 секунд позволит инструменту не отказываться от поиска решения из-за отсутствия приемлемого значения в целевой ячейке в течение 3600 секунд. Таким образом, **Поиск решения** сможет продолжать работать, даже если вы отойдете от компьютера.

Теперь применим инструмент **Поиск решения (Solver)** 2019-го года к двум интереснейшим задачам размещения складов.

## Ответы на вопросы

**❓ В каком месте в США транспортная интернет-компания должна расположить единственный склад для минимизации суммарной протяженности транспортировки пакетов?**

Количество отправок (в тысячах), выполняемых каждый год в различные города, показано на рис. 36.6. (См. в файле **Warehouseloc.xlsx** лист **One warehouse**.)

Ключом к этой модели является формула приблизительного расстояния между двумя городами США, для которых указана широта и долгота: (широта1, долгота1) и (широта2, долгота2).

$$\text{Расстояние} = 69 \times \sqrt{(\text{широта1} - \text{широта2})^2 + (\text{долгота1} - \text{долгота2})^2}.$$

Сначала в ячейки **F4:G4** введите пробные значения для широты и долготы склада. Затем скопируйте формулу **=69\*КОРЕНЬ((C7-\$F\$4)^2+(D7-\$G\$4)^2)** из **F7** в **F8:F27** для вычисления приблизительного расстояния от каждого города до склада. Далее скопируйте формулу **=E7\*F7** из **G7** в **G8:G27** для вычисления расстояния транспортировки пакетов в каждый город. В ячейке **H5** формула **=СУММ(G7:G27)** вычисляет общее расстояние транспортировки всех пакетов. Цель — минимизировать значение в ячейке **H5** путем изменения значений в ячейках **F4:G4**. После выбора для решения метода ОПГ диалоговое окно **Параметры поиска решения** должно выглядеть так, как на рис. 36.7.

	B	C	D	E	F	G	H	I
3					Lat	Long		Mean
4				1	36.813439	92.48191		1125.827
5						Total	252185.2	
6	City	Lat	Long	Shipments	Distance	Shipped* Dist		
7	New York	40.7	73.9	15	1309.8969	19648.45		
8	Boston	42.3	71	8	1529.8326	12238.66		
9	Philadelphia	40	75.1	10	1219.3395	12193.39		
10	Charlotte	35.2	80.8	6	813.70342	4882.221		
11	Atlanta	33.8	84.4	11	595.15484	6546.703		
12	New Orleans	30	89.9	8	502.75019	4022.002		
13	Miami	25.8	80.2	13	1138.2724	14797.54		
14	Dallas	32.8	96.8	10	406.77005	4067.701		
15	Houston	29.8	95.4	12	524.14383	6289.726		
16	Chicago	41.8	87.7	14	476.71185	6673.966		
17	Detroit	42.4	83.1	11	753.42785	8287.706		
18	Cleveland	41.5	81.7	8	811.19304	6489.544		
19	Indy	39.8	86.1	7	486.18479	3403.294		
20	Denver	39.8	105	8	881.28021	7050.242		
21	Minneapolis	45	93.3	9	567.68613	5109.175		
22	Phoenix	33.5	112	11	1372.8197	15101.02		
23	Salt Lake City	40.8	112	10	1367.7932	13677.93		
24	LA	34.1	118	18	1798.1222	32366.2		
25	SF	37.8	123	12	2079.2628	24951.15		
26	SD	32.8	117	10	1721.0736	17210.74		
27	Seattle	41.6	122	13	2090.6012	27177.82		

Рис. 36.6. Данные к задаче с одним складом

Solver Parameters

Set Objective:

To: ☐ Max ☒ Min ☐ Value Of:

By Changing Variable Cells:

Subject to the Constraints:

☐ Make Unconstrained Variables Non-Negative

Select a Solving Method:

**Solving Method**  
 Select the GRG Nonlinear engine for Solver Problems that are smooth nonlinear. Select the LP Simplex engine for linear Solver Problems, and select the Evolutionary engine for Solver problems that are non-smooth.

Рис. 36.7. Диалоговое окно Параметры поиска решения для задачи с одним складом



Нажмите Найти решение (Solve). Оказывается, склад должен находиться в точке 36,81 градуса широты и 92,48 градуса долготы, что возле Спрингфилда, штат Миссури (см. координаты в ячейках F4:G4 на рис. 36.6).

**?** В каком месте в США транспортная интернет-компания должна расположить два склада для минимизации суммарной протяженности транспортировки пакетов?

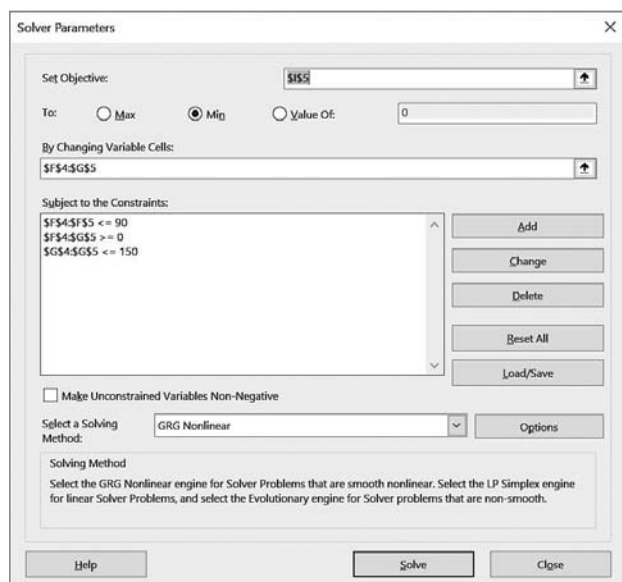
Результаты к решению этой задачи находятся в файле Warehouseloc.xlsx на листе Two warehouses (рис. 36.8).

	B	C	D	E	F	G	H	I
2								Mean dist
3					Lat	Long		501.811
4				1	38.1641	84.029		Total
5				2	34.9319	117.792		119676
6	City	Lat	Long	Shipment to	Distance to 1	Distance to 2	Min Distance	Dist*Shipped
7	New York	40.7	73.9	15	720.47	3054.56	720.47	10807.1
8	Boston	42.3	71	8	943.207	3268.4	943.207	7545.66
9	Philadelphia	40	75.1	10	628.987	2966.41	628.987	6289.87
10	Charlotte	35.2	80.8	6	302.436	2552.49	302.436	1814.61
11	Atlanta	33.8	84.4	11	302.206	2305.34	302.206	3324.27
12	New Orleans	30	89.9	8	693.856	1954.38	693.856	5550.85
13	Miami	25.8	80.2	13	893.092	2669.26	893.092	11610.2
14	Dallas	32.8	96.8	10	955.774	1455.87	955.774	9557.74
15	Houston	29.8	95.4	12	973.995	1585.08	973.995	11687.9
16	Chicago	41.8	87.7	14	356.515	2129.72	356.515	4991.2
17	Detroit	42.4	83.1	11	299.226	2448.56	299.226	3291.49
18	Cleveland	41.5	81.7	8	280.726	2531.22	280.726	2245.81
19	Indy	39.8	86.1	7	182.107	2212.37	182.107	1274.75
20	Denver	39.8	104.9	8	1444.52	950.829	950.829	7606.63
21	Minneapolis	45	93.3	9	794.796	1827.14	794.796	7153.16
22	Phoenix	33.5	112.1	11	1963.46	404.958	404.958	4454.54
23	Salt Lake City	40.8	111.9	10	1931.68	573.762	573.762	5737.62
24	LA	34.1	118.4	18	2388.12	71.1134	71.1134	1280.04
25	SF	37.8	122.6	12	2661.52	386.318	386.318	4635.82
26	SD	32.8	117.1	10	2311.72	154.647	154.647	1546.47
27	Seattle	41.6	122.4	13	2658.2	559.287	559.287	7270.74

**Рис. 36.8.** Модель для размещения двух складов

Сначала введите в ячейки F4:G5 пробные значения широты и долготы для складов. Затем скопируйте формулу  $=69*\text{КОРЕНЬ}((C7-\$F\$4)^2+(D7-\$G\$4)^2)$  из F7 в F8:F27 для вычисления расстояния от каждого города до склада 1. Расстояние от каждого города до склада 2 рассчитайте по формуле  $=69*\text{КОРЕНЬ}((C7-\$F\$5)^2+(D7-\$G\$5)^2)$ , скопировав ее из G7 в G8:G27. Поскольку транспортировка в каждый город будет осуществляться с ближайшего склада, вычислим расстояние от каждого города до ближайшего склада, скопировав формулу  $=\text{МИН}(F7;G7)$  из H7 в H8:H27. В ячейках I7:I27 вычисляется расстояние транспортировки всех отправок для каждого города по формуле  $=H7*E7$ , скопированной из I7 в I8:I27. В ячейке I5 вычисляется суммарная протяженность транспортировки всех отправок по формуле  $=\text{СУММ}(I7:I27)$ .

Теперь все готово для поиска оптимального расположения складов. Диалоговое окно **Параметры поиска решения (Solver Parameters)** показано на рис. 36.9.



**Рис. 36.9.** Поиск решения для размещения двух складов

Начнем с выбора метода ОПГ, затем используем заведомо неподходящее решение, задав для каждого склада нулевые значения широты и долготы. Это решение не подходит по двум причинам: оба склада оказываются в Африке и к тому же в одном и том же месте. Нажмите кнопку **Найти решение (Solver)**. В полученном решении оба склада рекомендуется разместить в одном месте. Безусловно, это субоптимальное решение. Проблема имеет два аспекта: функция **МИН** не имеет градиента, и возможно, что цель как функция четырех изменяемых ячеек имеет несколько максимумов и минимумов. Если цель имеет несколько экстремумов (в четырех измерениях), возможно, что неподходящие начальные данные находятся не возле абсолютно минимального значения, которое является оптимальным решением. Если предполагается, что у функции несколько экстремумов, идеально будет применить метод ОПГ с несколькими начальными точками, для каждой из которых выполняется поиск решения и выбирается наиболее удачное решение. Как правило, самое удачное из удачных решений будет оптимальным решением задачи.

Для использования нескольких начальных точек установите верхние и нижние границы для изменяемых ячеек. Для границ изменения широты я выбрал 0 и 90 градусов. Эти значения гарантируют, что склад будет находиться севернее экватора. Для границ изменения долготы я выбрал значения 0 и 150, что гарантирует размещение склада к западу от Гринвича, Великобритания, и к востоку от Анкориджа, штат Аляска.

В результате среднее расстояние транспортировки одной партии составило 502 мили. Расположение складов показано в ячейках F4:G5 на рис. 36.8. Первый склад расположен возле Лексингтона, штат Кентукки, второй — возле Ланкастера, штат Калифорния.

Для подтверждения оптимальности решения запустите эволюционный поиск решения. Вы увидите, что в оптимальное решение не будет внесено никаких изменений.

Допустим, для верхней границы долготы установлено значение 110 градусов. После запуска поиска решения окажется, что рекомендуемое значение долготы находится вблизи 110 градусов. Если для изменяемой ячейки установлены границы и при поиске решения в изменяемой ячейке оказывается значение, близкое к какой-либо границе, необходимо расширить соответствующую границу.

## Задания

1. Найдите оптимальное решение для задачи размещения трех складов.
2. Предположим, что необходимо расположить единственную комнату отдыха на минимальном возможном расстоянии от всех рабочих мест сотрудников компании. Сотрудники работают на заводе в четырех зонах:

Х	Y	Количество служащих
5	20	6
50	50	12
25	75	23
80	30	15

Пусть сотрудники по дороге в комнату отдыха и обратно всегда идут в направлении с севера на юг или с востока на запад. Где должна быть расположена комната отдыха?

3. Выполните задание 2 для двух комнат отдыха.
4. Коробка должна иметь объем 4 куб. фута. Необходимо задать высоту, длину и ширину коробки. Стоимость верхней и нижней стороны коробки составляет 30 центов за кв. фут, а остальных сторон коробки — 20 центов за кв. фут. Каковы размеры самой дешевой возможной коробки?
5. Предположим, вы толкаете ядро с высоты  $H$  с начальной скоростью  $V$  под углом  $\theta$ . Если  $H = 5$  футов, а скорость  $V = 100$  футов в секунду, какой угол обеспечит ядру полет на самое далекое расстояние? Исходите из предположения, что через заданное время  $time$  координата  $y$  ядра равна  $h + v * \sin(\theta) * time - 0,5 * g * (time^2)$ , а координата  $x$  равна  $v * time * \cos(\theta)$ .

## ГЛАВА 37

# Штрафы и эволюционный поиск решения

### Обсуждаемые вопросы

- Что является ключом к успешному использованию эволюционного поиска решения?
- Каким образом с помощью эволюционного поиска решения можно распределить 80 сотрудников финансового отдела Microsoft по четырем рабочим группам?

## Ответы на вопросы

### Что является ключом к успешному использованию эволюционного поиска решения?

Ранее я говорил, что Эволюционный поиск решения (Evolutionary Solver) предназначен для поиска решений оптимизационных задач, в которых целевая ячейка и/или изменяемые ячейки связаны с негладкими функциями: ЕСЛИ (IF), ABS, МАКС (MAX), МИН (MIN), СЧЁТЕСЛИ (COUNTIF), СЧЁТЕСЛИМН (COUNTIFS), СУММЕСЛИ (SUMIF), СУММЕСЛИМН (SUMIFS), СРЗНАЧЕСЛИ (AVERAGEIF) и СРЗНАЧЕСЛИМН (AVERAGEIFS). Перед решением задачи с помощью эволюционного поиска вы должны выполнить следующие действия в диалоговом окне **Параметры поиска решения (Solver Parameters)** (для этого сначала выберите на вкладке **Данные (Data)** в группе **Анализ (Analysis)** инструмент **Поиск решения (Solver)**).

- Нажмите кнопку **Параметры (Options)**, перейдите на вкладку **Эволюционный поиск решения (Evolutionary)** и увеличьте **Скорость изменения (Mutation Rate)** до 0,50.
- Измените значение параметра **Максимальное время без улучшения (Maximum Time Without Improvement)** на 3600 секунд.
- Установите разумные верхние и нижние границы для значений в изменяемых ячейках.

Все в жизни имеет плюсы и минусы, и **Эволюционный поиск решения** — не исключение. Преимущество эволюционного поиска решения — умение обрабатывать негладкие функции. К минусам следует отнести недостаточно четкую обработку ограничений, которые не являются линейными функциями изменяемых ячеек. При обработке большинства ограничений в эволюционном поиске решения не-

обходимо наложить на целевую ячейку штраф, который является наказанием за нарушение ограничения. Тогда правило «выживает наиболее приспособленный» заставит избегать нарушения любого ограничения. Применение штрафов при эволюционном поиске решения рассмотрено в ответе на следующий вопрос.

**?** Каким образом с помощью эволюционного поиска решения можно распределить 80 сотрудников финансового отдела Microsoft по четырем рабочим группам?

Требуется распределить 80 сотрудников по четырем рабочим группам. Руководитель каждой рабочей группы оценил квалификацию каждого сотрудника по шкале от 0 до 10 (10 означает наивысшую квалификацию). Каждый сотрудник оценил свою удовлетворенность распределением в каждую из групп (снова по шкале от 0 до 10). Например, сотруднику 1 присвоен рейтинг 9 для рабочей группы 1, и сотрудник 1 оценил рабочую группу 4 в 7 баллов.

Результаты решения этой задачи представлены в файле Assign.xlsx (рис. 37.1). В каждую группу вы хотите распределить от 18 до 22 сотрудников. Считается, что квалификация в два раза важнее удовлетворенности сотрудника. Каким образом следует распределить сотрудников по рабочим группам с учетом максимальной общей удовлетворенности (руководителей и сотрудников) и с учетом обеспечения каждой группы требуемым количеством сотрудников?

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1			Qual				Sat				576	499			
2	Division	Worker	1	2	3	4	1	2	3	4	Quality	Satisfaction			
3		4	1	9	8	6	8	1	2	6	7	8			
4		1	2	10	0	5	6	9	6	7	4	10	9		
5		3	3	5	8	10	5	1	7	7	3	10	7	Group	#assigned
6		1	4	4	0	5	2	9	1	0	3	4	9	1	19
7		2	5	9	10	4	5	9	8	8	3	10	8	2	18
8		3	6	5	2	7	3	2	8	1	5	7	1	3	22
9		1	7	8	3	1	2	1	8	2	2	8	1	4	21
10		3	8	2	2	9	2	8	3	1	6	9	1	Total pen	0
11		1	9	8	7	6	3	4	3	4	1	8	4		
12		4	10	7	0	1	8	4	1	5	4	8	4	Total	1651
13		3	11	8	1	6	6	2	0	9	3	6	9		
14		2	12	0	7	1	2	5	2	1	1	7	2		
15		1	13	9	0	5	4	3	0	7	8	9	3		
16		4	14	9	2	2	7	1	1	2	#	7	10		
17		3	15	1	3	8	4	9	8	6	8	8	6		
18		1	16	9	6	4	5	5	7	8	8	9	5		
19		1	17	8	0	5	0	5	7	2	4	8	5		
20		2	18	6	7	6	3	2	4	1	6	7	4		

Рис. 37.1. Данные для задачи распределения сотрудников

В ячейки A3:A82 я ввел пробное распределение сотрудников по рабочим группам. Приемлемым начальным решением является, например, назначение всех сотрудников в группу 1. Скопируйте формулу =ГПР(A3;Квал;B3+1) из K3 в K3:K82

для просмотра соответствия квалификации каждого сотрудника назначенной рабочей группе. Обратите внимание, что **Квал** — это имя диапазона **C2:F82**. Затем скопируйте формулу **=ГПР(А3;Удовл;В3+1)** из **L3** в **L3:L82** для просмотра удовлетворенности сотрудников назначением в группу. **Удовл** — это имя диапазона **G2:J82**.

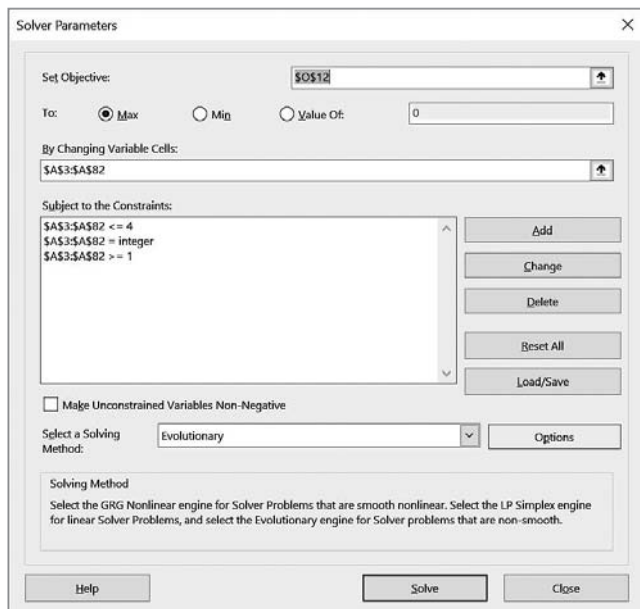
Для обеспечения требуемого количества сотрудников (от 18 до 22 в каждой группе) необходимо сосчитать число сотрудников, назначенных в каждую группу. Для этого скопируйте формулу **=СЧЁТЕСЛИ(\$A\$3:\$A\$82;M6)** из **N6** в **N7:N9**. Затем определите, достаточно ли сотрудников в группе, скопировав формулу **=ЕСЛИ(ИЛИ(N6<18;N6>22);1;0)** из **O6** в **O7:O9**.

Теперь я покажу, как вычислить целевую ячейку. В ячейках **K1:L1** вы вычисляете общую квалификацию и общее качество работы, копируя формулу **=СУММ(K3:K82)** из **K1** в **L1**. Для обеспечения требуемого количества сотрудников в каждой группе (от 18 до 22) можно наложить штраф на целевую ячейку. Я выбрал штраф 1000 баллов для каждой группы, в которой меньше 18 или больше 22 сотрудников. Четких правил для штрафа не существует. В данной ситуации средний рейтинг равен 5. В результате в целевой ячейке будет получено значение  $2 \times 400 + 400 = 1200$ . Таким образом, вполне вероятно, что неправильное число сотрудников в любой группе существенно отразится на значении в целевой ячейке, и правило «выживает наиболее приспособленный» забракует любое решение, при котором в группе слишком мало или слишком много сотрудников. Подходящий штраф не должен быть слишком большим (100 000), поскольку в таком случае Поиск решения (Solver) может проигнорировать реальный результат. Если штраф слишком мал, Поиск решения (Solver) не достигнет цели.

В ячейке **O10** я вычислил число рабочих групп с неправильным количеством сотрудников по формуле **=СУММ(O6:O9)**. Теперь мы готовы вычислить целевое значение в ячейке **O12** путем сложения удвоенных баллов за квалификацию и баллов удовлетворенности назначением и вычитания штрафа в 1000 баллов за каждую группу с неправильным количеством сотрудников. Целевая ячейка **O12** вычисляется по формуле **=2\*K1+L-1000\*O10**.

Теперь можно создать модель поиска решения задачи. Для решения необходим эволюционный поиск, поскольку функции **СЧЁТЕСЛИ** и **ЕСЛИ** в изменяемых ячейках являются негладкими. Модель представлена на рис. 37.2.

Выполняется поиск максимума для взвешенной суммы квалификации сотрудников и удовлетворенности назначением за вычетом штрафа за неправильное число сотрудников в группе (ячейка **O12**). Распределение сотрудников ограничивается значениями 1, 2, 3 или 4. Решение показано ранее на рис. 37.1. В каждую группу распределено правильное количество сотрудников; средняя квалификация сотрудника равняется 7,2 балла, а средняя удовлетворенность сотрудника — 6,3 балла. До распределения средняя квалификация всех 80 сотрудников составляла 4,4 балла, а средняя удовлетворенность — 5 баллов, так что по сравнению со случайным распределением условия работы существенно улучшились.



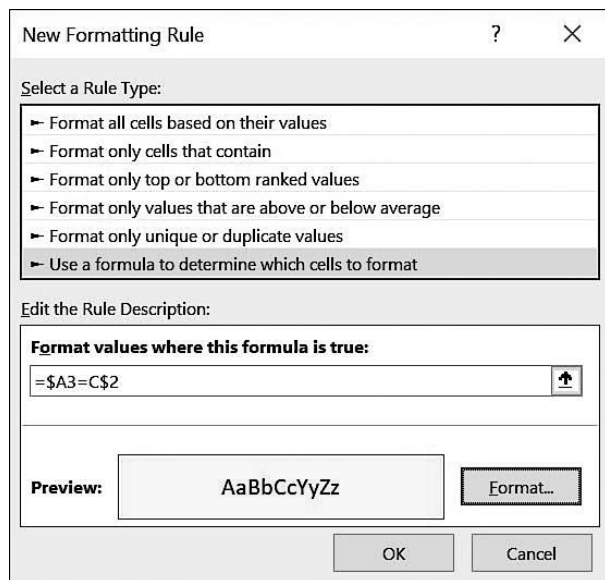
**Рис. 37.2.** Модель поиска решения для задачи распределения сотрудников

Если бы я применил метод ОПГ (даже с несколькими начальными точками), оптимальное решение не было бы найдено, поскольку модель включает негладкие функции. Еще один совет по применению эволюционного поиска решения: выбирайте как можно меньше изменяемых ячеек, и на поиск оптимального решения будет затрачено намного меньше времени.

## Условное форматирование для выделения рейтинга сотрудника

С помощью условного форматирования можно выделить желтым цветом квалификацию и удовлетворенность (на основе распределения по группам) для каждого сотрудника. Начиная с ячейки **C3** выделите диапазон ячеек **C3:J82**. На вкладке Главная (Home) из списка Условное форматирование (Conditional Formatting) выберите Создать правило (New Rule), затем Использовать формулу для определения форматлируемых ячеек (Use A Formula...) и заполните диалоговое окно, как показано на рис. 37.3.

По этой формуле форматирование желтым цветом применяется к ячейке **C3** тогда и только тогда, когда первый сотрудник распределен в группу 1. Вы можете скопировать эту формулу вправо и вниз, чтобы рейтинги квалификации и удовлетворенности каждого сотрудника выделялись только для рабочей группы, в которую он был распределен.



**Рис. 37.3.** Условное форматирование для выделения квалификации и удовлетворенности сотрудников

## Задания

1. Выполните задание 4 из главы 33 «Поиск решения для бюджетирования капиталовложений», используя эволюционный поиск решения.
2. Решите задачу с двумя складами из главы 36 «Расположение складов по методу ОПГ с несколькими начальными точками и согласно эволюционному поиску решения» при условии, что каждый склад может отгрузить не более 120 000 единиц товара.
3. В вымышленном штате Politicians Care about U.S. имеется восемь избирательных округов для выбора в конгресс. Каждый из 15 городов должен быть отнесен к одному из избирательных округов, и в каждом округе должно быть от 150 000 до 250 000 избирателей. Расстановка сил по городам приведена в таблице ниже. Распределите города по избирательным округам, обеспечив победу демократам в максимальном количестве округов.

Город	Республиканцы	Демократы
1	80	34
2	43	61
3	40	44
4	20	24



Город	Республиканцы	Демократы
5	40	114
6	40	64
7	70	34
8	50	44
9	70	54
10	70	64
11	80	45
12	40	50
13	50	60
14	60	65
15	50	70

- Решите пример с распределением сотрудников при условии, что удовлетворенность сотрудника работой в два раза важнее, чем оценки руководителей.
- Больница Cook County General составляет расписание работы для 20 медсестер. Каждая медсестра должна работать четыре дня подряд в соответствии с каким-нибудь из приведенных ниже расписаний:

Расписание	Дни работы
1	Понедельник — четверг
2	Вторник — пятница
3	Среда — суббота
4	Четверг — воскресенье
5	Пятница — понедельник
6	Суббота — вторник
7	Воскресенье — среда

Каждая медсестра в течение недели назначается на работу или в отделение интенсивной терапии, или в обычную палату. Удовлетворенность каждой медсестры своим назначением приведена в файле `Nursejackiedata.xlsx`. Например, если медсестра 5 назначена в реанимационное отделение, она дает оценку 10 расписанию № 3.

Каждый день реанимационному отделению требуется шесть медсестер, а в палаты больных — пять медсестер. Составьте расписание, максимально удовлетворяющее медсестер и соответствующее требованиям больницы.

Необходимо узнать, какие расписания удовлетворяют запросы медсестер в различные дни недели. Например, медсестры, приступающие к работе в понедельник, пятницу, субботу и воскресенье, будут работать в понедельник.

6. В файле `Problem6data.xlsx` находятся данные о тринадцати собраниях десяти человек продолжительностью в один час. В `A1` находится число людей, участвующих в каждом собрании. 6-е и 7-е собрания не могут проводиться в одно и то же время, и, конечно, никто не может быть на двух собраниях одновременно. Определите минимальное количество часов, необходимое, чтобы спланировать все собрания.

# Задача коммивояжера

Обсуждаемые вопросы:

- Как с помощью Excel решать задачи упорядочения?
- Как с помощью Excel решить задачу коммивояжера?

## Ответы на вопросы

### ❓ Как с помощью Excel решать задачи упорядочения?

Многие бизнес-задачи связаны с выбором оптимальной последовательности. Вот два примера.

- В каком порядке типография должна работать над десятью заказами для минимизации общего времени задержки сроков выполнения? Задачи такого типа называются *задачами календарного планирования*.
- Продавец живет в Бостоне и собирается посетить 10 других городов перед возвращением домой. В каком порядке он должен посещать города для максимального сокращения общего расстояния в пути? Это пример классической задачи коммивояжера.

А вот два примера задачи коммивояжера.

- Водитель курьерской службы должен сделать сегодня 20 доставок. В какой последовательности он должен доставлять пакеты, чтобы провести в пути минимум времени?
- Робот должен просверлить 10 отверстий для получения одной печатной платы. В каком порядке следует сверлить отверстия для минимизации общего времени, необходимого для производства печатной платы?

В Excel 2019 решение задач выбора последовательности с помощью инструмента Поиск решения (Solver) не вызывает затруднений. Перейдите в диалоговое окно Параметры поиска решения (Solver Parameters), выбрав инструмент Поиск решения (Solver) на вкладке Данные (Data) в группе Анализ (Analyze), и выберите Эволюционный поиск решения (Evolutionary Solver) из списка Выберите метод решения (Select a Solving Method). Затем выделите изменяемые ячейки, выберите раз (Dif) из списка сравнения, находящегося посередине диалогового окна Добавить ограничения (Add Constraint) при установке ограничений. Установка ограничений с помощью пара-

метра Все разные (All Different) гарантирует, что при наличии 10 изменяемых ячеек им автоматически будут присвоены значения 1, 2, ..., 10, и каждое значение будет встречаться только один раз. В общем, если значения в выделенном диапазоне из  $n$  изменяемых ячеек должны быть разными, Excel обеспечит присвоение каждого из значений 1, 2, ...,  $n$  только одной изменяемой ячейке. Рассмотрим, как применение параметра раз (Dif) позволяет легко решить задачу коммивояжера.

### ? Как с помощью Excel решить задачу коммивояжера?

Решите следующую задачу. Коммивояжер Уилли Лоумэн живет в Бостоне. Ему необходимо посетить каждый из городов в списке на рис. 38.1 и затем вернуться в Бостон. В какой очередности Вилли должен посещать города для максимального сокращения общего пути? Работаем с файлом Tsp.xlsx.

	E	F	G	H	I	J	K	L	M	N	O	P	Q	
3			Boston	Chicago	Dallas	Denver	LA	Miami	NY	Phoenix	Pittsburgh	SF	Seattle	
4		1	Boston	0	983	1815	1991	3036	1539	213	2664	792	2385	2612
5		2	Chicago	983	0	1205	1050	2112	1390	840	1729	457	2212	2052
6		3	Dallas	1815	1205	0	801	1425	1332	1604	1027	1237	1765	2404
7		4	Denver	1991	1050	801	0	1174	1332	1780	836	1411	1765	1373
8		5	LA	3036	2112	1425	1174	0	2757	2825	398	2456	403	1909
9		6	Miami	1539	1390	1332	1332	2757	0	1258	2359	1250	3097	3389
10		7	NY	213	840	1604	1780	2825	1258	0	2442	386	3036	2900
11		8	Phoenix	2664	1729	1027	836	398	2359	2442	0	2073	800	1482
12		9	Pittsburgh	792	457	1237	1411	2456	1250	386	2073	0	2653	2517
13		10	SF	2385	2212	1765	1765	403	3097	3036	800	2653	0	817
14		11	Seattle	2612	2052	2404	1373	1909	3389	2900	1482	2517	817	0
15		Order	Distance	City										
16		1	8	398	Phoenix									
17		2	3	1027	Dallas									
18		3	6	1332	Miami									
19		4	1	1539	Boston									
20		5	7	213	NY									
21		6	9	386	Pittsburgh									
22		7	2	457	Chicago									
23		8	4	1050	Denver									
24		9	11	1373	Seattle									
25		10	10	817	SF									
26		11	5	403	LA									
27		Total		8995										

Рис. 38.1. Данные для задачи коммивояжера

Следует отметить, что при моделировании этой задачи на листе любая последовательность или перестановка чисел от 1 до 11 представляет собой очередность посещения городов. Например, последовательность 2-4-6-8-10-1-3-5-7-9-11 можно рассматривать как путешествие из Бостона (город 1) в Даллас (город 3), в Лос-Анджелес (город 5) и, наконец, в Сан-Франциско (город 10) перед возвращением в Бостон. Поскольку последовательность рассматривается из местоположения города 1, у Уилли имеется  $10! = 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot \dots \cdot 2 \cdot 1 = 3\,628\,800$  возможных вариантов очередности посещения городов.

Сначала необходимо определить общее расстояние для любой заданной очередности посещения городов. В данной ситуации идеально подойдет функция ИНДЕКС

(INDEX). Как было показано в главе 4 «Функция ИНДЕКС», эта функция имеет синтаксис `ИНДЕКС(массив; номер_строки; номер_столбца)`. Функция `ИНДЕКС` выполняет просмотр диапазона ячеек и возвращает запись с указанным номером строки и номером столбца. В данном случае функция `ИНДЕКС` должна выполнять поиск общего расстояния при посещении всех городов.

Начнем с ввода последовательности целых чисел от 1 до 11 в диапазон `F16:F26`. Затем присвоим диапазону `G4:Q14` имя `distances` и введем формулу `=ИНДЕКС(distances;F26;F16)` в ячейку `G16`. Эта формула определяет расстояние между последним городом списка (в ячейке `F26`) и первым городом списка (в ячейке `F16`). Введем формулу `=ИНДЕКС(Distances;F16;F17)` в `G17` и скопируем ее в диапазон `G18:G26`. Эта формула вычисляет расстояние между первым и вторым городами списка, между вторым и третьим городами и т. д. Теперь можно вычислить целевое значение (общее расстояние) в ячейке `G27` по формуле `=СУММ(G16:G26)`.

Настал момент прибегнуть к эволюционному поиску решения. В диалоговом окне **Параметры поиска решения (Solver Parameters)** укажите **Минимум (Min)** для целевой функции в ячейке `G27` в диалоговом окне **Задать цель (Set Objective)**, нажмите **Добавить (Add)**, чтобы добавить ограничение, и выберите диапазон `F16:F26` в диалоговом окне **Добавить ограничение (Add Constraint)**. Выберите в среднем списке **раз (Dif)** для **Все разные (All Different)** и нажмите **ОК**, чтобы добавить ограничение. Это гарантирует, что при **Поиске решения** изменяемые ячейки останутся в пределах выбранного диапазона и им будут присвоены значения 1, 2, ..., 11. Каждое значение будет встречаться только один раз. Соответствующее диалоговое окно **Параметры поиска решения (Solver Parameters)** показано на рис. 38.2. Перед выполнением поиска увеличьте **Скорость изменения** до 0,5. Для этого нажмите кнопку **Параметры (Options)** справа от поля **Выберите метод решения (Select A Solving Method)** и затем в диалоговом окне **Параметры (Options)**, выбрав вкладку **Эволюционный поиск решения (Evolutionary)**, измените значение в поле **Скорость изменения (Mutation Rate)**.

Минимальное возможное расстояние составляет 8995 миль. Для просмотра очередности посещения городов начните со строки, в которой указано значение 1 (соответствующее родному городу Уилли — Бостону), и следуйте в указанном порядке номеров. Очередность посещения городов: Бостон—Нью-Йорк—Питтсбург—Чикаго—Денвер—Сиэтл—Сан-Франциско—Лос-Анджелес—Финикс—Даллас—Майами—Бостон. Существуют и другие последовательности посещения городов, которые дают в результате минимальное общее расстояние 8995 миль.

В заключение заметим, что иногда бывает так, что, наложив в модели **Поиска решения** на некоторые из изменяемых ячеек то ограничение, что они должны быть целыми числами, мы, однако, получаем дробные числа в результате. Чтобы этого избежать, выберите **Поиск решения (Solver)**, зайдите в **Параметры (Options)** и убедитесь, что снят флажок **Игнорировать целочисленные ограничения (Ignore Integer Constraints)**.

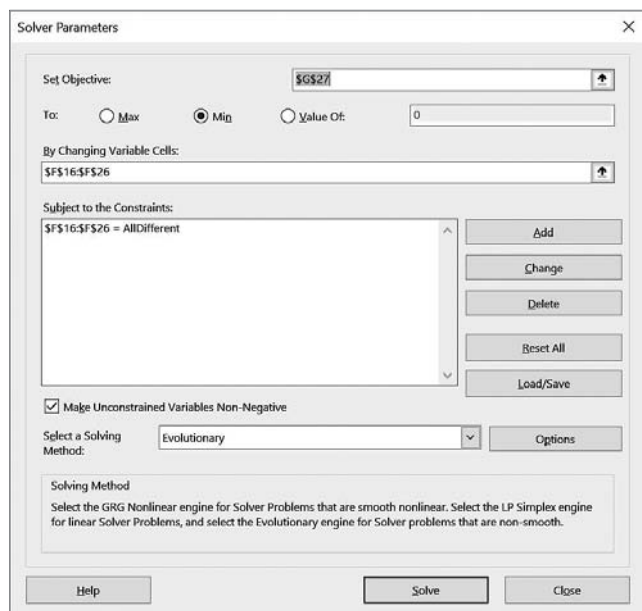


Рис. 38.2. Диалоговое окно Параметры поиска решения для задачи коммивояжера

## Задания

1. Небольшому ателье требуется запланировать выполнение шести заказов. Сроки исполнения и время, необходимое для выполнения каждого заказа, приведены ниже. В каком порядке следует запланировать выполнение заказов для сведения к минимуму общего времени задержек исполнения?

Заказ	Время выполнения	Срок исполнения в днях (с сегодняшнего дня)
1	9	32
2	7	29
3	8	22
4	18	21
5	9	37
6	6	28

2. В файле Nbamiles.xlsx указаны расстояния между всеми спортивными аренами НБА. Допустим, вы живете в Нью-Йорке и собираетесь посетить каждую аре-

ну по одному разу и вернуться в Нью-Йорк. В какой очередности вы должны посещать города для сведения к минимуму общего расстояния пути?

3. Теперь предположим, что вы живете в Атланте и берете в поездку по спортивным аренам НБА 29 генеральных управляющих. Каждый генеральный управляющий должен вернуться в свой родной город. При посещении каждой спортивной арены вы оставляете соответствующего генерального управляющего в его домашнем городе. В какой последовательности необходимо посещать арены для минимизации общего расстояния путешествия управляющих?
4. Предположим, что в задаче с Уилли Лоумэном необходимо сразу после Денвера посетить Нью-Йорк. Каким будет решение?
5. В магическом квадрате  $3 \times 3$  целые числа от 1 до 9 расположены так, что сумма чисел в каждом столбце, в каждой строке и в двух диагоналях равна 15. Используя эволюционный поиск решения, расставь те числа в магическом квадрате  $3 \times 3$ .

## ГЛАВА 39

# Импорт данных из текстового файла или документа

Обсуждаемый вопрос

- Как импортировать данные из текстового файла в Excel для анализа?

Джефф Сагарин, автор баскетбольного и футбольного рейтингов на сайте USA Today, и я разработали систему оценки игроков НБА, используемую в нескольких командах НБА, включая «Даллас Мэверикс» и «Нью-Йорк Никс». Каждый день в течение всего сезона программа Джеффа, написанная на языке программирования FORTRAN, обрабатывает и производит массу информации, в том числе рейтинги для каждого состава команды «Даллас Мэверикс» для каждой игры. Программа Джеффа выдает эти данные в виде текстового файла. В этой главе я покажу, как импортировать текстовый файл в Microsoft Excel для анализа данных.

## Ответ на вопрос

### Как импортировать данные из текстового файла в Excel для анализа?

Часто бывает необходимо импортировать данные из документа Microsoft Word или текстового файла (файла с расширением .txt) в Excel для количественного анализа. Перед импортом в Excel документ Microsoft Word необходимо сохранить как текстовый файл (.txt). Затем его можно импортировать с помощью инструмента Мастер текстов (импорт) (Text Import Wizard). Этот инструмент позволяет распределить данные из текстового файла по столбцам на основе одного из двух подходов.

- При выборе варианта фиксированной ширины (Fixed width) разбивка данных на столбцы происходит в Excel автоматически. Условия разбиения могут быть легко изменены.
- При выборе варианта с разделителями (Delimited) необходимо указать знак-разделитель (обычно это запятая, пробел или «плюс»). В этом случае данные разбиваются на столбцы в соответствии с выбранными знаками-разделителями.



Например, в файле `Lineupsch39temp.docx` в папке `Templates` этой главы содержатся данные об отрезках времени, сыгранных каждым составом команды Далласа в нескольких играх сезона 2002–2003 гг. В файле также хранятся рейтинги составов. Например, в первых двух строках указано, что против команды из Сакраменто состав, в который входили игроки Bell, Finley, LaFrentz, Nash и Nowitzki, находился на площадке 9,05 минуты и сыграл на уровне –19,79 очка (в пересчете на 48 минут) — хуже, чем в среднем по составам НБА. Образец исходного вида данных приведен в следующем абзаце:

```
Bell Finley LaFrentz Nash Nowitzki – 19.79 695# 9.05m SAC DAL* Finley Nash Nowitzki  
Van Exel Williams – 11.63 695# 8.86m SAC DAL* Finley LaFrentz Nash Nowitzki Van Exel  
102.98 695# 4.44m SAC DAL* Bradley Finley Nash Nowitzki Van Exel – 44.26 695# 4.38m  
SAC DAL* Bradley Nash Nowitzki Van Exel Williams 9.71 695# 3.05m SAC DAL* Bell Finley  
LaFrentz Nowitzki Van Exel – 121.50 695# 2.73m SAC DAL* Bell LaFrentz Nowitzki Van  
Exel Williams 39.35 695# 2.70m SAC DAL* Bradley Finley Nowitzki Van Exel Williams 86.87  
695# 2.45m SAC DAL* Bradley Nash Van Exel Williams Rigaudeau – 54.55 695# 2.32m  
SAC DAL*
```

Мы хотим импортировать данные в Excel с размещением в отдельных столбцах следующей информации:

- имя каждого игрока;
- число минут, сыгранных составом игроков;
- рейтинг состава игроков.

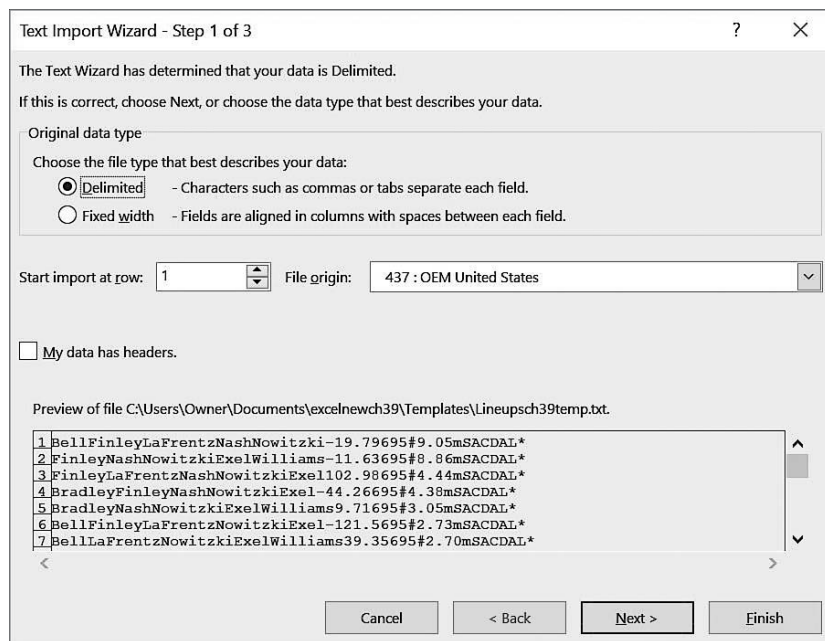
Имя игрока Van Exel (полное имя Nick Van Exel) представляет собой проблему. Если для разбивки по столбцам выбрать вариант со знаком-разделителем «пробел», то имя Van Exel займет два столбца. Числовые данные для составов команд, включающих игрока Van Exel, и для составов, не включающих игрока Van Exel, будут помещены в разные столбцы. Для решения этой проблемы замените в Word с помощью команды **Заменить** каждое вхождение текстовой строки Van Exel на Exel. Теперь при разбивке в Excel данных по столбцам с помощью знака-разделителя «пробел» для имени Excel потребуется только один столбец. Первые несколько строк данных теперь выглядят следующим образом:

```
Bell Finley LaFrentz Nash Nowitzki – 19.79 695# 9.05m SAC DAL* Finley Nash Nowitzki  
Exel Williams – 11.63 695# 8.86m SAC DAL* Finley LaFrentz Nash Nowitzki Exel 102.98  
69 5# 4.44m SAC DAL* Bradley Finley Nash Nowitzki Exel – 44.26 695# 4.38m SAC DAL*  
Bradley Nash Nowitzki Exel Williams 9.71 69 5# 3.05m SAC DAL* Bell Finley LaFrentz  
Nowitzki Exel – 121.50 695# 2.73m SAC DAL* Bell LaFrentz Nowitzki Exel Williams 39.35  
69 5# 2.70m SAC DAL* Bradley Finley Nowitzki Exel Williams 86.87 69 5# 2.45m SAC DAL*  
Bradley Nash Exel Williams Rigaudeau – 54.55 695# 2.32m SAC DAL*
```

Успешный импорт данных из файла Word или из текстового файла в Excel возможен при использовании инструмента **Мастер текстов (импорт)** (Text Import Wizard). Как я уже упоминал, сначала необходимо сохранить файл Word (`Lineupsch39temp.docx` в данном примере) как текстовый файл. Для этого откройте файл в програм-

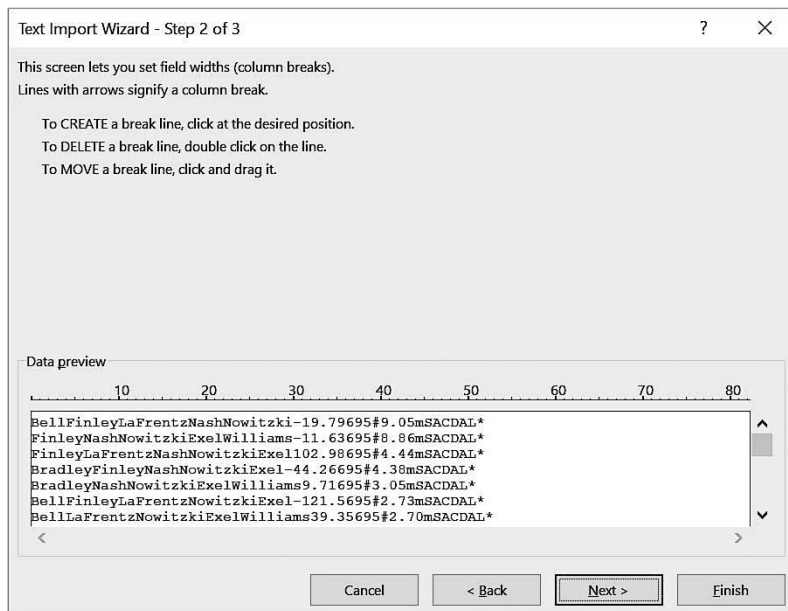
ме Microsoft Word, перейдите на вкладку **Файл (File)**, выберите **Сохранить как (Save As)** и затем в списке **Тип файла (Save As Type)** выберите **Обычный текст (Plain Text)**. В диалоговом окне **Преобразование файла (File Conversion)** установите переключатель в положение **Windows (по умолчанию) Windows (Default)** и нажмите **ОК**. Файл будет сохранен с именем `Lineupsch39temp.txt`. Закройте документ Word. В Excel откройте файл `Lineupsch39.txt` — выберите **Файл**, затем **Открыть**, затем щелкните на **Обзор**, перейдите в папку с файлом `.txt`, выберите **Все файлы (\*.\*)**. В списке типов файлов справа выберите файл, а затем щелкните **Открыть**. Вы увидите шаг 1 мастера импорта текста, показанный на рис. 39.1.

Очевидно, что в данном случае следует выбрать формат данных **с разделителями (Delimited)** и разбить данные по пробелам. Однако предположим, что выбран формат **фиксированной ширины (Fixed Width)**. Появится диалоговое окно **Мастер текстов (импорт) — шаг 2 из 3 (Text Import Wizard — Step 2 of 3)**, показанное на рис. 39.2. Как видите, точку разрыва можно вставить, удалить и переместить. Для многих операций импорта данных изменение положения точек разрыва может разделить данные невнопад.

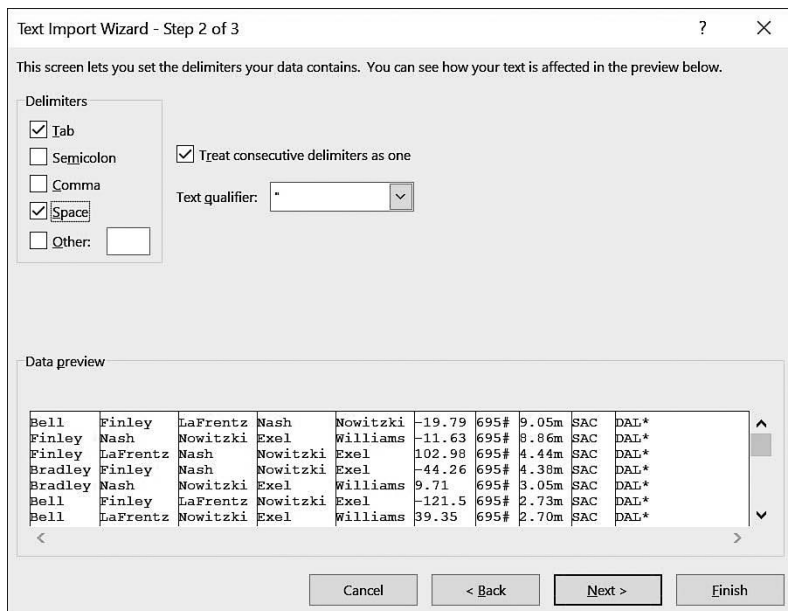


**Рис. 39.1.** Диалоговое окно Мастер текстов (импорт) — шаг 1 из 3

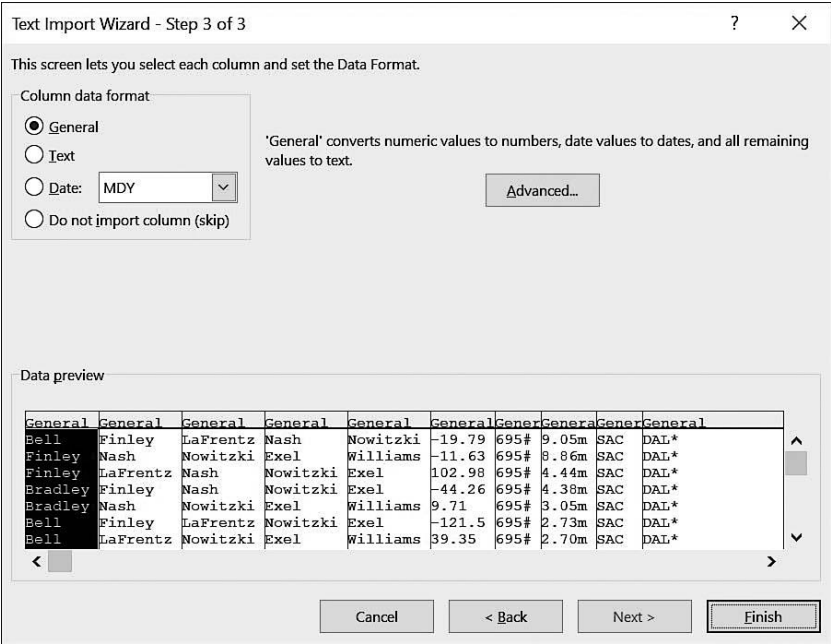
Если на шаге 1 выбрать формат данных **с разделителями (Delimited)**, появится диалоговое окно **Мастер текстов (импорт) — шаг 2 из 3 (Text Import Wizard — Step 2 of 3)**, показанное на рис. 39.3. Символ табуляции выбирается по умолчанию, и рекомендуется этот флажок не снимать, поскольку в противном случае многие надстройки



**Рис. 39.2.** Диалоговое окно Мастер текстов (импорт) — шаг 2 из 3 после выбора формата фиксированной ширины



**Рис. 39.3.** Диалоговое окно Мастер текстов (импорт) — шаг 2 из 3 после выбора формата с разделителями



**Рис. 39.4.** Диалоговое окно Мастер текстов (импорт) — шаг 3 из 3, в котором можно выбрать формат для импортируемых данных

	A	B	C	D	E	F	G	H	I	J
1	Bell	Finley	LaFrentz	Nash	Nowitzki	-19.79	695#	9.05m	SAC	DAL*
2	Finley	Nash	Nowitzki	Exel	Williams	-11.63	695#	8.86m	SAC	DAL*
3	Finley	LaFrentz	Nash	Nowitzki	Exel	102.98	695#	4.44m	SAC	DAL*
4	Bradley	Finley	Nash	Nowitzki	Exel	-44.26	695#	4.38m	SAC	DAL*
5	Bradley	Nash	Nowitzki	Exel	Williams	9.71	695#	3.05m	SAC	DAL*
6	Bell	Finley	LaFrentz	Nowitzki	Exel	-121.5	695#	2.73m	SAC	DAL*
7	Bell	LaFrentz	Nowitzki	Exel	Williams	39.35	695#	2.70m	SAC	DAL*
8	Bradley	Finley	Nowitzki	Exel	Williams	86.87	695#	2.45m	SAC	DAL*
9	Bradley	Nash	Exel	Williams	Rigaudeau	-54.55	695#	2.32m	SAC	DAL*
10	Finley	LaFrentz	Exel	Williams	Rigaudeau	-26.4	695#	1.73m	SAC	DAL*
11	Bradley	Finley	Nash	Nowitzki	Williams	91.89	695#	1.70m	SAC	DAL*
12	Bell	Finley	Nash	Nowitzki	Exel	34.18	695#	1.05m	SAC	DAL*
13	LaFrentz	Nash	Nowitzki	Exel	Williams	-50.9	695#	1.02m	SAC	DAL*
14	Bell	Bradley	Finley	Nash	Nowitzki	1.42	695#	1.00m	SAC	DAL*
15	Bradley	Finley	Exel	Williams	Rigaudeau	46.75	695#	0.93m	SAC	DAL*
16	Bell	Bradley	Nowitzki	Exel	Williams	-314.43	695#	0.60m	SAC	DAL*
17	Bell	Finley	LaFrentz	Nash	Nowitzki	123.62	686#	6.05m	UTA	DAL*
18	Finley	LaFrentz	Nash	Nowitzki	Exel	62.3	686#	5.80m	UTA	DAL*
19	LaFrentz	Nash	Nowitzki	Exel	Williams	-10.09	686#	5.68m	UTA	DAL*
20	Bell	Bradley	Finley	Nash	Nowitzki	-30.32	686#	5.60m	UTA	DAL*
21	Bell	Finley	Nowitzki	Exel	Williams	-42.93	686#	4.75m	UTA	DAL*

**Рис. 39.5.** Файл Excel с информацией по составам команд

Excel не будут работать должным образом. Установите в качестве символа-разделителя пробел (Space). Установка флажка **Считать последовательные разделители одним** (Treat consecutive delimiters as one) гарантирует, что последовательность пробелов будет интерпретироваться как одна точка разрыва.

После нажатия кнопки **Далее (Next)** появится диалоговое окно **Мастер текстов (импорт)** — шаг 3 из 3 (Text Import Wizard — Step 3 of 3), показанное на рис. 39.4. В нем можно выбрать формат для выделенного столбца. **Формат общий (General)** позволяет обрабатывать цифровые данные как числа, а остальные значения как текст.

После нажатия кнопки **Готово (Finish)** мастер импортирует данные в Excel, как показано на рис. 39.5.

Каждый игрок указан в отдельном столбце (столбцы A–E); столбец F содержит рейтинг каждого состава команд, в столбце G указан номер игры, в столбце H — число минут, проведенных на площадке каждым составом, и в столбцах I и J указаны две команды, принимавшие участие в игре. После сохранения файла как книги Excel (с расширением .xlsx) для анализа игры составов команды из Далласа можно задействовать все аналитические возможности Excel. Например, можно вычислить средние характеристики игры команды, когда Дирк Новицки (Dirk Nowitzki) находился на площадке и на скамейке запасных.

## Задания

1. В файле **Kingslineups.docx** хранятся данные об игре составов команды Sacramento Kings. Импортируйте данные в Excel.
2. В примере, рассмотренном в данной главе, время игры каждого состава (в столбце H) заканчивается символом **m**. Измените файл так, чтобы время, сыгранное каждым составом, было фактическим числом.

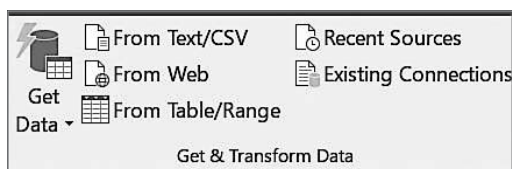
## ГЛАВА 40

# Инструмент Получить и преобразовать данные

### Обсуждаемые вопросы

- Как загрузить актуальный курс биткойна и сделать так, чтобы эти данные обновлялись каждый день?
- Как загрузить цифры текущего населения городов США?

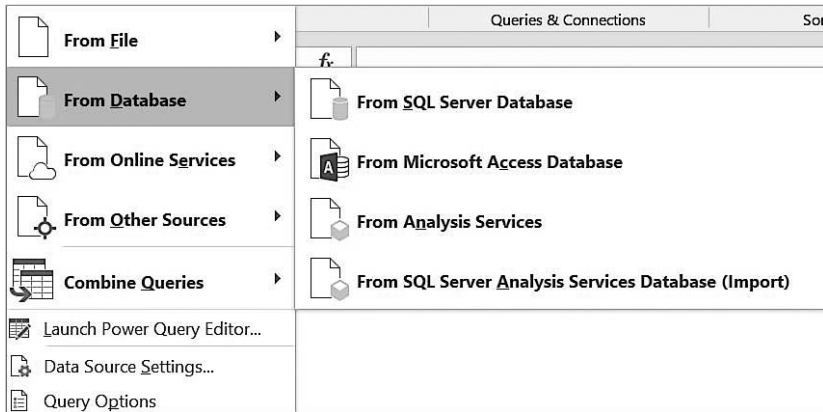
Бизнес-аналитикам нередко нужен простой способ импортировать в Excel данные из интернета, текстового файла, базы данных или другого источника. Эти данные нужно упорядочить или обработать. Наконец, импортированные данные должны обновляться, не отставать от изменений в их источнике. В этой главе мы познакомим читателя с потрясающими возможностями инструмента Excel 2019 «Получить и преобразовать данные» (Get & Transform), который позволяет аналитикам эффективно импортировать, по-новому упорядочивать и преобразовывать данные. Как видно из рис. 40.1, начиная с Excel 2016 инструмент «Получить и преобразовать данные» находится непосредственно на вкладке Данные (Data).



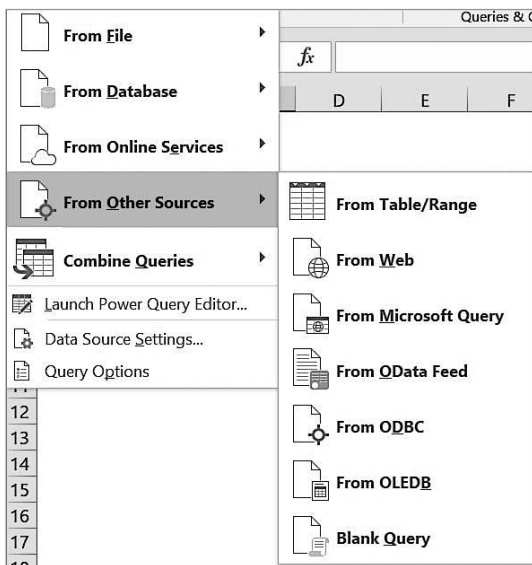
**Рис. 40.1.** Параметры инструмента Получить и преобразовать данные

Как показано на рис. 40.2 и 40.3, нажав на кнопку **Получить данные (Get Data)**, вы увидите подробный список источников данных, поддерживаемых инструментом **Получить и преобразовать данные**.

По причине ограниченности печатных площадей мы остановимся только на применении инструмента «Получить и преобразовать» к данным, добытым из интернета: как их импортировать, упорядочивать и преобразовывать.



**Рис. 40.2.** Список возможностей загрузки информации из баз данных для инструмента Получить и преобразовать данные



**Рис. 40.3.** Другие источники данных для инструмента Получить и преобразовать данные

## Ответы на вопросы

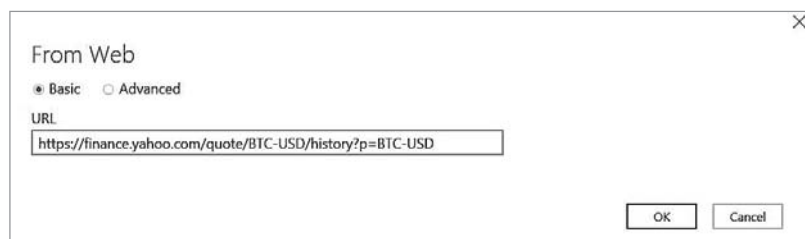
- ❓ Как загрузить актуальный курс биткойна и сделать так, чтобы эти данные обновлялись каждый день?

Многие люди заворуженно наблюдают за ежедневными вариациями в стоимости биткойна. В этой главе мы покажем, как импортировать свежие ежедневные



курсы биткойна в Excel. Наш рабочий лист будет отражать курс биткойна за последние 100 дней. Данные можно будет в любой момент обновить, чтобы отразить самые свежие.

Для начала нам нужен веб-адрес, по которому находятся ежедневные курсы биткойна. К счастью, эти сведения есть на Yahoo Finance. Нужный нам URL — <https://finance.yahoo.com/quote/BTC-USD/history/>. Если вам нужны сведения по акциям (например, «Microsoft»), просто замените текст после /quote условным биржевым обозначением (тикером). Например, цены на акции «Microsoft» можно импортировать с адреса <https://finance.yahoo.com/quote/MSFT/history/>. В пустой книге Excel нажмите **Получить данные (Get Data)** на вкладке **Данные (Data)**, в разделе **Получить и преобразовать данные (Get & Transform)** на ленте и выберите **Из других источников (Other Sources)**. Выбрав **Из Интернета (From Web)**, заполните диалоговое окно, как показано на рис. 40.4. Вот вы и создали поисковый запрос! Нажав **ОК**, вы увидите список всех таблиц, содержащихся по указанному веб-адресу (рис. 40.5). Щелкнув по таблице 2 (Table 2), вы увидите предпросмотр того, что будет импортировано. В нашем случае таблица 2 содержит необходимую информацию о курсе биткойна.



**Рис. 40.4.** Ввод веб-адреса источника данных

Теперь, если хотите, можете выбрать **Загрузить (Load)** и сразу же загрузить данные в свою книгу. Мы же, однако, вместо этого решили упорядочить импортируемые данные иным образом, поэтому выбираем **Преобразовать данные (Edit)**, что вызывает окно **Редактор Power Query (Power Query Editor)**, представленное на рис. 40.6.

Допустим, вам нужно импортировать только столбцы **Дата (Date)** и **Скорректированная цена закрытия (Adj Close)**. Тогда с помощью клавиши **Control** мы выделяем столбцы, которые хотим удалить. Щелкнув правой кнопкой мыши, выберите **Удалить столбцы (Remove Columns)**, и у вас останутся только столбцы **Дата (Date)** и **Скорректированная цена закрытия (Adj Close)**. Или, предположим, вам также надо импортировать неделю года. Для этого сначала выполните щелчок правой кнопкой мышки по столбцу **Дата (Date)** и выберите **Создать дубликат столбца (Duplicate Column)**. Выполнив на нем щелчок правой кнопкой мышки, выбираем **Переименовать (Rename)** и переименовываем дубликат столбца с датами именем **Неделя года (Week of Year)**.

Выделив столбец **Неделя года (Week Of Year)**, выберите из контекстного меню **Преобразование (Transform)**, затем **Неделя (Week)** и **Неделя года (Week Of Year)**. Как видим на рис. 40.7, теперь у нас есть столбец **Неделя года (Week Of Year)**.



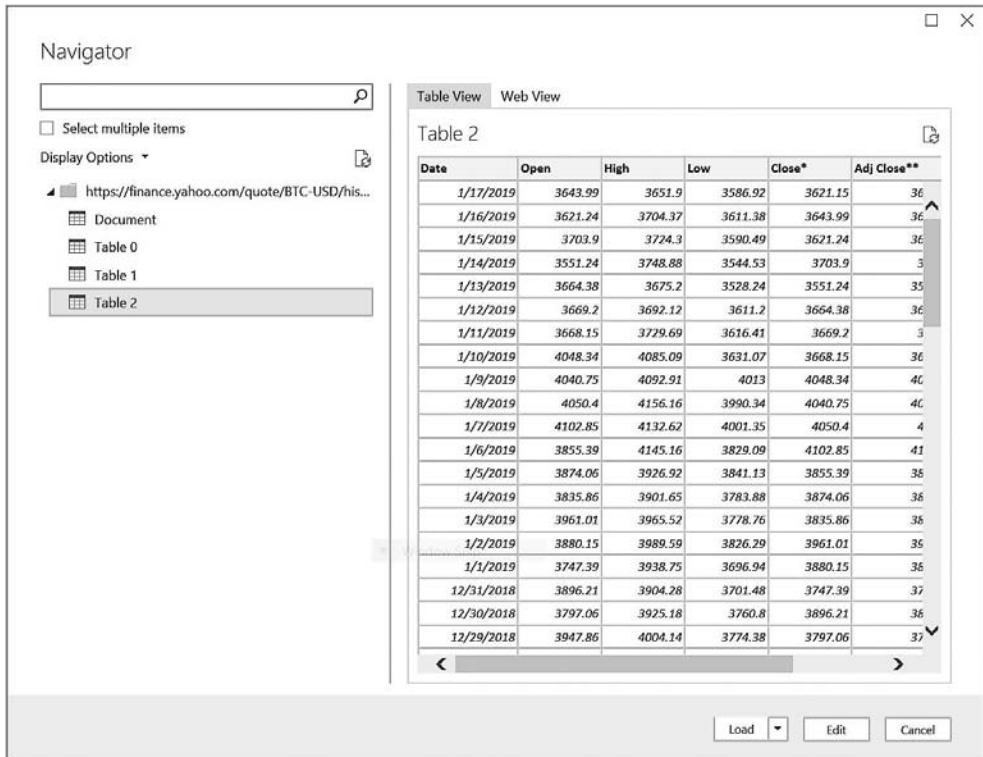


Рис. 40.5. Котировки биткойна

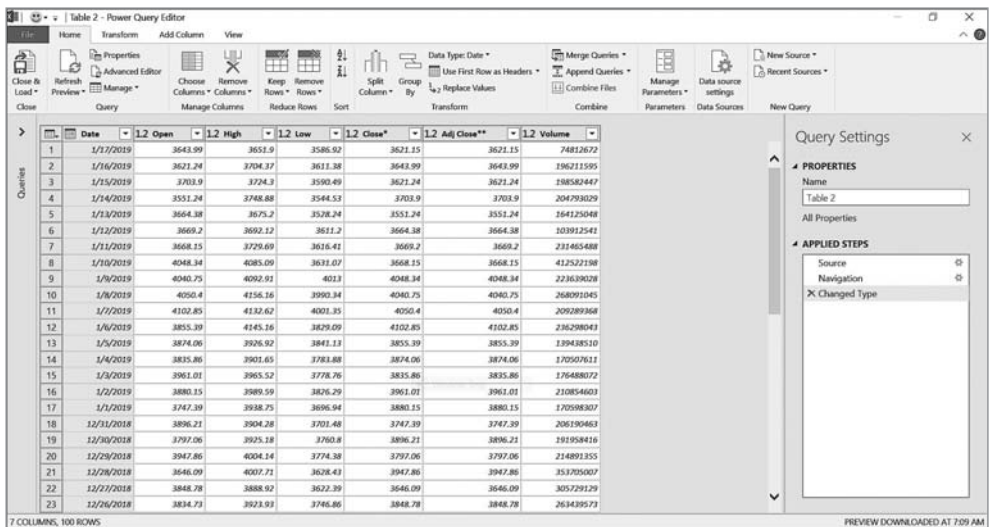


Рис. 40.6. Окно Редактор запроса (Power Query Editor)

Теперь мы готовы импортировать нужные нам данные в Excel. Просто выберите **Закрыть и загрузить (Close And Load)** на вкладке **Главная (Home)**. Вы увидите курс биткойна за 100 последние дней, как показано на рис. 40.8 и в файле **Bitcoinquery.xlsx**.

	Date	1.2 Adj Close**	1 <sup>2</sup> 3 Week of Year
1	1/17/2019	3621.15	3
2	1/16/2019	3643.99	3
3	1/15/2019	3621.24	3
4	1/14/2019	3703.9	3
5	1/13/2019	3551.24	3
6	1/12/2019	3664.38	2
7	1/11/2019	3669.2	2
8	1/10/2019	3668.15	2
9	1/9/2019	4048.34	2
10	1/8/2019	4040.75	2
11	1/7/2019	4050.4	2
12	1/6/2019	4102.85	2
13	1/5/2019	3855.39	1
14	1/4/2019	3874.06	1
15	1/3/2019	3835.86	1
16	1/2/2019	3961.01	1
17	1/1/2019	3880.15	1
18	12/31/2018	3747.39	53
19	12/30/2018	3896.21	53
20	12/29/2018	3797.06	52
21	12/28/2018	3947.86	52
22	12/27/2018	3646.09	52

**Рис. 40.7.** Создание столбца Неделя года (Week Of Year)

	A	B	C
1	Date	Adj Close**	Week of Year
2	1/17/2019	3623.85	3
3	1/16/2019	3643.99	3
4	1/15/2019	3621.24	3
5	1/14/2019	3703.9	3
6	1/13/2019	3551.24	3
7	1/12/2019	3664.38	2
8	1/11/2019	3669.2	2
9	1/10/2019	3668.15	2
10	1/9/2019	4048.34	2
11	1/8/2019	4040.75	2
12	1/7/2019	4050.4	2
13	1/6/2019	4102.85	2
14	1/5/2019	3855.39	1
15	1/4/2019	3874.06	1
16	1/3/2019	3835.86	1
17	1/2/2019	3961.01	1
18	1/1/2019	3880.15	1
19	12/31/2018	3747.39	53
20	12/30/2018	3896.21	53
21	12/29/2018	3797.06	52

**Рис. 40.8.** Курсы биткойна за последние 100 дней импортированы в столбцы A-C

Query Properties

Query name: Table 2

Description:

Usage Definition Used In

Refresh control

Last Refreshed:

☒ Enable background refresh

☒ Refresh every 60 minutes

☐ Refresh data when opening the file

☐ Remove data from the external data range before saving the workbook

☒ Refresh this connection on Refresh All

☐ Enable Fast Data Load

OLAP Server Formatting

Retrieve the following formats from the server when using this connection:

☐ Number Format ☐ Fill Color

☐ Font Style ☐ Text Color

OLAP Drill Through

Maximum number of records to retrieve: 1

Language

☐ Retrieve data and errors in the Office display language when available

OK Cancel

**Рис. 40.9.** Диалоговое окно Свойства запроса (Query Properties)

Чтобы в любой момент обновить данные, просто поместите курсор в ячейку внутри импортированных данных, выполните щелчок правой кнопкой мышки и выберите **Обновить (Refresh)**. Если вы хотите, чтобы информация обновлялась через заданные промежутки времени или при каждом открытии файла, то выберите **Обновить все (Refresh All)** на вкладке **Данные (Data)**, в группе **Запросы и подключения (Queries and Connections)**, выберите **Запросы и подключения** и затем на появившейся панели в **Запросах** щелкните по **Table 2** и из контекстного меню выберите **Свойства....** Теперь в диалоговом окне **Свойства запроса (Query Properties)** вы можете настроить параметры его обновления. Как показано на рис. 40.9, мы задали период обновления — каждые 60 минут.

### ❓ Как загрузить цифры текущего населения городов США?

Допустим, вы хотите импортировать в Excel население 100 крупнейших городов США. Необходимые сведения содержатся на сайте <http://worldpopulationreview.com/us-cities>. Применяя тот же метод, что мы использовали для загрузки курсов биткойна, мы выбрали **Таблицу 0 (Table 0)** и получили результаты, представленные на рис. 40.10.

Rank	City Name	State	2018 Population	2010 Population	2010 Census	% Change	2018 Density
1	New York	New York	8580013	8336722	8336722	0.25%	1,109/km²
2	Los Angeles	California	4030668	3976322	3782621	0.68%	1,322/km²
3	Chicago	Illinois	2687682	2704558	2695588	-0.32%	4,565/km²
4	Houston	Texas	2340814	2303482	2099451	0.80%	1,432/km²
5	Phoenix	Arizona	1679243	1615017	1445632	1.95%	1,252/km²
6	Philadelphia	Pennsylvania	1573688	1567872	1526006	0.19%	4,528/km²
7	San Antonio	Texas	1541456	1492510	1327407	1.61%	1,201/km²
8	San Diego	California	1438060	1406630	1307402	1.10%	1,710/km²
9	Dallas	Texas	1359133	1317929	1197816	1.54%	1,544/km²
10	San Jose	California	1030796	1025350	949842	0.26%	2,242/km²
11	Austin	Texas	983366	947890	790390	1.84%	1,184/km²
12	Jacksonville	Florida	907529	880619	821784	1.50%	469/km²
13	Fort Worth	Texas	893997	854113	741206	2.28%	1,001/km²
14	San Francisco	California	888653	870887	805251	1.01%	7,315/km²
15	Columbus	Ohio	880182	860090	787013	1.13%	1,555/km²
16	Charlotte	North Carolina	873363	842051	731424	1.81%	1,101/km²
17	Indianapolis	Indiana	869902	855184	820945	0.17%	902/km²
18	Seattle	Washington	746646	706312	608660	2.07%	1,436/km²
19	Denver	Colorado	731116	690960	600158	1.85%	1,811/km²
20	Washington Dc	District Of Columbia	709756	681170	601723	1.56%	4,438/km²
21	El Paso	Texas	692100	683080	649121	0.66%	1,038/km²
22	Boston	Massachusetts	687594	673184	617594	1.06%	5,491/km²
23	Nashville	Tennessee	673008	660388	602222	0.95%	546/km²

**Рис. 40.10.** Данные по населению городов США

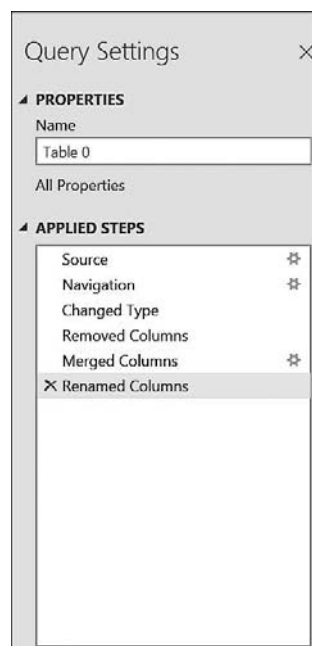
Предположим, мы хотим, чтобы город и штат находились в одном столбце, а население каждого города — в другом. Также предположим, что иные сведения нам не нужны, и мы не хотим их импортировать. Чтобы добиться этого, нажимаем **Преобразование (Transform)**, удалим с помощью клавиши **Control** последние четыре столбца. Затем выбираем столбцы **Город (City)** и **Штат (State)**. Выбираем **Объединить столбцы (Merge Columns)** из контекстного меню и выбираем запятую в качестве символа-разделителя между городом и штатом. Переименование получившегося столбца предложено выполнить здесь же — называем его **Город и штат**

(City and state). Теперь из меню Файл (File) мы сможем загрузить требующиеся нам данные и поисковый запрос в файл `USCityquery.xlsx`. Конечный результат представлен на рис. 40.11.

Если вы хотите увидеть, из каких шагов состояло выполнение нашего интернет-запроса, поместите курсор в ячейку внутри импортированных данных и выберите Запрос (Query) в правой части меню ленты. Щелкнув Редактировать (Edit) в правой части экрана, вы увидите шаги, которые потребовались для реализации вашего запроса (рис. 40.12). Панель Параметры запроса также, как правило, отображается сама — где мы видим Примененные шаги. Разумеется, выбрав Заккрыть и загрузить (Close And Load), вы вернетесь к книге Excel.

	A	B	C
1	Rank	City and State	2018 Population
2	1	New York, New York	8580015
3	2	Los Angeles, California	4030668
4	3	Chicago, Illinois	2687682
5	4	Houston, Texas	2340814
6	5	Phoenix, Arizona	1679243
7	6	Philadelphia, Pennsylvania	1573688
8	7	San Antonio, Texas	1541456
9	8	San Diego, California	1438060
10	9	Dallas, Texas	1359133
11	10	San Jose, California	1030796
12	11	Austin, Texas	983366
13	12	Jacksonville, Florida	907529
14	13	San Francisco, California	888653
15	14	Fort Worth, Texas	893997
16	15	Columbus, Ohio	880182
17	16	Indianapolis, Indiana	860902
18	17	Charlotte, North Carolina	873363
19	18	Seattle, Washington	746046
20	19	Denver, Colorado	719116
21	20	Washington Dc, District Of Columbia	702756
22	21	El Paso, Texas	692100
23	22	Boston, Massachusetts	687584
24	23	Detroit, Michigan	665713

**Рис. 40.11.** Информация по крупнейшим городам США



**Рис. 40.12.** Шаги, которые потребовались для реализации запроса

## Задания

1. Задайте запрос для того, чтобы загрузить свежие ежедневные цены на Facebook и ежедневные объемы продаж.
2. С помощью информации с веб-сайта [http://www.baruch.cuny.edu/nycdata/world\\_cities/largest\\_cities-world.htm](http://www.baruch.cuny.edu/nycdata/world_cities/largest_cities-world.htm) загрузите сведения о размере населения крупнейших 100 городов. Сделайте так, чтобы страна и город находились в одном столбце.

## ГЛАВА 41

# Типы данных «Акции» и «География»

### Обсуждаемые вопросы

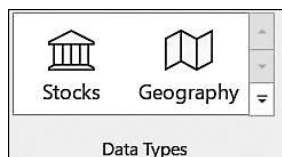
- Какие географические сведения можно использовать в формулах Excel?
- Какие показатели деятельности корпораций можно использовать в формулах Excel?

Новые типы данных, используемые в Office 365, означают, что в ячейках, содержащих географическое название или название корпорации, теперь содержится много различных актуальных географических данных об этом месте или экономических сведений о корпорации. Если у вас стоит Office 365, вы теперь можете писать формулы, возвращающие различные географические сведения. Например, если ячейка E4 содержит значение Париж (Paris), то формула E4.Population (Население), запросив «Википедию», возвратит величину текущего населения Парижа. Вы также можете писать формулы, которые возвратят показатели деятельности корпораций. Например, если в ячейке E8 содержится Microsoft, то формула E8.Price (Цена) возвратит текущую цену акций Microsoft.

## Ответы на вопросы

### ❓ Какие географические сведения можно использовать в формулах Excel?

Чтобы проиллюстрировать использование нового типа данных — географического, — мы ввели названия некоторых штатов в ячейки E3:E7 на листе States в файле Newdatatypestemp.xlsx. Выбрав ячейки E3:E7 и присвоив им тип География (Geography) на вкладке Данные (Data) (рис. 41.1), вы увидите красивые «визитки» для каждого местоположения, представленные на рис. 41.2. Щелчком мыши на визитке местоположения вы откроете информацию, содержащуюся в ячейке. Часть данных с визитки штата Индиана показана на рис. 41.2.



**Рис. 41.1.** Значки типов данных «География» и «Акции»





	E	F	G	H	I	J
1						
2	<b>States</b>	<b>Population</b>	<b>Area</b>			
3	🇺🇸Indiana	6,666,818	94,321	=E3.Population	=FIELDVALUE(E3,\$G\$2	
4	🇺🇸California	39,536,653	423,970	=E4.Population	=FIELDVALUE(E4,\$G\$2	
5	🇺🇸Texas	28,304,596	696,241	=E5.Population	=FIELDVALUE(E5,\$G\$2	
6	🇺🇸North Carolina	10,273,419	139,390	=E6.Population	=FIELDVALUE(E6,\$G\$2	
7	🇺🇸Idaho	1,716,943	216,632	=E7.Population	=FIELDVALUE(E7,\$G\$2	

**Рис. 41.4.** Население штатов

внешнюю ячейку. Например, как показано на рис. 41.4, скопировав из G3 в G4:G7 формулу =FIELDVALUE(E3;\$G\$2), мы получим Площадь (Area) пяти перечисленных штатов. Заметьте, что, если вы измените G2, скажем, на Столица (Capital), столбец G будет перечислять столицы штатов.

На рис. 41.5 показано, как, применяя новые типы данных, мы можем вернуть Штат и графство (State и County) для города в США, а на рис. 41.6 — как вернуть Телефонный код (Calling Code) и Среднюю продолжительность жизни (Life Expectancy) в выбранных странах.

	F	G	H	I	J
1					
2					
3	<b>Cities</b>	<b>State</b>	<b>County</b>		
4	🇺🇸Los Angeles	California	Los Angeles County	=F4.[Admin Division 1 (State/province/other)]	=F4.[Admin Division 2 (County/district/other)]
5	🇺🇸Houston	Texas	Harris County	=F5.[Admin Division 1 (State/province/other)]	=F5.[Admin Division 2 (County/district/other)]
6	🇺🇸Indianapolis	Indiana	Marion County	=F6.[Admin Division 1 (State/province/other)]	=F6.[Admin Division 2 (County/district/other)]
7	🇺🇸Charlotte	North Carolir	Mecklenburg County	=F7.[Admin Division 1 (State/province/other)]	=F7.[Admin Division 2 (County/district/other)]
8	🇺🇸Chicago	Illinois	Cook County	=F8.[Admin Division 1 (State/province/other)]	=F8.[Admin Division 2 (County/district/other)]

**Рис. 41.5.** Штат и графство для городов США

	E	F	G	H	I
3	<b>Country</b>	<b>Calling Code</b>	<b>Life Expectancy</b>		
4	🇯🇵Japan	81	84.0	=E4.[Calling code]	=E4.[Life expectancy]
5	🇨🇳China	86	76.3	=E5.[Calling code]	=E5.[Life expectancy]
6	🇫🇷France	33	82.3	=E6.[Calling code]	=E6.[Life expectancy]
7	🇳🇮Nigeria	234	53.4	=E7.[Calling code]	=E7.[Life expectancy]
8	🇦🇺Australia	61	82.5	=E8.[Calling code]	=E8.[Life expectancy]
9	🇧🇷Brazil	55	75.5	=E9.[Calling code]	=E9.[Life expectancy]
10	🇲🇽Mexico	52	77.1	=E10.[Calling code]	=E10.[Life expectancy]

**Рис. 41.6.** Телефонные коды и средняя продолжительность жизни по странам

## Замечания

Если Excel не может определить, какое местоположение вы имеете в виду, то вы увидите знак вопроса. Щелкнув на нем, вы сможете выбрать конкретное местоположение. Например, как мы видим в файле `Questions.xlsx` и на рис. 41.7, введя четыре упоминания Блумингтона, мы выбрали для них, соответственно, штаты Индиану, Миннесоту, Иллинойс и Калифорнию. Только уже затем мы извлекли штат и население для каждого города.

Как показано на листе `Tables.xlsx` в файле `Newdatatypes.xlsx` и на рис. 41.8, если вы делаете таблицу Excel с географическими пунктами и затем даете столбцу заголовок типа **Население (Population)**, то сразу же получаете цифры по населению в этих местах; при добавлении населенных пунктов вы увидите и их население.

	E	F	G
1			
2			
3		<b>State</b>	<b>Population</b>
4	Bloomington	Indiana	84,465
5	Bloomington	Minnesota	85,319
6	Bloomington	Illinois	78,005
7	Bloomington	California	23,879

**Рис. 41.7.** Штат и население для городов под названием Блумингтон

	F	G
1		
2	<b>State</b>	<b>Population</b>
3	California	39,536,653
4	Wyoming	579,315
5	Maine	1,335,907
6	Massachusetts	6,859,819
7	South Carolina	5,024,369
8	Ohio	11,658,609
9	Illinois	12,802,023
10	New Jersey	9,005,644

**Рис. 41.8.** Использование географического типа данных в таблице

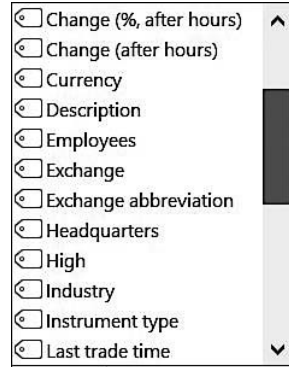
**?** Какие показатели деятельности корпораций можно использовать в формулах Excel?

Можно задать формулы для извлечения ключевых фактов о выбранной корпорации, если ввести ее название или условное биржевое обозначение. Выберите ячей-



ку, и на вкладке **Данные (Data)** выберите тип данных **Акции (Stocks)**. Например, на листе **Stock.xlsx** мы ввели названия корпораций вперемешку с биржевыми тикерами. Выделив диапазон **F4:F8**, выбираем на вкладке **Данные (Data)** тип **Акции (Stocks)**. Мы увидим визитку каждой корпорации. Щелкнув на ней, открываем сведения, имеющиеся по каждой корпорации. Если в ячейке **G4** напечатать **=F4**, то мы увидим перечень многочисленных сведений, доступных о Microsoft (некоторые показаны на рис. 41.9).

На рис. 41.10 представлены цены закрытия торгов по акциям различных корпораций на 12 ноября 2018 г. Открыв файл сами, вы, конечно, увидите более свежие данные.



**Рис. 41.9.** Перечень сведений, имеющихся о корпорации

	F	G	H
1			
2			
3	<b>Company</b>	<b>Price</b>	<b>Last Trade Time</b>
4	<b>Microsoft Corp</b>	<b>\$ 106.87</b>	<b>11/12/2018 22:29</b>
5	<b>Southwest Airlines Co</b>	<b>\$ 51.89</b>	<b>11/12/2018 22:16</b>
6	<b>Cisco Systems Inc</b>	<b>\$ 45.62</b>	<b>11/12/2018 22:24</b>
7	<b>General Motors Co</b>	<b>\$ 35.69</b>	<b>11/12/2018 22:21</b>
8	<b>Facebook Inc</b>	<b>\$ 141.55</b>	<b>11/12/2018 22:28</b>

**Рис. 41.10.** Биржевые цены на 12 ноября 2018 г.

## Задания

1. Загрузите список штатов США при помощи **Power Query** и, применив новые типы данных, извлеките средний доход и губернатора каждого штата.
2. С помощью **Power Query** загрузите список всех стран и извлеките из него численность их населения и площадь.
3. Извлеките коэффициент **Цена/Доходность** и количество сотрудников для акций фирм, перечисленных на листе **Stock.xlsx**. Сделайте так, чтобы эти данные автоматически извлекались при добавлении в список новых компаний.

## ГЛАВА 42

# Проверка достоверности данных

### Обсуждаемые вопросы

- Я ввожу результаты профессиональных баскетбольных матчей в Excel. Я знаю, что команда набирает от 50 до 200 очков за игру. Один раз я ввел 1000 очков вместо 100, что привело к неправильным результатам анализа. Существует ли в Excel способ предотвращения ошибок такого типа?
- Я ввожу даты и суммы деловых расходов на следующий год. В начале года я часто по ошибке ввожу в поле даты предыдущий год. Могут ли настройки Excel предотвратить ошибки такого типа?
- Я ввожу длинный список чисел. Могу ли я получить в Excel предупреждение, если введу нечисловое значение?
- Мой помощник при вводе многочисленных торговых сделок должен вводить сокращенные названия штатов. Можно ли создать список сокращенных названий штатов для сведения к минимуму ошибок ввода неправильных сокращений?

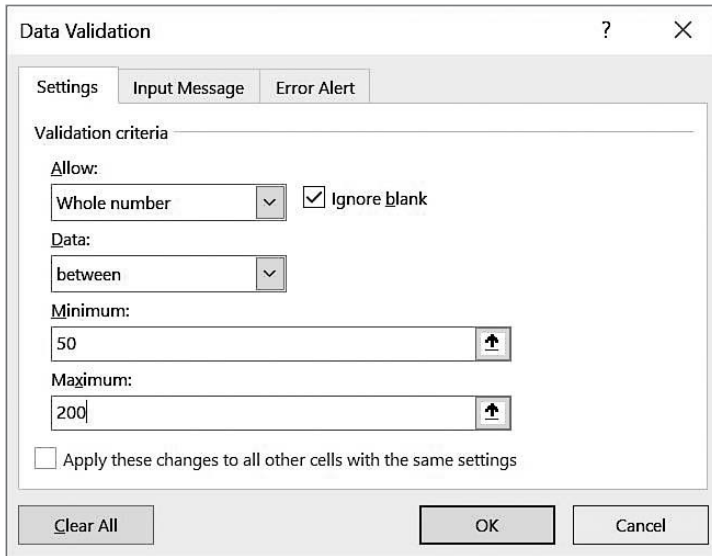
Рабочий процесс часто предполагает ввод большого количества однотипных данных. При вводе большого количества информации в Excel легко допустить ошибку. В Microsoft Excel 2019 использование критериев проверки достоверности данных значительно снижает вероятность совершения дорогостоящей ошибки. Для настройки проверки данных сначала выделите диапазон ячеек, который хотите проверить. На вкладке **Данные (Data)** выберите **Проверка данных (Data Validation)** и укажите критерии (см. примеры в этой главе), в соответствии с которыми Excel отметит любые неправильно введенные данные.

## Ответы на вопросы

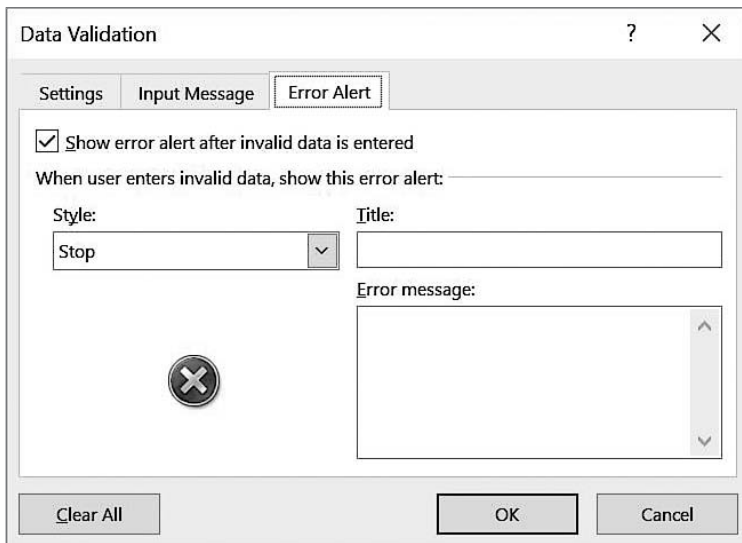
**?** Я ввожу результаты профессиональных баскетбольных матчей в Excel. Я знаю, что команда набирает от 50 до 200 очков за игру. Один раз я ввел 1000 очков вместо 100, что привело к неправильным результатам анализа. Существует ли в Excel способ предотвращения ошибок такого типа?

Пусть очки, набранные командой хозяев, вводятся в ячейки A2:A11, а очки команды гостей — в ячейки B2:B11. (Результаты решения этой задачи см. в файле Nbadvl.xlsx.) Каждое значение, введенное в диапазон A2:B11, должно быть целым числом от 50 до 200.

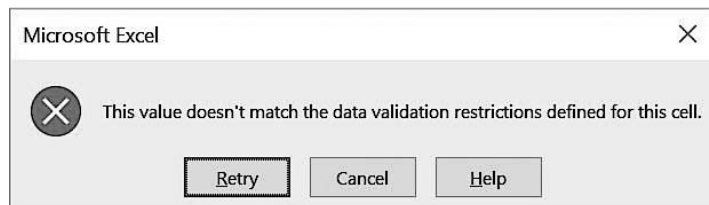
Сначала выделите диапазон A2:B11, затем на вкладке Данные (Data) выберите Проверка данных (Data Validation). В диалоговом окне Проверка вводимых значений (Data Validation) откройте вкладку Параметры (Settings) и в списке Тип данных (Allow) выберите Целое число (Whole Number), а затем введите данные, как показано на рис. 42.1.



**Рис. 42.1.** Вкладка Параметры в диалоговом окне Проверка вводимых значений для установки критериев проверки данных



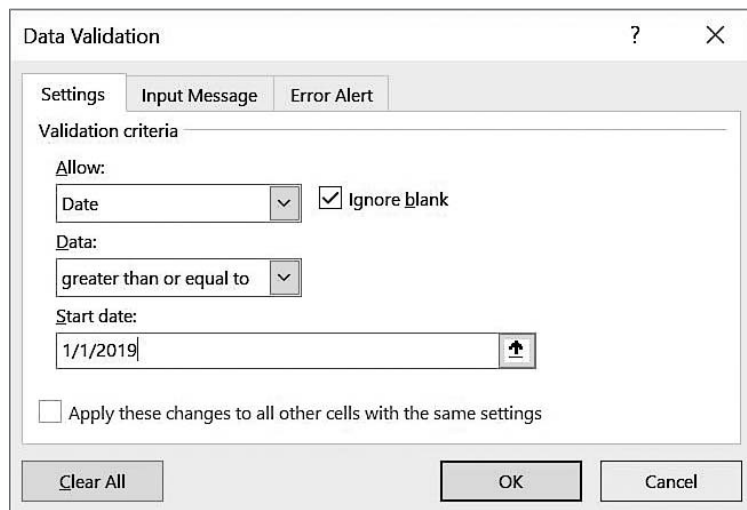
**Рис. 42.2.** Вкладка Сообщение об ошибке в диалоговом окне Проверка вводимых значений



**Рис. 42.3.** Сообщение об ошибке для примера проверки данных матчей

Ответ по умолчанию на ввод неправильных данных в Excel (так называемое *сообщение об ошибке*) — это сообщение «Введенное значение неверно. Набор значений, которые могут быть введены в ячейку, ограничен» (The value entered is not valid. A user has restricted values that can be entered into the cell). В диалоговом окне Проверка вводимых значений (Data Validation) на вкладке Сообщение об ошибке (Error Alert), показанной на рис. 42.2, можно изменить сообщение, включая значок и заголовок окна сообщения, а также сам текст сообщения. На вкладке Сообщение для ввода (Input Message) вы можете создать подсказку для пользователя о требуемом типе вводимых данных. Сообщение появляется в виде всплывающей подсказки в выбранной ячейке. Например, можно ввести сообщение об ошибке Введите целое число от 50 до 200 (Please enter a whole number between 50 and 200). После ввода в ячейку E5 числа (например, числа 34), не отвечающего критерию, появится сообщение, показанное на рис. 42.3.

**?** Я ввожу даты и суммы деловых расходов на следующий год. В начале года я часто по ошибке ввожу в поле даты предыдущий год. Могут ли настройки Excel предотвратить ошибки такого типа?



**Рис. 42.4.** Вкладка Параметры для проверки вводимых дат

Допустим, сейчас начало 2019 г., и вы вводите даты в ячейки A2:A20. (См. файл Datedv.xlsx.) Выделите диапазон A2:A20 и затем на вкладке Данные (Data) выберите инструмент Проверка данных (Data Validation) из группы Работа с данными (Data tools). В диалоговом окне Проверка вводимых значений (Data Validation) откройте вкладку Параметры (Settings) и заполните поля, как показано на рис. 42.4. Я выбрал Дата (Date) в поле Тип данных (Allow), выбрал из списка настройку Больше или равно (Grater Than Or Equal To), 1 января 2019 г. как начальную дату и нажал ОК. После этого ввести какую-либо дату ранее 1 января 2019 г. будет невозможно.

**❓ Я ввожу длинный список чисел. Могу ли я получить в Excel предупреждение, если введу нечисловое значение?**

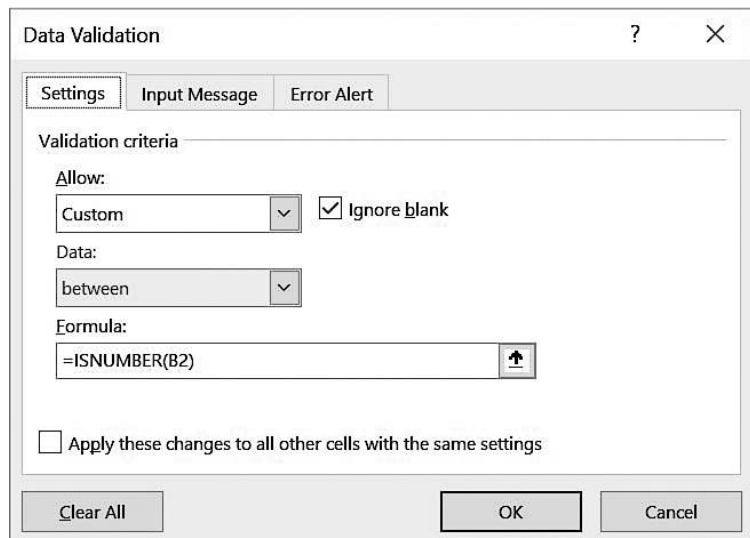
Полностью раскрыть возможности проверки данных позволяет пользовательский формат. Для этого в диалоговом окне Проверка вводимых значений (Data Validation) откройте вкладку Параметры (Settings), в списке Тип данных (Allow) выберите Другой (Custom) и введите формулу для определения достоверных данных (рис. 42.5). Эта формула работает так же, как формула условного форматирования, о котором шла речь в главе 24 «Условное форматирование». Введенная формула возвращает значение ИСТИНА тогда и только тогда, когда содержимое первой ячейки выделенного диапазона достоверно. После нажатия ОК в диалоговом окне Проверка вводимых значений (Data Validation) формула скопируется в остальные ячейки диапазона. Если после ввода значения в ячейку выделенного диапазона формула возвратит значение ЛОЖЬ, появится сообщение об ошибке.

Для иллюстрации применения пользовательского формата в ячейки B2:B20 необходимо вводить только числа. (См. файл Numberdv.xlsx.) Ключом к решению этой задачи является функция Excel=ЕЧИСЛО (ISNUMBER). Функция ЕЧИСЛО возвращает значение ИСТИНА, если функция ссылается на ячейку, содержащую числовые данные. Функция возвращает значение ЛОЖЬ, если ссылается на ячейку, содержащую нечисловое значение.

Установите курсор в ячейку B2 и выделите диапазон B2:B20. Откройте вкладку Данные (Data), в группе Работа с данными (Data Tools) выберите Проверка данных (Data Validation) и затем заполните диалоговое окно Проверка вводимых значений (Data Validation) на вкладке Параметры (Settings), как показано на рис. 42.5: я выбрал Другой (Custom) в поле Тип данных (Allow) и ввел в поле Формула (Formula) =ЕЧИСЛО(B2).

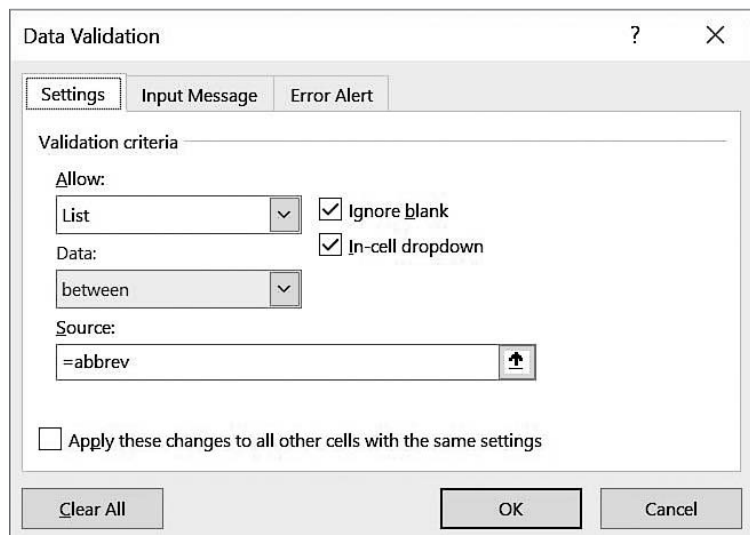
Теперь при попытке ввода нечисловых данных в диапазон B2:B20 будет появляться предупреждение. Например, если ввести в ячейку B3 текст John, появится сообщение об ошибке.

Если при работе в ячейке B3 выбрать инструмент Проверка данных (Data Validation), формула, показанная на рис. 42.5, примет вид=ЕЧИСЛО(B3). Это подтверждает, что введенная в ячейку B2 формула копируется правильно. При вводе значения John в ячейку B3 формула =ЕЧИСЛО(B3) возвратит значение ЛОЖЬ, поэтому появится сообщение об ошибке.



**Рис. 42.5.** Функция ЕЧИСЛО для проверки типа данных в диапазоне

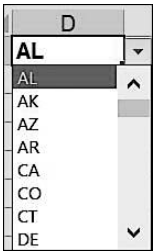
**?** Мой помощник при вводе многочисленных торговых сделок должен вводить сокращенные названия штатов. Можно ли создать список сокращенных названий штатов для сведения к минимуму ошибок ввода неправильных сокращений?



**Рис. 42.6.** Диалоговое окно Проверка вводимых значений для определения списка допустимых значений

Ключ к решению этой задачи — тип данных **Список (List)** в поле **Тип данных (Allow)** в области **Условия проверки (Validation criteria)**. Сначала введите список сокращенных названий штатов (см. файл **Statedv.xlsx**). В данном примере я присвоил диапазону **I6:I55** имя **abbrev**. Затем выделите диапазон, в который будете вводить аббревиатуры. В примере это диапазон **D5:D156**. Откройте диалоговое окно **Проверка вводимых значений (Data Validation)**, выбрав на вкладке **Данные (Data)** из списка инструмент **Проверка данных (Data Validation)**, и заполните его, как показано на рис. 42.6: я выбрал **Список (List)** в поле **Тип данных (Allow)**, ввел **=abbrev** в поле **Источник (Source)** и нажал **ОК**.

Теперь в любой ячейке диапазона **D5:D156** можно нажать стрелку раскрывающегося списка и выбрать сокращенное название штата (рис. 42.7). В этом диапазоне допустимы только сокращенные названия, указанные в списке. Если вы не будете выбирать из списка, а наберете с клавиатуры аббревиатуру с ошибкой, например, **ALK** для Аляски, появится сообщение об ошибке.



**Рис. 42.7.** Раскрывающийся список с аббревиатурами

## Замечания

- Если нажать **F5**, а затем в диалоговом окне **Переход (Go To)** щелкнуть на **Выделить (Special)** и установить переключатель в положение **Проверка данных (Data Validation)**, на листе автоматически будут выделены все ячейки, для которых выполняется проверка достоверности данных. Также после перехода по кнопке **Выделить (Special)** можно выделить все ячейки, содержащие условное форматирование.
- В версиях Excel до Microsoft Excel 2010 для работы раскрывающегося списка проверки данных на основе списка с исходными данными, расположенными на другом листе, следовало присваивать списку имя (как это было сделано в показанном примере). В версиях Microsoft Excel 2010 и выше это ограничение снято.
- При использовании динамического диапазона, описанного в главе 22 «Функция СМЕЩ», изменения в списке с исходными данными (добавление или удаление элементов) отображаются в раскрывающемся списке автоматиче-

ски. (См. задание 10 в данной главе.) Кроме того, если вы оформляете список с исходными данными как таблицу Excel (см. главу 26 «Таблицы»), изменения в исходных данных отображаются в раскрывающемся списке до тех пор, пока пользователь ссылается на диапазон списка и не пытается ввести имя таблицы.

- Скажем, вам нужно создать раскрывающийся список для выбора магазина, в который вы поставляете шоколадные плитки. Кроме того, вам необходим еще один список, из которого можно выбрать все плитки, продающиеся в конкретном магазине. Проблема в том, что для разных магазинов списки могут не совпадать. Как организовать выбор таких вложенных списков (часто называемый каскадным выбором)? Пусть поставки осуществляются в магазины Target и CVS; предположим, что имя Target присвоено диапазону со списком шоколадных плиток, продаваемых в магазине Target, а имя CVS — диапазону со списком шоколадных плиток, продаваемых в CVS. Если раскрывающийся список для выбора магазина находится, например, в ячейке A20, можно создать соответствующий раскрывающийся список в ячейке B20, введя в диалоговом окне Проверка вводимых значений (Data Validation) формулу =ДВССЫЛ(A20) для выбранного списка. Как мы обсудили в главе 23 «Функция ДВССЫЛ», если ячейка A20 содержит CVS, список будет включать диапазон CVS со всеми шоколадными плитками, продаваемыми в CVS, и т. д. См. задание 11.
- Для отказа от проверки данных в диапазоне выделите этот диапазон, откройте диалоговое окно Проверка вводимых значений (Data Validation) на вкладке Данные (Data) и нажмите Очистить все (Clear All).
- При наличии длинного списка можно воспользоваться функцией автозавершения (AutoComplete) в Excel. Если начать диапазон для раскрывающегося списка в ячейке непосредственно под списком (пустые ячейки не допускаются), функция автозавершения выполнит всю работу. См. файл Fruitlist.xlsx.
- Кроме того, проверка данных может использоваться для установки критериев на основе длины текста в ячейке (см. задание 4) или времени дня (см. задание 15).

## Задания

1. Вы вводите в диапазон C1:C20 целые неотрицательные числа. Установите параметры проверки данных, обеспечивающие ввод только целых неотрицательных чисел.
2. Вы вводите в диапазон C1:C15 даты сделок за июль 2004 г. Установите параметры проверки данных, обеспечивающие ввод только дат, приходящихся на июль 2004 г.
3. С помощью типа данных Список (List) в области Условия проверки (Validation criteria) можно создать сообщение об ошибке, если значение, введенное в про-



веряемый диапазон ячеек, отсутствует в списке. Допустим, вы вводите имена сотрудников в диапазон **A1:A10**. Сотрудниками компании являются Jen, Greg, Vivian, Jon и John. Обеспечьте отсутствие орфографических ошибок в именах с помощью типа данных **Список (List)**.

4. С помощью типа данных **Длина текста (Text Length)** в области **Условия проверки (Validation criteria)**, можно создать сообщение об ошибке, если число символов в ячейке не соответствует определенному числу. Обеспечьте с помощью типа данных **Длина текста (Text Length)** в ячейках **C1:C10** ввод значений, содержащих не более пяти символов (включая пробелы).
5. Вы вводите имена сотрудников в диапазон **A1:A10**. Настройте проверку данных так, чтобы имя любого сотрудника могло быть введено не более двух раз. (Подсказка: используйте тип данных **Другой (Custom)** и функцию **СЧЁТЕСЛИ**.)
6. Вы вводите идентификаторы продуктов в диапазон **A1:A15**. Идентификатор всегда должен заканчиваться символами хуз. Используйте проверку данных для гарантии ввода только тех идентификаторов, которые заканчиваются на хуз. (Подсказка: используйте типа данных **Другой (Custom)** и функцию **ПРАВСИМВ**.)
7. Пусть каждая запись в диапазоне **B2:B15** должна быть текстовой, а не числовой. С помощью проверки данных обеспечьте при вводе числовых данных появление сообщения об ошибке. (Подсказка: используйте функцию **ЕТЕКСТ**.)
8. Настройте процедуру проверки данных для ввода в столбец **E** чисел, содержащих только два знака после запятой. (Подсказка: используйте функции **ДЛСТР** и **НАЙТИ**.)
9. Файл **Latitude.xlsx** содержит формулу для вычисления расстояния между двумя городами по их широте и долготе. В файле также хранятся широта и долгота для различных городов США. Создайте раскрывающийся список так, что при выборе одного города в ячейке **P2** и другого города в ячейке **Q2** расстояние между городами вычислялось бы в ячейке **Q10**.
10. Сделайте так, чтобы при добавлении новых городов в список из задания 9 они включались в раскрывающийся список.
11. В файле **Candybardata.xlsx** содержится список магазинов, в которых вы продаете шоколадные плитки. На листе также указаны виды плиток, продаваемые в каждом из магазинов, и их цены. Настройте лист для ввода или выбора пользователями как магазина, так и вида плитки из раскрывающегося списка, а также для получения цены в ячейке **D19**.
  - Обеспечьте выбор магазина из раскрывающегося списка в ячейке **B19**.
  - Создайте в ячейке **C13** раскрывающийся список для выбора вида шоколадной плитки из видов, продающихся в выбранном магазине. (Подсказка: при определении списка используйте функцию **ДВССЫЛ**.)

- При смене магазина в ячейке **B19** в ячейке **C19** сначала может быть указан не тот вид шоколадной плитки, который продается в новом выбранном магазине. Обеспечьте для такого случая появление в ячейке **C19** сообщения **Make Selection Above** (Выберите выше). Например, если в ячейке **B19** указан магазин **CVS**, а в ячейке **C13** — **gumballs**, в ячейке **C19** должно появиться сообщение **Make Selection Above**.
12. Ваша смета расходов составляет \$100 000. Расходы вводятся в столбце **A** по мере возникновения. Создайте критерий проверки данных, который гарантирует, что общая сумма расходов, указанных в столбце **E**, не превысит смету.
  13. Создайте критерий проверки данных, который обеспечит ввод чисел в порядке убывания.
  14. Создайте критерий проверки данных, обеспечивающий ввод только тех дат , которые приходятся на дни недели с понедельника по пятницу . (Подсказка: функция **ДЕНЬНЕД** возвращает 1 для воскресенья, 2 для понедельника и т . д.)
  15. Создайте критерий проверки данных, обеспечивающий в ячейках **A1:A20** ввод временных значений, соответствующих утру.

## ГЛАВА 43

# Обобщение данных на гистограммах и диаграммах Парето

### Обсуждаемые вопросы

- Говорят, одна картина стоит тысячи слов. Как с помощью Excel можно создать изображение (называемое гистограммой), которое обобщает значения в наборе данных?
- Каковы самые распространенные типы гистограмм?
- Что можно узнать, сравнивая гистограммы для разных наборов данных?
- Как создать диаграмму Парето?

Возможность обобщения больших наборов данных очень важна. В Microsoft Excel для обобщения данных наиболее часто применяются три инструмента: гистограммы, описательная статистика и сводные таблицы. В этой главе рассматривается обобщение данных с помощью гистограмм. До Excel 2016 для создания гистограмм использовалась надстройка **Анализ данных (Data Analysis)**. Гистограммы, созданные с помощью надстройки **Анализ данных**, не изменялись вслед за добавлением новых данных. Начиная с Excel 2016, в Excel предоставлены новые возможности для гибкого создания гистограмм, автоматически изменяющихся при появлении новых данных. Описательная статистика обсуждается в главе 44 «Обобщение данных с помощью описательной статистики», а сводные таблицы — в главе 45 «Сводные таблицы и срезы для описания данных».

## Ответы на вопросы

**?** Говорят, одна картина стоит тысячи слов. Как с помощью Excel можно создать изображение (называемое гистограммой), которое обобщает значения в наборе данных?

Гистограмма очень часто используется для обобщения данных. По сути, гистограмма показывает, сколько *наблюдений* (другой термин для *точек данных*) попадает в различные интервалы значений. Например, гистограмма, созданная на

основе данных о ежемесячной доходности акций Cisco, показывает, какая часть ежемесячных доходов компании Cisco попадает в интервал от 0 до 10%, от 11 до 20% и т. д. Интервалы, по которым группируются данные, называются *карманами*, или *бинами*.

Давайте научимся создавать и интерпретировать гистограммы, обобщающие значения, на примере ежемесячной доходности для акций Cisco и GM в 1990–2000 гг. Эти данные (а также данные для других акций) находятся в файле **Stock.xlsx** (на листе **Stockprices**). Часть этих данных показана на рис. 43.1. Например, за март 1990 г. (строка 52) акции Cisco поднялись в цене на 1,075%.

	A	B	C	D	E	F
49				min	-0.240320429	-0.202509
50				max	0.276619107	0.338983
51	Date	Microsoft	GE	Intel	GM	CSCO
52	30-Mar-90	0.121518984	0.040485829	0.037267081	0.022284122	0.010753
53	30-Apr-90	0.047404062	-0.003891051	-0.053892214	-0.035422344	0.010638
54	31-May-90	0.258620679	0.083515622	0.221518993	0.115819208	0.042105
55	29-Jun-90	0.04109589	0.005444646	-0.025906736	-0.020565553	0.070707
56	31-Jul-90	-0.125	0.034296028	-0.05319149	-0.020997375	-0.037736
57	31-Aug-90	-0.075187966	-0.13438046	-0.25	-0.131367296	-0.029412
58	28-Sep-90	0.024390243	-0.113387093	-0.003745318	-0.088050313	-0.090909
59	31-Oct-90	0.011904762	-0.04587156	0.007518797	0.013793103	0.311111
60	30-Nov-90	0.13333334	0.052884616	0.119402982	0.013605442	0.338983
61	31-Dec-90	0.041522492	0.057260275	0.026666667	-0.05821918	0.136076
62	31-Jan-91	0.303986698	0.115468413	0.188311681	0.054545455	0.303621
63	28-Feb-91	0.057324842	0.070468754	0.043715846	0.100689657	-0.042735
64	28-Mar-91	0.022891566	0.023897059	-0.020942409	-0.044303797	-0.129464
65	30-Apr-91	-0.067137808	0.016157989	0.053475935	-0.052980132	0.220513
66	31-May-91	0.108585857	0.09908127	0.131979689	0.217482507	0.084034
67	28-Jun-91	-0.068906605	-0.042071197	-0.165919289	-0.055072464	-0.054264
68	31-Jul-91	0.078899086	-0.010135135	0.010752688	-0.024539877	0.286885
69	30-Aug-91	0.159863949	0.022184301	0.053191491	-0.033962265	0.156051
70	30-Sep-91	0.043988269	-0.066644408	-0.146464646	-0.016447369	-0.096419
71	31-Oct-91	0.054775283	-0.005405406	-0.03846154	-0.060200669	0.189024

**Рис. 43.1.** Ежемесячная доходность акций

При создании гистограмм в Excel необходимо определить карманы, самостоятельно или автоматически. Автоматически определенные карманы могут выглядеть странно, например, от -12,53% до 4,52%. По этой причине я предпочитаю определять их самостоятельно.

Эффективней всего начать определение карманов для гистограммы (определение карманов можно рассматривать как установку границ) с разделения входного интервала значений (от наименьшего до наибольшего) на 8–15 равных частей. Ежемесячная доходность акций Cisco составляет от -30% до 40%, поэтому выберем для границ карманов значения -30%, -20%, -10%, 0% и т. д. до 40%.

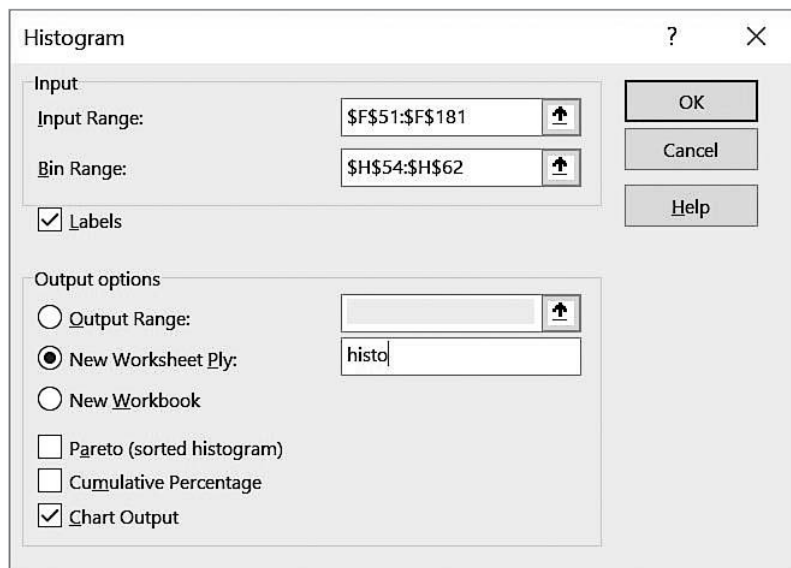
Для создания карманов введите CSCO 0,4 0,3 0,2, ..., -0,2, -0,3 (границы карманов) в ячейки H54:H62. Затем на вкладке **Данные (Data)** в группе **Анализ (Analysis)** выберите **Анализ данных (Data Analysis)**. Появится диалоговое окно **Анализ данных (Data Analysis)**,

в котором перечислены инструменты пакета анализа данных (Analysis ToolPak), содержащего массу возможностей статистической обработки данных в Excel.

### ПРИМЕЧАНИЕ

Если на вкладке Данные (Data) пакет Анализ данных (Data Analysis) отсутствует, перейдите на вкладку Файл (File), выберите Параметры (Options) и затем Надстройки (Add-Ins). В списке Управление (Manage) выберите Надстройки Excel (Excel Add-Ins) и затем щелкните на Перейти (Go). В диалоговом окне Надстройки (Add-Ins) включите флажок Пакет анализа (Analysis ToolPak) (это первый флажок, не флажок Пакет анализа — VBA (Analysis ToolPak — VBA)) и нажмите OK. Теперь на вкладке Данные (Data) в группе Анализ (Analysis) можно получить доступ к инструментам пакета Анализ данных (Data Analysis).

В диалоговом окне Анализ данных (Data Analysis) выберите Гистограмма (Histogram). Откроется диалоговое окно, как на рис. 43.2.



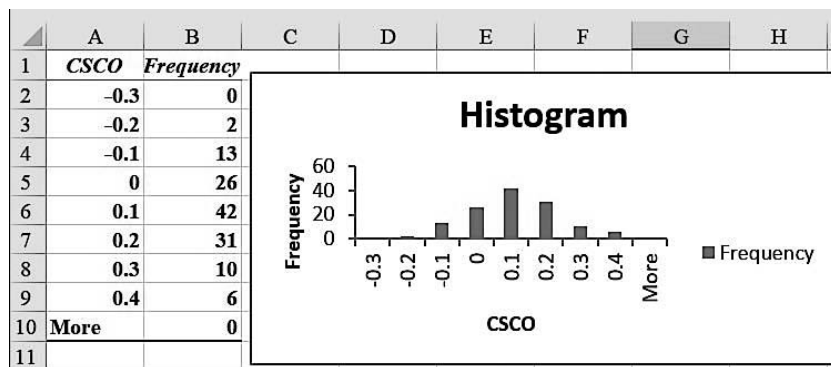
**Рис. 43.2.** Диалоговое окно Гистограмма для гистограммы доходности акций Cisco

Заполните диалоговое окно, как показано на рис. 43.2:

- Выделите входной интервал (диапазон F51:F181). (Для выделения диапазона F51:F181 можно выделить ячейку F51 и нажать комбинацию Ctrl+Shift+↓. Диапазон будет выделен до нижней заполненной данными строки в столбце.) Этот интервал включает все данные, необходимые для создания гистограммы. Включите в интервал метку CSCO в ячейке F51, поскольку в противном случае ось x гистограммы получит вместо названия число, что будет выглядеть странно.

- Выберите диапазон с данными карманов (H54:H62) (границы входят в интервалы). Будут созданы интервалы от  $-30$  до  $-20\%$ , от  $-20$  до  $-10\%$  и т. д. до интервала  $30-40\%$ .
- Установите флажок **Метки (Labels)**, поскольку первые строки и входного интервала, и интервала карманов содержат метки.
- В разделе **Параметры вывода (Output Options)** установите переключатель **Новый рабочий лист (New Worksheet Ply)** (и назовите его **Гистограмма**).
- Установите флажок **Вывод графика (Chart Output)**, или гистограмма не будет создана.

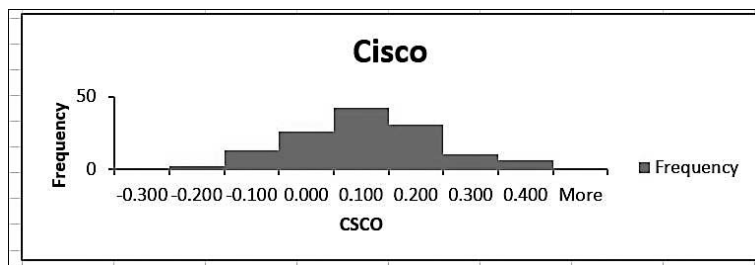
После этого в диалоговом окне **Гистограмма (Histogram)** нажмите ОК. Появится гистограмма по компании Cisco, показанная на рис. 43.3.



**Рис. 43.3.** Гистограмма Cisco, созданная с помощью инструмента из пакета анализа

Между столбцами созданной гистограммы имеются промежутки. Для их удаления щелкните правой кнопкой мыши по любому столбику гистограммы и выберите **Формат ряда данных (Format Data Series)**. На панели **Формат ряда данных (Format Data Series)**, в **Параметрах ряда (Series Options)**, установите на счетчике **Боковой зазор (Gap Width)** значение  $0\%$ . Кроме того, на гистограмме могут отсутствовать метки под столбцами. Если метки не появились, выберите гистограмму и с помощью боковых меток-манипуляторов (круглой формы, курсор видоизменяется в две стрелочки) увеличьте ее размеры. Или уменьшите размер шрифта. Для этого щелкните правой кнопкой мыши по оси графика и выберите **Шрифт (Font)**. Установите на счетчике **Размер (Size)** значение 5. Также можно изменить название гистограммы, выделив текст и введя новое название. После внесения некоторых из описанных изменений гистограмма будет выглядеть так, как на рис. 43.4.

Обратите внимание, что доход по акциям Cisco, вероятнее всего, составит от 0 до  $10\%$  в месяц, и что высота столбиков гистограммы на краях интервала намного меньше высоты самого высокого столбика. При построении в дополнение к самой гистограмме создается сводка по частоте заполнения карманов, показанная на рис. 43.5.



**Рис. 43.4.** Промежутки между столбцами гистограммы удалены

Например, как видно из сводки, в течение двух месяцев доход по акциям Cisco составлял от  $-30$  до  $-20\%$ ; в течение 13 месяцев ежемесячный доход был больше  $-20\%$  и меньше или равен  $-10\%$ .

Чтобы в гистограмме отобразилось добавление новых данных о ежемесячных доходах по акциям Cisco или внесение изменений в имеющиеся данные, необходимо перезапустить процедуру создания гистограммы в диалоговом окне **Анализ данных** (Data Analysis).

Excel 2019 предоставляет удобное средство создания эффектно выглядящих гистограмм, автоматически отображающих внесение изменений в данные. Проиллюстрируем использование этого инструмента создания гистограмм на файле *Iqtemp.xlsx* (в папке *Templates*), содержащем данные IQ для 1083 шестиклассников (рис. 43.6).

<i>CSCO</i>	<i>Frequency</i>
$-0.300$	0
$-0.200$	2
$-0.100$	13
0.000	26
0.100	42
0.200	31
0.300	10
0.400	6
More	0

**Рис. 43.5.** Сводка по частоте заполнения карманов

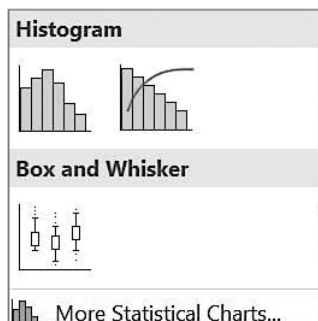
	E
4	IQ
5	95
6	105
7	93
8	103
9	103
10	129
11	95
12	98
13	94
14	106
15	102
16	96
17	112
18	106
19	106
20	78
21	132
22	111
23	107

**Рис. 43.6.** Данные IQ

Выбрав диапазон E4:E1177, нажмите Ctrl+T для представления данных в виде таблицы (пусть будет установлен флажок **Таблица с заголовками (My Table Has Headers)** в диалоговом окне **Создание таблицы (Create Table)**). Теперь, когда мы будем вносить изменения в данные, гистограмма будет отражать эти изменения. Затем выделите все данные (включая ячейку E4), и после этого на вкладке **Вставка (Insert)** в группе **Гистограмма (Charts)** нажмите стрелку вниз, чтобы перейти к опции **Вставить диаграмму статистики (Insert Statistical Chart)** (рис. 43.7), и выберите гистограмму, как показано на рис. 43.8.

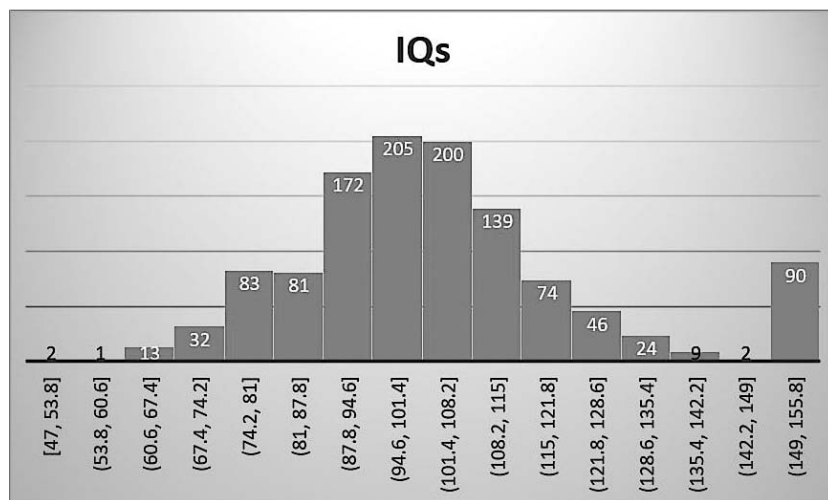


**Рис. 43.7.** Значок вставки статистической диаграммы



**Рис. 43.8.** Варианты статистических диаграмм

Вы получаете прекрасную гистограмму. На вкладке **Конструктор (Design)** вы можете настроить вид гистограммы. Например, в группе **Стили диаграмм (Chart Styles)** выберите третий слева вариант, тогда в каждом диапазоне будет отображено количество значений, попавших в этот диапазон (рис. 43.9).

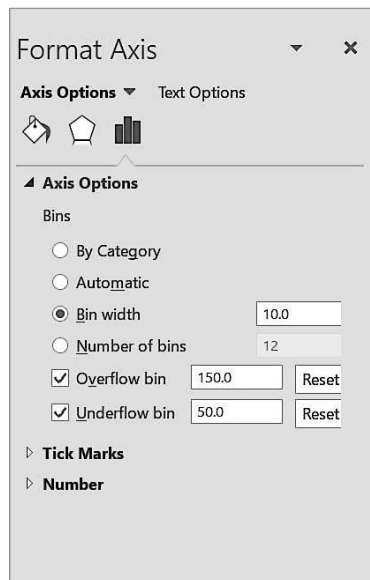


**Рис. 43.9.** Гистограмма распределения IQ

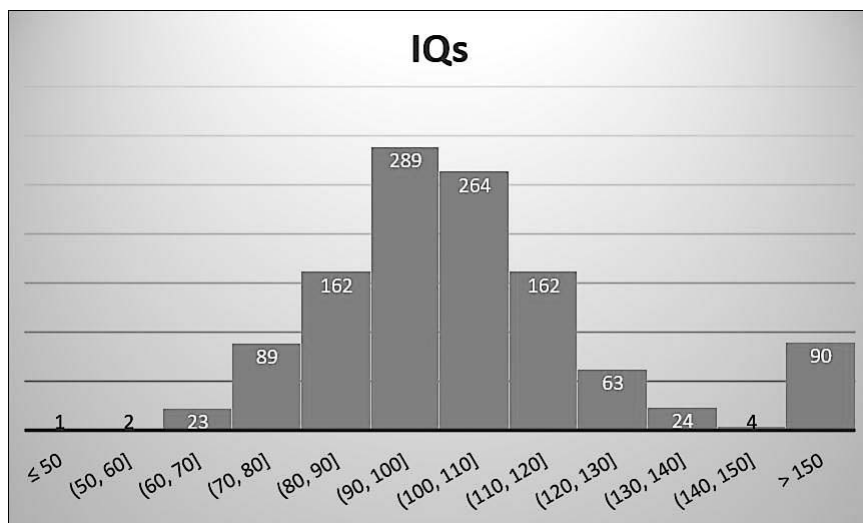


Щелкнув правой кнопкой мыши по оси и выбрав **Формат оси (Format Axis)**, вы можете изменить определения интервалов и задать нижнюю границу первого интервала и верхнюю границу последнего интервала (рис. 43.10).

Установите флажок **Выход за нижнюю границу интервала (Underflow Bin)** и значение нижней границы первого интервала равным 50. Для **выхода за верхнюю границу интервала (Overflow Bin)** установите значение верхней границы последнего интервала равным 150. Установите флажок **Длина интервала (Bin Width)** и значение каждого интервала 10 (вам может потребоваться прокрутить окно вправо). Щелкнув раскрывающийся список **ЧИСЛО (Number)**, вы можете изменить формат оси (например, для денежных данных — на формат **Денежный**). После внесения всех изменений гистограмма приобретет вид, приведенный на рис. 43.11.



**Рис. 43.10.** Изменение бинарных интервалов



**Рис. 43.11.** Гистограмма с измененными интервалами (шириной 10 и разрывами в 50 и 150)

Мы выяснили, например, что 90 учащихся имеют IQ выше 150. Очень толковые ребята! Обратите внимание, что при добавлении данных (например, 20 человек с IQ ниже 55), введенные данные автоматически будут учтены в гистограмме.

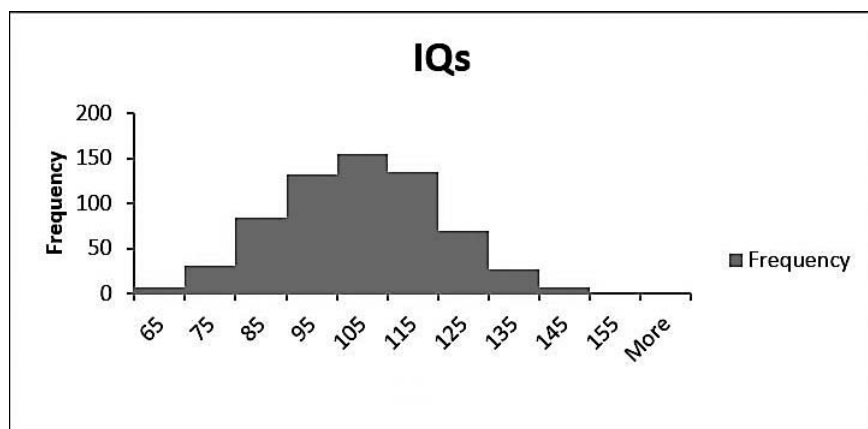
### ❓ Каковы самые распространенные типы гистограмм?

Гистограммы, созданные для большинства наборов данных, можно разделить на несколько типов:

- симметричные;
- с наклоном вправо (положительная асимметрия);
- с наклоном влево (отрицательная асимметрия);
- с несколькими пиками.

В приведенном далее списке каждый тип гистограммы рассмотрен более подробно. См. файл *Skewexamples.xlsx*.

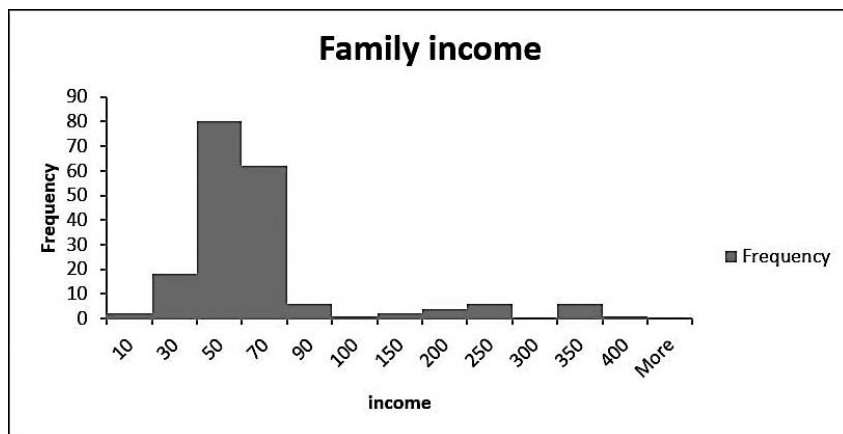
- **Симметричное распределение.** Гистограмма является *симметричной*, если имеет единственный пик и выглядит примерно одинаково слева и справа от пика. Симметричными часто бывают результаты тестов (например, тесты IQ). Скажем, гистограмма результатов теста IQ (см. ячейку Z42) может выглядеть так, как на рис. 43.12. Обратите внимание, что высота столбиков слева и справа от пика примерно одинакова, высота следующих столбиков слева и справа от пика также примерно одинакова и т. д. Столбик с меткой 105 представляет всех людей, прошедших тест с IQ больше 95 и не больше 105, столбик с меткой 65 — с IQ не больше 65 и т. д. Также отметим, что гистограмма ежемесячной доходности акций Cisco почти симметричная.



**Рис. 43.12.** Симметричная гистограмма

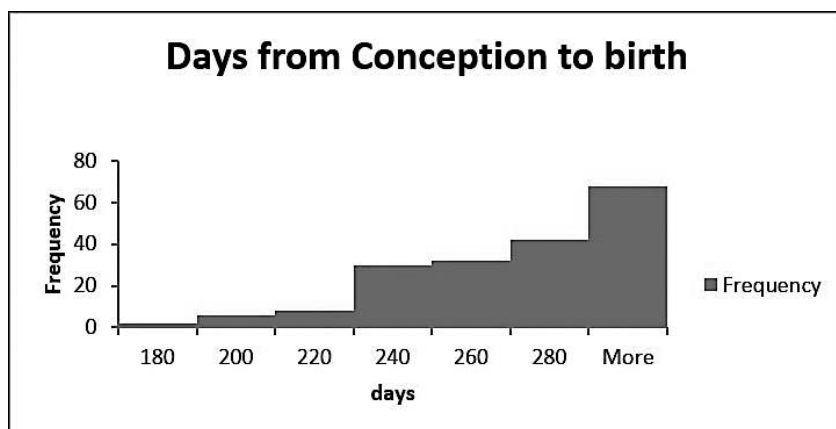
- **С наклоном вправо (положительная асимметрия).** Гистограмма имеет *наклон вправо* (положительную асимметрию), если она имеет единственный пик и набор значений справа от пика значительно шире набора значений слева от пика. Гистограммы для многих наборов экономических данных (таких как семейные доходы или доходы физических лиц) имеют положительную асимметрию. На

рис. 43.13 (см. ячейку T24) представлен пример гистограммы с положительной асимметрией, созданной на основе данных о семейных доходах.



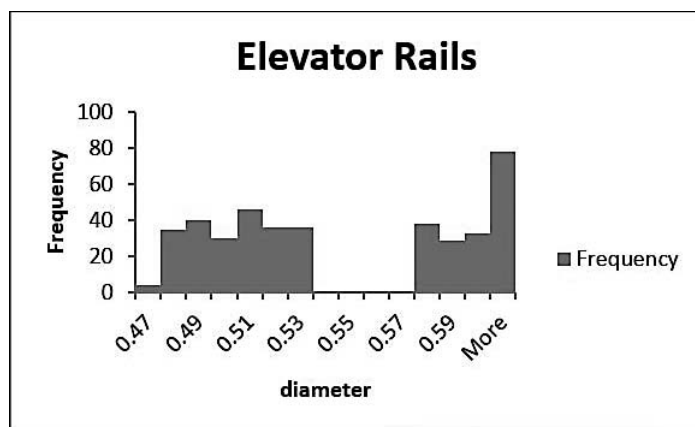
**Рис. 43.13.** Гистограмма с положительной асимметрией для данных о семейных доходах

- **С наклоном влево (отрицательная асимметрия).** Гистограмма имеет *наклон влево* (отрицательную асимметрию), если она имеет единственный пик и набор значений слева от пика значительно шире набора значений справа от пика. Гистограмма для количества дней, прошедших от зачатия до рождения, имеет отрицательную асимметрию. Пример приведен на рис. 43.14 (см. ячейку Q7). Высота каждого столбика представляет число женщин, у которых количество дней от зачатия до рождения попадает в указанный интервал. Например, две женщины родили меньше чем через 180 дней после зачатия.



**Рис. 43.14.** Гистограмма с отрицательной асимметрией для количества дней от зачатия до рождения

- **С несколькими пиками.** Если гистограмма имеет несколько пиков, как правило, это означает, что данные из нескольких совокупностей оказались на графике вместе. Например, гистограмма на рис. 43.15 представляет измерения диаметра рельсовых направляющих грузоподъемника, производимых на двух станках. (См. ячейку Q11 файла *Twinpeaks.xlsx*.) На этой гистограмме данные образуют две группы. По всей вероятности, каждая группа данных соответствует направляющим, производимым на одном из станков. Если предположить, что требуемый диаметр должен быть равен 0,55 дюйма, придется сделать вывод, что один станок производит слишком узкие направляющие, а другой — слишком широкие. В соответствии с интерпретацией этой гистограммы следует создать две гистограммы для направляющих, производимых на каждом станке. Этот пример поясняет почему гистограммы являются мощным инструментом контроля качества.



**Рис. 43.15.** Гистограмма с несколькими пиками

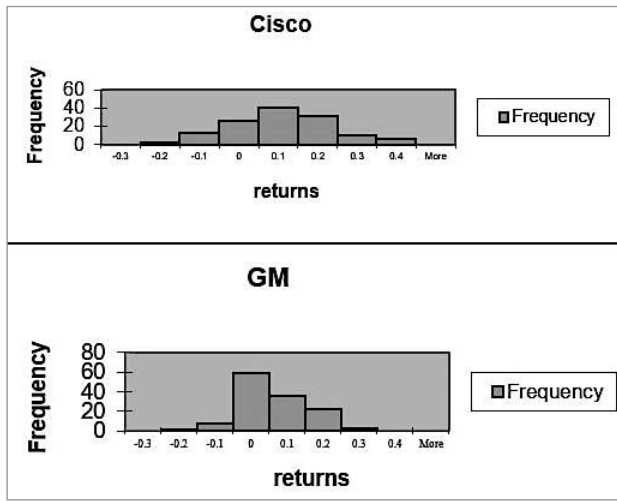
### **(?) Что можно узнать, сравнивая гистограммы для разных наборов данных?**

Аналитиков часто просят сравнить два различных набора данных. Например, требуется узнать, чем отличаются ежемесячные доходности акций GM и Cisco. Для ответа на такой вопрос можно создать гистограмму для акций GM с теми же интервалами, что и для акций Cisco, и затем поместить гистограммы друг над другом (рис. 43.16). См. в файле *Stock.xlsx* лист *Histograms*.

При сравнении этих двух гистограмм можно прийти к двум важным выводам.

- Обычно дела у Cisco идут лучше, чем у GM. Это видно из рисунка, поскольку самый высокий столбик на гистограмме доходов Cisco на один столбик правее, чем самый высокий столбик на гистограмме доходов GM. Кроме того, столбики на гистограмме доходов Cisco простираются вправо дальше, чем на гистограмме доходов GM.
- На гистограмме доходов Cisco изменчивость, или разброс вокруг среднего значения, шире, чем на гистограмме доходов GM. Обратите внимание, что пик на

гистограмме доходов GM содержит 59 месяцев, в то время как пик на гистограмме доходов Cisco — только 42 месяца. Это говорит о том, что для Cisco большая часть доходов находится вне интервала, который представляет наиболее вероятный доход от акций Cisco. Доходы Cisco имеют больший разброс, чем доходы GM.



**Рис. 43.16.** Гистограммы с одинаковыми интервалами для сравнения разных наборов данных

В главе 44 мы рассмотрим возможности описательной статистики и ящиков с усами (Boxplots) для более детального анализа различий между ежемесячными доходами Cisco и GM.

### ? Как создать диаграмму Парето?

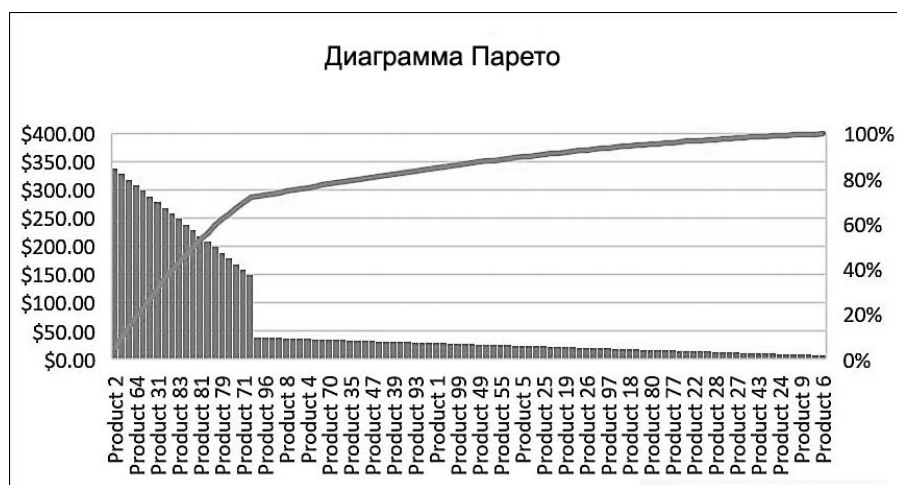
Диаграмма Парето (Pareto chart) — это тип диаграммы, который содержит столбцы и график. Отдельные значения отображаются в убывающем порядке столбцами, а сумма нарастающим итогом представлена в виде линии. Диаграмма Парето часто используется для иллюстрации знаменитого правила 80 : 20, впервые сформулированного выдающимся итальянским экономистом Вильфредо Парето (1848–1923). *Правило Парето* делает акцент на вкладе отдельных явлений в картину в целом. Например:

- 20% продуктов приносят 80% дохода;
- 20% людей принадлежат 80% доходов;
- 80% обращений за технической поддержкой порождаются 20% всех возникающих проблем;
- на 20% веб-сайтов приходится 80% посещений.

Проиллюстрируем создание диаграммы Парето в Excel 2019 на файле **Paretotemp.xlsx** (в папке **Templates**), в котором представлены доходы, получаемые компанией от выпускаемых ею 100 продуктов (рис. 43.17).

	E	F
3	<b>Product</b>	<b>Revenue</b>
4	<b>Product 1</b>	<b>\$30.00</b>
5	<b>Product 2</b>	<b>\$340.00</b>
6	<b>Product 3</b>	<b>\$11.60</b>
7	<b>Product 4</b>	<b>\$37.20</b>
8	<b>Product 5</b>	<b>\$25.20</b>
9	<b>Product 6</b>	<b>\$8.40</b>
10	<b>Product 7</b>	<b>\$38.00</b>
11	<b>Product 8</b>	<b>\$38.40</b>
12	<b>Product 9</b>	<b>\$9.60</b>
13	<b>Product 10</b>	<b>\$29.20</b>
14	<b>Product 11</b>	<b>\$14.80</b>
15	<b>Product 12</b>	<b>\$10.00</b>
16	<b>Product 13</b>	<b>\$22.40</b>
17	<b>Product 14</b>	<b>\$29.60</b>

**Рис. 43.17.** Данные для диаграммы Парето

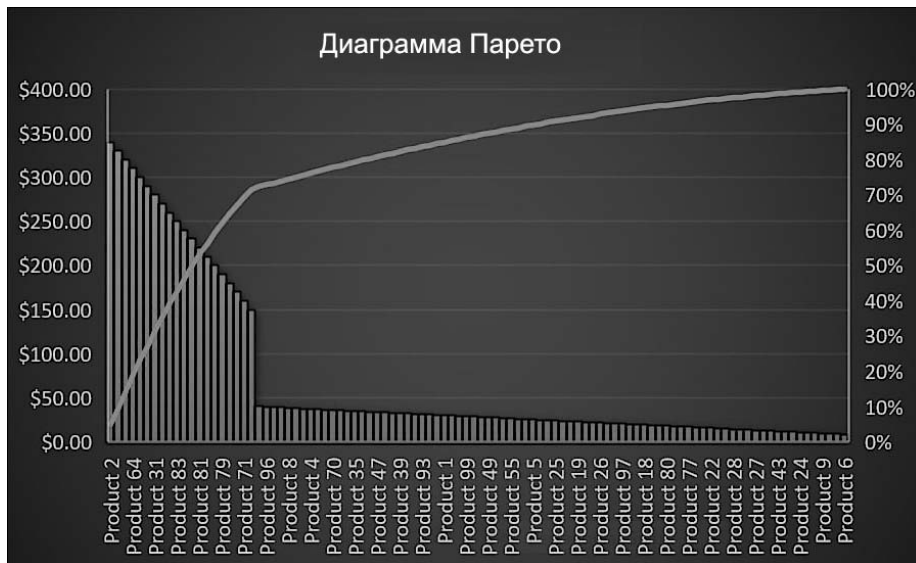


**Рис. 43.18.** Диаграмма Парето

Выделив данные (диапазон **E3:F103**), выберите значок **Вставить диаграмму статистики** (**Insert Statistic Chart**) на вкладке **Вставка** (**Insert**) в группе **Диаграммы** (**Charts**) и затем выберите диаграмму Парето, вторую в перечне видов гистограмм (см. рис. 43.8

выше). Мы получили диаграмму Парето, представленную на рис. 43.18. Продукты перечислены в порядке снижения сумм продаж. Линия отображает суммарный процент продаж этих продуктов. Мы видим, что 10% лучше всего продаваемых продуктов обеспечивает около 80% всех продаж. И, как мы помним, если представить исходные данные в виде таблицы, то новые данные будут автоматически включены в диаграмму.

Теперь выделите созданную диаграмму Парето и выберите пятый слева стиль в группе **Стили диаграмм (Chart Styles)** на вкладке **Конструктор (Design)**. Вы получите диаграмму, представленную на рис. 43.19.



**Рис. 43.19.** Диаграмма Парето после применения стиля из вкладки Конструктор

## Задания

1. Используйте данные файла **Stock.xlsx** для создания гистограммы ежемесячных доходов от акций GE и Intel.
2. Используйте данные файла **Historicalinvest2009.xlsx** для создания гистограммы ежегодных доходов по акциям и облигациям. Затем сравните ежегодные доходы по этим акциям и облигациям.
3. В файле **Deming.xlsx** содержатся измерения диаметра (в дюймах) для 500 стержней, произведенных компанией Rodco, согласно отчету руководителя производства. Допустимый размер диаметра стержня составляет не менее 1 дюйма. Раньше диаметрам стержней, производимых компанией Rodco, соответствовала симметричная гистограмма. Задание следующее:

- Постройте гистограмму для этих измерений.
- Прокомментируйте любые необычные аспекты этой гистограммы.

Можно ли догадаться, почему у гистограммы необычный вид? (Подсказка: один из 14 принципов управления качеством Эдварда Деминга гласит: «Искорените страх».)

4. В файле **Unemployment.xlsx** находятся данные об уровне безработицы в США. Создайте на их основе гистограмму. Симметрична или асимметрична гистограмма уровня безработицы?
5. В файле **Teams.xlsx** содержатся данные о засчитанных пробежках (нарушение правил) в высшей баскетбольной лиге за сезон. Создайте на их основе гистограмму. Симметрична или асимметрична гистограмма, представляющая засчитанные пробежки?
6. В файле **NFLpoints.xlsx** содержатся данные об очках, заработанных командами НФЛ за сезон. Создайте на их основе гистограмму. Симметрична или асимметрична гистограмма, представляющая заработанные очки?
7. Опираясь на данные из файла **Problem7data.xlsx**, создайте гистограмму, обобщающую данные о росте американских мужчин.
8. Данные в файле **Problem8data.xlsx** представляют очки, заработанные каждой из футбольных университетских команд первого дивизиона за сезон 2015 г. Создайте гистограмму, обобщающую эти данные. Является ли эта гистограмма симметричной?
9. Данные в файле **Problem9data.xlsx** представляют доходы семей в г. Смоллтаун, США. Создайте диаграмму Парето, обобщающую эти доходы.



## ГЛАВА 44

# Обобщение данных с помощью описательной статистики

### Обсуждаемые вопросы

- Как определяется типичное значение набора данных?
- Как измерить, насколько набор данных отклонился от своего типичного значения?
- Как среднее значение и стандартное отклонение характеризуют набор данных?
- Как с помощью описательной статистики сравнить наборы данных?
- Как для заданной точки данных можно найти величину ее процентиля в наборе данных? Например, как найти девяностый процентиль набора данных?
- Как найти второе наибольшее или второе наименьшее число в наборе данных?
- Как можно ранжировать числа в наборе данных?
- Что такое урезанное среднее набора данных?
- Существует ли быстрый способ получить для выбранного диапазона ячеек различные статистики, описывающие данные в этих ячейках?
- Почему финансовые аналитики часто определяют средний доход по акциям с помощью среднего геометрического?
- Как использовать ящики с усами (boxplots) для характеристики и сравнения наборов данных?

В главе 43 «Обобщение данных на гистограммах и диаграммах Парето» я показал, как наборы данных можно описать с помощью гистограмм. В этой главе рассматривается описание наборов данных с помощью особых характеристик, таких как среднее, медиана, стандартное отклонение и дисперсия, — показателей, сгруппированных в Microsoft Excel 2019 под названием описательная статистика. Для получения доступа к описательной статистике набора данных выберите на вкладке **Данные (Data)** в группе **Анализ (Analysis)** пакет **Анализ данных (Data Analysis)**, доступный как надстройка, и в нем инструмент **Описательная статистика (Descriptive Statistics)**. После ввода имеющихся данных и нажатия **ОК** появится вся описательная статистика по соответствующему набору данных. Также описательная статистика доступна через функции Excel. В конце главы будет показано, как для характеристики и сравнения наборов данных пользоваться ящиками с усами.

## Ответы на вопросы

### ? Как определяется типичное значение набора данных?

Чтобы понять, как можно использовать описательную статистику, вернемся к данным о ежемесячной доходности акций Cisco и GM в файле `Stock.xlsx`. Для создания показателей описательной статистики этого набора данных на вкладке **Данные** (Data) в группе **Анализ** (Analysis) выберите пакет **Анализ данных** (Data Analysis) и в нем инструмент **Описательная статистика** (Descriptive Statistics). Заполните диалоговое окно **Описательная статистика** (Descriptive Statistics), как показано на рис. 44.1.

The screenshot shows the 'Descriptive Statistics' dialog box. The 'Input' section has 'Input Range' set to '\$F\$51:\$F\$181' and 'Grouped By' set to 'Columns'. The 'Labels in first row' checkbox is checked. The 'Output options' section has 'Output Range' set to '\$I\$53'. The 'Summary statistics' checkbox is checked. The 'Confidence Level for Mean' is set to '95 %'. The 'Kth Largest' and 'Kth Smallest' options are both set to '1'.

**Рис. 44.1.** Диалоговое окно Описательная статистика

### ПРИМЕЧАНИЕ

Если инструмент **Анализ данных** (Data Analysis) не отображается на вкладке **Данные** (Data), перейдите на вкладку **Файл** (File), выберите **Параметры** (Options) и выберите **Надстройки** (Add-Ins) в левой панели. В списке **Управление** (Manage) в диалоговом окне **Параметры Excel** (Excel Options) выберите **Надстройки Excel** (Excel Add-Ins) и нажмите **Перейти** (Go). В открывшемся диалоговом окне **Надстройки Excel** (Excel Add-Ins) установите флажок **Пакет анализа** (Analysis ToolPak) (он идет первым в списке, не **Пакет анализа VBA**) и нажмите **OK**.

Входной интервал представляет данные о ежемесячной доходности акций Cisco и GM, хранящиеся в диапазоне `E51:F181` (включая метки в строке 51). Оставшуюся

ся часть диалогового окна **Описательная статистика (Descriptive Statistics)** я заполнил, как показано на рис. 44.1:

- переключатель **Группирование (Grouped By)** установлен в положение **по столбцам (Columns)**, поскольку каждый набор данных находится в отдельном столбце;
- флажок **Метки в первой строке (Labels In First Row)** установлен потому, что первая строка диапазона данных содержит метки, а не данные;
- выбран **Интервал для вывода (Output Range)** в **Параметрах вывода (Output options)** и ячейка **I53** текущего листа выбрана как первая ячейка выходного интервала;
- флажок **Итоговая статистика (Summary Statistics)** установлен для получения самых распространенных статистических показателей для ежемесячной доходности акций GM и Cisco.

После нажатия **ОК** на листе появится описательная статистика, представленная на рис. 44.2.

	I	J	K	L
53	<b>GM</b>		<b>CSCO</b>	
54				
55	<b>Mean</b>	<b>0.009</b>	<b>Mean</b>	<b>0.056</b>
56	<b>Standard Error</b>	<b>0.008</b>	<b>Standard Error</b>	<b>0.011</b>
57	<b>Median</b>	<b>-0.005</b>	<b>Median</b>	<b>0.050</b>
58	<b>Mode</b>	<b>#N/A</b>	<b>Mode</b>	<b>0.051</b>
59	<b>Standard Deviation</b>	<b>0.090</b>	<b>Standard Deviation</b>	<b>0.122</b>
60	<b>Sample Variance</b>	<b>0.008</b>	<b>Sample Variance</b>	<b>0.015</b>
61	<b>Kurtosis</b>	<b>0.475</b>	<b>Kurtosis</b>	<b>-0.320</b>
62	<b>Skewness</b>	<b>0.224</b>	<b>Skewness</b>	<b>0.105</b>
63	<b>Range</b>	<b>0.517</b>	<b>Range</b>	<b>0.541</b>
64	<b>Minimum</b>	<b>-0.240</b>	<b>Minimum</b>	<b>-0.203</b>
65	<b>Maximum</b>	<b>0.277</b>	<b>Maximum</b>	<b>0.339</b>
66	<b>Sum</b>	<b>1.206</b>	<b>Sum</b>	<b>7.224</b>
67	<b>Count</b>	<b>130.000</b>	<b>Count</b>	<b>130.000</b>

**Рис. 44.2.** Описательная статистика для акций Cisco и GM

Теперь интерпретируем описательную статистику, которая определяет типичное значение (или средний уровень) для ежемесячной доходности акций Cisco. Результаты содержат три измерения среднего уровня: среднее (или средняя величина), медиана и мода.

- **Среднее.** Среднее значение набора данных обозначается как  $\bar{x}$  и является средним значением всех наблюдений в выборке. Для значений данных  $x_1, x_2, \dots, x_n$  среднее вычисляется по формуле:

$$\bar{x} = \frac{1}{n} \sum x_i.$$

Здесь  $n$  — число наблюдений в выборке, а  $x_i$  —  $i$ -е наблюдение в выборке. Как видно из рисунка, средний доход по акциям Cisco составляет 5,6% в месяц.

Сумма отклонений всех значений от среднего всегда равна 0. Таким образом, среднее набора данных можно рассматривать как точку равновесия данных. Безусловно, среднее значение выборки в Excel можно получить и без инструмента **Описательная статистика (Descriptive Statistics)**, применив функцию **СРЗНАЧ (AVERAGE)** к соответствующему диапазону ячеек.

- **Медиана.** Это *срединное* наблюдение для ряда данных, упорядоченных от наименьшего к наибольшему. Если выборка содержит нечетное число наблюдений, то медианой является наблюдение, слева и справа от которого находится одинаковое число наблюдений. Таким образом, для выборки из девяти наблюдений медианой будет являться пятое наименьшее (или пятое наибольшее) наблюдение. Если выборка включает четное число наблюдений, можно просто взять среднее значение от двух срединных наблюдений. По сути, медиана является 50-м процентилем данных. Например, медиана для ежемесячного дохода по акциям Cisco составляет 5%. Эту информацию также можно получить при помощи функции **МЕДИАНА (MEDIAN)**.
- **Мода.** Это наиболее часто встречающееся значение в выборке. Если ни одно значение не встречается более одного раза, моды не существует. Для акций GM ежемесячный доход за 1990–2000 гг. ни разу не повторился, поэтому для этих данных мода не существует. Для акций Cisco мода равна приблизительно 5,14%. В версиях Excel до 2010 года для вычисления моды применялась также функция **МОДА (MODE)**. Если значения встречались только один раз, функция **МОДА** возвращала ошибку #Н/Д.

Проблема в том, что набор данных может иметь не одну моду, а функция **МОДА** просто возвращала первую найденную моду. Поэтому в Microsoft Excel 2010 были представлены две функции: **МОДА.ОДН (MODE.SNGL)** и **МОДА.НСК (MODE.MULT)**. (См. файл **Modelfunctions.xlsx**.)

Функция **МОДА.ОДН** работает точно как функция **МОДА** в более ранних версиях Excel. Но в более ранних версиях Excel функция **МОДА.ОДН** не распознается.

Функция **МОДА.НСК** является функцией, возвращающей массив результатов (см. главу 91 «Формулы массива и функции, возвращающие массив»). Для использования функции **МОДА.НСК** (или любой другой функции, возвращающей массив) сначала необходимо выделить диапазон ячеек, в который функция возвратит значения. Затем следует ввести функцию или формулу. Но не торопитесь жать на **Enter**; нужно удерживать **Ctrl** и **Shift** (**Shift** нажата вслед за **Ctrl**) и только затем нажать **Enter** (или **Tab**). Сложность в применении функции **МОДА.НСК** состоит в том, что заранее неизвестно, сколько мод имеет набор данных, и поэтому неизвестен размер диапазона, который требуется выделить.

В файле **Modelfunctions.xlsx** (рис. 44.3) показаны все три функции, связанные с модой.

	C	D	E	F	G
3		<b>3 and 5 are both modes</b>			
4					
5			3		
6			4		
7			5		
8			3		
9			2		
10			1		
11			5		
12					
13				<b>3 =MODE(D5:D11)</b>	
14	<b>3 cells selected</b>			<b>3 =MODE.SNGL(d5:d11)</b>	
15	<b>Mode.Mult</b>				
16		3		<b>3 Mode.Mult</b>	
17		5		<b>5 two cells selected</b>	
18	<b>#N/A</b>				
19			<b>Mode.mult one cell selected</b>		
20			3		

**Рис. 44.3.** Примеры использования функций МОДА

Данные, представленные на рис. 44.3, имеют две моды (3 и 5). В ячейке E13 введите старую функцию МОДА, формулу =МОДА(D5:D11). Excel вернет значение первой найденной моды (3). В ячейке E14 формула =МОДА.ОДН (D5:D11) вернет тот же результат.

После выделения диапазона ячеек E16:E17 введите формулу =МОДА.НСК(D5:D11) (способом для массива, Ctrl+Shift+Enter), и Excel вернет значения обеих найденных мод (3 и 5). Выбрав диапазон C16:C18 (включающий три ячейки), введите вышеописанным способом для массива ту же формулу, и тогда вы увидите в ячейке C18 значение #N/Д (#N/A), так как в вашем наборе данных нет третьей моды, значение которой могло бы быть помещено в эту ячейку. Наконец, выделите диапазон из одной ячейки E20 и введите вышеописанным способом для массива ту же формулу. Так как был выделен диапазон из одной ячейки, Excel вернет только первое найденное значение моды (3).

Мода редко используется как измерение среднего уровня. Однако следует отметить, что для симметричных наборов данных среднее, медиана и мода равны.

Возникает естественный вопрос: что является наиболее точной мерой среднего уровня — среднее или медиана? Следует отдать предпочтение среднему, если набор данных не проявляет чрезмерную асимметрию. В противном случае в качестве меры среднего уровня следует использовать медиану. Если набор данных асимметричен, экстремальные значения искажают среднее. Тогда медиана является более точной мерой типичного значения набора данных. Например, в отчетах правительства США указывается медианный семейный доход, а не средний

семейный доход, поскольку семейный доход имеет сильную положительную асимметрию.

*Показатель асимметричности* в результатах работы инструмента *Описательная статистика (Descriptive Statistics)* определяет, насколько асимметричным является набор данных.

- Асимметричность больше +1 указывает на высокую степень положительной асимметрии.
- Асимметричность меньше -1 указывает на высокую степень отрицательной асимметрии.
- Асимметричность от -1 до +1 включительно указывает на относительную симметричность набора данных.

Таким образом, ежемесячная доходность акций GM и Cisco демонстрирует небольшую положительную асимметрию. Поскольку показатель асимметричности для каждого набора данных меньше +1, среднее является более точной мерой типичной доходности, чем медиана. Для вычисления асимметрии набора данных служит также функция *СКОС (SKEW)*.

Следует отметить, что эксцесс (*kurtosis*), который звучит как название болезни, — не самая важная мера, хотя он и присутствует в описательной статистике, приведенной на рис. 44.2. Значение эксцесса около 0 означает, что набор данных демонстрирует пиковость, близкую к нормальной (или стандартной колоколообразной) кривой. (Нормальная кривая рассмотрена в главе 72 «Нормальная случайная величина и Z-оценка») *Положительный эксцесс* означает, что набор данных имеет более острый пик, чем нормальная случайная величина, в то время как *отрицательный эксцесс* означает, что данные имеют менее острый пик, чем нормальная случайная величина. Ежемесячная доходность акций GM имеет более острый максимум, чем нормальная кривая, в то время как ежемесячная доходность акций Cisco имеет менее острый максимум по сравнению с нормальной кривой. Для вычисления эксцесса набора данных служит также функция *ЭКСЦЕСС (KURT)*.

### ❓ Как измерить, насколько набор данных отклонился от своего типичного значения?

Рассмотрим две инвестиции, каждая из которых приносит в среднем 20% в год. Прежде чем принять решение, какой из них отдать предпочтение, мы должны узнать, каков разброс, или рискованность, инвестиции. Наиболее важными мерами разброса (или отклонения от среднего) для набора данных являются дисперсия выборки, стандартное отклонение выборки и интервал.

Обсудим *дисперсию выборки и стандартное отклонение выборки* вместе. Дисперсия выборки  $s^2$  определяется по следующей формуле:

$$s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2.$$

Дисперсию выборки можно рассматривать как среднеквадратическое отклонение данных от среднего. Интуитивно кажется, что для вычисления истинного среднеквадратического отклонения следует делить на  $n$ , но по техническим причинам необходимо делить на  $n - 1$ .

Деление суммы квадратов отклонений на  $n - 1$  гарантирует, что дисперсия выборки является несмещенной мерой истинной дисперсии совокупности, из которой выбраны данные.

Стандартное отклонение выборки  $s$  — это просто корень квадратный из  $s^2$ .

Далее приведен пример вычисления этих мер для трех чисел: 1, 3 и 5.

$$s^2 = \frac{1}{3-1}[(1-3)^2 + (3-3)^2 + (5-3)^2] = 4.$$

Очевидно, что:  $s = \sqrt{4} = 2$ .

В примере с акциями стандартное отклонение выборки ежемесячных доходов по акциям Cisco составляет 12,2% с дисперсией выборки 0,015%<sup>2</sup>. Естественно, %<sup>2</sup> трудно интерпретировать, поэтому, как правило, мы рассматриваем стандартное отклонение выборки. Для акций GM стандартное отклонение выборки составляет 8,97%.

В версиях Excel 2007 и ранее дисперсия выборки для набора данных вычислялась по функции ДИСП (VAR), а стандартное отклонение выборки — по функции СТАНДОТКЛОН (STDEV). Эти функции по-прежнему можно использовать в Excel 2019, но начиная с Excel 2010 были также добавлены равноценные функции ДИСП.В (VAR.S) и СТАНДОТКЛОН.В (STDEV.P). (В расшифровывается как *выборка*.) Функции ДИСП.Г (VAR.P) и СТАНДОТКЛОН.Г (STDEV.P) вычисляют дисперсию и стандартное отклонение по генеральной совокупности. Для вычисления дисперсии или стандартного отклонения по генеральной совокупности просто замените  $n - 1$  на  $n$  в знаменателе формулы вычисления  $s^2$ .

*Интервал разброса данных* — это наибольшее число в наборе данных минус наименьшее число. Здесь интервал для ежемесячной доходности акций Cisco составляет 54%, а интервал для ежемесячной доходности акций GM — 52%.

### ❓ Как среднее значение и стандартное отклонение характеризуют набор данных?

Предположим, что гистограмма отражает *Гауссово*, или *нормальное*, распределение; тогда для такой гистограммы *эмпирическое* правило (основанное на соответствующих математических правилах) гласит:

- приблизительно 68% всех наблюдений приходятся на интервал между  $x - s$  и  $x + s$ ;
- приблизительно 95% всех наблюдений приходятся на интервал между  $x - 2s$  и  $x + 2s$ ;

О приблизительно 99,7% всех наблюдений приходится на интервал между  $x - 3s$  и  $x + 3s$ .

Например, можно ожидать, что приблизительно 95% всех ежемесячных доходов по акциям Cisco составят от  $-19$  до  $30\%$ , как показано ниже:

среднее  $- 2s = 0,056 - 2 \times (0,122) = -19\%$  и среднее  $+ 2s = 0,056 + 2 \times (0,122) = 30\%$ .

Любое наблюдение, отличающееся более чем на два стандартных отклонения от среднего, называется *выбросом*. Для данных Cisco из 130 наблюдений выбросами являются 9 (или примерно 7% всех доходов). Обычно эмпирическое правило дает меньшую точность для наборов данных со скошенным распределением, но в целом оно достаточно точно для относительно симметричных наборов данных, даже если данные не относятся к нормально распределенной совокупности.

Часто ценную информацию можно получить, обнаружив причины выбросов. Компании должны быть заинтересованы в том, чтобы «благоприятные» выбросы возникали более часто, а «неблагоприятные» — менее часто.

## Выделение выбросов с помощью условного форматирования

Как показывает опыт, в наборе данных часто требуется выделить все выбросы. Пример такого выделения представлен на рис. 44.4. (См. в файле `Stock.xlsx` лист `Stockprices`.)

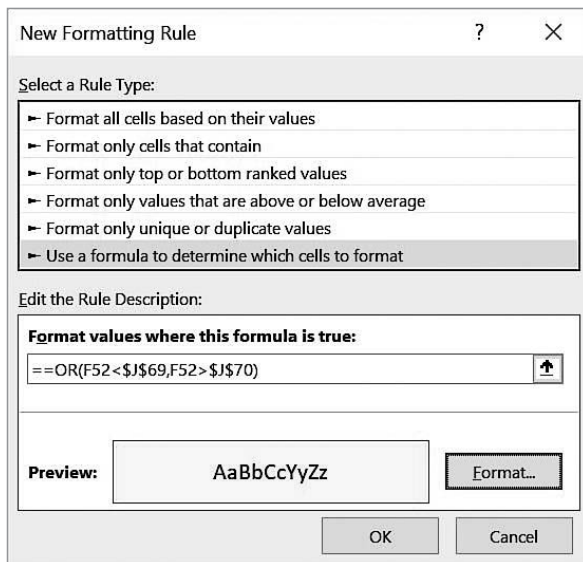
Например, для выделения выбросов в данных Cisco сначала вычислите нижний предел для выброса (среднее  $- 2s$ ) в ячейке `J69` и верхний предел (среднее  $+ 2s$ ) в ячейке `J70`. Затем выделите весь диапазон доходов Cisco (ячейки `F52:F181`) начиная с ячейки `F52`. Далее на вкладке Главная (Home) в группе Стили (Styles) в списке Условное форматирование (Conditional Formatting) выберите Создать правило (New Rule). В диалоговом окне Создание правила форматирования (New Formatting Rule) выберите тип правила Использовать формулу для определения форматируемых ячеек (Use a formula...) и затем заполните остальную часть диалогового окна, как показано на рис. 44.5: в поле Форматировать значения, для которых следующая формула является истинной (Format Values Where This Formula Is True) введите формулу `=ИЛИ(F52<=$J$69,F52>=$J$70)`.

Введенное условие гарантирует, что если значение в ячейке `F52` на  $2s$  больше или меньше среднего ежемесячного дохода по акциям Cisco, выбранный формат (в данном случае красный шрифт) будет применен к ячейке `F52`. Условие форматирования автоматически копируется в выделенный диапазон, и все выбросы будут представлены красным шрифтом.



	E	F
48		
49	-0.240	-0.203
50	0.277	0.339
51	GM	CSCO
52	0.022	0.011
53	-0.035	0.011
54	0.116	0.042
55	-0.021	0.071
56	-0.021	-0.038
57	-0.131	-0.029
58	-0.088	-0.091
59	0.014	0.311
60	0.014	0.339
61	-0.058	0.136
62	0.055	0.304
63	0.101	-0.043
64	-0.044	-0.129

**Рис. 44.4.** Выбросы для доходности акций Cisco выделены с помощью условного форматирования



**Рис. 44.5.** Правило условного форматирования для выделения выбросов

### ? Как с помощью описательной статистики сравнить наборы данных?

Описательная статистика помогает обобщить информацию о различиях между наборами данных, например, между ежемесячной доходностью акций Cisco и GM. На основе мер уровня среднего и разброса можно сделать следующие выводы.

- Обычно (если судить по среднему и медиане), ежемесячная доходность акций Cisco выше, чем доходность акций GM.
- Ежемесячные доходы по акциям Cisco более изменчивы (если судить по стандартному отклонению, дисперсии и интервалу), чем доходы по акциям GM.
- И доходность акций Cisco, и доходность акций GM демонстрирует небольшую положительную асимметрию. Ежемесячная доходность акций GM имеет более острый максимум, чем нормальная кривая, в то время как ежемесячная доходность акций Cisco имеет менее острый максимум, чем нормальная кривая.

Далее в этой главе я покажу, как использовать ящики с усами для визуального сравнения наборов данных.

### ? Как для заданной точки данных можно найти величину ее процентиля в наборе данных? Например, как найти девяностый процентиль набора данных?

До Excel 2010 для определения относительного положения наблюдения в наборе данных применялись функции ПРОЦЕНТИЛЬ (PERCENTILE) и ПРОЦЕНТРАНГ

(PERCENTRANK). В Excel 2010 добавилось четыре новых функции: ПРОЦЕНТИЛЬ.ВКЛ (PERCENTILE.INC), ПРОЦЕНТИЛЬ.ИСКЛ (PERCENTILE.EXC), ПРОЦЕНТРАНГ.ВКЛ (PERCENTRANK.INC) и ПРОЦЕНТРАНГ.ИСКЛ (PERCENTRANK.EXC). Функции ПРОЦЕНТИЛЬ.ВКЛ и ПРОЦЕНТРАНГ.ВКЛ возвращают результаты, идентичные старым функциям ПРОЦЕНТИЛЬ и ПРОЦЕНТРАНГ. Обратите внимание, что в предыдущих версиях Excel эти новые функции не распознаются. Примеры работы новых функций представлены в файле Percentile.xlsx и на рис. 44.6.

	C	D	E	F	G	H
2		<b>RANK</b>	<b>RANK</b>			
3	<b>Data</b>	<b>EXC</b>	<b>INC</b>			
4	<b>10</b>	<b>0.062</b>	<b>0</b>	<b>Percentile</b>	<b>EXC</b>	<b>INC</b>
5	<b>20</b>	<b>0.125</b>	<b>0.071</b>	<b>0.1</b>	<b>16</b>	<b>24</b>
6	<b>30</b>	<b>0.187</b>	<b>0.142</b>	<b>0.2</b>	<b>32</b>	<b>38</b>
7	<b>40</b>	<b>0.25</b>	<b>0.214</b>	<b>0.3</b>	<b>48</b>	<b>52</b>
8	<b>50</b>	<b>0.312</b>	<b>0.285</b>	<b>0.4</b>	<b>64</b>	<b>66</b>
9	<b>60</b>	<b>0.375</b>	<b>0.357</b>	<b>0.5</b>	<b>80</b>	<b>80</b>
10	<b>70</b>	<b>0.437</b>	<b>0.428</b>	<b>0.6</b>	<b>96</b>	<b>94</b>
11	<b>80</b>	<b>0.5</b>	<b>0.5</b>	<b>0.7</b>	<b>112</b>	<b>108</b>
12	<b>90</b>	<b>0.562</b>	<b>0.571</b>	<b>0.8</b>	<b>128</b>	<b>122</b>
13	<b>100</b>	<b>0.625</b>	<b>0.642</b>	<b>0.9</b>	<b>204</b>	<b>136</b>
14	<b>110</b>	<b>0.687</b>	<b>0.714</b>			
15	<b>120</b>	<b>0.75</b>	<b>0.785</b>			
16	<b>130</b>	<b>0.812</b>	<b>0.857</b>			
17	<b>140</b>	<b>0.875</b>	<b>0.928</b>			
18	<b>300</b>	<b>0.937</b>	<b>1</b>			

**Рис. 44.6.** Пример использования функций ПРОЦЕНТИЛЬ и ПРОЦЕНТРАНГ

Функции ПРОЦЕНТИЛЬ, ПРОЦЕНТИЛЬ.ВКЛ и ПРОЦЕНТИЛЬ.ИСКЛ возвращают процентиль для указанного набора данных. Синтаксис у всех этих функций такой же, как у функции ПРОЦЕНТИЛЬ.ВКЛ(массив;k), которая возвращает  $k$ -й процентиль для информации в диапазоне ячеек, указанном в аргументе массив.

Рассмотрим набор данных, состоящий из  $n$  элементов. Функции ПРОЦЕНТИЛЬ и ПРОЦЕНТИЛЬ.ВКЛ возвращали  $p$ -й процентиль ( $0 < p < 1$ ) как элемент с рангом  $1 + (n - 1) \times p$  в наборе данных. Например, в ячейке H13 формула =ПРОЦЕНТИЛЬ.ВКЛ(C4:C18;F13) вычисляет 90-й процентиль для данных в ячейках C4:C18 как  $1 + (15 - 1) \times (0,9)$ , что в результате дает элемент с рангом 13,6. Таким образом, при условии, что данные отсортированы по возрастанию, вычисляется число, находящееся между тринадцатым (130) и четырнадцатым элементом данных (140) с учетом 60%. В результате получается 136.

Функция ПРОЦЕНТИЛЬ.ИСКЛ вычисляет  $k$ -й процентиль как элемент с рангом  $(n + 1) \times p$  в наборе данных. Функция ПРОЦЕНТИЛЬ.ИСКЛ вычисляет 90-й процен-

тиль для данных как  $(15 + 1) \times (0,9)$ , что в результате дает элемент с рангом 14,4. Таким образом, 90-й процентиль (при условии, что данные отсортированы по возрастанию) вычисляется как число между четырнадцатым (140) и пятнадцатым элементом данных (300) с учетом 40%. В результате получается  $(0,60) \times (140) + (0,40) \times (300) = 204$ . Как мы видим, эти две функции возвращают совершенно разные результаты. Если считать, что имеющиеся данные были отобраны из большого набора данных, следует предположить, что элемент данных может быть больше 136 с вероятностью, намного превышающей 10%. В конце концов, 2 из 15 элементов данных больше 130, поэтому не кажется разумным полагать, что 90-й процентиль для этих данных равен 136. Следовательно, 90-й процентиль, равный 204, может оказаться более приемлемым. Я крайне рекомендую выполнять вычисления с помощью функций с окончаниями **.ИСКЛ**, а не с окончаниями **.ВКЛ**. Обратите внимание, что функции с **.ИСКЛ** не вычисляют процентиль для 0 и 1. Окончание **.ИСКЛ** означает, что функция **ПРОЦЕНТИЛЬ.ИСКЛ** исключает нулевой и сотый процентиля.

Функции **ПРОЦЕНТРАНГ**, **ПРОЦЕНТРАНГ.ВКЛ** и **ПРОЦЕНТРАНГ.ИСКЛ** возвращают ранг наблюдения относительно всех значений в наборе данных. Синтаксис на примере одной из функций: **ПРОЦЕНТРАНГ.ИСКЛ(массив;значение)**. Функции **ПРОЦЕНТРАНГ** и **ПРОЦЕНТРАНГ.ВКЛ** вычисляют процентильный ранг  $k$ -го наименьшего числа в наборе данных как  $(k - 1)/(n - 1)$ . Таким образом, как показано в ячейке E4, функция **ПРОЦЕНТРАНГ** или **ПРОЦЕНТРАНГ.ВКЛ** вычисляет ранг 0 для 10, поскольку для этой точки данных  $k = 1$ . Функция **ПРОЦЕНТРАНГ.ИСКЛ** вычисляет ранг  $k$ -й наименьшей точки данных как  $k/(n + 1)$ . В ячейке D4 функция **ПРОЦЕНТРАНГ.ИСКЛ** возвращает ранг  $1/16 = 0,0625$ . Процентильный ранг 6,25% кажется более реалистичным, чем ранг 0%, поскольку нет причин полагать, что значение 10 является наименьшей точкой данных в большом наборе данных, из которого была сделана эта выборка.

## ПРИМЕЧАНИЕ

Функции **ПРОЦЕНТИЛЬ** и **ПРОЦЕНТРАНГ** легко перепутать. Упрощенно говоря, функция **ПРОЦЕНТИЛЬ** возвращает значение данных, а функция **ПРОЦЕНТРАНГ** возвращает процентное отношение.

## ? Как найти второе наибольшее или второе наименьшее число в наборе данных?

Формула **=НАИБОЛЬШИЙ(массив;k)** возвращает  $k$ -е наибольшее число в диапазоне ячеек.  $k$ -е наименьшее число в диапазоне ячеек вычисляет формула **=НАИМЕНЬШИЙ(массив;k)**. Например, в файле **Trimmean.xlsx** в ячейке H1 формула **=НАИБОЛЬШИЙ(C4:C62;2)** вычисляет второе наибольшее число в диапазоне C4:C62 (99), а в ячейке H2 формула **=НАИМЕНЬШИЙ(C4:C62;2)** вычисляет второе наименьшее число в диапазоне C4:C62 (80) — рис. 44.7.

	C	D	E	F	G	H
1					2nd highest score	99
2					2nd lowest score	80
3	Scores	Rank(ties)	Average Rank		10% trimmed mean	90.04
4	93	20	21.5		5% trimmed mean	90.02
5	84	48	49			
6	88	38	39			
7	100	1	1			
8	86	45	45.5			
9	86	45	45.5		93 is 20th-23rd	
10	95	12	14		94 17th-19th	
11	92	24	24.5			
12	88	38	39			
13	94	17	18			
14	97	5	6.5			
15	91	26	27			
16	92	24	24.5			
17	95	12	14			
18	93	20	21.5			
19	80	56	57.5			
20	89	32	34.5			
21	98	3	3.5			
22	98	3	3.5			
23	90	29	30			

**Рис. 44.7.** Примеры функций НАИБОЛЬШИЙ, НАИМЕНЬШИЙ, РАНГ, РАНГ.СР и УРЕЗСРЕДНЕЕ

### ? Как можно ранжировать числа в наборе данных?

Функция РАНГ (RANK) ранжирует числа в наборе данных. Синтаксис этой функции: РАНГ(число;массив;0). В Excel 2010 была введена функция РАНГ.РВ (RANK.EQ), возвращающая результаты, аналогичные результатам функции РАНГ. Эта функция вычисляет ранг числа в заданном массиве, в котором наибольшему числу присвоен ранг 1, второму наибольшему числу — ранг 2 и т. д. По формуле =РАНГ(число;массив;1) или =РАНГ.РВ(число;массив;1) ранг 1 присваивается наименьшему числу в массиве, ранг 2 — второму наименьшему числу и т. д. В файле Trimmean.xlsx (см. рис. 44.7) по формуле=РАНГ.РВ(C4;\$C\$4:\$C\$62;0), скопированной из D4 в D5:D62, вычисляется ранг каждого балла. Например, 100 баллов в ячейке C7 являются наибольшим значением, а 98 баллов, повторяющиеся в ячейках C21 и C22, — третьим по величине значением. Обратите внимание, что функция РАНГ.РВ возвратила ранг 3 для обоих значений 98 баллов.

Функция РАНГ.СР (RANK.AVG) имеет тот же синтаксис, что и другие функции РАНГ, но в случае повторов функция РАНГ.СР возвращает средний ранг для всех повторяющихся точек данных. Например, поскольку 98 баллов имеют ранг третий и четвертый, функция РАНГ.СР возвращает ранг 3,5 для каждого. Средние ранги я вычислил по формуле =РАНГ.СР(C4;\$C\$4:\$C\$62;0), скопировав ее из E4 в E5:E62.

### ? Что такое урезанное среднее набора данных?

Чрезмерная асимметрия набора данных может исказить среднее значение набора данных. В таких ситуациях, как правило, в качестве меры среднего уровня ис-

пользуется медиана. Однако на медиану многие изменения данных не оказывают влияния. Например, сравните следующие два набора данных:

- набор 1: -5, -3, 0, 1, 3, 5, 7, 9, 11, 13, 15;
- набор 2: -20, -18, -15, -10, -8, 5, 6, 7, 8, 9, 10.

Эти наборы данных имеют одинаковые медианы (5), но набор 2 должен иметь более низкий уровень среднего, чем первый. Функция **УРЕЗСРЕДНЕЕ** отбрасывает наибольшие и наименьшие значения из набора данных. Функция **УРЕЗСРЕДНЕЕ (TRIMMEAN)** менее подвержена искажениям экстремальными значениями, чем функция **СРЗНАЧ (AVERAGE)**, но она больше зависит от экстремальных значений, чем медиана. По формуле **=УРЕЗСРЕДНЕЕ(массив;доля)** вычисляется среднее набора данных после отбрасывания половины доли наибольших и половины доли наименьших значений точек данных. Например, функция **УРЕЗСРЕДНЕЕ** с долей 10% пересчитывает среднее после удаления 5% наибольших и 5% наименьших данных. В файле *Trimmean.xlsx* в ячейке Н3 формула **=УРЕЗСРЕДНЕЕ(С4:С62;0,10)** вычисляет среднее для баллов в ячейках С4:С62 после отбрасывания трех наибольших и трех наименьших значений. (Урезанное среднее равно 90,04.) В ячейке Н4 по формуле **=УРЕЗСРЕДНЕЕ(С4:С62;0,05)** вычисляется среднее (90,02) для баллов в ячейках С4:С62 после отбрасывания наибольшего и наименьшего значения. Отбрасываются два наблюдения, поскольку  $0,05 \times 59 = 2,95$  означает удаление 1,48 наибольших наблюдений и 1,48 наименьших. Округление 1,48 приводит к отбрасыванию только наибольшего и наименьшего значения (см. рис. 44.7).

**?** Существует ли быстрый способ получить для выбранного диапазона ячеек различные статистики, описывающие данные в этих ячейках?

Для решения проблемы выделите в файле *Trimmean.xlsx* диапазон С4:С8. В правом нижнем углу экрана в строке состояния Excel есть статистика, отображающая числа в выделенном диапазоне ячеек (рис. 44.8). Например, в диапазоне ячеек С4:С8 содержатся пять чисел, среднее значение равно 90,2, наименьшее значение равно 84, а наибольшее — 100. Отображаемый набор статистик можно изменить, если щелкнуть правой кнопкой мыши по строке состояния. Выберите, например, **Минимум (Minimum)** и **Максимум (Maximum)**, и эти значения добавятся к статистикам по умолчанию, представленным на рис. 44.8.

Average: 90.2	Count: 5	Min: 84	Max: 100	Sum: 451
---------------	----------	---------	----------	----------

**Рис. 44.8.** Статистики в строке состояния

**?** Почему финансовые аналитики часто определяют средний доход по акциям с помощью среднего геометрического?

Файл *Geommean.xlsx* содержит годовые доходы по двум фиктивным акциям (рис. 44.9).

	B	C	D	E
4		Stock 1	Stock 2	
5	Year 1	0.05	-0.5	
6	Year 2	0.05	0.7	
7	Year 3	0.05	-0.5	
8	Year 4	0.05	0.7	
9	Average	0.05	0.1	
10				
11		1+return		
12		1.05	0.5	
13		1.05	1.7	
14		1.05	0.5	
15		1.05	1.7	
16	geometric means	0.05	-0.07805	=GEOMEAN(D12:D15)-1

**Рис. 44.9.** Среднее геометрическое

В ячейках C9 и D9 указаны средние годовые доходы: по акции 1 — 5% и по акции 2 — 10%. Казалось бы, акция 2 является более выгодной инвестицией. Однако если внимательно посмотреть, с акцией 2 в первый год вы теряете 50% и на следующий год получаете 70%. Это означает, что каждые два года \$1,00 превращается в  $1 \times (1,7) \times (0,5) = \$0,85$ . Поскольку с акцией 1 вы никогда не теряете в деньгах, очевидно, что она является более выгодной инвестицией. Среднее геометрическое как мера среднего годового дохода позволяет сделать правильный вывод о выгоде инвестирования в первую акцию. Среднее геометрическое  $n$  чисел — это просто корень  $n$ -й степени из произведения этих чисел (центральное число в геометрической прогрессии, в которой вы перемножаете входящие в нее числа и потом извлекаете квадратный корень, если это были два числа; кубический, если это были три числа, и т. д.). Например, среднее геометрическое для 1 и 4 равно квадратному корню из 4, то есть 2, а среднее геометрическое для 1, 2 и 4 равно корню кубическому из 8 (также 2). Для использования среднего геометрического при вычислении среднего годового дохода от инвестиций необходимо добавить 1 к каждому годовому доходу и взять среднее геометрическое от получившихся в результате чисел. Затем вычесть 1 из результата и получить оценку среднего годового дохода по акции.

Функция СРГЕОМ (GEOMEAN) возвращает среднее геометрическое для чисел в заданном диапазоне. Таким образом, для оценки среднего годового дохода по каждой акции необходимо выполнить следующее:

1. Вычислить каждую сумму  $(1 + \text{годовой доход})$  путем копирования формулы  $=1+C5$  из C12 в C12:D15.
2. Скопировать формулу  $=\text{СРГЕОМ}(C12:C15)-1$  из C16 в D16.

Средний годовой доход по первой акции оценивается в 5%, а средний годовой доход по второй акции — в -7,8%. Обратите внимание, что если вторая акция приносит средний доход -7,8% два года подряд, \$1 превращается в  $1 \times (1 - 0,078) \times 2 = \$0,85$ .

❓ **Как использовать ящики с усами (boxplots) для характеристики и сравнения наборов данных?**

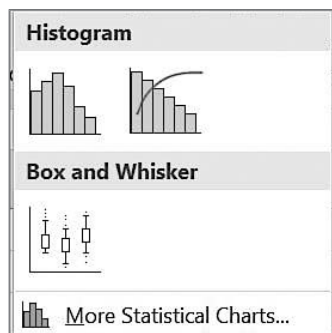
В главе 43 говорилось о том, что *гистограмма* — это диаграмма, показывающая частоту попадания значений данных в различные диапазоны. *Ящик с усами* (ящичковая диаграмма) — это диаграмма, графически представляющая пять важных характеристик множества данных. Это следующие характеристики:

- минимальное значение;
- максимальное значение;
- первый квартиль (25-й процентиль);
- медиана (50-й процентиль);
- третий квартиль (75-й процентиль).

Пример ящика с усами приведен на рис. 44.11. Длина ящика в ящичковой диаграмме — это интерквартильный диапазон (interquartile range, IQR) = 75-й процентиль минус 25-й процентиль.

Используя данные из файла *Boxplottemp.xlsx* (в папке *Templates*), создадим ящичковую диаграмму для анализа классического набора данных, представляющих собой номера, выпавшие в лотерею по выбору призывников на службу в вооруженные силы США в 1969 г. Для проведения лотереи 1969 г. в контейнер были помещены номера 1–366, которые были затем предположительно полностью перемешаны. Первый выпавший шар соответствовал первому января (номер 305), следующий — второму января (номер 159) и т. д.

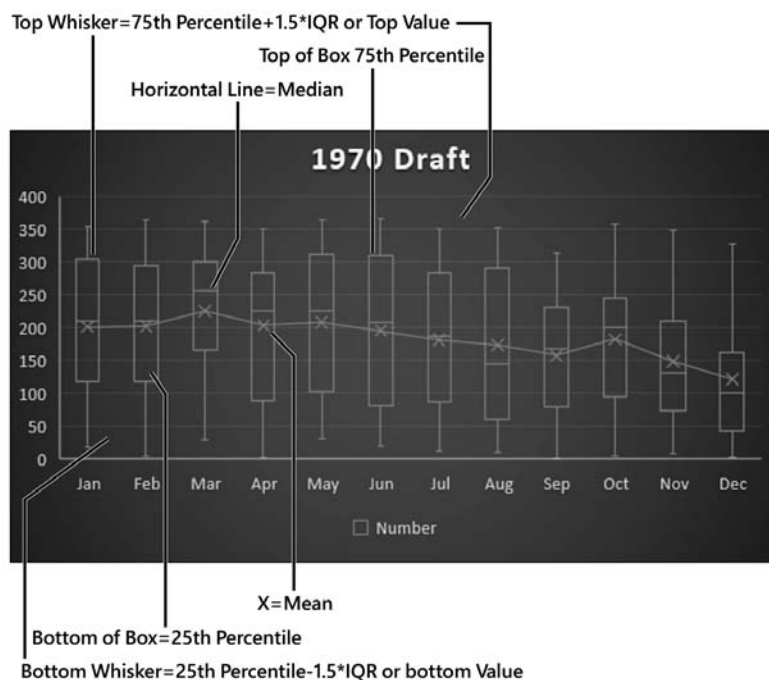
Мужчины с датой рождения, соответствующей номеру 1, были призваны первыми, затем мужчины с призывным номером 2 и т. д. Чем ниже был призывной номер, тем больше была вероятность, что мужчина будет призван. В диапазоне ячеек **A7:B373** содержатся номера военной лотереи для каждой даты рождения (столбец **B**) и месяца года (столбец **A**). Выделив этот диапазон, на вкладке **Вставка** (**Insert**) выберите **Вставить диаграмму статистики** (**Insert Statistic Chart**) (в группе **Диаграммы** (**Charts**)) и выберите опцию **Ящик с усами** (**Box And Whisker**), как показано на рис. 44.10.



**Рис. 44.10.** Выбор диаграммы Ящик с усами



Для создания ящичной диаграммы, приведенной на рис. 44.11, выберите на вкладке Конструктор (Design) четвертую опцию в группе Стили диаграмм (Chart Styles) — с черным фоном. Затем наведите мышь на область построения диаграммы и, когда высветится подсказка Ряд "Number", щелкните правой кнопкой мыши и выберите Формат ряда данных (Format Data Series). В панели Формат ряда данных (Format Data Series) установите флажок Показать точки выбросов (Show Outlier Points), так как любое значение, отстоящее более чем на  $1,5 \cdot (IQR)$  от любого из краев ящика, объявляется выбросом. Затем установите флажки Показать средние метки (Show Mean Markers) (значки X на диаграмме) и Показать среднюю линию (Show Mean Line) (линия, соединяющая значки X).



**Рис. 44.11.** Ящичная диаграмма для данных призывной лотереи

Для призывных номеров каждого месяца представлена следующая информация (на рис. 44.12 приведены расчеты за январь):

- Верх ящика (305) — 75-й процентиль.
- Низ ящика (118) — 25-й процентиль.
- Горизонтальная линия (211) — медиана, или 50-й процентиль.
- Верхнее значение верхнего уса (355) — максимум (максимальное значение, 75-й процентиль +  $1,5 \cdot IQR$ ).



- Нижнее значение нижнего уса (17) — минимум (минимальное значение, 25-й процентиль –  $1,5 \cdot \text{IQR}$ ).

	D	E	F	G
1				
2	Jan 1 was 305th ball chosensn			
3	Jan 2 Was 159th ball chosen etc.			
4				
5				
6				
7				
8				
9	Jan Median	211	=MEDIAN(Jan)	
10	Jan Mean	201.1613	=AVERAGE(Jan)	
11	Jan 25%ile	118	=QUARTILE.EXC(Jan,1)	
12	Jan 75%ile	305	=QUARTILE.EXC(Jan,3)	
13	Jan max	355	=MAX(Jan)	
14	Jan min	17	=MIN(Jan)	
15				
16	IQR	187	=F12-F11	
17	1.5*IQR	280.5	=1.5*F16	
18	Upper Outlier Cutoff	585.5	=F12+F17	
19	Lower Outlier Cutoff	-162.5	=F11-F17	

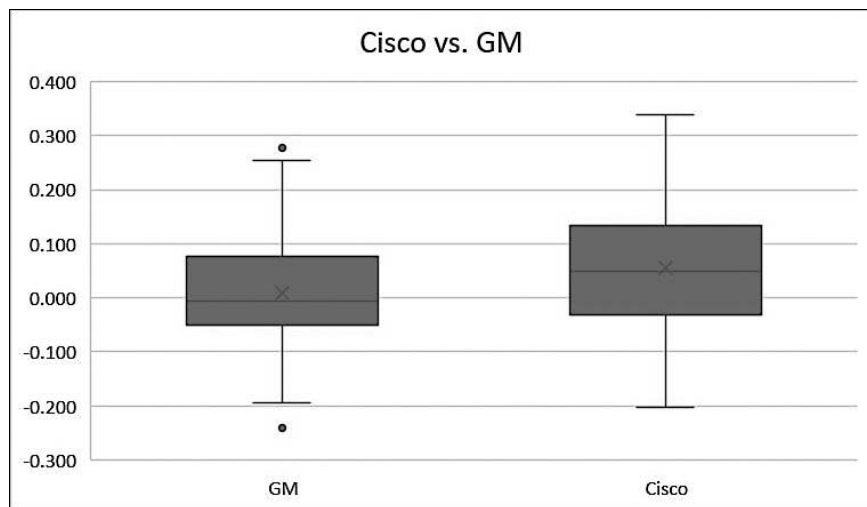
**Рис. 44.12.** Расчеты для ящика с данными за январь

Главный вывод, который можно сделать, глядя на эту ящичную диаграмму, — что средние и медианы убывают в течение календарного года. Это означает, что более поздние даты имеют тенденцию иметь меньшие призывные номера, и выбрасывание чисел не было случайным. Более мощные статистические методы, такие как повторная выборка (см. главу 81), подтверждают, что лотерея проводилась с нарушением случайности выборки. Наиболее распространенное объяснение этого нарушения случайности в выборке — что шары с меньшими номерами были положены первыми и затем шары были плохо перемешаны. Это привело к тому, что более поздние даты получили меньшие призывные номера. В 1970 г. шары были более тщательно перемешаны и нарушений случайности выявлено не было.

В качестве второго примера мощи ящичковых диаграмм на рис. 44.13 (см. файл *Stocksandboxplots.xlsx*) приведена диаграмма, позволяющая сравнить доходность акций компаний Cisco и GM, которые обсуждались ранее в этой главе.

Эта диаграмма позволяет сразу же сделать следующие три вывода:

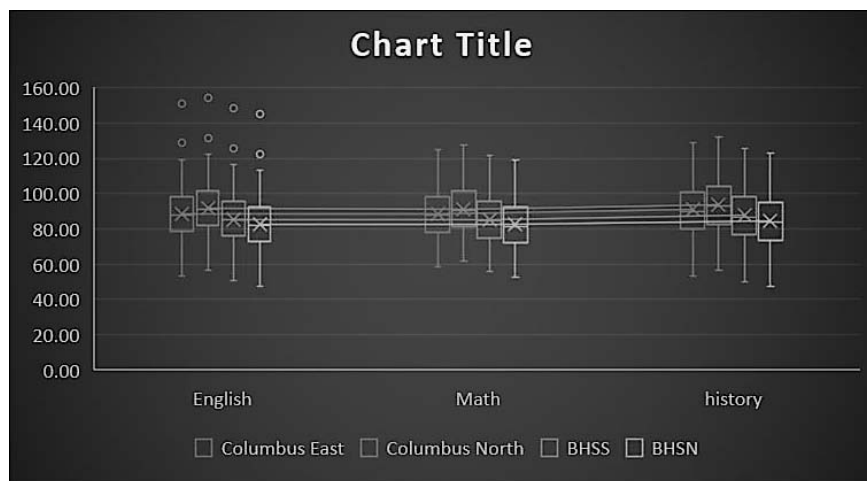
- Ящик с данными по Cisco расположен выше, чем ящик с данными по GM, следовательно, доходы по Cisco в среднем выше, чем по GM.
- Ящик с данными по Cisco имеет большую длину (выше), чем ящик с данными по GM, и усы в данных по Cisco длиннее, чем усы в данных по GM, следовательно, данные по Cisco имеют больший разброс, чем данные по GM.



**Рис. 44.13.** Ящичковая диаграмма, отражающая доходность акций Cisco и GM

- Верхние и нижние усы каждого из ящиков имеют приблизительно равную длину, а среднее и медиана практически совпадают. Это означает, что наборы данных GM и Cisco симметричны.

Файл `Boxplotmultiple.xlsx` (рис. 44.14) позволяет увидеть, как можно использовать ящичковые диаграммы для сравнения нескольких совокупностей данных, зависящих от нескольких переменных.



**Рис. 44.14.** Ящичковая диаграмма для сравнения экзаменационных баллов студентов

В столбцах представлены данные по университетам, соответствующие оценкам за экзамены по английскому языку, математике и истории. На диаграмме сразу же обращает на себя внимание следующее:

- Данные по каждому университету имеют два выброса (вверх) по баллам за экзамен по английскому языку.
- Студенты Columbus North в целом показали лучшие оценки по каждому из тестов, а студенты Bloomington North в целом показали худшие оценки по каждому из тестов.
- По каждому из экзаменов все университеты имеют ящики и усы сходных размеров, что свидетельствует о том, что разброс баллов по каждому из тестов одинаков для всех университетов.

## Задания

1. Используйте данные из файла **Stock.xlsx** для описательной статистики для акций Intel и GE.
2. Сравните ежемесячные доходы по акциям Intel и GE на основе ответа на задание 1.
3. Компания City Power&Light производит оборудование для регулировки напряжения в Нью-Йорке и отправляет его в Чикаго. Регулятор напряжения считается годным, если он поддерживает напряжение от 25 до 75 вольт. Напряжение, поддерживаемое каждым устройством, измеряется в Нью-Йорке перед отправкой. После доставки оборудования в Чикаго напряжение измеряется снова. Результаты замеров напряжения в каждом городе приведены в файле **Citypower.xlsx**.
  - С помощью описательной статистики прокомментируйте то, что вы узнали о поддерживаемом регуляторами напряжении до и после доставки.
  - Каков процент годных устройств до и после доставки?
  - Выскажите предложения по повышению качества регуляторов напряжения, производимых компанией City Power&Light.
  - Какое значение напряжения было превышено десятью процентами регуляторов при измерении в Нью-Йорке?
  - Какое значение напряжения не было превышено пятью процентами регуляторов при измерении в Нью-Йорке?
4. В файле **Decadeincome.xlsx** приведены примеры доходов (в тысячах долларов на 1980 г.) для нескольких семей в 1980 и 1990 гг. Предположим, что эти семьи являются репрезентативными для США. Республиканцы утверждают, что в 1990 г. страна стала богаче, чем в 1980 г., поскольку средний доход семьи вырос. Согласны ли вы с этим утверждением?

5. С помощью описательной статистики сравните годовые доходы по акциям, казначейским векселям и корпоративным облигациям. Данные см. в файле `Historicalinvest.xlsx`.
6. В 1969 и 1970 г. комплектование вооруженных сил США проводилось на основе лотереи. Лотерейный номер определялся по дню рождения. В общей сложности 366 шаров, по одному на каждую возможную дату рождения, поместили в контейнер и перемешали. Первый выбранный шар получил в лотерее номер 1, второй шар — номер 2 и т. д. Мужчины, чьи дни рождения соответствовали меньшим номерам, были призваны первыми. В файле `Draftlottery.xlsx` содержатся фактические результаты лотерей 1969 и 1970 гг. Например, дата 1 января в лотерее 1969 г. получила номер 305. С помощью описательной статистики продемонстрируйте, что выбор шаров в лотерее набора на воинскую службу в 1969 г. не был случайным, а в 1970 г. — был случайным. (Подсказка: вычислите среднее и медиану лотерейного номера для каждого месяца с помощью функций `СРЗНАЧ` и `МЕДИАНА`.)
7. В файле `Jordan.xlsx` приведены начальные зарплаты (гипотетические) всех выпускников географического факультета 1984 г. Университета Северной Каролины. Какова, по вашему мнению, типичная начальная зарплата выпускника? В действительности в 1984 г. самая высокая средняя начальная зарплата была у географов, потому что великий баскетболист Майкл Джордан в университете специализировался в географии!
8. С помощью функции `НАИБОЛЬШИЙ (LARGE)` или `НАИМЕНЬШИЙ (SMALL)` отсортируйте ежегодные доходы по акциям в файле `Historicalinvest.xlsx`. В чем преимущество этого метода сортировки по сравнению с использованием кнопки `Сортировка (Sort)`?
9. Сравните среднее, медиану и урезанное среднее (с урезанием 10% данных) для акций, казначейских векселей и корпоративных облигаций в файле `Historicalinvest.xlsx`.
10. С помощью геометрического среднего оцените средний годовой доход по акциям, облигациям и казначейским векселям в файле `Historicalinvest.xlsx`.
11. В файле `Dow.xlsx` содержатся данные о ежемесячных доходах по 30 акциям крупнейших корпораций США за последние 20 лет. С помощью этих данных определите три акции с наибольшим средним ежемесячным доходом.
12. Для данных файла `Dow.xlsx` определите три акции с самым высоким риском или изменчивостью.
13. Для данных файла `Dow.xlsx` определите три акции с самой большой асимметрией.
14. Для данных файла `Dow.xlsx` определите, насколько урезанные средние доходы (с урезанием 10% доходов) отличаются от общих средних доходов?
15. В файле `Incomedata.xlsx` содержатся доходы репрезентативной выборки американцев в 1975, 1985, 1995 и 2005 г. Опишите, как изменились личные доходы американцев со временем.

16. В файле **Coltsdata.xlsx** содержится количество ярдов в засчитанных попытках выполнить вынос и пас, набранных командой Indianapolis Colts в 2006 г. Опишите различие итогов в игре на вынос и на пас.
17. В файле **Problem17datat.xlsx** представлены ежедневные доходы по акциям Facebook. Используя эти данные, ответьте на следующие вопросы:
- Имеют ли доходы по акциям Facebook значимую асимметрию?
  - Определите все выбросы (используя эмпирическое правило). Соответствует ли количество выбросов эмпирическому правилу?
  - Существует 1%-ная вероятность, что дневной доход по акциям Facebook превысит \_\_\_\_\_.
18. Одна из теорий науки о мозге гласит, что уровень дофамина в организме человека определяет, будет ли человек проявлять психотическое поведение. В файле **Problem18data.xlsx** приведены данные уровня дофамина для 10 взрослых с психотическими расстройствами и 14 взрослых без психотических расстройств. Используйте дескриптивную статистику и ящичную диаграмму для сравнения и противопоставления распределения дофамина у людей с психотическими расстройствами и без них.
19. В эксперименте, в котором участвовали 508 человек, у людей спрашивали, как они оценивают процент афроамериканцев в США. Перед тем как они отвечали на вопрос, перед ними прокручивали «колесо фортуны» с числами от 1 до 100. Они не знали, что колесо было настроено на выбрасывание значений 25 или 65. В файле **Problem19data.xlsx** приведены ответы участников эксперимента. Опишите, как результат, выброшенный колесом, повлиял на даваемые ответы.

## ГЛАВА 45

# Сводные таблицы и срезы для описания данных

### Обсуждаемые вопросы

- Что такое сводная таблица?
- Как с помощью сводной таблицы обобщить данные о продажах товаров в нескольких продуктовых магазинах?
- Какие макеты сводных таблиц доступны в Microsoft Excel 2019?
- Почему сводные таблицы называются *сводными*?
- Как быстро изменить формат в сводной таблице?
- Как свернуть и развернуть поля?
- Как сортировать и фильтровать поля сводной таблицы?
- Как обобщить данные в сводной таблице с помощью сводной диаграммы?
- Для чего предназначена область ФИЛЬТРЫ в сводной таблице?
- Как работают срезы сводной таблицы?
- Как добавить пустые строки или скрыть промежуточные итоги в сводной таблице?
- Как применить к сводной таблице Условное форматирование?
- Как обновить вычисления при добавлении новых данных?
- Я работаю в небольшой турфирме и должен разослать проспект о ней потенциальным клиентам. Мой бюджет ограничен, поэтому послать проспект необходимо тем людям, которые больше остальных тратят денег на путешествия. У меня есть случайная выборка из 925 человек, о которых известна следующая информация: возраст, пол и количество денег, потраченных на путешествия за прошлый год. Как с помощью этих данных определить влияние пола и возраста на расходы на путешествия? Какие выводы можно сделать о типе потенциального клиента, которому следует отправить проспект по почте?
- Я провожу маркетинговое исследование автомобилей-универсалов. Мне необходимо определить, какие факторы влияют на вероятность покупки семьей автомобиля-универсала. У меня есть большая выборка семей с информацией о ее размере (большая или нет) и доходе (высокий или низкий). Как определить влияние размера семьи и семейного дохода на вероятность покупки автомобиля-универсала?
- Я работаю в компании, продающей микрочипы по всему миру. Ежемесячно я получаю данные о фактических и прогнозируемых объемах продаж чипа 1, чипа 2

и чипа 3 в Канаде, Франции и США. Я также получаю расхождение, или разницу, между фактической выручкой и выручкой, заложенной в бюджет. Для каждого месяца и каждой комбинации страны и продукта мне нужно отобразить следующие данные: фактическая выручка, выручка, заложенная в бюджет, фактическое расхождение, фактическая выручка как процент от годового дохода и расхождение как процент от выручки, заложенной в бюджет. Каким образом можно отобразить эту информацию?

- Что такое вычисляемое поле?
- Что использовать: фильтр отчета или срез?
- Как сгруппировать элементы в сводной таблице?
- Что такое вычисляемый объект?
- Что такое детализация?
- Мне часто приходится использовать конкретные данные в сводной таблице для вычисления прибыли, например апрельских продаж чипа 1 во Франции. К сожалению, при добавлении новых данных в сводную таблицу эти перемещаются. Существует ли в Excel функция, позволяющая всегда извлекать из сводной таблицы объем апрельских продаж чипа 1 во Франции?
- Как применить временную шкалу для обобщения данных за разные периоды времени?
- Как с помощью сводной таблицы получить нарастающий итог общего объема продаж для каждого периода года?
- Как с помощью сводной таблицы сравнить объем продаж за месяц с объемом продаж за тот же месяц предыдущего года?
- Как создать сводную таблицу на основе данных из нескольких местоположений?
- Как создать сводную таблицу на основе уже имеющейся сводной таблицы?
- Как использовать фильтр отчета для создания нескольких сводных таблиц?
- Как быстро изменить настройки сводной таблицы по умолчанию?

## Ответы на вопросы

### ❓ Что такое сводная таблица?

Во многих ситуациях в сфере бизнеса требуется проанализировать данные или создать срез данных, для того чтобы сделать важные выводы. Представьте себе, что вы продаете различные продовольственные товары в различных магазинах в разное время. Вам нужно отслеживать тысячи точек данных. Обобщить такие данные так быстро, насколько это вообще возможно, позволяют сводные таблицы. Например, для данных по продуктовым магазинам с помощью сводной таблицы можно быстро определить следующее:

- сумму, оставленную покупателями за год в каждом магазине за каждый продукт;
- объем выручки для каждого магазина;
- объем выручки для каждого года.

Для турфирмы (еще один пример) можно создать срез данных, позволяющий определить, зависит ли средняя сумма, затрачиваемая на путешествия, от возраста, пола или обоих факторов. При анализе ситуации с покупкой автомобиля можно сравнить долю многодетных семей, покупающих автомобиль-универсал, с долей небольших семей, приобретающих такой автомобиль. Если вы производитель микрочипов, то сможете определить общий объем продаж чипа 1 во Франции, например, в апреле и т. д. Сводная таблица является очень эффективным инструментом для получения среза данных. Самый простой способ изучить сводные таблицы — разобрать несколько подробных примеров, так что за работу! Начнем с вводного примера, а затем рассмотрим многие дополнительные свойства сводных таблиц на последующих примерах.

### ❓ Как с помощью сводной таблицы обобщить данные о продажах товаров в нескольких продуктовых магазинах?

В файле `Groceriespttemp.xlsx` на листе `data` содержатся более 900 строк данных о продажах (рис. 45.1). Каждая строка содержит количество проданных единиц товара, выручку, полученную от продажи продукта в магазине, месяц и год продажи. Также указана группа продуктов (фрукты, молоко, хлопья или мороженое).

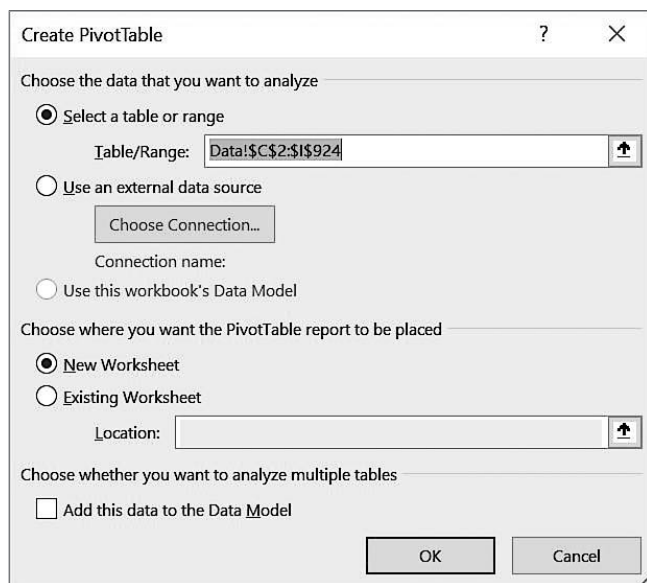
	C	D	E	F	G	H	I
2	Year	Month	Store	Group	Product	Units	Revenue
3	2007	August	south	milk	low fat	805	\$ 3,187.80
4	2007	March	south	ice cream	Edies	992	\$ 3,412.48
5	2007	January	east	milk	skim	712	\$ 1,808.48
6	2006	March	north	ice cream	Edies	904	\$ 2,260.00
7	2006	January	south	ice cream	Edies	647	\$ 2,076.87
8	2005	September	west	fruit	plums	739	\$ 1,707.09
9	2006	March	east	milk	low fat	974	\$ 2,181.76
10	2007	June	north	fruit	apples	615	\$ 1,894.20
11	2007	July	west	fruit	cherries	714	\$ 1,856.40
12	2006	May	south	cereal	Special K	703	\$ 1,553.63
13	2005	June	west	ice cream	Edies	528	\$ 2,064.48
14	2006	October	east	cereal	Raisin Bran	644	\$ 1,809.64
15	2005	June	south	fruit	grapes	919	\$ 2,196.41
16	2007	May	west	milk	skim	767	\$ 1,932.84
17	2007	June	west	cereal	Raisin Bran	984	\$ 1,987.68
18	2005	March	south	cereal	Raisin Bran	744	\$ 2,217.12
19	2007	September	east	ice cream	Edies	693	\$ 2,189.88
20	2006	October	east	milk	chocolate	658	\$ 1,895.04

**Рис. 45.1.** Данные для создания сводной таблицы по продовольственным товарам



Вы хотите получить разбивку данных о продажах за каждый год по каждой группе продуктов и каждому продукту в каждом магазине. Т акже вы хотите отобразить эту разбивку для любых указанных месяцев в заданном году (например, продажи с января по июнь).

Перед созданием сводной таблицы необходимо поместить заголовки в первую строку данных. Обратите внимание, что данные о продовольственных товарах содержат заголовки (Год, Месяц, Магазин, Группа, Продукт, Количество и Выручка) в строке 1. Установите курсор в какую-либо ячейку данных и затем на вкладке **Вставка (Insert)** в группе **Таблицы (Tables)** выберите **Сводная таблица (PivotTable)**. Откроется диалоговое окно **Создание сводной таблицы (Create PivotTable)**, показанное на рис. 45.2, с предполагаемым диапазоном данных. (В нашем случае автоматически был определен правильный диапазон данных **C2:I924**.) Положение переключателя **Использовать внешний источник данных (Use an external data source)** позволяет указать базу данных как источник данных для сводной таблицы. **Модель данных**, добавленная в Microsoft Excel 2013, подробно описана в главе 46 «Модель данных».



**Рис. 45.2.** Диалоговое окно Создание сводной таблицы

Мы помещаем сводную таблицу на новый лист. На текущем листе ее также можно разместить — для этого нужно выбрать верхний левый угол диапазона, в котором будет находиться сводная таблица. После нажатия **ОК** вы увидите панель **Поля сводной таблицы (PivotTable Fields)**, показанную на рис. 45.3.

Вы заполняете список полей сводной таблицы, перетаскивая заголовки сводной таблицы или полей в нужную область. Этот шаг нельзя опустить, так как он обес-



**Рис. 45.3.** Список полей сводной таблицы

печивает обобщение и отображение в том виде, который вам нужен. Вот описание всех четырех областей.

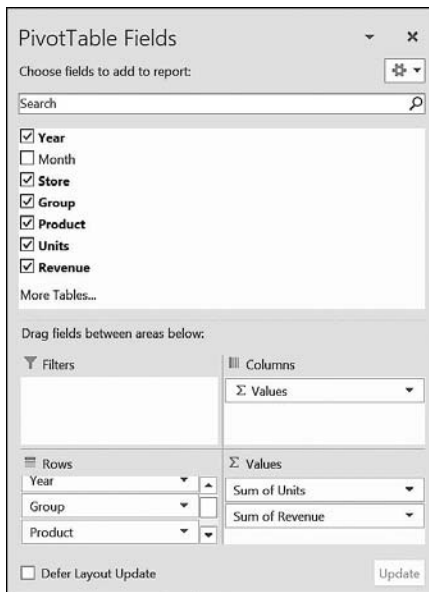
- **СТРОКИ (Rows).** Перетаскиваемые поля перечисляются с левой стороны таблицы в том порядке, в каком они добавлялись. Например, я перетащил поля **Год**, **Группа**, **Продукт** и **Магазин** в указанном порядке в область **СТРОКИ (Rows)**. В результате этого Excel обобщил данные сначала по годам, потом по каждой группе продуктов в заданном году, затем по продуктам в каждой группе и, наконец, по магазинам для каждого продукта. Вы можете в любое время перетащить поле в другую область или переупорядочить поля внутри области, перетащив поле вверх или вниз либо нажав стрелку справа от имени поля.
- **СТОЛБЦЫ (Columns).** Значения для полей, перетаскиваемых в эту область, перечисляются в верхней строке сводной таблицы. В нашем примере полей в области **СТОЛБЦЫ** нет.
- **ЗНАЧЕНИЯ (Values).** Значения для полей, перетаскиваемых в эту область, в таблице суммируются математически. Я перетащил в эту область поля **Количество** и **Выручка** (в таком порядке). Excel пытается угадать, какой тип расчета вы хотите выполнить в поле. В этом примере Excel считает, что поля **Количество** и **Выручка** нужно просуммировать, и не ошибается. Если метод расчета для поля данных необходимо заменить на другой, например **Среднее**, **Максимум** и т. д., щелкните на поле данных и выберите **Параметры полей значений (Value Field Settings)**. Примеры изменения параметров полей значений рассмотрены далее в этой главе, в задачах про автомобиль-универсал и компьютерные чипы.

- **ФИЛЬТРЫ (Filters).** Для полей, перетаскиваемых в эту область, можно выбрать любое подмножество значений, так что в сводной таблице будут отображены только расчеты на основе выбранного подмножества. В этом примере в область **ФИЛЬТРЫ (FILTERS)** перемещено поле **Месяц**. Теперь можно выбрать любое подмножество месяцев, например с января по июнь, и расчеты будут выполнены с учетом только этих месяцев. На сегодняшний день, с появлением срезов данных, фильтры отчетов практически вышли из употребления.

Заполненные области с полями сводной таблицы показаны на рис. 45.4. Получившая сводная таблица представлена на рис. 45.5 и на листе **All Rows Fields** в файле **Groceriespt.xlsx**. В строке 6 показано, что 243 228 единиц продукта были проданы за \$728 218,68 в 2005 г.

## ПРИМЕЧАНИЕ

Для просмотра списка полей необходимо поместить курсор в поле сводной таблицы. Если список на листе не виден, щелкните правой кнопкой мыши по любой ячейке в сводной таблице и выберите в контекстном меню команду **Показать список полей (Show Field List)**.



**Рис. 45.4.** Заполненный список полей в сводной таблице

	A	B	C
4	<b>Values</b>		
5	<b>Row Labels</b>	<b>Sum of Units</b>	<b>Sum of Revenue</b>
6	<b>2005</b>	<b>243228</b>	<b>728218.68</b>
7	<b>cereal</b>	<b>63689</b>	<b>192172.93</b>
8	Cheerios	11163	32993.1
9	west	4614	13732.83
10	south	3265	9216.47
11	north	1639	4988
12	east	1645	5055.8
13	<b>Raisin Bran</b>	<b>35797</b>	<b>105793.04</b>
14	west	12124	37401.27
15	south	8989	26069.35
16	north	8458	22822.23
17	east	6226	19500.19
18	<b>Special K</b>	<b>16729</b>	<b>53386.79</b>
19	west	6902	21092.75
20	south	3366	11339.47
21	north	4172	14956.36
22	east	2289	5998.21
23	<b>fruit</b>	<b>60047</b>	<b>182813.88</b>
24	<b>apples</b>	<b>14535</b>	<b>48127.74</b>
25	west	5255	15759.59
26	south	4317	14763.88

**Рис. 45.5.** Сводная таблица по продовольственным товарам в сжатой форме

## СОВЕТЫ

Далее приведено несколько советов по переходу между листами книги.

- Ctrl+PageUp — переход назад на один лист.
- Ctrl+PageDn — переход вперед на один лист.
- Если щелкнуть правой кнопкой мыши на стрелках слева от имени первого листа, появится список имен листов, из которого можно перейти на любой лист книги.

## ? Какие макеты сводных таблиц доступны в Microsoft Excel 2019?

Макет сводной таблицы, показанный на рис. 45.5, называется сжатой формой. В сжатой форме поля строк показываются одно над другим. Для изменения макета поместите курсор внутри таблицы, а затем на вкладке **Конструктор (Design)** в группе **Макет (Layout)** откройте список **Макет отчета (Report Layout)** и выберите один из следующих макетов: **Показать в сжатой форме (Show In Compact Form)**, как на рис. 45.5, **Показать в форме структуры (Show In Outline Form)**, как на рис. 45.6 и листе **Outline Form**, или **Показать в табличной форме (Show In Tabular Form)**, как на рис. 45.7 и листе **Tabular Form**.

	A	B	C	D	E	F
2	Month	(All)		Outline form		
3						
4					Values	
5	Year	Group	Product	Store	Sum of Units	Sum of Revenue
6	2005				243228	728218.68
7		cereal			63689	192172.93
8			Cheerios		11163	32993.1
9				west	4614	13732.83
10				south	3265	9216.47
11				north	1639	4988
12				east	1645	5055.8
13			Raisin Bran		35797	105793.04
14				west	12124	37401.27
15				south	8989	26069.35
16				north	8458	22822.23
17				east	6226	19500.19
18			Special K		16729	53386.79
19				west	6902	21092.75
20				south	3366	11339.47
21				north	4172	14956.36
22				east	2289	5998.21
23		fruit			60047	182813.88
24			apples		14535	48127.74
25				west	5255	15759.59
26				south	4317	14763.88

**Рис. 45.6.** Макет Показать в форме структуры

	A	B	C	D	E	F
2	Month	(All)		Tabular form		
3						
4					Values	
5	Year	Group	Product	Store	Sum of Units	Sum of Revenue
6	2005	cereal	Cheerios	west	4614	13732.83
7				south	3265	9216.47
8				north	1639	4988
9				east	1645	5055.8
10			Cheerios Total		11163	32993.1
11			Raisin Bran	west	12124	37401.27
12				south	8989	26069.35
13				north	8458	22822.23
14				east	6226	19500.19
15			Raisin Bran Total		35797	105793.04
16			Special K	west	6902	21092.75
17				south	3366	11339.47
18				north	4172	14956.36
19				east	2289	5998.21
20			Special K Total		16729	53386.79
21		cereal Total			63689	192172.93
22		fruit	apples	west	5255	15759.59
23				south	4317	14763.88
24				north	3734	13631.83

Рис. 45.7. Макет Показать в табличной форме

### ❓ Почему сводные таблицы называются сводными?

Сводная таблица (по-английски pivot — «вертушка») называется так потому, что ее поля можно быстро повернуть вокруг оси, то есть переместить из строки в столбец и наоборот, и создать другой макет. Например, после перетаскивания поля Год в область СТОЛБЦЫ (COLUMNS) получится макет сводной таблицы, показанный на рис. 45.8 (см. лист Years Column).

	A	B	C	D	E	F	G	H	I
2	Month	(All)							
3									
4									
5		Column Labels	2005	2006	2007	Total Sum of Units	Total Sum of Revenue		
6	Row Labels	Sum of Units	Sum of Revenue	Sum of Units	Sum of Revenue	Sum of Units	Sum of Revenue		
7	cereal	63689	192172.93	52489	150710	58671	172828.96	174849	515711.89
8	Cheerios	11163	32993.1	16142	46657.49	13652	38617.12	40957	118267.71
9	west	4614	13732.83	1454	4696.16	1586	4633.76	7654	23062.75
10	south	3265	9216.47	6424	18450.16	3064	8635.3	12753	36301.93
11	north	1639	4988	3027	10199.31	4207	11409.85	8873	26597.16
12	east	1645	5055.8	5237	13311.86	4795	13938.21	11677	32305.87
13	Raisin Bran	35797	105793.04	24056	69391.29	27715	81254.09	87568	256438.42
14	west	12124	37401.27	4515	14147.68	10115	27038.27	26754	78587.22
15	south	8989	26069.35	6015	16593.97	5329	16632.9	20333	59296.22
16	north	8458	22822.23	7505	23099.35	8366	25008.01	24329	70929.59
17	east	6226	19500.19	6021	15550.29	3905	12574.91	16152	47625.39
18	Special K	16729	53386.79	12291	34661.22	17304	52957.75	46324	141005.76
19	west	6902	21092.75	2585	7715.78	2328	6860.4	11815	35668.93
20	south	3366	11339.47	3436	9125.62	5826	17647.74	12628	38112.83
21	north	4172	14956.36	2570	8665.32	3789	12736.82	10531	36358.5
22	east	2289	5998.21	3700	9154.5	5361	15712.79	11350	30865.5

Рис. 45.8. Поле Год перенесено в область СТОЛБЦЫ

### ? Как быстро изменить формат в сводной таблице?

Если вы хотите изменить формат всего столбца, дважды щелкните на заголовке столбца и в появившемся диалоговом окне **Параметры поля значений (Value Field Settings)** нажмите **Числовой формат (Number Format)**. Затем примените желаемый формат и нажмите **OK** в обоих диалоговых окнах. Например, на листе **Formatted \$s** я назначил полю **Revenue (Выручка)** денежный формат, дважды щелкнув по заголовку **Sum Of Revenue (Сумма по полю Выручка)**. Также вы можете изменить формат поля значений следующим способом: на панели **Поля сводной таблицы (PivotTable Fields)** в области **ЗНАЧЕНИЯ (VALUE)** нажмите стрелку справа от поля значений. Выберите **Параметры полей значений (Value Field Settings)**, нажмите **Числовой формат (Number Format)** и переформатируйте столбец по своему желанию.

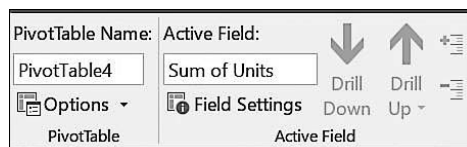
Из любой ячейки в сводной таблице можно открыть на ленте вкладку **Конструктор (Design)** и выбрать один из множества стилей сводной таблицы.

### ? Как свернуть и развернуть поля?

Сворачивание и разворачивание полей (функция, введенная в Microsoft Excel 2007) является большим преимуществом сводных таблиц. На рис. 45.5 можно увидеть знак «-» рядом с каждым годом, группой продуктов и продуктом. При нажатии на «-» поле сворачивается, и «-» заменяется на «+». При нажатии на «+» поле

	A	B	C
1	Month	(All)	
2			
3		Values	
4	Row Labels	Sum of Units	Sum of Revenue
5	2005	243228	728218.68
6	cereal	63689	192172.93
7	fruit	60047	182813.88
8	apples	14535	48127.74
9	east	1229	3972.44
10	north	3734	13631.83
11	south	4317	14763.88
12	west	5255	15759.59
13	cherries	11083	32042.39
14	east	1646	4051.22
15	north	3701	11087.14
16	south	3277	9092.92
17	west	2459	7811.11
18	grapes	20005	60126.15
19	east	4811	13052.68
20	north	4865	14698.63
21	south	6268	20474.65
22	west	4061	11900.19
23	plums	14424	42517.6
24	east	2216	7497.52
25	north	1515	5055.55

**Рис. 45.9.** Поле cereal свернуто



**Рис. 45.10.** Кнопки Развернуть поле и Свернуть поле

разворачивается. Например, если нажать на «-» рядом с группой продуктов **cereal** (хлопья) в любой из таких ячеек столбца **A**, то для каждого года информация о хлопьях будет сжата до одной строки, и конкретные продукты и магазины не будут показаны (рис. 45.9 и лист **Cerealcollapse**). Если вы нажмете на «+», в ячейке **A6** снова появится подробная информация о продажах по каждому виду хлопьев.

Аналогично можно развернуть или свернуть все поле. Перейдите в любую строку, содержащую элемент этого поля, и на вкладке **Анализ (Analyze)** в группе **Активное поле (Active Field)** нажмите **Развернуть поле (Expand Field)**, кнопка помечена зеленым знаком плюс) или **Свернуть поле (Collapse Field)**, кнопка помечена красным знаком минус) (рис. 45.10).

Пусть вам необходимо просмотреть только объемы продаж по группам продуктов за каждый год. Выберите любую ячейку, содержащую имя группы (например, **A6**), выберите вкладку **Инструменты анализа сводных таблиц (PivotTable Tools Analyze)** и в группе **Активное поле (Active Field)** нажмите **Свернуть поле (Collapse Field)**. Результат представлен на рис. 45.11 (и на листе **Groups Collapsed**). Нажав кнопку **Развернуть поле (Expand Field)**, вы вернете таблице прежний вид.

	A	B	C
1	Month	(All)	
2			
3	Values		
4	Row Labels	Sum of Units	Sum of Revenue
5	2005	243228	728218.68
6	cereal	63689	192172.93
7	fruit	60047	182813.88
8	ice cream	56518	174378.59
9	milk	62974	178853.28
10	2006	216738	637719.85
11	cereal	52489	150710
12	fruit	53910	157192.37
13	ice cream	56222	167211.04
14	milk	54117	162606.44
15	2007	233161	702395.82
16	cereal	58671	172828.96
17	fruit	61816	189616.27
18	ice cream	55693	169327.53
19	milk	56981	170623.06
20	Grand Total	693127	2068334.35

**Рис. 45.11.** Поля группы свернуты

### ? Как сортировать и фильтровать поля сводной таблицы?

На рис. 45.5 продукты перечислены в каждой группе в алфавитном порядке. Например, **Cheerios** — первый в группе хлопьев. Если вы хотите перечислить продукты в обратном порядке, установите курсор в какую-нибудь ячейку с названием продукта (например, на листе **Cerealcollapse** ячейка **A7** содержит название **Cheerios**)

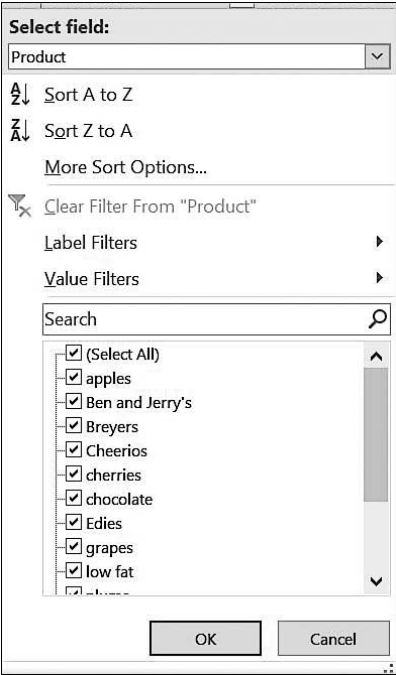


и нажмите стрелку раскрывающегося списка **Названия строк (Row Labels)** в ячейке A4. Появится список параметров фильтрации, показанный на рис. 44.12. Выберите **Сортировка от Я до А (Sort Z To A)**. В списке хлопьев первыми окажутся хлопья **Special K**, в группе **Milk** — цельное молоко (**whole milk**), в группе **Fruit** — сливы (**plums**) и т. д.

Первоначально наша сводная таблица отображалась за 2005 г., затем за 2006 г. и, наконец, за 2007 г. Если нам нужно сначала просмотреть данные за 2007 г., подвиньте курсор в любую ячейку, в которой указан год (например, в A5), щелкните правой кнопкой мыши и выберите команду **Сортировка (Sort)**, затем **Сортировка по убыванию (Sort Largest To Smallest)**.

Любое подмножество продуктов, информацию о котором требуется отобразить, также можно выбрать в нижней части диалогового окна фильтрации. Сначала снимите флажок **Выделить все (Select All)** и затем выберите желаемые продукты.

Другой пример фильтрации находится в файле **Ptcustomers.xlsx** на листе **Data**, показанном на рис. 45.13. На листе для каждой транзакции содержится номер клиента, уплаченная сумма и указан квартал, в котором была получена оплата. Перетащите поле **Клиент** в область **СТРОКИ (Rows)**, поле **Квартал** в область **СТОЛБЦЫ (Columns)** и поле **Оплата** в область **ЗНАЧЕНИЯ (Values)**. Появится сводная таблица, показанная на рис. 45.14. (См. в файле **Pcustomers.xlsx** лист **Ptable**.)



**Рис. 45.12.** Параметры фильтрации сводной таблицы для поля Продукт

	F	G	H
4	Customer	Paid	Quarter
5	20	8048	4
6	6	7398	4
7	10	5280	2
8	28	3412	3
9	8	3316	1
10	17	821	2
11	4	7024	3
12	20	1379	1
13	27	1924	2
14	23	631	3
15	28	9743	4
16	8	8192	2
17	19	875	1
18	3	9803	4
19	24	7344	3
20	13	6114	1
21	9	6728	4
22	2	4554	1
23	16	8230	4
24	25	1296	1

**Рис. 45.13.** Данные для сводной таблицы по клиентам



	A	B	C	D	E	F
3	Sum of Paid	Column Labels				
4	Row Labels		1	2	3	4 Grand Total
5	1	30965	42039	57790	43417	174211
6	2	96038	121118	59089	45355	321600
7	3	57419	33589	61960	97548	250516
8	4	48947	79352	63052	59520	250871
9	5	57270	86555	69517	33471	246813
10	6	75639	71976	55212	78644	281471
11	7	53130	65768	49064	89018	256980
12	8	33289	74001	45219	43512	196021
13	9	61611	99009	61075	50945	272640
14	10	31785	71213	60417	63835	227250
15	11	59127	35567	62130	107832	264656
16	12	71862	21670	67312	63558	224402
17	13	100626	56058	39500	75109	271293
18	14	74240	63023	36217	77218	250698
19	15	30612	62277	45561	52567	191017
20	16	41870	71490	64909	57120	235389
21	17	61811	85706	46978	40802	235297
22	18	24456	44916	55519	81421	206312
23	19	89591	53157	37558	38247	218553
24	20	68349	104140	35083	69424	276996
25	21	77336	37476	51815	57065	223692

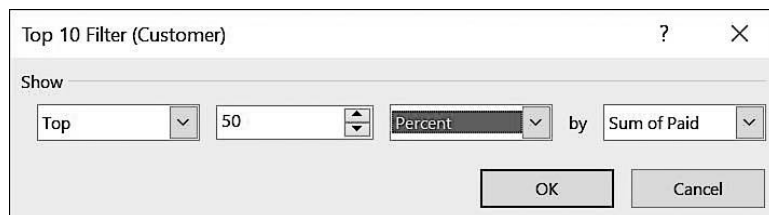
Рис. 45.14. Сводная таблица по клиентам

Вполне вероятно, вам захочется получить список только из 10 основных клиентов. Для этого нажмите стрелку раскрывающегося списка Названия строк (Row Labels) и выберите Фильтры по значению (Value Filters). Затем выберите Первые 10 (Top 10 Items) для создания макета, показанного на рис. 45.15. (См. лист Top 10 cus.) Разумеется, выбрав Очистить фильтр (Clear Filter), вы вернетесь к исходному макету

	A	B	C	D	E	F
3	Sum of Paid	Column Labels				
4	Row Labels		1	2	3	4 Grand Total
5	2	96038	121118	59089	45355	321600
6	6	75639	71976	55212	78644	281471
7	9	61611	99009	61075	50945	272640
8	11	59127	35567	62130	107832	264656
9	13	100626	56058	39500	75109	271293
10	20	68349	104140	35083	69424	276996
11	22	31149	77333	104364	65664	278510
12	23	87124	56387	63290	71953	278754
13	27	45214	89826	56302	71285	262627
14	28	53737	69938	73471	69135	266281
15	Grand Total	678614	781352	609516	705346	2774828

Рис. 45.15. Десять основных клиентов

Допустим, необходимо просмотреть список клиентов, приносящих 50% дохода. Нажмите стрелку раскрывающегося списка **Названия строк (Row Labels)** и выберите **Фильтры по значению (Value Filters)**. Затем выберите **Первые 10 (Top 10 Items)** и заполните данными диалоговое окно, показанное на рис. 45.16, изменив выбор в полях **Показать (Show)** на **Первые 50 процентов (Top 50 Percent)**.



**Рис. 45.16.** Диалоговое окно фильтра **Первые 10 (Клиент)** для отображения списка клиентов, приносящих 50% дохода

Получившаяся сводная таблица находится на листе **Top Half** и показана на рис. 45.17. Как видите, свыше половины дохода приносят 14 клиентов.

	A	B	C	D	E	F
3	Sum of Paid	Column Labels				
4	Row Labels	1	2	3	4	Grand Total
5	2	96038	121118	59089	45355	321600
6	3	57419	33589	61960	97548	250516
7	4	48947	79352	63052	59520	250871
8	6	75639	71976	55212	78644	281471
9	7	53130	65768	49064	89018	256980
10	9	61611	99009	61075	50945	272640
11	11	59127	35567	62130	107832	264656
12	13	100626	56058	39500	75109	271293
13	14	74240	63023	36217	77218	250698
14	20	68349	104140	35083	69424	276996
15	22	31149	77333	104364	65664	278510
16	23	87124	56387	63290	71953	278754
17	27	45214	89826	56302	71285	262627
18	28	53737	69938	73471	69135	266281
19	Grand Total	912350	1023084	819809	1028650	3783893

**Рис. 45.17.** Основные клиенты, приносящие половину дохода

Теперь предположим, что необходимо отсортировать клиентов по доходу за первый квартал. (См. лист **Sorted q1**.) Щелкните правой кнопкой мыши в какой-либо ячейке столбца с оплатой за квартал 1, выберите **Сортировка (Sort)**, затем **Сортировка по убыванию (Sort Largest To Smallest)**. Получившаяся сводная таблица показана

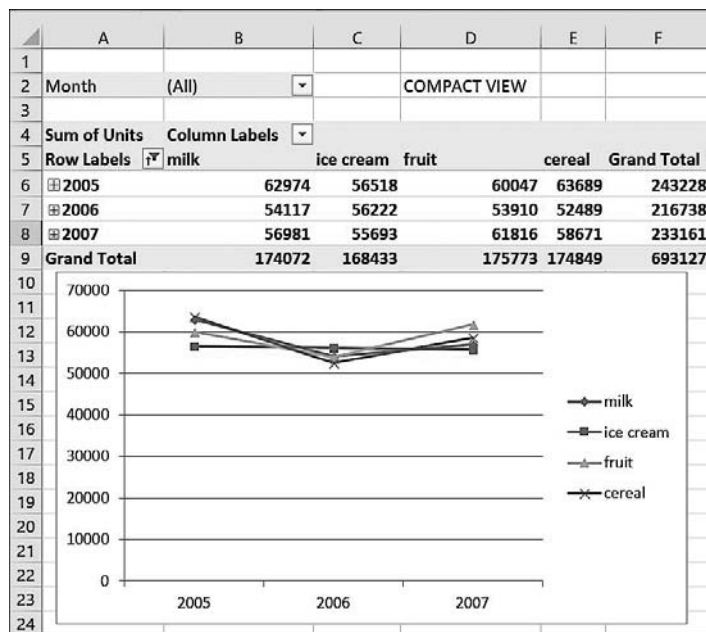
на рис. 45.18. Обратите внимание, что в первом квартале больше всех заплатил клиент 13, вторым был клиент 2 и т. д.

	A	B	C	D	E	F
3	Sum of Paid	Column Labels				
4	Row Labels	1	2	3	4	Grand Total
5	13	100626	56058	39500	75109	271293
6	2	96038	121118	59089	45355	321600
7	19	89591	53157	37558	38247	218553
8	23	87124	56387	63290	71953	278754
9	21	77336	37476	51815	57065	223692
10	6	75639	71976	55212	78644	281471
11	14	74240	63023	36217	77218	250698
12	12	71862	21670	67312	63558	224402
13	20	68349	104140	35083	69424	276996
14	17	61811	85706	46978	40802	235297
15	9	61611	99009	61075	50945	272640
16	26	59994	70594	50446	44050	225084
17	30	59599	64192	44335	42944	211070
18	11	59127	35567	62130	107832	264656
19	3	57419	33589	61960	97548	250516
20	5	57270	86555	69517	33471	246813
21	28	53737	69938	73471	69135	266281
22	7	53130	65768	49064	89018	256980
23	4	48947	79352	63052	59520	250871
24	25	46960	64764	53394	33256	198374

**Рис. 45.18.** Клиенты отсортированы по доходам за первый квартал

### ❓ Как обобщить данные в сводной таблице с помощью сводной диаграммы?

Excel позволяет визуально обобщить данные сводных таблиц с помощью сводных диаграмм. Для задания требуемого представления данных в сводной диаграмме требуется отсортировать данные и свернуть/развернуть поля. Пусть в примере с продовольственными товарами требуется отследить во времени тенденции изменения объемов продаж для товаров каждой группы. (См. в файле Groceriespt.xlsx лист Chart 2.) Для этого следует перетащить поле Год в область СТРОКИ (ROWS) и удалить поле Выручка из области ЗНАЧЕНИЯ (VALUES). Также необходимо свернуть все поле Группа в столбце Названия строк (Row Labels) и затем на панели Поля сводной таблицы (PivotTable Fields) перетащить поле Группа в область СТОЛБЦЫ (COLUMNS). Теперь все готово для создания сводной диаграммы. Установите курсор в какую-либо ячейку сводной таблицы и на вкладке Анализ (Analyze) выберите Сводная диаграмма (PivotChart). Затем в диалоговом окне Вставить диаграмму (Insert Chart) выберите тип создаваемой диаграммы. В этом примере выбран четвертый вариант График (Line Graph), показанный на рис. 45.19. Как видно из графика, продажи молока были самыми высокими в 2005 г и самыми низкими в 2006 г. При желании, щелкнув правой кнопкой мыши по любой диаграмме Excel, вы сможете изменить тип диаграммы.



**Рис. 45.19.** Сводная диаграмма тенденций изменения объемов продаж для групп продуктов

### ❓ Для чего предназначена область ФИЛЬТРЫ в сводной таблице?

Напомню, что я поместил поле Месяц в область ФИЛЬТРЫ (FILTERS) сводной таблицы. Рассмотрим применение фильтров на следующем примере. Допустим, вам нужно обобщить данные о продажах с января по июнь. На листе **First 6 months** нажмите стрелку раскрывающегося списка в ячейке B2 и выберите месяцы с января по июнь. Получившаяся сводная таблица показана на рис. 45.20, в ней обобщены данные продаж по продуктам, группам и годам для месяцев с января по июнь.

### ❓ Как работают срезы сводной таблицы?

При работе с фильтрами возникает следующая проблема: при режиме просмотра сводной таблицы, как на рис. 45.20, невозможно сразу

	A	B	C
1			
2	Month	(Multiple Iter	
3			
4		Values	
5	Row Labels	Sum of Units	Sum of Revenue
6	2007	115258	346295.58
7	milk	30069	89222.68
8	chocolate	4875	14077.99
9	west	2014	5839.06
10	south	736	2141.76
11	north	1437	4136.37
12	east	688	1960.8
13	low fat	9447	27341.25
14	west	3285	10732.35
15	south	531	1062
16	north	3905	10212.44
17	east	1726	5334.46
18	skim	10182	30450.57
19	west	2156	6169.53
20	south	2778	8881.2
21	north	2521	7726.56
22	east	2727	7673.28
23	whole	5565	17352.87
24	west	735	1764
25	south	1562	4495.15
26	east	3268	11093.72
27	ice cream	22336	69108.5

**Рис. 45.20.** Сводная таблица с данными продаж с января по июнь

увидеть, что в ней обобщены данные продаж с января по июнь. Проблема решается с помощью такого инструмента, как срез (появившегося в Excel 2010). Для создания среза по любому столбцу данных, формирующему сводную таблицу установите курсор внутрь таблицы и на вкладке **Вставка (Insert)** в группе **Фильтры (Filters)** выберите **Срез (Slicer)** (см. в файле *Groceriespt.xlsx* лист *Slicers*). В диалоговом окне **Вставка срезов** выберите **Месяц** и **Продукт** для создания срезов по этим полям. В каждом заданном срезе вы можете выбрать любое подмножество возможных значений для вставки в сводную таблицу. В срезе **Месяц** выбраны месяцы с января по июнь (по одному с нажатым **Ctrl**). Для среза **Продукт** я фильтр не устанавливал, так что данные представлены для всех продуктов с января по июнь. Срезы показаны на рис. 45.21.

	A	B	C	D	E	F
1						
2	Month	(Multiple Items)				
3						
4		Values				
5	Row Labels	Sum of Units	Sum of Revenue			
6	2007	115258	346295.58			
7	milk	30069	89222.68			
8	chocolate	4875	14077.99			
9	west	2014	5839.06			
10	south	736	2141.76			
11	north	1437	4136.37			
12	east	688	1960.8			
13	low fat	9447	27341.25			
14	west	3285	10732.35			
15	south	531	1062			
16	north	3905	10212.44			
17	east	1726	5334.46			
18	skim	10182	30450.57			
19	west	2156	6169.53			
20	south	2778	8881.2			
21	north	2521	7726.56			
22	east	2727	7673.28			
23	whole	5565	17352.87			
24	west	735	1764			
25	south	1562	4495.15			
26	east	3268	11093.72			
27	ice cream	22336	69108.5			
28	Breyers	11446	36341.78			
29	west	3377	10496.44			

**Рис. 45.21.** Срезы для полей Месяц и Продукт

Если выделить срез, то на ленте, на вкладке **Параметры (Options)**, можно выбрать параметры форматирования, позволяющие менять его внешний вид. Например, можно изменить высоту и ширину, а также количество столбцов в срезе. Кроме того, размеры среза можно быстро изменить, потянув за его стороны или углы, удерживая **Ctrl**.

### ❓ Как добавить пустые строки или скрыть промежуточные итоги в сводной таблице?

Если вы хотите добавить пустую строку между элементами группы, щелкните на ячейке сводной таблицы и на вкладке **Конструктор сводных таблиц (PivotTable Tools)**

Design) в группе Пустые строки (Blank Rows) выберите Вставить пустую строку после каждого элемента (Insert Blank Line After Each Item). Если необходимо скрыть промежуточные или общие итоги, на вкладке Конструктор сводных таблиц (PivotTable Tools Design) в группе Промежуточные итоги (Subtotals) или в группе Общие итоги (Grand Totals) выберите соответствующие строки. После добавления пустых строк и скрытия всех итогов получится таблица, представленная в файле Groceriespt.xlsx на листе Blank rows no totals (рис. 45.22). Щелкните правой кнопкой мыши в любой ячейке сводной таблицы и выберите команду Параметры сводной таблицы (PivotTable Options). Откроется диалоговое окно Параметры сводной таблицы (PivotTable Options), в котором в разделе Формат (Format) вкладки Разметка и Формат (Layout & Format) вы можете заменить пустые ячейки символом, который зададите в поле Для пустых ячеек отображать (For Empty Cells Show), например, подчеркивание ( \_ ) или 0.

	A	B	C
2	Month	(All)	
3			
4			
5	Row Labels	Sum of Units	Sum of Revenue
6	2007		
7	milk		
8	chocolate		
9	west	4379	12668.95
10	south	1545	4528.31
11	north	2322	7579.02
12	east	2184	5791.1
13			
14	low fat		
15	west	4668	15042.24
16	south	2431	7606.76
17	north	7957	23490.32
18	east	2517	7762.83
19			
20	skim		
21	west	3571	9951.37
22	south	2778	8881.2
23	north	3594	10792.09
24	east	4839	14609.52
25			
26	whole		
27	west	3252	8311.84
28	south	1562	4495.15
29	north	2621	7640.88
30	east	6761	21471.48

Рис. 45.22. Сводная таблица без итогов

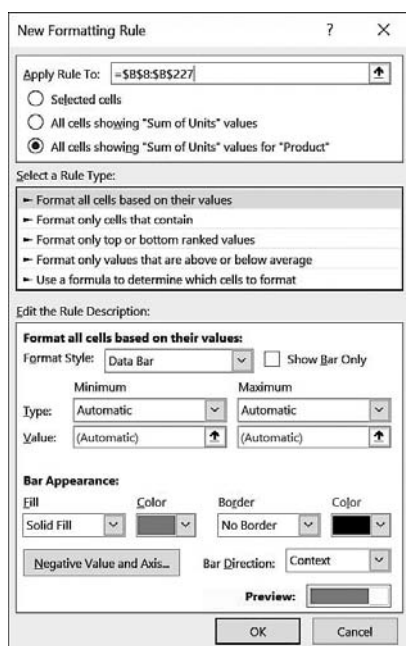
### ? Как применить к сводной таблице Условное форматирование?

Допустим, вы хотите применить гистограмму к столбцу Количество в сводной таблице продуктовых товаров. Одна из проблем, которые вам предстоит преодолеть, — промежуточным и общим итогам соответствуют слишком большие столбики гистограммы, из-за чего столбики, соответствующие остальным данным, меньше, чем должны быть. В таком случае гистограмму было бы лучше применить только к ячейкам, содержащим количество проданных продуктов, но не к промежуточным и общим итогам. (См. в файле Groceriespt.xlsx лист Cond form.) Для этого



сначала установите курсор в любую ячейку с продажами какого-либо продукта (например, шоколадного молока в ячейке B8). На вкладке Главная (Home) выберите Условное форматирование (Conditional Formatting), затем Гистограммы (Data Bars) и Другие правила (More Rules). Появится диалоговое окно Создание правила форматирования (New Formatting Rule), как на рис. 45.23.

Положение переключателя ко всем ячейкам, содержащим значения «Сумма по полю Количество» для «Продукт», и выбор диапазона \$B\$8:\$B\$227 обеспечат применение гистограммы только к тем ячейкам, в которых указано количество проданных продуктов (см. лист Cond form и рис. 45.24).



**Рис. 45.23.** Диалоговое окно Создание правила форматирования для условного форматирования сводной таблицы

	A	B	C
4		Values	
5	Row Labels	Sum of Units	Sum of Revenue
6	2007	233161	702395.82
7	milk	56981	170623.06
8	chocolate	10430	30567.38
9	west	4379	12668.95
10	south	1545	4528.31
11	north	2322	7579.02
12	east	2184	5791.1
13	low fat	17573	53902.15
14	west	4668	15042.24
15	south	2431	7606.76
16	north	7957	23490.32
17	east	2517	7762.83
18	skim	14782	44234.18
19	west	3571	9951.37
20	south	2778	8881.2
21	north	3594	10792.09
22	east	4839	14609.52
23	whole	14196	41919.35
24	west	3252	8311.84
25	south	1562	4495.15
26	north	2621	7640.88
27	east	6761	21471.48
28	ice cream	55693	169327.53

**Рис. 45.24.** Данные для столбцов сводной таблицы

## ❓ Как обновить вычисления при добавлении новых данных?

Если данные в исходных строках изменились, для обновления сводной таблицы щелкните правой кнопкой мыши по ней и в контекстном меню выберите Обновить (Refresh). Группу Обновить (Refresh) можно также найти на вкладке Анализ (Analyze) в группе Данные (Data).

Если при обновлении нужно автоматически включать добавляемые данные в расчеты сводной таблицы, нужно определить исходный набор данных как таблицу, выделив ее Ctrl+T. (См. главу 26 «Таблицы».)

Для изменения диапазона данных, используемого при создании сводной таблицы, можно в группе **Данные (Data)** на вкладке **Анализ (Analyze)** выбрать **Источник данных (Data Source)** и указать в диалоговом окне новый диапазон. Кроме того, можно переместить таблицу в другое место, выбрав **Переместить сводную таблицу (Move PivotTable)** на вкладке **Анализ (Analyze)** в группе **Действия (Actions)**.

**?** Я работаю в небольшой турфирме и должен разослать проспект о ней потенциальным клиентам. Мой бюджет ограничен, поэтому послать проспект необходимо тем людям, которые больше остальных тратят денег на путешествия. У меня есть случайная выборка из 925 человек, о которых известна следующая информация: возраст, пол и количество денег, потраченных на путешествия за прошлый год. Как с помощью этих данных определить влияние пола и возраста на расходы на путешествия? Какие выводы можно сделать о типе потенциального клиента, которому следует отправить проспект по почте?

Для понимания этих данных необходимо сделать следующие разбивки:

- средняя сумма, потраченная на путешествия, с учетом пола;
- средняя сумма, потраченная на путешествия, в каждой возрастной группе;
- средняя сумма, потраченная на путешествия, в каждой возрастной группе с учетом пола.

На рис. 45.25 представлена часть данных из файла **Traveldata.xlsx** (см. лист **Data**). Например, первым в списке идет 44-летний мужчина, потративший на поездки \$997.

Откройте лист **Data** в файле **Traveldatatempl.xlsx** в папке **Templates**. Сначала определим разбивку расходов по половой принадлежности. На вкладке **Вставка (Insert)** выберите **Сводная таблица (PivotTable)**. Excel выделит в поле **Выделить таблицу или диапазон (Select A Table Or Range)** диапазон **A2:D927**. Нажмите **ОК**. Откроется новый лист, содержащий в правой части панель **Поля сводной таблицы (PivotTable Fields)**. Затем перетащите поле **Пол** в область **СТРОКИ (Rows)** (из списка полей в верхней части панели) и поле **Израсходованная сумма** в область **ЗНАЧЕНИЯ (Values)**. Получившаяся сводная таблица представлена на рис. 45.26.

## ПРИМЕЧАНИЕ

Панель **Список полей (Fields List)** была видоизменена в Excel 2016 — она стала панелью, поддерживающей технологию перетаскивания мышью, и получила название **Поля сводной таблицы (PivotTable Fields)**.

Как видно из заголовка, **Суммы, затрачиваемые на путешествия (Sum Of Amount Spent On Travel)**, — это общие суммы, истраченные на путешествия, а нам требуется узнать средние суммы, истраченные на путешествия мужчинами и женщинами. Для вычисления этих величин дважды щелкните по заголовку **Суммы, затрачиваемые на путешествия** в сводной таблице (не в списке полей!) и затем в диалоговом окне **Параметры поля значений (Value Field Settings)** из списка **Суммировать значения по (Summarize Value Field By)** выберите **Среднее (Average)**, как показано на рис. 45.27.



	A	B	C
2	Amount Spent on Travel	Age	Gender
3	997	44	M
4	850	39	F
5	997	43	M
6	951	41	M
7	993	50	F
8	781	39	F
9	912	45	F
10	649	59	M
11	1265	25	M
12	680	38	F
13	800	41	F
14	613	32	F
15	993	46	F
16	1059	38	M
17	939	42	F
18	841	44	F
19	828	38	F
20	1004	50	F
21	983	48	F
22	837	46	M
23	924	42	M
24	852	48	M
25	963	39	M

**Рис. 45.25.** Данные турфирмы о возрасте, поле и суммах, затраченных на путешествия

Row Labels	Sum of Amount Spent on Travel
F	413632
M	426387
<b>Grand Total</b>	<b>840019</b>

**Рис. 45.26.** Сводная таблица общих расходов на путешествия с учетом пола

Value Field Settings?×

Source Name: Amount Spent on Travel

Custom Name: Average of Amount Spent on Travel

Summarize Values By

Show Values As

Summarize value field by

Choose the type of calculation that you want to use to summarize data from the selected field

Sum

Count

Average

Max

Min

Product

Number Format

OK

Cancel

**Рис. 45.27.** Выбор другой обобщающей функции в диалоговом окне Параметры поля значений

Нажмите ОК. Вы получите результат, показанный на рис. 45.28.

Row Labels	Average of Amount Spent on Travel
F	901.1590414
M	914.9935622
<b>Grand Total</b>	<b>908.1286486</b>

**Рис. 45.28.** Средние расходы на путешествия с учетом пола

В среднем люди тратят на путешествия \$908,13. Женщины тратят в среднем \$901,16, а мужчины — \$914,99. Эта сводная таблица показывает, что пол мало влияет на любовь к путешествиям. Нажав на стрелку раскрывающегося списка **Названия строк (Row Labels)**, можно с помощью фильтра отобразить результаты только по мужчинам или только по женщинам.

Теперь нам нужно узнать, как расходы на путешествия зависят от возраста. Для разбивки по возрасту из области **СТРОКИ (ROWS)** удалите поле **Пол** (щелкнув на нем и выбрав **Удалить поле (Remove Field)**) и перетащите в область **СТРОКИ (ROWS)** поле **Возраст**. Получившаяся сводная таблица представлена на рис. 45.29.

	A	B	C	D
4	Row Labels	F	M	Grand Total
5	25	482.3846154	1305.764706	948.9666667
6	26	526.6923077	1281.583333	889.04
7	27	532.5714286	1266.722222	1061.16
8	28	584	1243.666667	960.952381
9	29	564.5	1229.833333	814
10	30	578	1176.666667	877.3333333
11	31	604.3333333	1212.733333	1038.904762
12	32	636.6153846	1160.818182	876.875
13	33	661.6153846	1146.928571	913.2592593
14	34	674.0909091	1128.615385	920.2916667
15	35	705.5555556	1089.25	886.1176471
16	36	722.4285714	1071.888889	859.173913
17	37	746.9166667	1061.416667	904.1666667
18	38	764.25	1051.846154	913.8
19	39	789.9375	1007.923077	887.6551724

**Рис. 45.29.** Сводная таблица расходов на путешествия с учетом возраста

По-видимому, возраст мало влияет на расходы на путешествия. Фактически эта сводная таблица бесполезна в своем нынешнем состоянии. Необходимо сгруппировать данные по возрасту для поиска каких-либо закономерностей. Для группировки результатов по возрасту щелкните правой кнопкой мыши в любом месте столбца, в котором указан возраст (столбец **A**), и выберите **Группировать (Group)**.

В диалоговом окне Группирование (Grouping), показанном на рис. 45.30, можно указать интервалы возрастных групп. Если укажете шаг 10 лет, то получите сводную таблицу, как на рис. 45.31.

The image shows the 'Grouping' dialog box with the following settings:

- Auto** (selected)
- ☒ **Starting at:** 25
- ☒ **Ending at:** 64
- By:** 10
- OK** and **Cancel** buttons.

**Рис. 45.30.** Диалоговое окно параметров группировки

В среднем 25–34-летние люди тратят на путешествия \$935,84, 55–64-летние — \$903,57 и т. д. Эта таблица более информативна, но она все еще показывает, что люди всех возрастов, как правило, тратят примерно одинаковые суммы на путешествия. Такой вид данных не позволяет решить, кому следует отправить проспект.

Наконец, получим разбивку средних расходов на путешествия по возрасту для мужчин и женщин отдельно. Для этого перетащите поле **Пол** в область **СТОЛБЦЫ** (COLUMNS). Получившаяся таблица приведена на рис. 45.31. (См. лист **Final Table** в файле **Traveldata.xlsx**.)

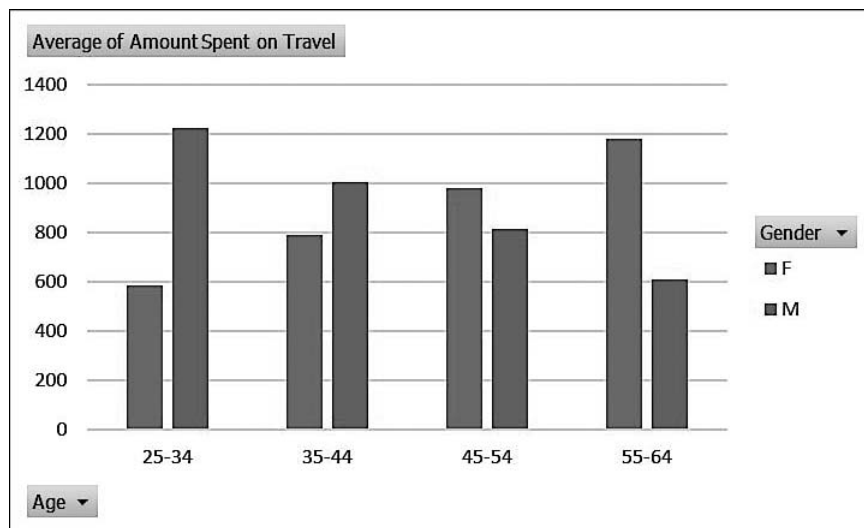
	A	B	C	D
4	Row Labels	F	M	Grand Total
5	25-34	585.4752475	1221.209677	935.8355556
6	35-44	790.1652174	1004.098214	895.7180617
7	45-54	979.4782609	813.5765766	897.9955752
8	55-64	1179.609375	606.6470588	903.5668016
9	<b>Grand Total</b>	<b>901.1590414</b>	<b>914.9935622</b>	<b>908.1286486</b>

**Рис. 45.31.** Разбивка расходов на путешествия по возрасту и полу

То, что нужно! Как видите, с возрастом женщины начинают тратить на путешествия больше, а мужчины — меньше. Теперь понятно, кому следует рассылать проспекты: пожилым женщинам и молодым мужчинам. Как сказал один студент: «Ничего себе круиз бы получился!»

Итоги анализа удобно представить в виде диаграммы. Установите курсор внутри сводной таблицы и на вкладке **Анализ** (Analyze) выберите **Сводная диаграмма** (PivotChart) в группе **Инструменты** (Tools). Появится меню рекомендуемых диаграмм

(диалоговое окно **Вставить диаграмму (Insert Chart)**). Выберите первый вариант **Гистограмма (Column)** и — **Гистограмма с группировкой (Clustered Column)**. Результат показан на рис. 45.32. Если вы хотите редактировать диаграмму дальше, откройте вкладку **Конструктор (Design)** среди вкладок **Сводная таблица (PivotTable)** для сводной диаграммы. На ней можно выбрать макет, добавить названия для диаграммы и осей, а также внести другие изменения. Для этого щелкните **Добавить элементы диаграммы (Add Chart Element)** в группе **Макеты диаграмм (Chart Layouts)**.



**Рис. 45.32.** Сводная диаграмма для разбивки расходов на путешествия по возрасту и полу

Каждая возрастная группа тратит на путешествия примерно одинаковую сумму, но с возрастом женщины начинают тратить больше мужчин. Для изменения типа диаграммы щелкните правой кнопкой мыши по диаграмме и в контекстном меню выберите **Изменить тип диаграммы (Change Chart Type)**. Затем выберите другой тип в диалоговом окне **Изменить тип диаграммы (Change Chart Type)** и нажмите **ОК**.

Обратите внимание, что высота столбиков, отображающих расходы мужчин, с повышением возраста уменьшается, а высота столбиков с расходами женщин с повышением возраста увеличивается. Вы видите, почему сводные таблицы, отображавшие данные только по возрасту или по полу, не позволили заметить эту закономерность. Мы выяснили, что средняя сумма затрат не зависит от возраста, поскольку выборка включает поровну мужчин и женщин. (Следует отметить, что средняя высота обоих столбиков для каждого возраста практически одинаковая.) Оказалось также, что средняя сумма расходов мужчин и женщин примерно одинакова. Это видно, поскольку усредненные по всем возрастам синие и красные столбики приблизительно одинаковы по высоте. Срез данных одновременно по возрасту и полу позволил понять, как обстоят дела в реальности.

Следует также отметить, что щелкнув (на сводной диаграмме) на раскрывающемся списке **Возраст (Age)**, можно отфильтровать диаграмму по возрасту, а щелкнув на списке **Пол (Gender)** — по полу.

**?** Я провожу маркетинговое исследование автомобилей-универсалов. Мне необходимо определить, какие факторы влияют на вероятность покупки семей автомобиля-универсала. У меня есть большая выборка семей с информацией об их размере (большая или нет) и доходе (высокий или низкий). Как определить влияние размера семьи и семейного дохода на вероятность покупки автомобиля-универсала?

В файле *Station.xlsx* содержится следующая информация:

- ☐ семья большая или небольшая (*Large* или *Small*);
- ☐ уровень дохода высокий или низкий (*High* или *Low*);
- ☐ купила ли семья автомобиль-универсал (*Yes* или *No*).

Часть выборки данных показана на рис. 45.33 (см. лист **Data**). Например, первая указанная семья — это небольшая семья с высоким уровнем дохода, не купившая автомобиль-универсал.

	B	C	D
2	Station Wagon?	Family Siz	Salary
3	No	Small	High
4	Yes	Large	High
5	Yes	Large	High
6	Yes	Large	High
7	Yes	Large	High
8	No	Small	High
9	Yes	Large	High
10	Yes	Large	High
11	Yes	Large	Low
12	Yes	Large	High
13	Yes	Large	Low
14	No	Small	Low
15	No	Small	Low
16	No	Small	High
17	Yes	Large	High
18	Yes	Large	High
19	No	Small	High
20	Yes	Large	High
21	No	Small	High

**Рис. 45.33.** Данные об уровне дохода, размере семьи и покупке автомобиля-универсала

Мы хотим определить, как размер семьи и уровень дохода влияют на вероятность покупки автомобиля-универсала. Для этого нужно проследить влияние уровня дохода на покупку универсала для каждого размера семьи и влияние размера семьи на покупку универсала для каждого уровня дохода.

Сначала на вкладке **Вставка (Insert)** выберите **Сводная таблица (Pivot Table)** в группе **Таблицы (Table)**, а затем выделите данные (диапазон B2:D345) и нажмите **ОК** в диалоговом окне **Создание сводной таблицы (Create PivotTable)**. На панели **Поля сводной таблицы (PivotTable Fields)** перетащите поля **Размер семьи** и **Уровень дохода** в область **СТРОКИ (ROWS)**, поле **Универсал? —** в область **СТОЛБЦЫ (COLUMNS)** и любое из трех полей в область **ЗНАЧЕНИЯ (VALUES)**. Сводная таблица с результатами представлена на рис. 45.34 (см. лист 1st Table). Обратите внимание, что Excel обобщил данные правильно, подсчитав число наблюдений в каждой категории. Например, 34 большие семьи с высоким уровнем дохода не купили универсал, а 100 больших семей с высоким уровнем дохода — купили.

	A	B	C	D
3	Count of Station Wagon?	Column Labels		
4	Row Labels	No	Yes	Grand Total
5	Large	48	138	186
6	High	34	100	134
7	Low	14	38	52
8	Small	147	10	157
9	High	104	8	112
10	Low	43	2	45
11	Grand Total	195	148	343

**Рис. 45.34.** Обобщение данных о покупке автомобиля-универсала по размеру семьи и уровню дохода

В каждой строке сводной таблицы необходимо указать процент семей, купивших автомобиль-универсал. Для отображения данных в таком формате щелкните правой кнопкой мыши по любой ячейке сводной таблицы, содержащей значение, и затем в контекстном меню выберите **Параметры полей значений (Value Fields Settings)**. Появится диалоговое окно **Параметры поля значений (Value Field Settings)**. В нем на вкладке **Дополнительные вычисления (Show Values As)** в раскрывающемся списке **Дополнительные вычисления (Show Values As)** выберите **% от суммы по строке (% Of Row)**. Сводная таблица примет вид, как на рис. 45.35. (См. лист 1st Percent Breakdown.)

	A	B	C	D
3	Count of Station Wagon?	Column Labels		
4	Row Labels	No	Yes	Grand Total
5	Large	25.81%	74.19%	100.00%
6	High	25.37%	74.63%	100.00%
7	Low	26.92%	73.08%	100.00%
8	Small	93.63%	6.37%	100.00%
9	High	92.86%	7.14%	100.00%
10	Low	95.56%	4.44%	100.00%
11	Grand Total	56.85%	43.15%	100.00%

**Рис. 45.35.** Процентная разбивка данных о покупке автомобиля-универсала по уровню дохода для больших и маленьких семей

Как видно из рис. 45.35, и в больших, и в маленьких семьях влияние уровня дохода на вероятность покупки автомобиля-универсала незначительно. Теперь необходимо определить влияние размера семьи на предпочтение автомобиля-универсала в семьях с высоким и низким уровнем дохода. Для этого на панели Поля сводной таблицы (Pivot-Table Fields) внутри области СТРОКИ (ROWS) поместите поле Уровень дохода над полем Размер семьи, что даст нам сводную таблицу, как на рис. 45.36. (См. лист Final Percent Breakdown.)

	A	B	C	D
3	Count of Station Wagon?	Column Labels		
4	Row Labels	No	Yes	Grand Total
5	High	56.10%	43.90%	100.00%
6	Large	25.37%	74.63%	100.00%
7	Small	92.86%	7.14%	100.00%
8	Low	58.76%	41.24%	100.00%
9	Large	26.92%	73.08%	100.00%
10	Small	95.56%	4.44%	100.00%
11	Grand Total	56.85%	43.15%	100.00%

**Рис. 45.36.** Процентная разбивка данных о покупке автомобиля-универсала по размеру семьи для высокого и низкого уровня дохода

Как видно из таблицы, для семей с высоким уровнем дохода покупка автомобиля-универсала намного вероятнее в больших семьях, чем в маленьких. Аналогично для семей с низким уровнем дохода: покупка универсала также более вероятна в больших семьях. Значит, размер семьи оказывает большее влияние на вероятность покупки автомобиля-универсала, чем уровень дохода.

❓ Я работаю в компании, продающей микрочипы по всему миру Ежемесячно я получаю данные о фактических и прогнозируемых объемах продаж чипа 1, чипа 2 и чипа 3 в Канаде, Франции и США. Я также получаю расхождение, или разницу, между фактической выручкой и выручкой, заложенной в бюджет. Для каждого месяца и каждой комбинации страны и продукта мне нужно отобразить следующие данные: фактическая выручка, выручка, заложенная в бюджет, фактическое расхождение, фактическая выручка как процент от годового дохода и расхождение как процент от выручки, заложенной в бюджет. Каким образом можно отобразить эту информацию?

В этом сценарии вы — финансовый директор в компании, производящей микрочипы. Вы продаете продукцию в различные страны/регионы и в различное время. Сводные таблицы позволяют обобщить эти данные в простом для понимания формате.

Файл Ptableexample.xlsx содержит данные о ежемесячных фактических продажах и продажах, заложенных в бюджет, в 1997 г. для чипов Chip 1, Chip 2 и Chip 3 в Канаде, Франции и США. В файле также указана дисперсия, или разница, между



фактической выручкой и выручкой, заложенной в бюджет. Часть этих данных представлена на рис. 45.37. (См. лист **Data**.) Например, в январе в США объемы продаж для чипа 1 составили \$4000, а прогнозировался объем продаж в \$5454. Разница составила –\$1454.

	A	B	C	D	E	F
1	Month	Product	Country	Revenue	Budget	Var
2	January	Chip 1	US	4000	5454	-1454
3	January	Chip 1	Canada	3424	5341	-1917
4	January	Chip 1	US	8324	1232	7092
5	January	Chip 1	France	5555	3424	2131
6	January	Chip 1	Canada	5341	8324	-2983
7	January	Chip 1	US	1232	5555	-4323
8	January	Chip 1	France	3424	5341	-1917
9	January	Chip 1	Canada	8324	1232	7092
10	January	Chip 1	US	5555	3424	2131
11	January	Chip 1	France	5341	8324	-2983
12	January	Chip 1	Canada	1232	5555	-4323
13	January	Chip 1	US	3424	5341	-1917
14	January	Chip 1	Canada	8383	5454	2929
15	January	Chip 1	France	8324	1232	7092
16	January	Chip 1	Canada	5555	3424	2131
17	January	Chip 1	US	5341	8324	-2983
18	January	Chip 1	France	1232	5555	-4323
19	January	Chip 1	France	3523	9295	-5772
20	February	Chip 2	Canada	5555	3424	2131
21	February	Chip 2	US	5454	4000	1454

**Рис. 45.37.** Данные о продаже чипов в разных странах по месяцам: фактическая выручка, выручка, заложенная в бюджет, и разница

Для каждого месяца и каждой комбинации страны и продукта необходимо отобразить следующие данные:

- ☐ фактическая выручка;
- ☐ выручка, заложенная в бюджет;
- ☐ фактическая разница;
- ☐ фактическая выручка как процент от годовой выручки;
- ☐ разница как доля от выручки, заложенной в бюджет.

Сначала выделите ячейку в диапазоне данных (не забудьте, что первая строка должна содержать заголовки) и затем на вкладке **Вставка (Insert)** выберите **Сводная таблица (PivotTable)**. Требуемый диапазон данных (A1:F208) будет определен автоматически. Нажмите **ОК** в диалоговом окне **Создание сводной таблицы (Create PivotTable)**.

Перетащите поле **Месяц** в область **СТРОКИ (ROWS)**, поле **Страна** в область **СТОЛБЦЫ (COLUMNS)** и поле **Выручка** в область **ЗНАЧЕНИЯ (Values)**. Появится таблица с общей



выручкой по странам для каждого месяца. Добавление какого-либо поля в область **ФИЛЬТРЫ (FILTERS)** позволит фильтровать данные в сводной таблице по значениям в этом поле. Например, добавление поля **Продукт** в область **ФИЛЬТРЫ (FILTERS)** позволит просмотреть продажи только чипа 1 по месяцам для каждой страны. Для отображения любой комбинации страны и продукта перетащите поле **Месяц** в область **СТРОКИ (ROWS)** и поля **Страна** и **Продукт** в область **ФИЛЬТРЫ (FILTERS)**. Затем перетащите поля **Разница**, **Выручка** и **Бюджет** (в этом порядке!) в область **ЗНАЧЕНИЯ (VALUES)**. Вы создали таблицу, как на рис. 45.38. (См. лист 1st Table.)

	A	B	C	D
1	Product	(All)		
2	Country	(All)		
3				
4	Values			
5	Row Labels	Sum of Var	Sum of Revenue	Sum of Budget
6	January	-4297	8.53%	91831
7	February	2843	8.81%	87534
8	March	-1389	8.67%	90377
9	April	-2774	8.28%	87756
10	May	-423	8.24%	84982
11	June	-548	8.19%	84559
12	July	2366	8.42%	84011
13	August	-2843	8.14%	86377
14	September	1389	8.27%	83534
15	October	-4318	7.85%	84923
16	November	3406	8.19%	80605
17	December	2366	8.42%	84011
18	Grand Total	-4222	100.00%	1030500

**Рис. 45.38.** Сводка по месяцам для выручки, заложенной в бюджет, фактической выручки и разницы

Например, в январе общая выручка составила \$87 534, а в бюджет было заложено \$91 831. Таким образом, фактические продажи оказались ниже запланированных на \$4297 (налицо разница). Пусть требуется определить процент выручки для каждого месяца. Снова перетащите поле **Выручка** из списка полей в область **ЗНАЧЕНИЯ (VALUES)**. Поле получит имя **Сумма по полю Выручка2**. Щелкните правой кнопкой мыши на столбце с этими данными и выберите в контекстном меню **Параметры полей значений (Value Fields Settings)**. В диалоговом окне **Параметры поля значений (Value Field Settings)** откройте вкладку **Дополнительные вычисления (Show Values As)**. В раскрывающемся списке **Дополнительные вычисления (Show Values As)** выберите **% от суммы по столбцу (% Of Column)**; вы также можете дать этому полю свое имя в поле **Пользовательское имя (Custom Name)**, как показано на рис. 45.39.

Теперь вы получили сводную таблицу как на рис. 45.40. (См. лист 2nd Table.) Январские продажи составляют 8,53% от выручки. Общая выручка за год — \$1 026 278 (рис. 45.38).

Value Field Settings ? X

Source Name: Revenue

Custom Name: Sum of Revenue2

Summarize Values By Show Values As

**Show values as**

% of Column Total

Base field:

Product  
Country  
Revenue  
Budget  
Var

Base item:

Number Format OK Cancel

**Рис. 45.39.** Создание ежемесячного процента годовой выручки

	A	B	C	D	E
1	Product	(All)			
2	Country	(All)			
3					
4	Values				
5	Row Labels	Sum of Var	Sum of Revenue	Sum of Budget	Sum of Revenue2
6	January	-4297	87534	91831	8.53%
7	February	2843	90377	87534	8.81%
8	March	-1389	88988	90377	8.67%
9	April	-2774	84982	87756	8.28%
10	May	-423	84559	84982	8.24%
11	June	-548	84011	84559	8.19%
12	July	2366	86377	84011	8.42%
13	August	-2843	83534	86377	8.14%
14	September	1389	84923	83534	8.27%
15	October	-4318	80605	84923	7.85%
16	November	3406	84011	80605	8.19%
17	December	2366	86377	84011	8.42%
18	Grand Total	-4222	1026278	1030500	100.00%

**Рис. 45.40.** Разбивка выручки по месяцам

### ? Что такое вычисляемое поле?

Теперь вам требуется определить разницу для каждого месяца как процент от средних показателей продаж. Для этого вы можете создать *вычисляемое поле*. Выделите в области данных сводной таблицы какую-нибудь ячейку и затем на вкладке **Анализ**

(Analyze) в группе Вычисления в списке Поля, элементы и наборы (Fields, Items & Sets) выберите **Вычисляемое поле (Calculated Field)**. Откроется диалоговое окно **Вставка вычисляемого поля (Insert Calculated Field)**. Введите имя поля и формулу, как показано на рис. 45.41. Формулу  $\text{=Разница/Бюджет}$  для этого примера можно ввести целиком или воспользоваться списком полей и кнопкой **Добавить поле (Insert Field)** для вставки поля в формулу. После нажатия кнопок **Добавить (Add)** и **OK** появится сводная таблица, как на рис. 45.42. (См. лист Calc Field в файле Ptableexample.xlsx.)

**Insert Calculated Field**

Name: Var Percentage of Revenue

Formula: = Var/ Budget

Fields:

- Month
- Product
- Country
- Revenue
- Budget**
- Var
- Month2
- Var percentage of Budget

Buttons: Add, Delete, Insert Field, OK, Close

**Рис. 45.41.** Создание вычисляемого поля

	A	B	C	D	E	F
1	Product	(All)				
2	Country	(All)				
3						
4		Values				
5	Row Labels	Sum of Var	Sum of Budget	Sum of Revenue	Sum of Revenue2	Sum of Var percentage of Budget
6	January	-4297	91831	87534	8.53%	-0.046792477
7	February	2843	87534	90377	8.81%	0.032478808
8	March	-1389	90377	88988	8.67%	-0.015368954
9	April	-2774	87756	84982	8.28%	-0.031610374
10	May	-423	84982	84559	8.24%	-0.004977525
11	June	-548	84559	84011	8.19%	-0.006480682
12	July	2366	84011	86377	8.42%	0.028162979
13	August	-2843	86377	83534	8.14%	-0.032913854
14	September	1389	83534	84923	8.27%	0.01662796
15	October	-4318	84923	80605	7.85%	-0.050846061
16	November	3406	80605	84011	8.19%	0.042255443
17	December	2366	84011	86377	8.42%	0.028162979
18	Grand Total	-4222	1030500	1026278	100.00%	-0.00409704

**Рис. 45.42.** Сводная таблица с вычисляемым полем для расчета доли разницы

Таким образом, в январе продажи были на 4,7% ниже, чем заложено в бюджет. Открыв диалоговое окно **Вставка вычисляемого поля** (Insert Calculated Field) еще раз, вы можете внести изменения или удалить вычисляемое поле.

### ? Что использовать: фильтр отчета или срез?

Чтобы просмотреть показатели продаж чипа 2 во Франции, можно перетащить поля **Продукт** и **Страна** в область фильтров. Когда вы выберете **Chip 2** и **France**, вы получите сводную таблицу, как на рис. 45.43. На рис. 45.44 показано, как создать такую же таблицу с помощью срезов (см. лист **Slicers**).

	A	B	C	D	E
1	Product	Chip 2			
2	Country	France			
3					
4	Values				
5	Row Labels	Sum of Var	Sum of Revenue	Sum of Budget	Sum of Revenue2
6	February	-3846	29108	32954	23.90%
7	May	3318	35363	32045	29.04%
8	August	2769	33432	30663	27.45%
9	November	0	23876	23876	19.61%
10	Grand Total	2241	121779	119538	100.00%

**Рис. 45.43.** Продажи чипа Chip 2 во Франции

	A	B	C	D	E	F	G	H
1	Product	Chip 2						
2	Country	France						
3								
4	Values							
5	Row Labels	Sum of Var	Sum of Revenue	Sum of Budget	Sum of Revenue2			
6	February	-3846	29108	32954	23.90%			
7	May	3318	35363	32045	29.04%			
8	August	2769	33432	30663	27.45%			
9	November	0	23876	23876	19.61%			
10	Grand Total	2241	121779	119538	100.00%			
11								
12								
13								
14								
15								

**Country**

Canada

**France**

US

**Product**

Chip 1

**Chip 2**

Chip 3

**Рис. 45.44.** Представление продаж чипа Chip 2 во Франции со помощью срезов

Для создания такой таблицы с помощью срезов щелкните по сводной таблице на листе **Slicers** и на вкладке **Вставка** (Insert) в группе **Фильтры** (Filters) выберите **Срез** (Slicers). Для создания срезов, показанных на рис. 44.44, выберите поля **Продукт**

и Страна. Выбор Chip 2 из среза Продукт и France из среза Страна приводит к соответствующим расчетам для всех сделок, связанных со вторым чипом во Франции. Если вам требуется изменить размеры окна со срезом, потяните за метку-манипулятор на границе окна. Для пропорционального изменения размеров окна сделайте это, удерживая Ctrl.

### ? Как сгруппировать элементы в сводной таблице?

Часто бывает нужно сгруппировать заголовки в сводной таблице. Например, можно объединить продажи за январь — март (см. группу на листе). Для создания группы выделите элементы, которые хотите сгруппировать (нажмите Ctrl и выберите январь, февраль и март в ячейках A6:A8), щелкните правой кнопкой мыши по выделенной области и выберите Группировать (Group). После изменения имени Группа 1 на Январь-март и удаления дублирующего поля Месяц из области Строки (Rows) получится сводная таблица (рис. 45.45).

	A	B	C	D	E
1	Product	(All)			
2	Country	(All)			
3					
4	Values				
5	Row Labels	Sum of Var	Sum of Revenue	Sum of Revenue2	Sum of Budget
6	Jan-March	-2843	26.01%	266899	269742
7	April	-2774	8.28%	84982	87756
8	May	-423	8.24%	84559	84982
9	June	-548	8.19%	84011	84559
10	July	2366	8.42%	86377	84011
11	August	-2843	8.14%	83534	86377
12	September	1389	8.27%	84923	83534
13	October	-4318	7.85%	80605	84923
14	November	3406	8.19%	84011	80605
15	December	2366	8.42%	86377	84011
16	Grand Total	-4222	100.00%	1026278	1030500

Рис. 45.45. Группирование элементов для января, февраля и марта

## Замечания по группированию

- Вы можете разгруппировать данные, щелкнув правой кнопкой мыши по группе и выбрав Разгруппировать (Ungroup).
- Вы можете сгруппировать несмежные элементы, удерживая нажатым Ctrl при выборе несмежных строк и столбцов.
- Числовые данные или даты вы можете сгруппировать в произвольные интервалы. Например, можно создать группы для возрастных диапазонов и затем вычислить средний доход для возраста 25–34 года.

### ? Что такое вычисляемый объект?

*Вычисляемый объект* работает, как и вычисляемое поле, только вместо столбца создается строка. Для создания вычисляемого объекта сначала следует выделить элемент в столбце **Названия строк (Row Labels)** сводной таблицы (а не элемент в теле сводной таблицы). Затем на вкладке **Анализ (Analyze)** в группе **Вычисления (Calculations)** в списке **Поля, элементы и наборы (Fields, Items & Sets)** выберите **Вычисляемый объект (Calculated Item)**.

Рассмотрим файл *Calculateditem.xlsx*. На листе *Data* (рис. 45.46) показаны данные продаж автомобилей различных марок. Мы хотим обобщить продажи по странам (Япония, Германия и США).

Сначала создадим сводную таблицу суммирующую продажи по маркам автомобилей (см. лист *PT1*). Перетащите поле **Марка** в область **СТРОКИ (ROWS)**, а поле **Продажи** в область **ЗНАЧЕНИЯ (VALUES)**. После изменения **Количество** на **Сумма** (щелкните по **Количество** продаж в области **ЗНАЧЕНИЯ (VALUES)**, выберите **Параметры полей значений (Value Field Settings)**, выберите операцию **Сумма** в списке операций и нажмите **OK**) вы получите таблицу, показанную на рис. 45.47. (Напоминаю, чтобы открыть панель **Поля сводной таблицы (PivotTable Fields)**, на вкладке **Анализ (Analyze)**, в группе **Показать или скрыть (Show)**, щелкните **Список полей (Field List)**.)

	Н	И
8	Brand	Sales
9	Ford	3
10	Nissan	2
11	Ford	6
12	VW	2
13	VW	4
14	Nissan	4
15	Chrysler	2
16	VW	2
17	BMW	6
18	Honda	2
19	VW	4
20	Honda	5
21	BMW	6
22	Honda	4
23	Ford	5
24	Ford	5
25	Ford	2
26	Nissan	6

**Рис. 45.46.** Данные для создания вычисляемого объекта

Row Labels	Sum of Sales
BMW	359
Chrysler	286
Ford	277
GM	239
Honda	283
Nissan	219
VW	323
<b>Grand Total</b>	<b>1986</b>

**Рис. 45.47.** Сводная таблица, обобщающая продажи автомобилей по маркам

Перед созданием вычисляемого объекта целесообразно скрыть общие итоги. Для этого на вкладке **Конструктор (Design)** в группе **Общие итоги (Grand Totals)** выберите **Отключить для строк и столбцов (Off for Rows and Off for Columns)**. В противном случае после создания вычисляемого объекта каждая продажа будет учтена в общем итоге дважды.

Для создания вычисляемого объекта для Японии установите курсор в столбце **Названия строк (Row Labels)** сводной таблицы и на вкладке **Анализ (Analyze)** в группе **Вычисления** в списке **Поля, элементы и наборы (Fields Items and Sets)** выберите **Вычисляемый объект (Calculated Item)**. Выбрав **Вычисляемый объект (Calculated Item)**, заполните диалоговое окно — введите **Япония (Japan)** в поле **Имя (Name)**, формулу **=Honda+Nissan** в поле **Формула (Formula)**, выберите **Марка (Brand)** в списке **полей (Fields)** и **Nissan** в списке **элементов (Items)**. Заполненное диалоговое окно показано на рис. 45.48, осталось нажать **ОК**.

Insert Calculated Item in "Brand"

Name: Japan

Formula: = Honda+ Nissan

Fields:

- Brand
- Sales

Items:

- BMW
- Chrysler
- Ford
- Germany
- GM
- Honda
- Japan
- Nissan

Buttons: Modify, Delete, Insert Field, Insert Item, OK, Close

**Рис. 45.48.** Создание вычисляемого объекта Японии

Вычисляемый объект **Япония** создается путем суммирования продаж автомобилей **Honda** и **Nissan**. Аналогичным образом вы можете создать вычисляемые объекты **Германия** и **США**. Получившаяся сводная таблица представлена на листе **Calc Item** и на рис. 45.49.

Если вы хотите удалить вычисляемый объект или вычисляемое поле, в списке **Поля, элементы и наборы (Fields, Items & Sets)** выберите **Вычисляемый объект (Calculated Item)** или **Вычисляемое поле (Calculated Field)**. В диалоговом окне в списке **Имя (Name)** выберите объект или поле, которое требуется удалить, и нажмите **Удалить (Delete)**.



	A	B
1		
2		
3	Row Labels	Sum of Sales
4	BMW	359
5	Chrysler	286
6	Ford	277
7	GM	239
8	Honda	283
9	Nissan	219
10	VW	323
11	Japan	502
12	Germany	682
13	US	802

**Рис. 45.49.** Вычисляемые объекты

При необходимости можно скрыть продажи отдельных марок путем филь трации по меткам строк и оставив только продажи для Японии, Германии и США.

См. задание 11 этой главы в качестве примера создания вычисляемого объекта. В сводной таблице примера с производством микрочипов невозможно создать вычисляемый объект из-за нескольких копий поля **Выручка**.

### ? Что такое детализация?

Детализация — это отображение всех данных, обобщенных в поле, с помощью двойного щелчка на ячейке в сводной таблице. Так, в файле примера с микрочипами для отображения данных, связанных с продажами за март дважды щелкните на любой ячейке, содержащей **March**.

? Мне часто приходится использовать конкретные данные в сводной таблице для вычисления прибыли, например апрельских продаж чипа 1 во Франции. К сожалению, при добавлении новых данных в сводную таблицу эти перемещаются. Существует ли в Excel функция, позволяющая всегда извлекать из сводной таблицы объем апрельских продаж чипа 1 во Франции?

Да, такая функция существует. Нам подойдет функция **ПОЛУЧИТЬ.ДААННЫЕ.СВОДНОЙ.ТАБЛИЦЫ (GETPIVOTDATA)**. Пусть необходимо извлечь показатели продаж чипа 1 во Франции за апрель из сводной таблицы на листе **Data** в файле **Getpivotdata.xlsx** (рис. 45.50 и лист **Get Pivot Data**). Введенная в ячейку **E2** формула **=ПОЛУЧИТЬ.ДААННЫЕ.СВОДНОЙ.ТАБЛИЦЫ(A4;"April France Chip 1 Сумма по полю Выручка")** дает правильный ответ (**\$37 600**) даже после добавления в дальнейшем новых данных о продуктах и странах. Этот же результат можно получить, если просто указать в соответствующей ячейке знак **=** и затем выбрать ячейку сводной таблицы, содержащую показатели продаж чипа 1 во Франции за апрель (ячейка **D24**).



Первым аргументом функции является ячейка в левом верхнем углу сводной таблицы (ячейка A4). Далее вы вводите требуемые заголовки сводной таблицы (разделенные пробелами), заключив их в кавычки. Последним должно быть задано поле данных, но остальные заголовки могут следовать в произвольном порядке. Таким образом заданная, эта формула означает «Для сводной таблицы, левый верхний угол которой находится в ячейке A4, найти Сумму по полю Выручка для чипа 1 во Франции за апрель». Эта формула возвращает правильный ответ, даже если показатели продаж для чипа 1 во Франции за апрель будут перемещены в другое место сводной таблицы.

	A	B	C	D	E	F
1					April Chip 2 France	total revenue
2					37600	1026278
3						
4	Values					
5	Row Labels	Sum of Var	Sum of Budget	Sum of Revenue	Sum of Revenue2	Sum of Var percentage of Budget
6	January	-4297	91831	87534	8.53%	-0.046792477
7	Chip 1	-4297	91831	87534	8.53%	-0.046792477
8	Canada	2929	29330	32259	3.14%	0.099863621
9	France	-5772	33171	27399	2.67%	-0.174007416
10	US	-1454	29330	27876	2.72%	-0.049573815
11	February	2843	87534	90377	8.81%	0.032478808
12	Chip 2	2843	87534	90377	8.81%	0.032478808
13	Canada	3318	32045	35363	3.45%	0.103541894
14	France	-3846	32954	29108	2.84%	-0.116708139
15	US	3371	22535	25906	2.52%	0.149589527
16	March	-1389	90377	88988	8.67%	-0.015368954
17	Chip 3	-1389	90377	88988	8.67%	-0.015368954
18	Canada	-10733	35363	24630	2.40%	-0.303509318
19	France	11529	20784	32313	3.15%	0.554705543
20	US	-2185	34230	32045	3.12%	-0.063832895
21	April	-2774	87756	84982	8.28%	-0.031610374
22	Chip 1	-2774	87756	84982	8.28%	-0.031610374
23	Canada	1054	19289	20343	1.98%	0.054642542
24	France	-54	37654	37600	3.66%	-0.001434111
25	US	-3774	30813	27039	2.63%	-0.122480771

**Рис. 45.50.** Функция ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ для определения показателей продаж чипа 1 во Франции за апрель

Для получения общей выручки (\$1 026 278) введите формулу =ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ(A4;"Сумма по полю Выручка") в ячейку F2.

Чтобы увидеть функцию ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ во всей красе, просуммируйте показатели продаж каждого продукта по странам для указанного месяца в таблице, представленной на рис. 45.51.

Я начал с того, что создал выпадающий список, который позволит мне вводить месяц года в ячейке J9. Введем в ячейки I13:I15 названия стран и в ячейки J12:L12 названия продуктов. В ячейку J13 скопируем формулу ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ из ячейки E2 и отредактируем ее как =ЕСЛИОШИБКА(ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ("Сумма по полю Выручка"; \$A\$4; "Месяц"; \$J\$9; "Продукт"; J\$12;

	I	J	K	L	M	N	O	P	Q	R
7										
8										
9	Month	April								
10										
11										
12		Chip 1	Chip 2	Chip 3						
13	Canada	20343	0	0						
14	France	37600	0	0						
15	US	27039	0	0						
16										
17	J13 formula									
18	=IFERROR(GETPIVOTDATA("Sum of Revenue",\$A\$4,"Month",\$J\$9,"Product",\$J\$12,"Country",\$I13),0)									

**Рис. 45.51.** Извлечение показателей продаж каждого продукта в апреле в каждой стране из J13:L15 с помощью функции ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ

"Страна"; \$I13); 0). Копирование этой формулы из J13 в J13:L15 позволит извлечь из сводной таблицы показатели продаж каждого продукта за апрель. При копировании формулы в ячейки справа изменяется продукт, при копировании формулы в нижние ячейки — страна. В каждой ячейке месяц выбирается из ячейки J9. При отсутствии данных о продажах продукта в стране в выбранном месяце функция ЕСЛИОШИБКА выдает 0, а не сообщение об ошибке. Только представьте, насколько это удобно при продаже 1000 продуктов в 200 странах и регионах!

Зачастую функция ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ только мешает. В таких случаях ее следует отключить. Допустим, вы хотите сослаться на данные в ячейках B5:B11 сводной таблицы где-либо в другом месте книги. Вероятно, вы воспользуетесь формулой =B5, скопируете ее в диапазон B6:B11 и будете ждать, что это действие извлечет значения из ячеек B6, B7, ..., B11. Однако если функция ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ активна, вместо этого во всех ячейках будет получена функция ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ, ссылающаяся на одну и ту же ячейку. Для отключения функции ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ откройте вкладку Файл (File) и выберите Параметры (Options). На панели параметров Excel выберите Формулы (Formulas) и в группе Работа с формулами (Working With Formulas) снимите флажок Использовать функции GetPivotData для ссылок в сводной таблице (GetPivotData Function For PivotTable References). После этого при выборе ячейки в сводной таблице будет появляться формула типа =B6, а не функция ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ. Функцию ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ также можно отключить для отдельной таблицы, если на вкладке Анализ (Analyze) в группе Сводная таблица (PivotTable) в раскрывающемся списке Параметры (Options) снять флажок Создать GetPivotData (Generate GetPivotData). Если вы просто введете в ячейку =B5, не при помощи указателя мыши, то сможете перетащить формулу =B5 вниз, и можно не применять функцию GetPivotData.

Наконец, следует отметить, что для извлечения различных данных из сводной таблицы вы можете комбинировать функции ПОИСКПОЗ (MATCH) и СМЕЩ (OFFSET),

описанные в главах 5 «Функция ПОИСКПОЗ» и 22 «Функция СМЕЩ» соответственно.

### ❓ Как применить временную шкалу для обобщения данных за разные периоды времени?

В Excel 2013 появился прекрасный новый инструмент **Временная шкала (Timeline)** для фильтрации сводной таблицы по временным периодам. Вы можете выбрать любое подмножество последовательных календарных лет, кварталов, месяцев или дней, а функция **Timeline** обеспечит включение во все расчеты в сводной таблице строк только из выбранного периода времени.

Для примера рассмотрим данные в файле **Makeuptimeline.xlsx**. Лист **Data** содержит информацию по 1900 сделкам с косметикой: **Продавец** (столбец **Name**, **H**), **Дата** (столбец **I**), **Продукт** (столбец **J**), **Количество** (столбец **K**) и **Выручка** (столбец **Dollars**, **L**). На листе **Pivot Table** находится сводная таблица, суммирующая продажи каждого продукта по каждому продавцу. Щелкните в пределах таблицы, откройте вкладку **Вставка (Insert)** и в группе **Фильтры (Filters)** выберите инструмент **Временная шкала (Timeline)**. В диалоговом окне **Вставка временных шкал (Insert Timelines)** выберите **Дата (Date)** и нажмите **ОК**. Появится временная шкала, показанная на рис. 45.52.

	A	B	C	D	E	F	G
3	Sum of Dollars	Column Labels					
4	Row Labels	eye liner	foundation	lip gloss	lipstick	mascara	Grand Total
5	Ashley	\$5,844.95	\$4,186.06	\$6,053.68	\$3,245.44	\$6,617.10	\$25,947.24
6	Betsy	\$6,046.53	\$8,276.84	\$5,683.91	\$3,968.61	\$4,827.25	\$28,803.15
7	Cici	\$5,982.82	\$6,198.25	\$5,199.95	\$3,148.84	\$7,060.71	\$27,590.57
8	Colleen	\$3,389.63	\$6,834.77	\$5,573.32	\$2,346.41	\$6,746.53	\$24,890.66
9	Cristina	\$5,397.27	\$5,290.99	\$5,297.98	\$2,401.67	\$5,461.65	\$23,849.56
10	Emilee	\$7,587.39	\$5,492.80	\$5,270.25	\$2,189.14	\$4,719.30	\$25,258.87
11	Hallagan	\$6,964.62	\$7,256.20	\$5,603.12	\$3,177.87	\$5,703.35	\$28,705.16
12	Jen	\$7,010.44	\$5,747.95	\$5,461.61	\$3,953.30	\$6,877.23	\$29,050.53
13	Zaret	\$8,270.19	\$6,451.65	\$5,690.81	\$2,448.71	\$3,879.95	\$26,741.31
14	Grand Total	56493.84331	55735.50749	49834.6434	26879.98743	51893.06545	240837.0471
15			Date				
16			All Periods				
17			QUARTERS				
18			2004				
19			2005				
20			2006				
21			Q1	Q2	Q3	Q4	Q1
22							
23							

**Рис. 45.52.** Временная шкала для данных о продажах декоративной косметики

Для обобщения показателей продаж можно выделить смежные кварталы, нажав **Shift**. Например, на листе **Timeline** (рис. 45.53) просуммированы продажи за 2004 г. и два первых квартала 2005 г. Для восстановления сводной таблицы по всем периодам можно нажать кнопку с воронкой **Очистить фильтр (Clear Filter)**.

	A	B	C	D	E	F	G	
3	Sum of Dollars	Column Labels						
4	Row Labels	eye liner	foundation	lip gloss	lipstick	mascara	Grand Total	
5	Ashley	\$2,259.97	\$2,064.62	\$3,387.84	\$1,191.92	\$4,773.72	\$13,678.07	
6	Betsy	\$2,714.91	\$4,195.09	\$2,862.84	\$2,223.66	\$2,591.12	\$14,587.61	
7	Cici	\$3,310.71	\$3,313.91	\$2,076.46	\$1,676.03	\$3,026.36	\$13,403.47	
8	Colleen	\$2,122.76	\$4,064.70	\$3,190.32	\$1,454.87	\$3,163.44	\$13,996.09	
9	Cristina	\$3,890.14	\$2,047.07	\$2,844.16	\$1,531.34	\$2,708.86	\$13,021.57	
10	Emilee	\$4,086.75	\$2,754.36	\$2,808.22	\$1,180.96	\$2,462.89	\$13,293.17	
11	Hallagan	\$4,195.40	\$3,622.60	\$2,857.48	\$2,133.18	\$2,309.15	\$15,117.81	
12	Jen	\$3,364.89	\$2,830.40	\$1,875.47	\$1,762.88	\$3,866.27	\$13,699.91	
13	Zaret	\$3,188.38	\$2,695.03	\$3,040.78	\$1,659.86	\$1,828.48	\$12,412.53	
14	Grand Total	29133.89079	27587.77037	24943.58511	14814.69089	26730.28606	123210.2232	
15			Date					
16								
17			Q1 2004 - Q2 2005				QUARTERS	
18			2004		2005		2006	
19			Q1	Q2	Q3	Q4	Q1	Q2
20								
21								
22								
23								

**Рис. 45.53.** Временная шкала для продаж за 2004 г. и первые два квартала 2005 г.

❓ Как с помощью сводной таблицы получить нарастающий итог общего объема продаж для каждого периода года?

Суммировать общие показатели продаж для каждого периода года можно с помощью параметров полей значений — Value Field Settings. Для примера рассмотрим данные в файле MonthtoMonth.xlsx. На листе Data содержатся Год, Месяц и Выручка для некоторого количества сделок. Сначала на листе Year to Date мы суммировали ежемесячные продажи для каждого года, перетаскив поле Месяц в область СТРОКИ (ROWS), поле Год в область СТОЛБЦЫ (COLUMNS) и поле Выручка в область ЗНАЧЕНИЯ (VALUES). Сводная таблица представлена на рис. 45.54.

Sum of Revenue	Column Labels				
Row Labels		2009	2010	2011	Grand Total
January		10453	84058	45615	140126
February		47996	74896	32943	155835
March		113126	41689	19821	174636
April		69613	59910	39770	169293
May		65155	49345	29600	144100
June		54814	42355	61331	158500
July		34930	87516	38863	161309
August		51588	33060	47287	131935
September		60835	48963	45559	155357
October		85607	36243	30805	152655
November		72602	17010	48207	137819
December		30043	73381	79733	183157
Grand Total		696762	648426	519534	1864722

**Рис. 45.54.** Итоговые продажи по месяцам и годам

На этом рисунке продажи разбиты по месяцам и годам. Щелкните правой кнопкой мыши в пределах ячеек B5:E17 сводной таблицы и в контекстном меню выберите **Параметры полей значений (Value Field Settings)**. В появившемся диалоговом окне на вкладке **Дополнительные вычисления (Show Values As)** в раскрывающемся списке выберите **С нарастающим итогом в поле (Running Total In)**. Выделите поле **Месяц**. Нажмите **ОК**. Получившаяся таблица, приведенная на рис. 45.55 (см. лист **Year to Date**), отображает продажи для каждого года с нарастающим итогом в каждом месяце. Например, объем продаж к февралю 2009 г. составил \$58 449.

	A	B	C	D	E
1					
2					
3	Sum of Revenue Column Labels				
4	Row Labels	2009	2010	2011	Grand Total
5	January	10453	84058	45615	140126
6	February	58449	158954	78558	295961
7	March	171575	200643	98379	470597
8	April	241188	260553	138149	639890
9	May	306343	309898	167749	783990
10	June	361157	352253	229080	942490
11	July	396087	439769	267943	1103799
12	August	447675	472829	315230	1235734
13	September	508510	521792	360789	1391091
14	October	594117	558035	391594	1543746
15	November	666719	575045	439801	1681565
16	December	696762	648426	519534	1864722
17	Grand Total				

**Рис. 45.55.** Объемы продаж с нарастающим итогом

**?** Как с помощью сводной таблицы сравнить объем продаж за месяц с объемом продаж за тот же месяц предыдущего года?

Снова воспользуемся данными из файла **MonthtoMonth.xlsx**. Создадим таблицу на листе **Previous Year**, для чего перетащим поле **Месяц** в область **СТРОКИ (ROWS)**, поле **Год** в область **СТОЛБЦЫ (COLUMNS)** и поле **Выручка** в область **ЗНАЧЕНИЯ (VALUES)**. Щелкните правой кнопкой мыши в диапазоне B5:E17 и в контекстном меню выберите **Параметры полей значений (Value Field Settings)**. Заполните диалоговое окно, как показано на рис. 45.56; создастся сводная таблица, представленная на рис. 45.57: в поле **Пользовательское имя (Custom Name)** напишите **Сумма дохода (Sum of Revenue)**, также выберите **Представить значения в виде таблицы (Show Values As tab)**, выберите **%-ную разницу (%Difference)** в списке **Дополнительные вычисления (Show Value As)**, **Год (Year)** в списке **Основное поле (Base Field)**, **Предыдущий (Previous)** в списке **Основная единица (Base Item)** и нажмите **ОК**. На этом рисунке продажи за каждый месяц сравниваются с продажами того же месяца прошлого года. Например, продажи в январе 2010 г. возросли на 704,15% по сравнению с продажами в январе 2009 г.

**Рис. 45.56.** Параметры для сравнения продаж за месяц с продажами за тот же месяц предыдущего года

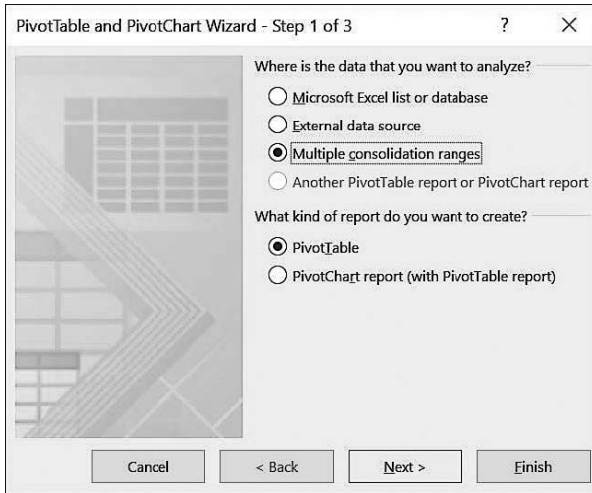
	A	B	C	D	E
3	Sum of Revenue	Column Labels			
4	Row Labels	2009	2010	2011	Grand Total
5	January		704.15%	-45.73%	
6	February		56.05%	-56.02%	
7	March		-63.15%	-52.46%	
8	April		-13.94%	-33.62%	
9	May		-24.27%	-40.01%	
10	June		-22.73%	44.80%	
11	July		150.55%	-55.59%	
12	August		-35.92%	43.03%	
13	September		-19.52%	-6.95%	
14	October		-57.66%	-15.00%	
15	November		-76.57%	183.40%	
16	December		144.25%	8.66%	
17	Grand Total		-6.94%	-19.88%	

**Рис. 45.57.** Сравнение продаж за месяц с продажами за тот же месяц предыдущего года

### ? Как создать сводную таблицу на основе данных из нескольких местоположений?

Зачастую данные, необходимые для создания сводной таблицы, находятся на разных листах или в разных книгах (файлах). Чтобы создать сводную таблицу из данных в разных местоположениях, нужно нажать комбинацию **Alt+D+P** и вызвать классический Мастер сводных таблиц и диаграмм (PivotTable And PivotChart Wizard), представленный на рис. 45.58. (Может потребоваться нажать комбинацию **Alt+D+P** несколько раз.)





**Рис. 45.58.** Классический Мастер сводных таблиц и диаграмм

Научимся создавать сводную таблицу на основе данных из разных диапазонов на примере файлов *East.xlsx* и *West.xlsx* из папки *Templates*. На вкладке Вид (View) выберите Упорядочить все (Arrange All) в группе Окно (Window) и в диалоговом окне Расположение окон (Arrange Windows) задайте расположение окон рядом друг с другом, как показано на рис. 45.59. Данные представляют собой продажи за январь, февраль и март в регионах Восток и Запад. Требуется создать сводную таблицу, обобщающую продажи каждого продукта по месяцам.

Product	January	February	March
A	205	263	20
B	164	-17	146
C	278	177	179
D	156	214	240
D	72	134	48
D	7	256	104
A	141	87	148
A	2	-15	135
A	-44	47	72

Product	January	February	March
A	173	1	256
A	208	201	224
B	176	33	350
B	190	249	215
D	162	74	156
D	90	150	170
D	112	284	141
G	154	217	113
G	152	200	275

**Рис. 45.59.** Два файла будут обобщены с помощью сводной таблицы

Сначала нажмите **Alt+D+P**, чтобы открыть Мастер сводных таблиц и диаграмм (PivotTable And PivotChart Wizard), и установите переключатель в нескольких диапазонах консолидации (Multiple Consolidation Ranges) на шаге 1. Нажмите кнопку Далее

(Next), на шаге 2а установите переключатель в положение **Создать одно поле страницы** (Create A Single Page Field For Me), на шаге 2б выберите данные о продажах на листе **East**, **EAST!\$A\$1:\$D\$18** (рис. 45.60) и нажмите **Добавить** (Add) для добавления их в диапазон данных, который будет использоваться при создании сводной таблицы.

**Рис. 45.60.** Добавление данных листа East

Теперь удалите ссылку на данные в файле **East** из поля **Диапазон (Range)**; выберите данные в файле **West**, **WEST!\$A\$1:\$D\$24** и нажмите **Добавить** (Add) для добавления их в **Список диапазонов (All Ranges)**. Нажмите **Далее (Next)** и (шаг 3) примите решение о размещении окончательной сводной таблицы — на новом или текущем листе. В данном случае был выбран новый лист. Нажмите **Готово (Finish)**. Появится сводная таблица, представленная на рис. 45.61 и на листе **PT** в файле **West.xlsx**.

	A	B	C	D	E
1	Page1	(All)			
2					
3	Sum of Value	Column			
4	Row	January	February	March	Grand Total
5	A	1323	1317	1445	4085
6	B	890	335	812	2037
7	C	1231	922	843	2996
8	D	767	1424	1199	3390
9	E	579	483	371	1433
10	F	597	577	327	1501
11	G	570	850	811	2231
12	H	131	71	266	468
13	Grand Total	6088	5979	6074	18141

**Рис. 45.61.** Сводная таблица, обобщающая данные о продажах в регионах Восток и Запад



Как видите, общий объем продаж продукта А в феврале составил 1317 единиц. Данные можно отфильтровать по продуктам с помощью раскрывающегося списка в ячейке А4 или по месяцам с помощью списка в ячейке В3. Раскрывающийся список в ячейке В1 позволяет использовать при фильтрации сводной таблицы данные только для региона Восток или Запад. Команда **Обновить (Refresh)** в контекстном меню обновляет сводную таблицу и показывает измененные данные. Срезы и временные шкалы в сводных таблицах, созданных из нескольких диапазонов, не работают.

Если комбинация **Alt+D+P** вам не нравится, можно добавить **Мастер сводных таблиц и диаграмм (PivotTable And PivotChart Wizard)** на Панель быстрого доступа. Для этого откройте вкладку **Файл (File)**, выберите команду **Параметры (Options)**, затем среди параметров Excel выберите **Панель быстрого доступа (Quick Access Toolbar)** и в раскрывающемся списке **Выбрать команды из** выберите **Команды не на ленте (Commands Not on the Ribbon)** и затем **Мастер сводных таблиц и диаграмм (PivotTable and PivotChart Wizard)**. После этого нажмите **Добавить (Add)**, чтобы эта команда добавилась на панель быстрого доступа, и нажмите **ОК**, чтобы закрыть диалоговое окно **Параметры Excel (Excel Options)**.

При создании сводной таблицы из нескольких диапазонов заголовки (в данном случае это **Январь, Февраль и Март**) в этих диапазонах должны быть идентичны. В главе 46 мы рассмотрим инструмент, появившийся в Excel 2013, — **Модель данных (Data Model)**, позволяющий создавать сводные таблицы даже в том случае, если заголовок в одном из исходных диапазонов не содержится в других исходных диапазонах.

### **Как создать сводную таблицу на основе уже имеющейся сводной таблицы?**

Зачастую нам нужно создать сводную таблицу на основе уже имеющейся. Это позволяет просмотреть несколько сводных таблиц, основанных на одних и тех же данных. Например, в файле **Makeuptimeline.xls** (см. рис. 45.52) была создана сводка по продажам косметики с перечислением продавцов в столбце и продуктов в строке. Допустим, мы хотим создать сводную таблицу с перечислением продуктов в столбце и продавцов в строке. Установите курсор внутри сводной таблицы на листе **Pivot Table** файла **Makeuptimeline.xls**, вызовите **Мастер сводных таблиц и диаграмм (PivotTable and PivotChart Wizard)** нажатием комбинации **Alt+D+P**. На шаге 1 установите переключатель в положение **в другой сводной таблице или сводной диаграмме (Another PivotTable or PivotChart Report)** и нажмите **Далее (Next)**. На следующем шаге выберите сводную таблицу, из которой требуется создать новую таблицу. Выбираем нашу исходную сводную таблицу: **PivotTable10**. Нажмите **Готово (Finish)**. Откроется панель **Поля сводной таблицы (PivotTable Fields)**, и вы сможете создать новую сводную таблицу без разрушения старой. Перетащите поле **Продукт** в область **СТРОКИ (ROWS)**, а поле **Продавец** в область **СТОЛБЦЫ (Columns)**. Создастся сводная таблица, представленная на рис. 45.62.

	H	I	J	K	L	M	N	O	P	Q	R
24	PIVOTTABLE BASED ON ANOTHER PIVOTTABLE										
25											
26	Sum of Dollars	Column Labels									
27	Row Labels	Ashley	Betsy	Cici	Colleen	Cristina	Emilee	Hallagan	Jen	Zaret	Grand Total
28	eye liner	\$5,844.95	\$6,046.53	\$5,982.82	\$3,389.63	\$5,397.27	\$7,587.39	\$6,964.62	\$7,010.44	\$8,270.19	\$56,493.84
29	foundation	\$4,186.06	\$8,276.84	\$6,198.25	\$6,834.77	\$5,290.99	\$5,492.80	\$7,256.20	\$5,747.95	\$6,451.65	\$55,735.51
30	lip gloss	\$6,053.68	\$5,683.91	\$5,199.95	\$5,573.32	\$5,297.98	\$5,270.25	\$5,603.12	\$5,461.61	\$5,690.81	\$49,834.64
31	lipstick	\$3,245.44	\$3,968.61	\$3,148.84	\$2,346.41	\$2,401.67	\$2,189.14	\$3,177.87	\$3,953.30	\$2,448.71	\$26,879.99
32	mascara	\$6,617.10	\$4,827.25	\$7,060.71	\$6,746.53	\$5,461.65	\$4,719.30	\$5,703.35	\$6,877.23	\$3,879.95	\$51,893.07
33	Grand Total	\$26,947.24	\$28,803.16	\$27,690.67	\$24,890.66	\$23,849.66	\$26,268.87	\$28,706.16	\$29,060.63	\$26,741.31	\$240,837.05

Рис. 45.62. Сводная таблица на основе другой сводной таблицы

**?** Как использовать фильтр отчета для создания нескольких сводных таблиц?

На листе Slicers в файле Ptableexample.xlsx мы создавали фильтры на основе Страна и Продукт. Предположим, нам требуется создать отдельную сводную страницу по каждой стране. Выберите ячейку внутри сводной таблицы. В левой части вкладки Анализ (Analyze) в группе Сводная таблица (PivotTable) выберите Параметры (Options), как показано на рис. 45.63. Затем выберите поле Страна (Country) для создания отдельной сводной таблицы для каждой страны и поле Продукт (Product) для создания сводной таблицы для каждого продукта. На рис. 45.64 представлена сводная таблица с продажами во Франции.

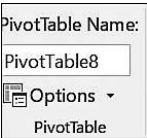


Рис. 45.63. Выберите Параметры для создания нескольких сводных таблиц с помощью фильтра отчетов

	A	B	C	D	E
1	Product	(All)			
2	Country	France			
3					
4	Values				
5	Row Labels	Sum of Var	Sum of Revenue	Sum of Budget	Sum of Revenue2
6	January	-5772	27399	33171	7.43%
7	February	-3846	29108	32954	7.89%
8	March	11529	32313	20784	8.76%
9	April	-54	37600	37654	10.19%
10	May	3318	35363	32045	9.59%
11	June	-3740	32855	36595	8.91%
12	July	-275	27239	27514	7.38%
13	August	2769	33432	30663	9.06%
14	September	-2983	29217	32200	7.92%
15	October	-1917	27300	29217	7.40%
16	November	0	23876	23876	6.47%
17	December	912	33171	32259	8.99%
18	Grand Total	-59	368873	368932	100.00%

Рис. 45.64. Сводная таблица с продажами во Франции, созданная с помощью фильтра по странам

### ? Как быстро изменить настройки для сводной таблицы по умолчанию?

Если у вас установлен Office 365, изменить настройки по умолчанию для всех ваших сводных таблиц легко. На вкладке **Файл (File)** выберите **Параметры (Options)** и затем — **Данные (Data)**. Выбрав **Редактировать вид по умолчанию (Edit Default Layout)** в **Настройках данных (Data Options)**, вы сможете изменить настройки по умолчанию для всех сводных таблиц. Существует также настройка **Вернуть настройки Excel по умолчанию (Reset To Excel Default)**, так что не бойтесь экспериментировать!

## Задания

1. Contoso, Ltd. производит микрочипы. При производстве было выявлено пять типов дефектов (обозначенных 1–5). Чипы производятся двумя операторами (А и В) на четырех станках (1–4). Вам предоставили следующие данные о дефектных чипах: тип дефекта, имя оператора, номер станка и день недели, когда дефект был обнаружен. На основе этих данных определите, что следует предпринять как можно быстрее, чтобы повысить качество продукции. Для разделения дефектов с учетом типа дефекта, дня недели, станка и оператора следует использовать сводные таблицы. Возможно, потребуется разбивка данных по станкам, операторам и т. д. Пусть каждый оператор и каждый станок произвели одинаковое количество продукции. Вы найдете данные в файле **Contoso.xlsx**.
2. Вы, владелец ресторана быстрого питания, произвели маркетинговое исследование в целях лучшего понимания потребностей клиентов. Для случайной выборки клиентов указаны их доход, пол и количество дней в неделю, когда они посещают рестораны быстрого питания. На основе этой информации определите, как пол и доход влияют на частоту посещения ресторанов быстрого питания. Вы найдете данные в файле **McDonalds.xlsx**.
3. Студенты Школы изящных искусств записались на курс или английского языка, или методов научного исследования. Вам поручили определить, есть ли в Школе изящных искусств дискриминация по половому признаку при приеме студентов на обучение с учетом их выбора. О студентах Школы изящных искусств имеются следующие данные:
  - женщина (Female) или мужчина (Male);
  - предпочитаемая дисциплина: английский (Eng) или методы научного исследования (Sci);
  - принят? (Yes или No).

При условии что женщины обладают равной с мужчинами квалификацией в каждой из дисциплин, свидетельствуют ли эти данные о дискриминации женщин? Используйте всю доступную информацию. Вы найдете данные в файле **Finearts.xlsx**.

4. Вам поручено оценить качество медицинской помощи, оказанной пациентам с инфарктом в больницах Emergency Room (ER) и Chicago Hope (CH). За последние месяцы получены следующие данные о пациентах:

- больница (ER или CH);
- группа риска (High или Low). Пациенты из группы высокого риска выживают реже, чем пациенты из группы низкого риска;
- исход для пациента (Live или Die).

На основе этих данных определите, в какой из больниц качество оказания медицинской помощи пациентам с инфарктом лучше. Подсказка: используйте все доступные данные. Вы найдете данные в файле **Hospital.xlsx**.

5. Дан ежемесячный уровень индекса Доу — Джонса для 1947–1992 гг. Отражают ли эти данные какие-либо необычные сезонные закономерности доходности акций? Подсказка: месяц (январь, февраль и т. д.) можно извлечь с помощью формулы `=ТЕКСТ(A4,"MMM")`, скопированной в любой столбец. Вы найдете данные в файле **Dow.xlsx**.

6. Файл **Makeupdb.xlsx** содержит информацию о продажах косметических продуктов. Для каждой сделки имеются следующие данные:

- имя продавца;
- дата продажи;
- продаваемый продукт;
- количество проданного;
- выручка.

Создайте сводную таблицу, обобщающую следующую информацию:

- число сделок для каждого продавца;
  - общая выручка по продукту для каждого продавца;
  - на основе предыдущего результата создайте функцию, которая всегда возвращает данные о продаже помады для Джен (Jen);
  - общая выручка для каждого продавца с разбивкой по регионам;
  - общая выручка по каждому продавцу и году Подсказка: необходимо сгруппировать данные по годам.
7. Для 1985–1992 гг. даны ежемесячные процентные ставки по облигациям, по которым деньги выплачиваются через год после покупки. Часто подразумевается, что высокие процентные ставки являются более подвижными. Соответствуют ли данные в файле **Intratevolvolatility.xlsx** этому утверждению? Подсказка: в сводных таблицах можно отобразить стандартные отклонения.
8. Для примера с продовольственными товарами подготовьте диаграмму, обобщающую тенденцию продаж по годам для каждого магазина.
9. Для примера с продовольственными товарами создайте вычисляемое поле для расчета средней цены единицы продукции для каждого продукта.

10. Для примера с продовольственными товарами создайте сводную диаграмму, суммирующую продажи каждого продукта в каждом магазине за 2005 и 2006 г.
11. На основе данных в файле `Calcitemdata.xlsx` создайте вычисляемые объекты, суммирующие продажи десертов (пирожные + пудинги) и фруктов (яблоки + виноград).
12. Для примера с микрочипами создайте сводную таблицу, суммирующую ежемесячные продажи чипа 1 и чипа 3 во Франции и в США.
13. Для примера со сводной таблицей клиентов покажите 15 самых крупных клиентов в одной таблице и 5 самых мелких клиентов в другой таблице.
14. Файл `Ptablepartsdta.xlsx` содержит данные о продажах различных запчастей. Код для каждой запчасти начинается либо с **Part** (для комплектующих), либо с **Comp** (для компьютеров). Создайте сводную таблицу, отображающую продажи только для комплектующих. Подсказка: используйте фильтр по подписи.
15. На основе данных к заданию 14 просуммируйте общие продажи для комплектующих и компьютеров.
16. Файл `Cigarettedata.xlsx` содержит данные выборки для американцев: возраст, информацию о том, курят ли они и что именно, сигареты или сигары, а также умерли ли они в текущем году. Какие выводы можно сделать на основании этих данных?
17. В файле `Collegedata.xlsx` находится следующая информация о студентах, подавших заявления в аспирантуру университета Kelley University: пол, желаемая специализация, было ли заявление принято или отклонено. Если вы создадите сводную таблицу правильно, то увидите, что женщин было принято меньше, чем мужчин. Есть ли в этом университете дискриминация женщин?
18. В файле `AnalyzeSurveydata.xlsx` содержатся ответы по шкале от 1 до 7 на различные вопросы анкеты о качестве преподавания. Создайте сводную диаграмму, отображающую долю каждой оценки (1–7) в ответе на каждый вопрос. Отфильтруйте диаграмму для демонстрации разбивки для любого набора из трех вопросов по вашему выбору.
19. На основе данных из файла `MonthtoMonth.xlsx` создайте сводную таблицу, которая показывала бы отличие каждой ежемесячной продажи от продажи за предыдущий месяц 2010 г.
20. Более 1000 зарегистрированных избирателей были опрошены об их политических взглядах (либеральные, умеренные или консервативные) и о том, поддерживают они или нет DREAM Act. Опираясь на данные из файла `Problem20data.xlsx`, постройте сводную таблицу и сводную диаграмму, показывающую, как отношение к DREAM Act зависит от политических взглядов.
21. В файле `Problem21data.xlsx` приведена следующая информация по репрезентативной выборке населения США:
  - Выписываете ли вы Garden and Gun Magazine?

- Уровень дохода (низкий, средний или высокий).
- Место жительства (город, сельская местность, пригород).

Используйте сводную таблицу, чтобы определить категорию людей, которые наиболее вероятно выписывают Garden and Gun Magazine.

22. Файл **Problem22data.xlsx** содержит информацию о родах, принятых в Dylan General Hospital в течение 1995 г. Создайте сводную таблицу, обобщающую данные о рождениях по дням недели. Затем создайте другую сводную таблицу или сводную диаграмму, которая позволит вам интерпретировать закономерности в рождениях по дням недели.
23. Штат Happyville состоит из девяти округов по выборам в конгресс. В файле **Problem23data.xlsx** содержатся данные об избирателе, округе, к которому он относится, и о том, демократ он или республиканец. Создайте сводную таблицу и сводную диаграмму, отражающую процент демократов и республиканцев в каждом округе. Предполагая, что партия, имеющая в округе большинство голосов избирателей, побеждает на выборах, выясните, в скольких округах победит каждая партия. Подсказка: эта задача показывает, почему республиканцы обычно контролируют палату представителей конгресса, хотя демократы обычно получают больше голосов избирателей в совокупности по стране.
24. Файл **Problem24data.xlsx** содержит позиции на игровом поле и зарплаты игроков команды НФЛ Carolina Panther в 2016 г. Создайте выпадающее меню, из которого можно выбрать игрока. Затем напишите формулы, чтобы показать позицию и зарплату игрока. Создайте круговую диаграмму на основе сводной таблицы, сравнивающую в процентах зарплаты, выплачиваемые на каждой позиции. При добавлении данных и обновлении сводной таблицы ваша диаграмма и формулы должны автоматически обновляться.
25. Файл **Problem25data.xlsx** содержит данные об игроках НБА в 2016 г. Процент забитых голов (EFG, effective field goal percentage) игрока рассчитывается как  $EFG = (FG + .5 * 3P) / FGA$ , где FG = забитые на поле голы, 3P = голы, принесшие три очка, а FGA = попытки забить гол на поле. С помощью вычисляемого поля определите EFG каждой команды НБА.
26. Файл **Problem26data.xlsx** содержит информацию по продажам для сети универсальных магазинов. Дан коэффициент налогообложения для каждого штата, а с благотворительных продаж налог не платится. Создайте сводную таблицу, показывающую сумму уплаченного налога с продаж для каждой группы продуктов в каждом штате, и сделайте срезы, при помощи которых будет легче фильтровать вашу сводную таблицу.
27. Файл **Problem27data.xlsx** содержит ежедневные курсы биткойна. Создайте сводную таблицу, обобщающую ежедневные доходы по годам, и еще одну сводную таблицу, обобщающую ежедневные доходы по дням недели. Для каждой сводной таблицы сделайте диаграмму. Подсказка: для начала вам, возможно, захочется изменить формат в столбце с датами.

## ГЛАВА 46

# Модель данных

### Обсуждаемые вопросы

- Что представляет собой инструмент Модель данных и зачем он нужен?
- Как добавить данные в модель данных?
- Как удалить данные из модели данных?
- Как создать связь в модели данных?
- Как с помощью модели данных создать сводную таблицу?
- Как добавить новые данные в модель данных?
- Как с помощью модели данных создать новую сводную таблицу?
- Как изменять и удалять связи?
- Как работает операция Число различных элементов (DISTINCT COUNT)?

## Ответы на вопросы

### ? Что представляет собой инструмент Модель данных и зачем он нужен?

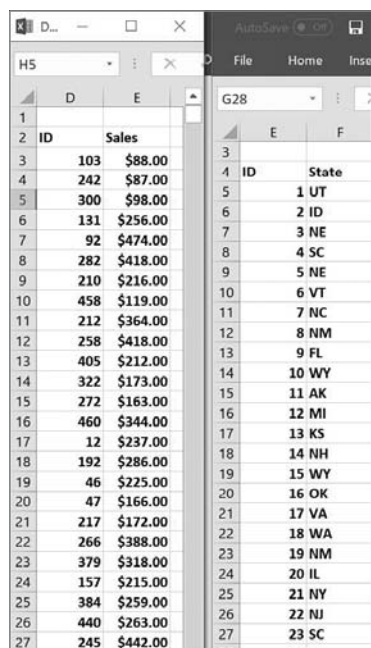
Модель данных — это инструмент, добавленный в Excel 2013. (В Microsoft Excel 2010 была надстройка **Модель данных**.) Модель данных обеспечивает простой способ загрузки данных, превышающих по объему обычные возможности Microsoft Excel 2019 (1 048 576 строк данных), и создание на их основе сводных таблиц. Модель данных также позволяет комбинировать данные, поступающие в Excel из внешних источников (включая Access и SQL Server), с данными Excel. Наконец, четкое представление о модели данных упрощает изучение возможностей Power Pivot (см. главу 47 «Power Pivot»).

В главе 45 «Сводные таблицы и срезы для описания данных» мы подробно разобрали, как использовать сводные таблицы для обобщения данных. Я показал, как создавать сводные таблицы на основе данных из разных местоположений, но этот подход требует, чтобы данные имели одинаковые заголовки столбцов. Довольно часто это условие не соблюдается. Например, файл **Datamodeltemp.xlsx** в папке **Templates** содержит два листа. На листе **Reps** (показанном на рис. 46.1 справа) перечислены идентификационные номера торговых представителей и штаты, в которых они продают продукцию. На листе **Sales** (показанном на рис. 46.1 слева) указаны объемы продаж для каждого сотрудника. Разумеется, мы хотим использовать



сводную таблицу для обобщения продаж по штатам. Но проблема в том, что информация на листе **Sales** не включает штат, с которым работает каждый продавец. Для вставки соответствующего штата на лист **Sales** можно было бы добавить столбец с помощью функции **ВПР (VLOOKUP)**. Однако если у нас сотни тысяч строк данных, формулы с функцией **ВПР** заметно снизят производительность электронных таблиц. Модель данных Excel позволяет обойти формулы с **ВПР** путем создания связи между каждым продавцом и соответствующим штатом без необходимости обращения к медленным функциям **ВПР**.

Наконец, если вы хотите свести воедино данные из различных источников, таких как базы данных, интернет, текстовые файлы и таблицы Excel, то вам понадобится **Power Pivot** (описанный в главе 47). После того как вы освоите инструмент **Модель данных (Data Model)**, понять **Power Pivot** будет намного проще. **Power Pivot** доступен во всех версиях Excel 2019.



ID	Sales
103	\$88.00
242	\$87.00
300	\$98.00
131	\$256.00
92	\$474.00
282	\$418.00
210	\$216.00
458	\$119.00
212	\$364.00
258	\$418.00
405	\$212.00
322	\$173.00
272	\$163.00
460	\$344.00
12	\$237.00
192	\$286.00
46	\$225.00
47	\$166.00
217	\$172.00
266	\$388.00
379	\$318.00
157	\$215.00
384	\$259.00
440	\$263.00
245	\$442.00

ID	State
1	UT
2	ID
3	NE
4	SC
5	NE
6	VT
7	NC
8	NM
9	FL
10	WY
11	AK
12	MI
13	KS
14	NH
15	WY
16	OK
17	VA
18	WA
19	NM
20	IL
21	NY
22	NJ
23	SC

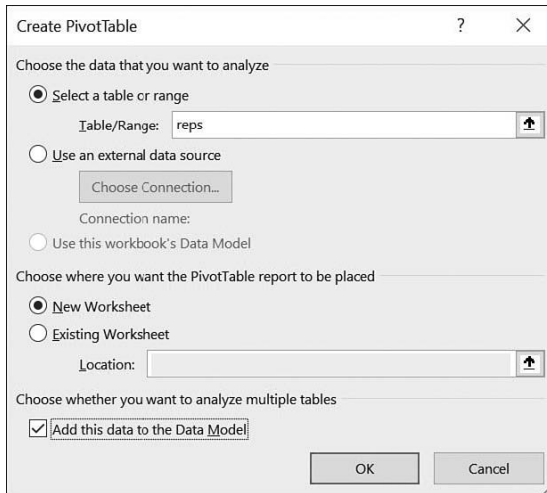
**Рис. 46.1.** Пример данных для модели данных

### ❓ Как добавить данные в модель данных?

Перед тем как добавить какие-либо данные в модель данных, необходимо создать из них таблицу (см. главу 26 «Таблицы»). Сначала на листе **Reps** выделите данные, включая заголовки столбцов, и создайте таблицу комбинацией **Ctrl+T**. На открывшейся вкладке **Конструктор (Работа с таблицами) (Table Tools Design)** в группе **Свойства (Properties)** в поле **Имя таблицы (Table Name)** присвойте таблице имя **Reps**. Затем на вкладке **Вставка (Insert)** выберите **Сводная таблица (PivotTable)** и в появившемся



диалоговом окне включите флажок **Добавить эти данные в модель данных** (Add this data to the Data Model), как на рис. 46.2. Таким образом, данные на листе **Reps** будут включены в модель данных. Аналогично выделите данные на листе **Sales** и присвойте таблице имя **Sales**. Добавьте таблицу **Sales** в модель данных.



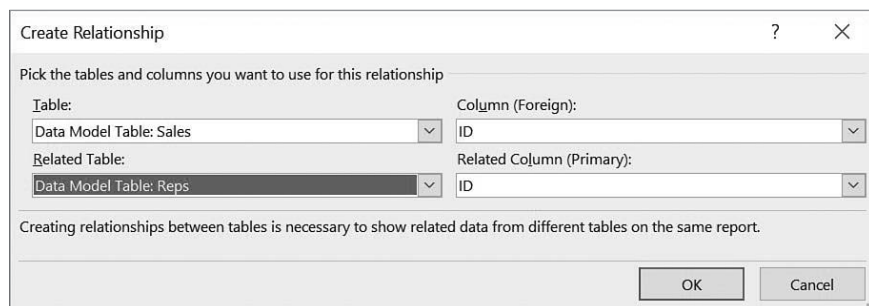
**Рис. 46.2.** Добавление данных из таблицы **Reps** в модель данных

### ? Как удалить данные из модели данных?

Для удаления данных из модели данных просто перейдите в таблицу Excel, содержащую эти данные, и на вкладке **Конструктор (Работа с таблицами)** (Table Tools Design) выберите **Преобразовать в диапазон** (Convert To Range).

### ? Как создать связь в модели данных?

Если мы хотим обобщить продажи по штатам — у нас проблема. В настоящий момент Excel не может знать, с каким штатом работает каждый продавец. Требуется создать каким-либо образом связь, позволяющую определить штат для каждой строки данных в таблице **Sales**. После добавления данных листа **Sales** в модель данных такую связь можно создать, если на вкладке **Данные (Data)** в группе **Работа с данными (Data Tools)** выбрать **Отношения (Relationships)** и нажать **Создать (New)**. Откроется диалоговое окно **Создание связи (Create Relationship)**, показанное на рис. 46.3. В поле **Связанный столбец (первичный ключ) (Primary Column)** должен быть указан столбец данных, обеспечивающий взаимно однозначное соответствие, которое может быть отображено в поле **Столбец (чужой) (Foreign Column)**. Поскольку штат для каждого продавца указан в таблице **Reps**, первичным ключом должен быть столбец с идентификатором (ID) из таблицы **Reps**. В качестве чужого столбца выберите столбец с кодом из листа **Sales**. Эти параметры показаны на рис. 46.3. Теперь при создании сводной таблицы, суммирующей продажи по штатам, для каждого идентификатора листа **Sales** будет извлечен правильный штат с листа **Reps**.

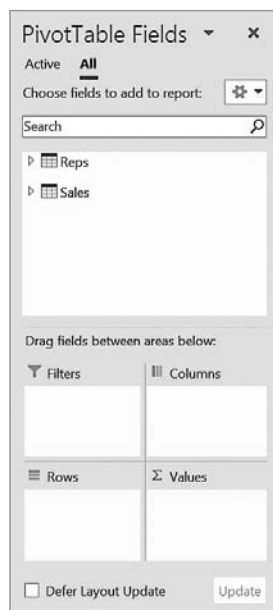


**Рис. 46.3.** Создание связи между листами

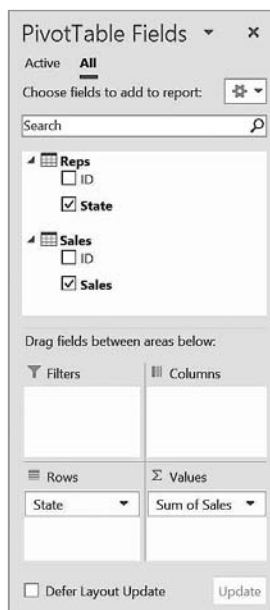
### ❓ Как с помощью модели данных создать сводную таблицу?

После добавления таблицы **Sales** в модель данных появился список полей сводной таблицы. Выбрав вариант **BCE (All)**, можно просмотреть список всех таблиц, добавленных в модель данных (рис. 46.4).

Щелкнув на треугольниках слева от значков таблиц **Reps** и **Sales**, вы увидите список всех столбцов из этих листов. Для создания сводной таблицы, вычисляющей общие показатели продаж в каждом штате, перетащите поле **Штат** в область **Строки (Rows)**, а поле **Продажи** в область **Значения (Values)**, как показано на рис. 46.5. У вас получится сводная таблица, как на рис. 46.6.



**Рис. 46.4.** Список полей сводной таблицы, включающий модель данных



**Рис. 46.5.** Создание сводной таблицы, определяющей продажи в каждом штате

	A	B
3	Row Labels	Sum of Sales
4	AK	10846
5	AL	23517
6	AR	24948
7	AZ	10532
8	CA	16646
9	CO	17447
10	CT	15350
11	DE	27179
12	FL	15424
13	GA	24147
14	HI	15966
15	IA	9172
16	ID	16656
17	IL	21525
18	IN	4706
19	KS	13875
20	KY	16254
21	LA	13465
22	MA	11816
23	MD	10899
24	ME	14865

**Рис. 46.6.** Объемы продаж по штатам

Например, на Аляске было продано 10 846 единиц товара, в штате Джорджия — 24 147 единиц и т. д.

Если бы вы попытались создать сводную таблицу до создания связи, показанной на рис. 46.3, Excel не смог бы понять, какой штат соответствует каждой строке листа Sales. Тогда он показал бы вам сообщение о том, что необходимо создать связь.

### ❓ Как добавить новые данные в модель данных?

Если требуется добавить новый набор данных в модель данных, просто создайте из данных таблицу и на вкладке **Вставка (Insert)** выберите **Сводная таблица (PivotTable)**. Установите флажок **Добавить эти данные в модель данных (Add this data to the Data Model)**. Если теперь в списке полей сводной таблицы выбрать **Все (All)** вместо **Активная (Active)**, то можно использовать новые данные, добавленные в сводную таблицу.

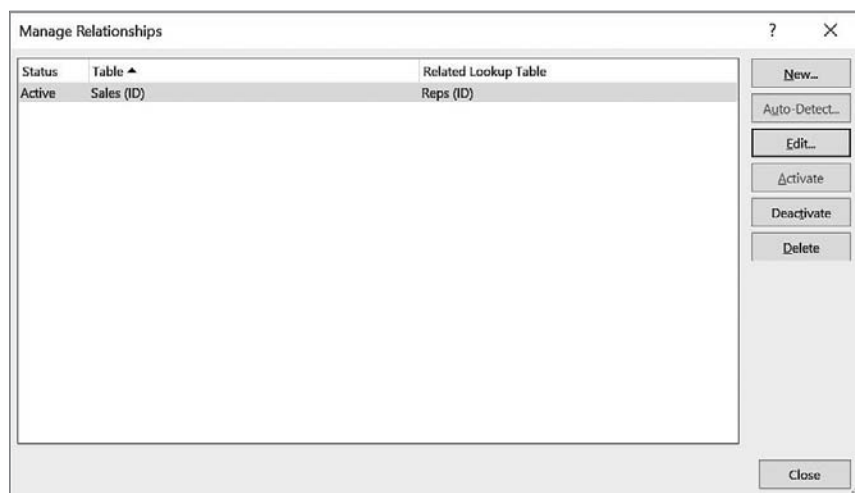
### ❓ Как с помощью модели данных создать новую сводную таблицу?

Щелкните на ячейке таблицы, выберите **Сводная таблица (PivotTable)** на вкладке **Вставка (Insert)**, в диалоговом окне **Создание сводной таблицы (Create PivotTable)** выберите **Использовать модель данных этой книги (Use This Workbook's Data Model)** и нажмите **ОК**. Затем выберите **Все (All)**, чтобы получить доступ ко всему, что было добавлено в модель данных. Щелчки по треугольничкам, показанным на рис. 46.5,

открывают доступ к полям модели. Теперь можно создать новую сводную таблицу опирающуюся на заданные вами связи.

### ? Как изменять и удалять связи?

Для изменения или удаления связей на вкладке **Данные (Data)** в группе **Работа с данными (Data Tools)** выберите **Отношения (Relationships)**. Откроется диалоговое окно **Управление связями (Manage Relationships)**, показанное на рис. 46.7. В этом диалоговом окне можно изменять и удалять связи (**Edit** или **Deactivate**), а также создавать новые связи.



**Рис. 46.7.** Диалоговое окно Управление связями

### ? Как работает функция Число различных элементов (DISTINCT COUNT)?

В файле **Distinctcounttemp.xlsx** в папке **Templates** находится список спортсменов (1–100) и их заработные платы (рис. 46.8). Поскольку многим спортсменам зарплата выплачивалась неоднократно, они встречаются в списке более одного раза. Также для каждого спортсмена указан вид спорта, которым он занимается.

Допустим, вам нужно узнать количество спортсменов в каждом виде спорта. Сначала я создал из диапазона **E3:F466** таблицу с именем **Деньги (Money)**, а из диапазона **K6:L106** таблицу с именем **Activity**. Добавил каждую таблицу в модель данных. Затем я создал связь с первичным ключом **Спортсмен (Person)** из таблицы **Спорт** и чужим столбцом **Спортсмен** из таблицы **Деньги**. Далее я создал сводную таблицу (см. лист **Count of Person** в файле **Distinctcount.xlsx**), перетащив поле **Вид спорта (Sport)** в область **СТРОКИ (ROWS)** и поле **Число спортсменов (Count of Person)** из таблицы **Деньги** в область **ЗНАЧЕНИЯ (VALUES)**. Эти действия приведут к созданию сводной таблицы, как на рис. 46.9.

	E	F	G	H	I	J	K	L
1								
2	Person	Salary						
3	32	6						
4	66	4						
5	15	4						
6	53	7					Person	Sport
7	26	9					1	Basketball
8	56	9					2	Hockey
9	80	7					3	Lacrosse
10	29	7					4	Football
11	63	4					5	Football
12	84	5					6	Lacrosse
13	79	4					7	Lacrosse
14	82	6					8	Baseball
15	69	6					9	Hockey
16	60	8					10	Lacrosse
17	59	8					11	Lacrosse
18	66	4					12	Football
19	6	10					13	Lacrosse
20	91	6					14	Baseball
21	27	9					15	Basketball
22	72	5					16	Hockey
23	51	10					17	Basketball

**Рис. 46.8.** Данные к примеру операции  
Число различных элементов (DISTINCT COUNT)

	A	B
1		
2		
3	Row Labels	Count of Person
4	Baseball	85
5	Basketball	96
6	Football	84
7	Hockey	72
8	Lacrosse	127
9	Grand Total	464

**Рис. 46.9.** Сводная таблица,  
в которой спортсмены учтены  
несколько раз

Value Field Settings

Source Name: Person

Custom Name: Distinct Count of Person

Summarize Values By

Show Values As

Summarize value field by

Choose the type of calculation that you want to use to summarize data from the selected field

Min

StdDev

StdDevp

Var

Varp

Distinct Count

Number Format

OK

Cancel

**Рис. 46.10.** Выбор функции Число различных элементов  
для получения количества спортсменов в каждом виде спорта

Поскольку в списке было только 100 спортсменов, очевидно, что каждый спортсмен был посчитан несколько раз. Для решения этой проблемы щелкните правой кнопкой мыши в любом месте сводной таблицы и выберите в контекстном меню **Параметры поля значений (Value Field Settings)**. Прокрутите список до конца и выберите, как показано на рис. 46.10, последний вариант **Число разных элементов (Distinct Count)**. Этот выбор гарантирует, что каждый спортсмен будет учтен только один раз. Полученная в результате сводная таблица представлена на рис. 46.11 и на листе **Distinct Count** в файле **Distinctcount.xlsx**. В этой таблице показано действительное количество спортсменов в каждом виде спорта.

	A	B	C
3	Row Labels	Distinct Count of Person	Count of Person
4	Baseball	20	85
5	Basketball	18	96
6	Football	20	84
7	Hockey	14	72
8	Lacrosse	26	127
9	Grand Total	98	464

**Рис. 46.11.** Количество спортсменов в каждом виде спорта

## Задания

В файле **Faberu.xlsx** содержатся зарплаты, командировочные расходы, а также коды преподавателей, коды кафедр и коды должностей профессорско-преподавательского состава для всех бизнес-школ факультета. На основе этих данных решите следующие задачи. (Использовать функции **ВПР** нельзя!)

1. Используйте модель данных для создания сводной таблицы, вычисляющей среднюю зарплату с разбивкой по кафедрам.
2. Используйте модель данных для создания сводной таблицы, вычисляющей среднюю зарплату с разбивкой по должностям и кафедрам.
3. Используйте модель данных для создания сводной таблицы, вычисляющей средние командировочные расходы с разбивкой по кафедрам.

# Power Pivot

### Обсуждаемые вопросы

- Как загрузить данные в Power Pivot?
- Как с помощью Power Pivot создать сводную таблицу?
- Как использовать срезы с Power Pivot?
- Что такое функции DAX и вычисляемые столбцы?
- Как работает функция RELATED?
- Как работает функция CALCULATE и что такое вычисляемая мера?

Power Pivot доступна (или скоро будет доступна) во всех версиях Microsoft Excel начиная с 2016 и старше. Чтобы добавить Power Pivot, на вкладке **Файл (File)** выберите **Параметры (Options)** и затем **Надстройки (Add-Ins)**. В списке **Управление (Manage)** выберите **Надстройки COM (COM Add-Ins)** и затем нажмите кнопку **Перейти (Go)**. В диалоговом окне **Надстройки COM (COM Add-Ins)** установите флажок рядом с **Power Pivot** и нажмите **OK**. На ленте появится вкладка **Power Pivot**. Надстройка Power Pivot позволяет выполнить перечисленные далее действия.

- Сохранять и запрашивать большие объемы данных (сотни миллионов строк) эффективным способом с возможностью объединения различных форматов данных из различных источников. Например, часть данных может поступать из базы данных Microsoft Access, часть — из текстового файла, часть — из нескольких файлов Excel и часть — из оперативных данных, импортируемых с веб-сайта. Это легкое и удобное создание сводных таблиц или диаграмм из данных, занимающих сотни миллионов строк.
- Создавать вычисляемые столбцы (помимо определения полей). Например, если каждая строка исходных данных содержит доходы и расходы, можно создать вычисляемый столбец: **прибыль = доходы – расходы**. В состав Power Pivot входит множество функций языка выражений анализа данных (DAX, Data Expression Language), упрощающих создание вычисляемых столбцов.
- Создавать вычисляемые поля, агрегирующие данные из разных строк. Например, в вычисляемом поле можно рассчитать общую прибыль в виде суммы прибылей от каждой сделки.
- Создавать меры Power Pivot, которые можно применять в сводных таблицах или диаграммах.

- Создавать ключевые показатели эффективности (Key Performance Indicator , KPI), позволяющие организации отследить соответствие результатов поставленным целям. Например, с помощью ключевых показателей эффективности Power Pivot для каждого года и торгового представителя можно сравнить фактические продажи и квоты продаж.
- Power Pivot может служить источником данных (сводные таблицы, диаграммы, функции куба и т. д.) для отчетов по книге.
- Отчеты Power Pivot можно опубликовать в Microsoft SharePoint для автоматического обновления данных, упрощения совместного использования и для возможности мониторинга ИТ-инфраструктуры. После публикации они могут применяться как источники данных для получения другой аналитической или отчетной информации (а публикация в SharePoint может быть частью процесса развертывания).

## Ответы на вопросы

### ? Как загрузить данные в Power Pivot?

После активации Power Pivot вкладка **Power Pivot** появляется на ленте Excel. Откройте эту вкладку (рис. 47.1) и нажмите кнопку **Управление (Manage)**. На рис. 47.2 показано диалоговое окно **Power Pivot для Excel** с открытой вкладкой **В начало (Home)**.

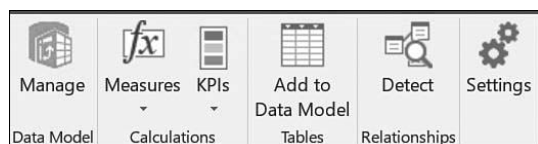


Рис. 47.1. Вкладка Power Pivot

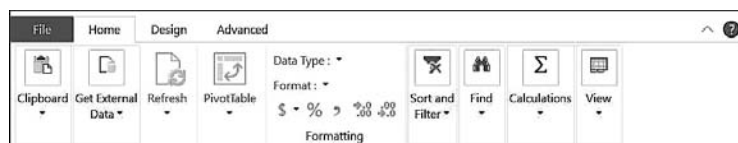


Рис. 47.2. Окно Power Pivot для Excel с вкладкой В начало

Если вы выбрали **Получение внешних данных (Get External Data)**, вы можете импортировать данные из различных источников, представленных на рис. 47.3.

Например, если выбрать **Из базы данных (From Database)**, Power Pivot загрузит данные из базы данных Access или SQL Server . Если выбрать **Из других источников (From Other Sources)**, можно загрузить данные из всех перечисленных выше источников данных, а также из файлов Excel, текстовых файлов, а также других типов баз данных, например Oracle или Teradata.





**Рис. 47.3.** Команда Получение внешних данных предлагает несколько расположений данных для выбора

Скопировав данные в Excel, вы можете выбрать **Вставить (Paste)** в группе **Буфер обмена (Clipboard)** на вкладке **В начало (Home)** в окне Power Pivot для вставки данных в Power Pivot.

Продemonстрируем загрузку данных в Power Pivot из нескольких источников на примере текстового файла **Storesales.txt**, в котором содержится информация о продажах в 20 магазинах. Часть данных представлена на рис. 47.4. Как видно из рисунка, для каждой сделки указан номер магазина, проданный продукт, дата продажи и выручка. Требуется суммировать эти данные по штатам, но штат для каждого магазина указан в другом файле — **States.xlsx**. Данные из этого файла по местоположению каждого магазина показаны на рис. 47.5.

	A	B	C	D	E
1	Store	Product	Date	Units	Revenue
2	20	food	3/15/2011	37	\$16.28
3	10	cds	11/29/2014	70	\$19.60
4	6	dvds	8/11/2010	68	\$39.44
5	4	dvds	5/11/2012	120	\$92.40
6	6	food	7/8/2011	129	\$101.91
7	6	books	9/7/2011	16	\$2.24
8	5	magazines	1/4/2014	101	\$72.72
9	18	cds	9/27/2011	67	\$8.04
10	15	magazines	9/29/2014	132	\$105.60
11	13	food	1/27/2011	38	\$23.18
12	7	books	4/13/2014	103	\$17.51
13	3	books	2/17/2010	127	\$21.59
14	16	food	6/6/2014	129	\$29.67
15	18	dvds	11/30/2011	149	\$80.46
16	17	cds	2/5/2010	64	\$9.60
17	2	books	6/18/2011	147	\$26.46
18	8	toys	6/22/2014	72	\$12.24
19	18	dvds	1/27/2013	85	\$66.30
20	12	food	5/15/2010	32	\$15.68

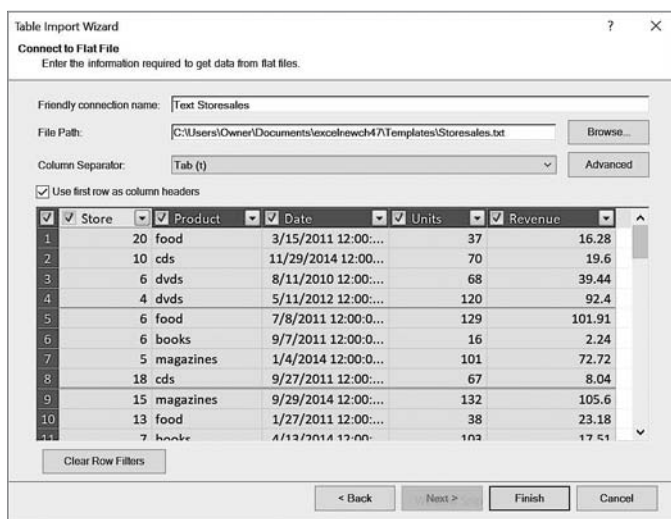
**Рис. 47.4.** Данные о продажах для загрузки в Power Pivot

Необходимо создать сводную таблицу, которая позволит получить срез данных по продажам каждого продукта в каждом штате. Сначала откройте вкладку **PowerPivot** и нажмите кнопку **Управление (Manage)**. Откроется диалоговое окно **Power Pivot для Excel**. Рассмотрим вкладку **В начало (Home)**, показанную на рис. 47.2. Так как в задаче требуется импортировать текстовый файл, выберите **Получение внешних дан-**

ных (Get External Data), а затем в группе Из других источников (From Other Sources). В диалоговом окне Мастер импорта таблиц (Table Import Wizard) выберите Текстовый файл (Text File; это последняя строка в списке) и нажмите Далее (Next). Как видно из рис. 47.6, здесь выбран файл Storesales.txt. Файл автоматически получает внутреннее (необязательное) имя. Поскольку первая строка данных содержит заголовки столбцов, установите флажок Использовать первую строку в качестве заголовков столбцов (Use first row as column headers). В раскрывающемся списке Разделитель столбцов (Column Separator) выберите Табуляция (Tab), так как разделителями полей данных в текстовом файле являются знаки табуляции. Нажмите кнопку Готово (Finish) и затем кнопку Заккрыть (Close) для завершения процесса импорта данных из текстового файла в Power Pivot.

	F	G
7	Store	State
8	1	IND
9	2	IND
10	3	ILL
11	4	ILL
12	5	ILL
13	6	ILL
14	7	ILL
15	8	ILL
16	9	MICH
17	10	MICH
18	11	MICH
19	12	MICH
20	13	MICH
21	14	MICH
22	15	KY
23	16	KY
24	17	KY
25	18	KY
26	19	IOWA

**Рис. 47.5.** Местоположение магазинов



**Рис. 47.6.** Настройка импорта текстового файла в Power Pivot

Результат импорта текстовых данных в Power Pivot представлен на рис. 47.7. На рисунке показано подмножество данных.

Далее необходимо импортировать файл States.xlsx с возможностью установления его связи с данными о продажах. Для импорта файла States.xlsx вернитесь в Excel, щелкнув на вкладке Power Pivot по значку Переключиться в книгу (Switch To Workbook Excel) в левом верхнем углу (над вкладкой Файл (File)). При закрытом файле States.xlsx на вкладке Управление (Manage) выберите Из других источников (From Other Sources) в разделе ленты Получение внешних данных (Get External Data), выберите Файл Excel (Excel File) в списке Из других источников (From Other Sources) и найдите файл States.xlsx. Выбрав Использовать первую строку в качестве заголовков столбцов (Use First Row As Column Headers), выбираем Лист1 (Sheet1) и нажимаем Заккрыть. Как показано на рис. 47.8, данные из файла States.xlsx будут импортированы в Power Pivot. Внизу экрана мы поменяли название вкладки с Лист1 на States. Напомним,

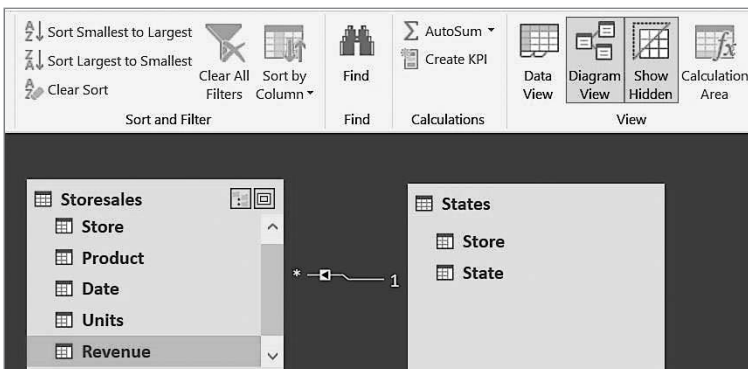
	Store	Product	Date	Units	Revenue
1	5	magazines	1/4/2...	101	72.72
2	15	magazines	9/29/...	132	105.6
3	19	magazines	8/7/2...	142	75.26
4	17	magazines	9/5/2...	146	65.7
5	17	magazines	7/5/2...	82	31.98
6	2	magazines	2/10/...	114	33.06
7	1	magazines	11/29...	107	53.5
8	3	magazines	2/2/2...	148	56.24
9	4	magazines	7/16/...	27	20.52
10	13	magazines	5/12/...	120	82.8
11	11	magazines	10/15...	108	73.44
12	19	magazines	11/4/...	103	71.07
13	2	magazines	9/25/...	107	51.36
14	5	magazines	12/7/...	115	77.05
15	20	magazines	9/20/...	29	20.59

**Рис. 47.7.** Подмножество данных, импортированных из файла Storesales.txt

	Store	State	Add Column
1	1	IND	
2	2	IND	
3	3	ILL	
4	4	ILL	
5	5	ILL	
6	6	ILL	
7	7	ILL	
8	8	ILL	
9	9	MICH	
10	10	MICH	
11	11	MICH	
12	12	MICH	
13	13	MICH	
14	14	MICH	
15	15	KY	

**Рис. 47.8.** Диалоговое окно Просмотр вставки

что в задаче требуется проанализировать продажи в различных штатах. Проблема заключается в том, что в настоящее время в Power Pivot не установлено соответствие между местоположением магазинов из файла States.xlsx с магазинами, перечисленными в текстовом файле. Для решения этой проблемы необходимо создать связь между двумя источниками данных. Для создания связи в окне Power Pivot откройте вкладку Конструктор (Design) (см. рис. 47.10), выберите Создание связи (Create Relationship) и выполните действия, описанные в главе 46 «Модель данных». Но также можно на вкладке В начало (Home) в группе Просмотр (View) выбрать Представление диаграммы (Diagram View) и определить необходимую связь путем перетаскивания соответствующего столбца (рис. 47.9).



**Рис. 47.9.** Схема связи между двумя источниками данных

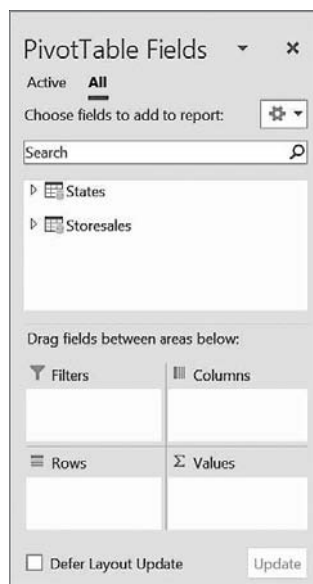
Для возврата к просмотру данных в группе **Просмотр (View)** щелкните на значке **Представление данных (Data View)**, показанной на рис. 47.9.

### ? Как с помощью Power Pivot создать сводную таблицу?

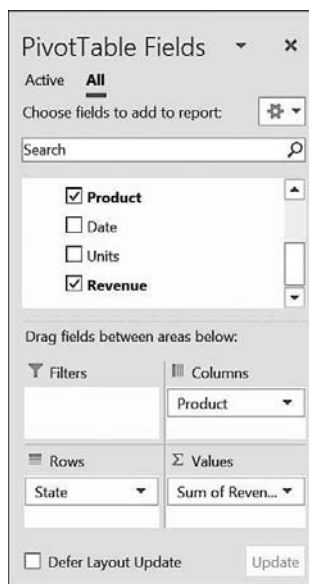
Теперь все готово для обобщения данных о продажах в Power Pivot посредством сводной таблицы. Щелкнув на значке **Представление данных (Data View)**, выберите **Сводная таблица (PivotTable)** в элементе **Получение внешних данных (Get External Data)** на ленте Power Pivot. Появляется диалоговое окно **Создать сводную таблицу (Create Pivot Table)** с запросом выбора расположения таблицы на новом листе или на существующем листе. Выберите новый лист. Затем появляется панель **Поля сводной таблицы (PivotTable Fields)** (вам может понадобиться щелкнуть правой кнопкой мыши и выбрать **Показать список полей (Show Field List)**), приведенная на рис. 47.10.

Выберите **Все (All)** и щелкните на треугольниках слева от таблиц **States** и **Storesales** для получения доступа ко всем импортированным столбцам (которые здесь называются полями). Цель — получить разбивку по штатам и продуктам для подсчета общей выручки и количества проданных единиц продукции. Для суммирования выручки и количества проданных единиц продукции перетащите поля **Выручка** и **Количество** в область **ЗНАЧЕНИЯ (VALUES)**. Перетащите поле **Штат** в область **СТРОКИ (ROWS)** и поле **Продукт** в область **СТОЛБЦЫ (COLUMNS)** — рис. 47.11. Обратите внимание, что поля, выбранные для сводной таблицы, помечены.

Часть полученной сводной таблицы представлена на рис. 47.12. Она показывает доход по каждому продукту, разбитый по штатам. Как видно из рисунка, в Илли-



**Рис. 47.10.** Панель Поля сводной таблицы



**Рис. 47.11.** Распределение полей по областям для создания сводной таблицы

нойсе (ILL) было продано DVD на общую сумму \$2295,76. Обратите внимание, что для правильного форматирования столбца Сумма по полю Выручка потребовалось щелкнуть на нем правой кнопкой мыши, выбрать Параметры полей значений (Value Field Settings), нажать кнопку Числовой формат (Number Format) и выбрать формат Денежный (Currency), после чего закрыть диалоговые окна, нажимая ОК. Так мы получаем выражение дохода в денежных единицах.

Sum of Revenue		Column Labels					
Row Labels	books	cds	dvds	food	magazines	toys	Grand Total
ILL	\$1,542.74	\$1,880.44	\$2,295.76	\$2,178.69	\$2,904.28	\$1,807.78	\$12,609.69
IND	\$697.08	\$799.09	\$874.89	\$644.58	\$1,012.93	\$953.91	\$4,982.48
IOWA	\$518.61	\$651.49	\$480.66	\$364.39	\$744.91	\$574.49	\$3,334.55
KY	\$1,654.53	\$1,051.28	\$1,952.70	\$1,224.88	\$1,716.58	\$1,485.99	\$9,086.96
MICH	\$2,071.94	\$1,680.84	\$1,320.06	\$2,108.54	\$2,260.14	\$2,334.86	\$11,776.38
Grand Total	\$6,484.90	\$6,063.14	\$6,924.07	\$6,521.08	\$8,638.84	\$7,158.03	\$41,790.06

**Рис. 47.12.** Сводная таблица с разбивкой продаж продукции по штатам

### ? Как использовать срезы с Power Pivot?

В главе 45 «Сводные таблицы и срезы для описания данных» описывалось применение срезов для получения подробной информации и различных ракурсов при анализе сводных таблиц. Здесь также можно создать срезы для обобщения данных по любому набору продуктов и магазинов. Для этого щелкните мышью внутри сводной таблицы и выберите Вставка среза (Insert Slicer) на вкладке Анализ (Analyze) в группе Фильтр (Filter). Срезы для полей Магазин и Продукт показаны на рис. 47.13 (см. файл Pivotwithslicers.xlsx). Внешний вид срезов можно изменить с помощью вкладки Инструменты для среза — Параметры (Slicer Tools Options). Например, для среза Магазин можно определить размещение списка в четырех столбцах (поле Столбцы (Columns) в группе Buttons). Как было сказано в главе 45, удер-

Sum of Revenue		Column Labels	
Row Labels	books	food	Grand Total
ILL	\$652.88	\$588.04	\$1,240.92
MICH	\$971.28	\$1,014.01	\$1,985.29
Grand Total	\$1,624.16	\$1,602.05	\$3,226.21

Product			Store		
books	cds	dvds	1	2	3
food	magazines	toys	4	5	6
			7	8	9
			10	11	12
			13	14	15
			16	17	18
			19	20	

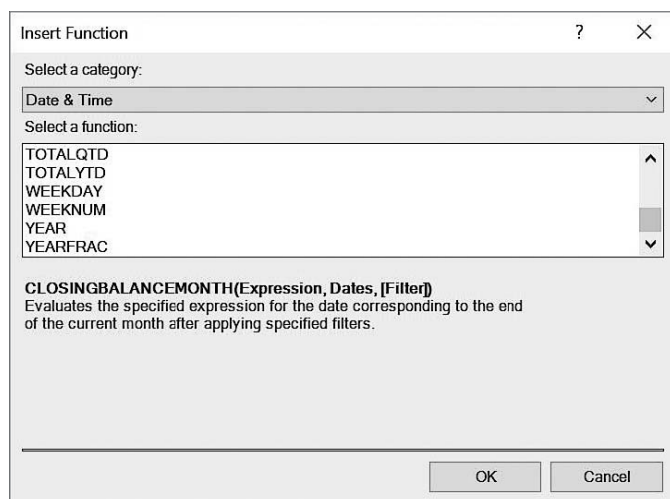
**Рис. 47.13.** Срезы Продукт и Магазин

живая **Ctrl**, можно выбрать несколько продуктов и/или магазинов. Кроме того, с помощью **Shift** можно выбрать в срезе любой непрерывный диапазон. Удерживая клавишу **Ctrl**, размеры среза можно быстро изменить, потянув за его стороны или углы. В сводной таблице, представленной на рис. 47.13, просуммированы общая выручка и количество проданных книг и продуктов питания в магазинах с 7 по 11. Поскольку все эти магазины находятся в Иллинойсе или Мичигане, то в общей сводной таблице показаны только эти два штата.

## ? Что такое функции DAX и вычисляемые столбцы?

Как было сказано в главе 45, в сводной таблице можно создавать новые формулы с помощью вычисляемых объектов или вычисляемых полей. После импорта данных в Power Pivot можно создавать новые вычисляемые столбцы посредством языка выражений анализа данных (DAX, Data Analysis eXpressions), что значительно повышает эффективность сводных таблиц.

Продemonстрируем некоторые формулы DAX на примере создания вычисляемых столбцов для года, месяца и дня в месяце. Сначала в окне **Power Pivot** откройте вкладку **Storesales** (в нижней части окна) и выберите первый пустой столбец. Нажмите под вкладкой **POWER PIVOT** кнопку **fx** для просмотра списка функций DAX. Многие из этих функций (например, **YEAR**, **MONTH** и **DAY**) основаны на функциях Excel. После выбора категории **Дата и время (Date&Time)** появится список, приведенный на рис. 47.14. Язык DAX включает множество других мощных функций. Например, функция **DISTINCT** возвращает таблицу уникальных значений в указанном столбце.



**Рис. 47.14.** Список функций DAX



Для распределения года, месяца и дня каждой сделки по разным столбцам введите в первой ячейке первого пустого столбца формулу `=YEAR(st)`. Появится подсказка со списком столбцов из источников данных Power Pivot, к которым можно применить функцию `YEAR`. Выберите столбец с датой и завершите ввод формулы `=YEAR(storesales[Дата])`. Столбец заполнится значениями соответствующих годов для каждой сделки. Щелкнув правой кнопкой мыши на заголовке столбца, можно переименовать его в **Год**. В следующем столбце вычислите месяц по формуле `=MONTH(storesales[Дата])`. И, наконец, вычислите день месяца по формуле `=DAY(storesales[Дата])`. Щелкните правой кнопкой мыши заголовки этих столбцов и присвойте им имена **Месяц** и **День месяца** соответственно. Таблица с вычисляемыми столбцами представлена на рис. 47.15. Полностью проделанная работа представлена в файле `PowerPivotexample1.xlsx`.

Теперь на основе вычисляемых столбцов можно создавать различные информативные сводные таблицы и диаграммы. Например, можно просуммировать продажи для каждого штата по годам (см. задание 1).

	Store	Product	Date	Units	Revenue	Year	Month	Day of Month
1	5	magazines	1/4/2014 12:00:00 AM	101	72.72	2014	1	4
2	15	magazines	9/29/2014 12:00:00 AM	132	105.6	2014	9	29
3	19	magazines	8/7/2012 12:00:00 AM	142	75.26	2012	8	7
4	17	magazines	9/5/2012 12:00:00 AM	146	65.7	2012	9	5
5	17	magazines	7/5/2013 12:00:00 AM	82	31.98	2013	7	5
6	2	magazines	2/10/2013 12:00:00 AM	114	33.06	2013	2	10
7	1	magazines	11/29/2013 12:00:00 AM	107	53.5	2013	11	29
8	3	magazines	2/2/2014 12:00:00 AM	148	56.24	2014	2	2
9	4	magazines	7/16/2012 12:00:00 AM	27	20.52	2012	7	16
10	13	magazines	5/12/2013 12:00:00 AM	120	82.8	2013	5	12
11	11	magazines	10/15/2014 12:00:00 AM	108	73.44	2014	10	15
12	19	magazines	11/4/2014 12:00:00 AM	103	71.07	2014	11	4
13	2	magazines	9/25/2012 12:00:00 AM	107	51.36	2012	9	25
14	5	magazines	12/7/2012 12:00:00 AM	115	77.05	2012	12	7
15	20	magazines	9/20/2013 12:00:00 AM	29	20.59	2013	9	20

**Рис. 47.15.** Вычисляемые столбцы Год, Месяц и День, созданные с помощью формул DAX

## ❓ Как работает функция RELATED?

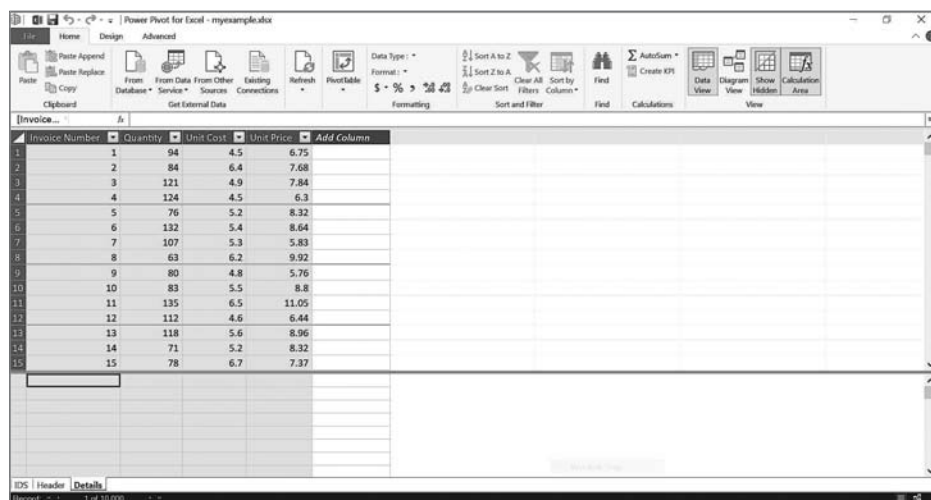
Чтобы продемонстрировать, как работают более сложные функции DAX, обсудим еще один пример. Покажем их применение к базе данных по продажам небольшой компании (скажем, Saleco). Нужные сведения содержатся в файле `Myexample.xlsx`. А именно:

- На листе **IDs** перечислены штаты, где проживают покупатели.
- На листе **Invoice Header** содержатся сведения о дате и паспортные данные (Customer ID) для 10 000 счетов-фактур.
- На листе **Invoice Details** приведены номер каждого счета-фактуры, количество купленного, себестоимость за единицу и цена за единицу.

С помощью этих данных мы хотим получить ответы на следующие вопросы:

- Сколько единиц товара было продано в каждом штате?
- Сколько единиц товара было продано в каждом штате в течение каждого года?

Для начала нам нужно преобразовать данные с каждого листа в таблицу. Таблицам были даны такие названия, как **Паспортные данные (IDs)**, **Заголовок (Header)** и **Детали (Details)**. Чтобы добавить данные из трех таблиц в Power Pivot, щелкните внутри каждой из таблиц и на ленте **Power Pivot** выберите **Добавить в модель данных (Add To Data Model)**. Внизу экрана вы увидите вкладку для каждой таблицы (рис. 47.16).

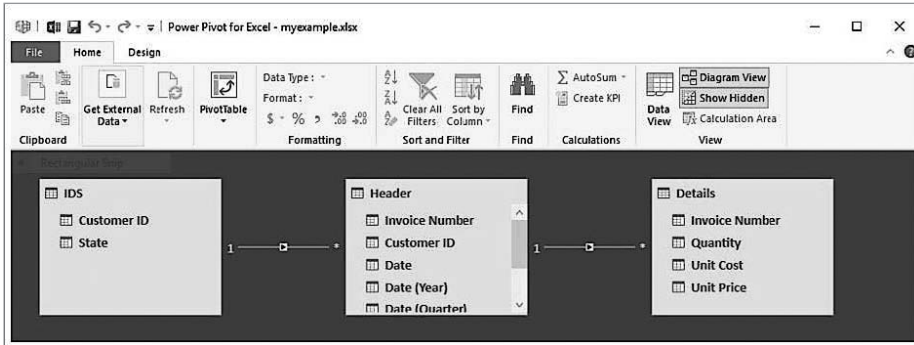


Invoice Number	Quantity	Unit Cost	Unit Price	Add Column
1	94	4.5	6.75	
2	84	6.4	7.68	
3	121	4.9	7.84	
4	124	4.5	6.3	
5	76	5.2	8.32	
6	132	5.4	8.64	
7	107	5.3	5.83	
8	63	6.2	9.92	
9	80	4.8	5.76	
10	83	5.5	8.8	
11	135	6.5	11.05	
12	112	4.6	6.44	
13	118	5.6	8.96	
14	71	5.2	8.32	
15	78	6.7	7.37	

**Рис. 47.16.** Данные, вводимые в Power Pivot

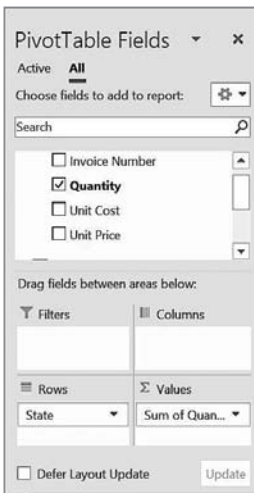
Предположим, нам надо определить количество единиц, проданных в течение каждого года в каждом штате. Чтобы выполнить такой расчет, Power Pivot нужно, чтобы лист **Заголовок (Header)** «знал» штат каждого покупателя. Для этого мы должны установить взаимоотношение, связывающее паспортные данные (IDs) на первых двух листах. Power Pivot также нужно «понимать», что номера счетов-фактур в листе **Заголовки** и **Детали** — одни и те же. Чтобы Power Pivot «знала» об этих отношениях, выбираем на ленте **Power Pivot** и создаем нужные отношения через **Вид диаграммы (Diagram View)**, просто проведя линии (как показано на рис. 47.17) между заголовками соответствующих таблиц. Вы сможете управлять этими отношениями (редактировать или удалять), щелкнув на вкладке **Конструктор (Design)**, как показано на рис. 47.17. В сущности, созданные нами отношения заменяют многие из функций ВПР. Даже несмотря на тот факт, что листы Excel могут содержать до 1 048 576 строк данных, отношения, установленные в Power Pivot (а не через Excel), позволяют Power Pivot быстро создавать сводные таблицы на основе сотен миллионов строк данных.





**Рис. 47.17.** Взаимоотношения в компании Saleco

Создав эти взаимосвязи, вы теперь можете создавать самые разные сводные таблицы. Для начала сосчитаем общее число единиц, проданных в каждом штате. Выбрав на ленте форму Power Pivot, щелкните **Управление (Manage)**. Теперь вы можете выбрать на ленте **Сводная таблица (Pivot Table)**. Выберем из множества разнообразных вариантов **Сводная таблица (Pivot Table)**, а затем выберите для нее расположение. Выбрав **Все (All)**, вы сделаете все столбцы источниками данных для сводной таблицы. Заполнив список полей, как показано на рис. 47.18, мы получим сводную таблицу, представленную на рис. 47.19.



**Рис. 47.18.** Список полей для сводной таблицы, рассчитывающей число единиц проданного продукта по штатам

Row Labels	Sum of Quantity
Alabama	17032
Arizona	18941
Arkansas	7287
California	124876
Colorado	19150
Connecticut	14557
Florida	52620
Georgia	34933
Illinois	56451
Indiana	13256
Iowa	7190
Kansas	17723
Kentucky	14854
Louisiana	12239
Maine	8733
Maryland	27509
Massachusetts	21893
Michigan	38600

**Рис. 47.19.** Число единиц проданного продукта по штатам

Как и в нашем примере с продажами универмага (Store Sales), мы применим формулы DAX по Году и Месяцу к вкладке **Заголовок**, чтобы создать новые столб-

цы, содержащие месяц и год для каждой сделки. Я также использовал формулу =формат(Заголовок[Дата], "MMM"), чтобы поместить название месяца, в котором был выставлен счет-фактура, в отдельный столбец. На рис. 47.20 (см. файл Calccolumns1.xlsx) показана вкладка **Заголовки (Headers)** после вставки вычисляемых столбцов.

	Invoice Number	Customer ID	Date	Year	Month	Name of Month
1	1	256	10/3/2018 ...	2018	10	Oct
2	2	60	12/10/201...	2019	12	Dec
3	3	45	7/6/2018 1...	2018	7	Jul
4	4	138	8/16/2018 ...	2018	8	Aug
5	5	5	11/30/201...	2017	11	Nov
6	6	139	11/17/202...	2020	11	Nov
7	7	194	10/14/201...	2019	10	Oct
8	8	59	3/16/2021 ...	2021	3	Mar
9	9	195	4/18/2020 ...	2020	4	Apr
10	10	12	3/13/2019 ...	2019	3	Mar
11	11	194	4/21/2021 ...	2021	4	Apr
12	12	29	12/11/201...	2017	12	Dec
13	13	168	12/25/201...	2018	12	Dec
14	14	185	11/9/2019 ...	2019	11	Nov
15	15	171	5/11/2019 ...	2019	5	May

**Рис. 47.20.** Функции DAX извлекают Год и Месяц для каждого счета-фактуры

На вкладке **Подробности (Details)** мы можем рассчитать общую стоимость по каждому счету-фактуре с помощью формулы, показанной на рис. 47.21. Выделив последние три столбца, щелчком на **денежном (\$)** формате, установив для них формат в долларах США. В Power Pivot (то есть не так, как при работе с обычной сводной таблицей) форматирование, выполненное на вкладках с данными, применяется и в сводных таблицах.

[Total Cost]		fx =Details[Quantity]*Details[Unit Cost]			
Invoice Number	Quantity	Unit Cost	Unit Price	Total Cost	
1	94	\$4.50	\$6.75	\$423.00	
2	84	\$6.40	\$7.68	\$537.60	
3	121	\$4.90	\$7.84	\$592.90	
4	124	\$4.50	\$6.30	\$558.00	
5	76	\$5.20	\$8.32	\$395.20	
6	132	\$5.40	\$8.64	\$712.80	
7	107	\$5.30	\$5.83	\$567.10	
8	63	\$6.20	\$9.92	\$390.60	
9	80	\$4.80	\$5.76	\$384.00	
10	83	\$5.50	\$8.80	\$456.50	
11	135	\$6.50	\$11.05	\$877.50	
12	112	\$4.60	\$6.44	\$515.20	
13	118	\$5.60	\$8.96	\$660.80	
14	71	\$5.20	\$8.32	\$369.20	
15	78	\$6.70	\$7.37	\$522.60	

**Рис. 47.21.** Общая стоимость рассчитывается для каждого счета-фактуры, и все столбцы, имеющие отношение к стоимости, имеют денежный формат данных (в долларах)

Предположим, компания Saleco предоставляет 5%-ную скидку с цены во всех майских счетах-фактурах. Нам нужно рассчитать прибыль по каждому счету-фактуре, а для этого нам надо определить доход от каждого счета. Прежде чем считать этот доход, мы должны сосчитать на вкладке **Подобности (Details)** скидку (либо 0, либо 5%) по каждому счету. Вычисляемый столбец **Скидка (Discount)** зависит от информации (это столбец **Месяц**) на вкладке **Заголовки**. К счастью, мы можем применить функцию **RELATED**, чтобы извлечь **Месяц** из каждого счета-фактуры, из вкладки **Заголовки** и переместить на вкладку **Подобности**. На вкладке **Подобности** создадим новый столбец, содержащий **Скидку** по каждому счету, при помощи следующей формулы: **=Если(Related(Заголовки[Месяц])=5,.05,0)**. Заполняя вкладку **Заголовки**, вы увидите, что **Excel AutoComplete** предлагает свои названия столбцов. Это очень облегчает ввод формул! Далее, мы вводим формулу **= (1-[Скидка])\*Подобности[Цена за единицу]\*Подобности[Количество]** в новый вычисляемый столбец, чтобы рассчитать **Доход** по каждому счету-фактуре. Наконец, рассчитываем прибыль по каждому счету-фактуре путем ввода формулы **=Подобности[Доход]-Подобности[Общая стоимость]**. На рис. 47.22 показаны все столбцы с вкладки **Подобности**.

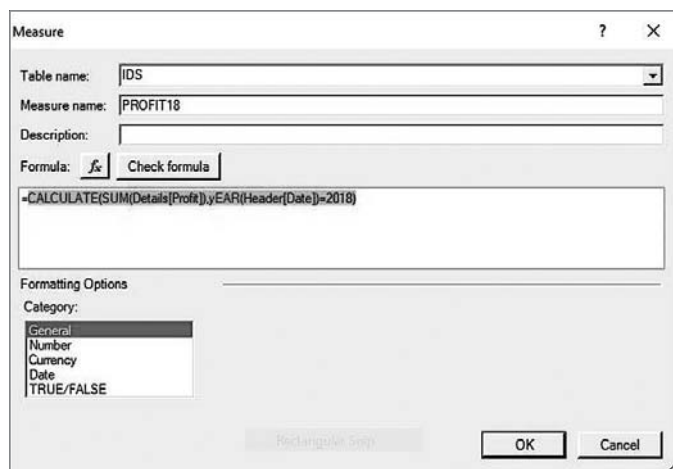
Теперь наши новые **Вычисляемые столбцы** можно использовать в сводной таблице. Например, на рис. 47.23 показано разбиение прибыли в каждом штате по **Годам**. Срезы и временные шкалы работают так, как описано в главе 45, так что мы добавили по срезу и по временной шкале к каждой из сводных таблиц.

### ? Как работает функция CALCULATE и что такое вычисляемая мера?

Power Pivot позволяет вам создавать вычисляемые меры, выполняющие расчеты, которые будут применяться в сводных таблицах. В пределах сводной таблицы значение вычисляемой меры рассчитывается на основе набора данных, разнесенного по строкам, столбцам и настройкам фильтрации («многомерное выражение»). Продemonстрируем применение вычисляемых мер следующим образом. Предположим, вам нужно рассчитать разницу между прибылью 2019-го и 2018 года по штатам. Придется создать три следующих вычисляемых меры:

- Прибыль 2019 г.
- Прибыль 2018 г.
- Разность «прибыль 2019 г. – прибыль 2018 г.»

Для создания многих вычисляемых мер используется функция **CALCULATE**, которая позволяет пользователю Power Pivot эмулировать использование функций Excel **СУММЕСЛИМН**, **СЧЁТЕСЛИМН** и **СРЗНАЧЕСЛИ** в Power Pivot. Чтобы создать меру вычисляющую прибыль 2018 г., выбираем **Меры (Measures)** на ленте Power Pivot, и затем выбираем **Создать меру (New Measure)**. Заполняем диалоговое окно **Мера (Measure)**, как показано на рис. 47.24.



**Рис. 47.24.** Создана мера для вычисления прибыли 2018 г.

Сходным образом создается мера **PROFIT19** для прибыли 2019 года. Затем мы создаем меру **PROFITCHANGE**, как показано на рис. 47.25, которая рассчитывает увеличение прибыли в 2019 г. по сравнению с 2018-м.

Measure

Table name: IDS

Measure name: PROFITCHANGE

Description:

Formula: 
$$=([PROFIT19]-[PROFIT18])$$

Formatting Options

Category:

- General
- Number
- Currency

OK Cancel

**Рис. 47.25.** Создание меры для изменения прибыли в 2019 г.

Рисунок 47.26 показывает сводную таблицу, обобщающую все три вычисляемых меры для разных штатов.

Row Labels	PROFIT18	PROFITCHANGE	PROFIT19
Alabama	\$7,301.69	-\$92.03	\$7,209.66
Arizona	\$9,350.69	-\$1,637.01	\$7,713.68
Arkansas	\$4,965.43	-\$2,679.71	\$2,285.72
California	\$64,971.05	-\$3,216.09	\$61,754.96
Colorado	\$9,150.68	-\$1,491.42	\$7,659.26
Connecticut	\$8,952.08	\$206.75	\$9,158.83
Florida	\$26,647.90	\$4,435.99	\$31,083.90
Georgia	\$18,696.18	-\$2,782.35	\$15,913.83
Illinois	\$32,756.67	-\$6,283.67	\$26,473.00
Indiana	\$7,684.94	-\$828.20	\$6,856.74
Iowa	\$2,724.70	-\$147.31	\$2,577.39
Kansas	\$9,324.27	-\$937.65	\$8,386.62
Kentucky	\$8,422.96	\$875.73	\$9,298.69

**Рис. 47.26.** Разбиение прибыли 2018 г., 2019 г. и изменения в прибыли по штатам

В файле Calccolumns3.xlsx содержится окончательный анализ по фирме Saleco.

## Задания

1. Вычислите общие продажи для каждого штата по годам, используя файл States.xlsx.
2. Вычислите общую выручку по магазинам и создайте срез для магазинов (поле Store), используя файл Storesales.txt.

## ГЛАВА 48

# 2D-картограммы и 3D-карты Power Map

### Обсуждаемые вопросы

- Как создать двумерную картограмму?
- Как создать 3D-карту Power Map?
- Как применить фильтр к карте Power Map?
- Как изменить способ отображения данных на карте Power Map?
- Как с помощью временной шкалы анимировать данные на карте Power Map?
- Можно ли обобщить данные с карты Power Map двумерной диаграммой?
- Как создать карту Power Map, содержащую круговые диаграммы с подписями?
- Как убедиться, что на карте Power Map правильно отображены геоточки?

Если у вас стоит Microsoft Excel 2019, то у вас есть возможность создавать 3D-карты Power Map. В этой главе вы узнаете, как создавать чудесные карты, использующие данные, привязанные к геоточкам. 2D-карты с заливкой (картограммы) в целом представляют из себя двумерные температурные карты (см. главу 24 «Условное форматирование»), привязанные к геоточкам. 3D-карты Power Map позволяют вам создавать столбчатые и пузырьковые диаграммы на основе данных, привязанных к геоточкам.

## Ответы на вопросы

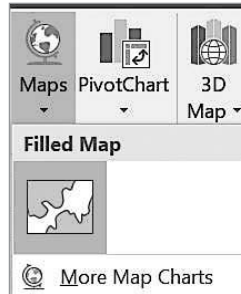
### ❓ Как создать двумерную картограмму?

Продemonстрируем создание двумерной карты с заливкой на примере данных о продажах из файла `Filledmaptemp.xlsx` (рис. 48.1).

Перед применением картографических возможностей Excel вы должны свести свои данные в таблицу Excel (см. главу 26 «Таблицы»). Выделите данные и выберите среди значков карт **Картограмма (Map)** в разделе **Диаграммы (Charts)** на вкладке **Вставка (Insert)**, как показано на рис. 48.2.

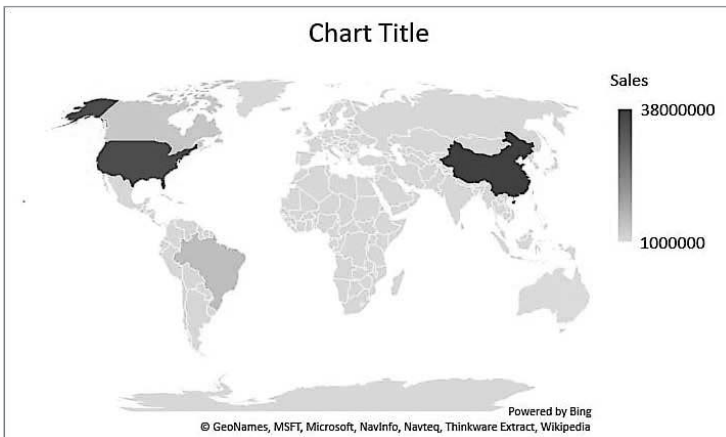
	A	B
1	Country	Sales
2	United States	34000000
3	Canada	5000000
4	Brazil	7000000
5	China	38000000
6	Australia	1000000

**Рис. 48.1.** Данные по странам для создания карты с заливкой



**Рис. 48.2.** Значки карт на ленте

Получится карта, показанная на рис. 48.3. Большой объем продаж показан более темным оттенком синего.



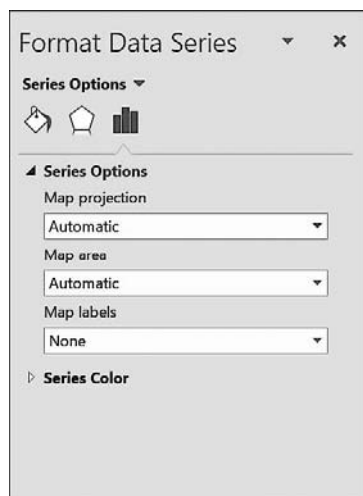
**Рис. 48.3.** Двумерная карта объема продаж по странам

Щелкнув правой кнопкой мыши на любой стране, выберите из контекстного меню **Формат ряда данных (Format Data Series)**, что вызовет панель, показанную на рис. 48.4. Вы можете выбрать из четырех картографических проекций. Вариант **Авто (Automatic)** предоставляет выбор проекции самой программе Excel. **Проекция Меркатора (Mercator)** — это стандартная проекция, примененная в 1569 г. Герардом Меркатором. **Проекция Миллера (Miller)** является разновидностью проекции Меркатора. Наконец, **Проекция Робинсона (Robinson)** — это карта мира.

В поле **Область карты (Map Area)** вы можете выбрать **Автоматически (Automatic)** и **Только области с данными (Only Regions With Data)** (отобразятся только те местоположения, по которым есть данные), либо **Мир (World)**, если вы хотите отобразить весь мир. Сам я предпочитаю вариант **Автоматически (Automatic)**.

Через Подписи на карте (Map Labels) можно управлять тем, какие подписи появятся на карте; здесь имеются варианты Нет (None), Только подходящие по размеру (Best Fit Only) и Показать все (Show All).

Цвет ряда (Series Color) позволяет вам выбрать двух- или трехцветную температурную карту и цветовую схему.



**Рис. 48.4.** Меню форматирования карты с заливкой

Мы добавили к карте заголовок. Выбрав для Проекция карты (Map Projection) и Область карты (Map Area) — Автоматически, а также Показать все и Трехцветная, мы получаем карту, показанную на рис. 48.5. Заметь те, что мы можем также настроить отображение подписей в окне Картограмма (Filled Map) на вкладке Конструктор



**Рис. 48.5.** Добавление подписей к карте с заливкой

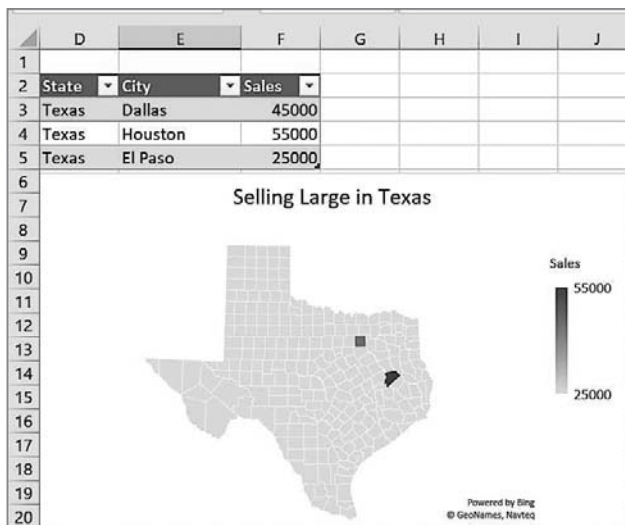


(Design) (она видна, когда курсор на карте). Выбрав **Метки данных (Labels)** в **Добавить элементы диаграммы (Add Chart Element)**, и затем выбрав **Дополнительные параметры подписей данных...** (**More Label Options**), вы можете добавить на карту с заливкой цифровые данные и географические названия.

С помощью данных из файла **USfilledmaps.xlsx** можно продемонстрировать, как Excel обращается с данными по графствам и городам США. На листе **counties** находятся данные о продажах хлеба в четырех графствах Индианы. Заметьте, что мы указали штат каждого графства. Из этих данных была создана карта с заливкой, показанная на рис. 48.6.



**Рис. 48.6.** Продажа хлеба в графствах Индианы



**Рис. 48.7.** Продажа хлеба в городах Техаса

Данные на листе **cities** отображают продажи хлеба в Хьюстоне, Далласе и Эль Пасо. На рис. 48.7 показана карта с заливкой, созданная на основе этих данных. Правильное отображение подписей требует создания карты, которая была бы слишком велика для показа на рисунке.

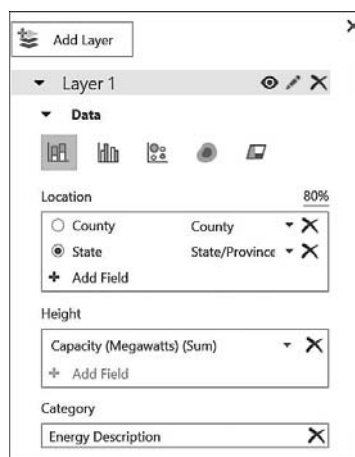
### ? Как создать 3D-карту Power Map?

Инструмент Excel 2019 3D Power Maps позволяет вам создавать по-настоящему замечательные карты на основе следующих географических признаков местоположения:

- Улица.
- Город.
- Штат/провинция.
- Индекс.
- Страна/регион.

В этом разделе я сделаю краткий обзор 3D-карт Excel. Как показано на рис. 48.2, трехмерные карты создаются из середины вкладки **Вставка (Insert)** в группе **Обзоры (Tours)**.

Чтобы посмотреть, как это работает, откройте файл **Powercapacitytemp.xlsx**. Он содержит информацию об энергомощностях США. Такие данные, как штат, графство, мощность завода и энергетические показатели, составят основу для создания нашей первой 3D-карты. Предположим, вы хотите отобразить при помощи накопительной столбчатой диаграммы мощности каждого штата по разным типам энергии. Начните с создания таблицы данных. Выделив данные в таблице, щелкните **3D-карта (3D Map)** и выберите **Открыть 3D-карты (Open 3D Maps)**. Затем заполните панель **Слой (Layer)** так, как показано на рис. 48.8.



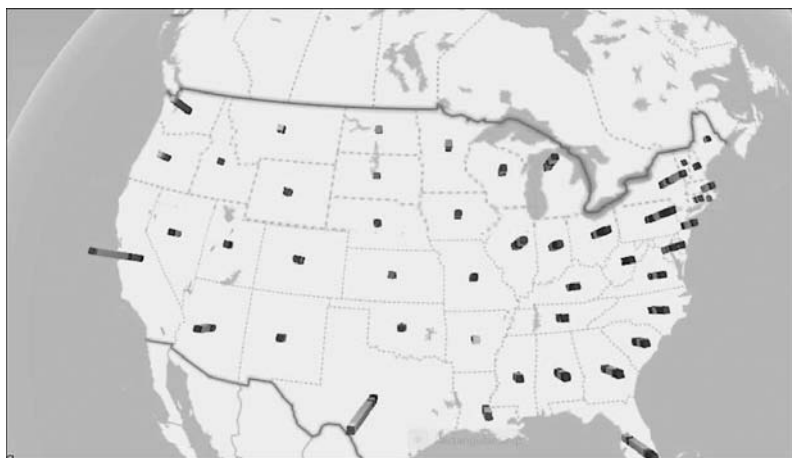
**Рис. 48.8.** Настройки для создания карты с накопительной столбчатой диаграммой

В разделе **Расположение** вы должны указать географические признаки местоположения (в данном случае **Штат**), на этой основе формируется картируемая область.

Выберите параметр **Накопительная (Capacity (Sum))**, чтобы контролировать высоту нашей 3D-карты. Высота столбика на штате будет показывать общую энергомощность штата.

Выбрав **Описание типа энергии (Energy Description)** в поле **Категория (Category)** и **Гистограмму с накоплением (Stacked-Column)** (первый вариант в списке наверху панели в разделе **Данные (Data)**), мы сделаем так, чтобы энергомощность штата отражалась по ее категориям. Возможны также варианты **Гистограмма с группировкой (Clustered Column Map)**, **Тепловая карта (Heat Map)**, **Пузырьковая диаграмма (Bubble Map)** или **Региональная карта (Region Map)**. (Варианты карт отражаются в разделе **Данные (Data)** наверху панели **Слой (Layer)**.)

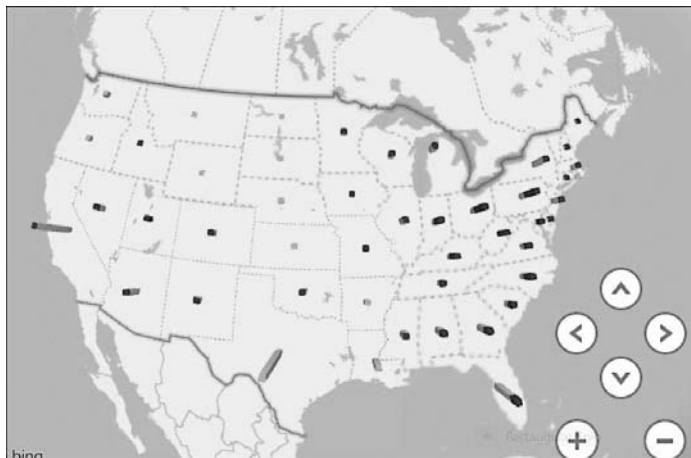
Полученная 3D-карта показана на рис. 48.9. Щелкнув по кнопке с + (или выполнив двойной щелчок по карте либо нажав клавишу +), вы увеличите карту. Щелкнув по кнопке – (или нажав клавишу –) — уменьшите. Стрелками вверх и вниз (кнопками или клавишами) можно приближаться к поверхности глобуса или отдаляться от него, а стрелками влево и вправо — вращать карту. Можно также воспользоваться колесиком мыши для увеличения и уменьшения, а также нажать и потянуть карту, чтобы повернуть глобус. Ваша карта сохраняется как **Обзор**, который можно воспроизвести в динамике. Если захотите, вы сможете начать новый обзор. Щелкнув по X в правом верхнем углу ленты, вы закроете карту и вернетесь в Excel. Щелкнув 3D-столбец любого штата, вы получите полезную информацию о том, откуда штат берет энергию. Например, Техас извлекает значительную часть своей энергии из природного газа, Вашингтон — из воды (то есть гидроэлектроэнергия), а Индиана — из угля.



**Рис. 48.9.** 3D-карта энергомощности, отражающая штат и тип энергии

### ? Как применить фильтр к карте Power Map?

Примерно таким же образом, как фильтр к сводной таблице: вы можете добавить фильтры на панели Слой. На рис. 48.10 показано применение фильтра, так чтобы карта отражала только мощности по природному газу и углю.



**Рис. 48.10.** Применение фильтра к карте с целью показать данные только по углю и природному газу

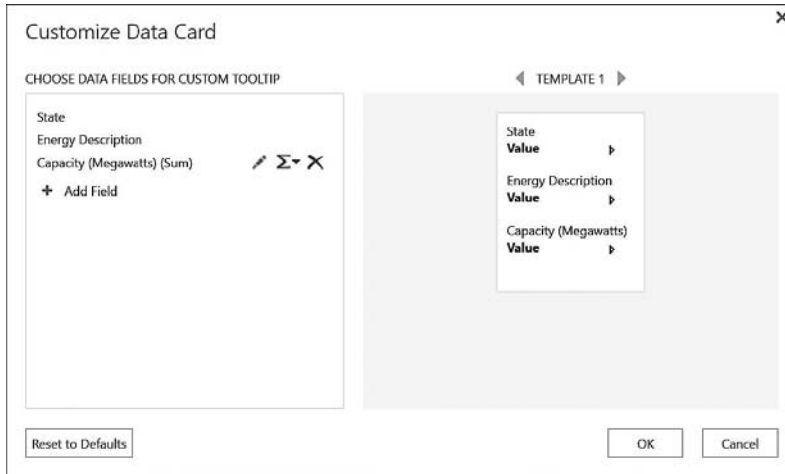
### ? Как изменить способ отображения данных на карте Power Map?

Наведя курсор на столбик по природному газу для Техаса, вы увидите информацию, показанную на рис. 48.11.

State	<b>TX</b>
Energy Description	<b>Natural Gas</b>
Capacity (Megawatts) (Sum)	<b>78335.8</b>

**Рис. 48.11.** Информация по природному газу для Техаса

Щелкнув по шестеренке, показанной на рис. 48.12, вы сможете изменить отображение «визитки» для каждого столбца данных. Щелкнув **Добавить поле + (Add Field +)**, вы можете добавлять, удалять или реорганизовывать поля, имеющиеся на визитке, которая появляется при наведении на столбец курсора мыши.



**Рис. 48.12.** Настройка вида информационной визитки

### ? Как с помощью временной шкалы анимировать карту Power Map?

Если ваши данные содержат даты или иные сведения о времени в каждой строке, то с помощью временной шкалы вы сможете добавить в **Power Map** анимацию, отражающую изменение карты со временем. Чтобы посмотреть, как это работает, откройте файл `Drugstoretimelinetemp.xlsx`. Как показано на рис. 48.13, этот файл содержит сведения о штате, годе (вводимом как 1 января такого-то года), категории товара и объеме продаж. Нам требуется создать 3D-карту со столбцами с накоплением и временной шкалой, которая будет отображать рост продаж в каждом штате со временем.

	B	C	D	E
3	State	Product	Year	Amount
4	Nebraska	drugs	1/1/2019	\$88.23
5	Texas	drugs	1/1/2016	\$70.60
6	Pennsylvania	drugs	1/1/2018	\$95.76
7	Florida	cosmetics	1/1/2016	\$55.90
8	Florida	cosmetics	1/1/2018	\$64.82
9	Florida	food	1/1/2015	\$28.70
10	California	cosmetics	1/1/2019	\$82.11
11	Virginia	drugs	1/1/2015	\$57.05
12	Arizona	drugs	1/1/2019	\$45.39
13	Minnesota	cosmetics	1/1/2017	\$41.52
14	Missouri	food	1/1/2018	\$87.64
15	Illinois	drugs	1/1/2016	\$35.30
16	North Carolina	drugs	1/1/2016	\$72.70
17	Ohio	food	1/1/2015	\$17.08
18	Delaware	cosmetics	1/1/2016	\$64.30
19	Michigan	cosmetics	1/1/2016	\$24.80
20	North Carolina	drugs	1/1/2016	\$13.10
21	Ohio	drugs	1/1/2018	\$81.62

**Рис. 48.13.** Данные по продажам в аптеках

Чтобы получилась карта со столбчато-накопительными диаграммами и временной шкалой, нужно добавить поле Год (Year) на панели Слой (Layer) в разделе Время (Time), как показано на рис. 48.14.



**Рис. 48.14.** Настройки на панели Слой, которые необходимо добавить для получения временной шкалы



**Рис. 48.15.** Отображение изменения аптечных продаж во времени

Как показано на рис. 48.15, временная шкала, позволяющая вам запускать изменения карты во времени, появляется внизу экрана.

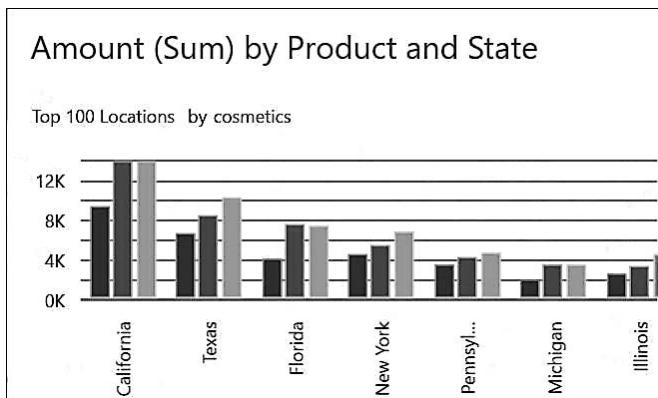
Можно просмотреть рост аптечных продаж по штатам, поочередно нажав кнопку Воспроизвести обзор (Play) либо перетаскив треугольник воспроизведения. Попад в начало года, вы каждый раз будете видеть увеличение размера каждого накопительного столбца. Как показано на рис. 48.16, нажатие на шестеренку на

временной шкале вызывает диалоговое окно **Параметры сцены (Scene Options)**, где вы можете настроить скорость передвижения по времени. Настройки, показанные на рис. 48.16, приведут к довольно высокой скорости.

**Рис. 48.16.** Диалоговое окно Параметры сцены

### ? Можно ли обобщить данные с карты Power Map двумерной диаграммой?

Если вы хотите взглянуть на ваши данные, обобщенные по местоположению, в виде простой двумерной диаграммы, то нажмите на значок **Двумерная диаграмма (2D Chart)**, расположенный в разделе **Вставить (Insert)** на ленте Power Map. Относительно данных по аптекам получится диаграмма, показанная на рис. 48.17. Вы можете поменять тип диаграммы с помощью значков, расположенных в правой верхней части диаграммы.



**Рис. 48.17.** Двумерная столбчатая диаграмма, обобщающая аптечные продажи

### ❓ Как создать карту Power Map, содержащую круговые диаграммы с подписями?

В разделе Карта (Map) на ленте Power Map можно настроить Подписи на карте (Map Labels). Если вы воспользуетесь этой настройкой для карты, показанной ранее на рис. 48.9, то, казалось бы, можно ожидать, что появятся подписи к каждому штату. Однако появляются только названия стран. Если вам нужны подписи к штатам, то нужно выбрать Плоская карта (Flat Map). Изменив тип диаграммы на Пузырьковая (Bubble), вы получите карту с подписями к штатам, показанную на рис. 48.18.



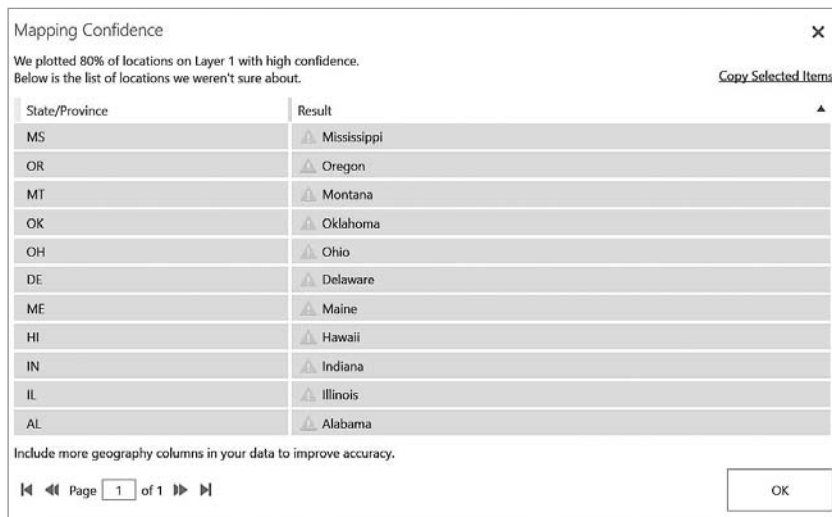
**Рис. 48.18.** Круговые диаграммы, отражающие энергомощности

### ❓ Как убедиться, что на карте Power Map правильно отображены геоточки?

Предположим, у вас есть данные по городам США, и в одной из строк вашего листа содержится Спрингфилд. Если вы смотрите мультсериал «Симпсоны», то знаете, что городов под названием Спрингфилд в США много, так что на карте Power Map может быть ошибка. Затем — Колумбия в Миссури и Колумбия в Южной Каролине, обе находятся в списке 300 крупнейших городов США по численности населения. Без информации о штате карта Power Map может скомпоноваться неправильно. На панели Слой (Layer) вы сможете без труда увидеть, как программа определила географическое местоположение. Щелкните на процентах, отразившихся в правом верхнем углу панели Слой — % (Открытие отчета о достоверности карты), и вы увидите Отчет о достоверности карты (Mapping Confidence Report) (рис. 48.19). В примере с энергомощностями вы увидите, что программа не уверена в правильности определения местоположения относительно 11 штатов.

Если программа сомневается, то вы можете добавить столбец Широта (Latitude) и Долгота (Longitude) к вашим данным (и то и другое в десятичном виде). Например, напечатать Columbia, Missouri, в Bing вы узнаете, что нужная широта — 38,95, а долгота — 92,33.





**Рис. 48.19.** Отчет о правильности карты

## Задания

В файле **Cellphonedata.xlsx** находятся данные о численности населения и количестве сотовых телефонов в 20 странах/регионах, где больше всего сотовых телефонов. Используйте этот файл для решения задач 1–2.

1. На основе данных в файле **Cellphonedata.xlsx** создайте карту с заполнением, обобщающую сведения о сотовых телефонах по стране/региону.
2. На основе данных в файле **Cellphonedata.xlsx** создайте столбчатую 3D-карту, обобщающую сведения о сотовых телефонах по стране/региону.
3. С помощью данных из файла **Seattle.xlsx** создайте карту, в которой высота каждого столбца будет представлять среднюю инспекционную оценку города.
4. Файл **Stateincome.xlsx** содержит медианный доход по каждому штату в 2010 г. Создайте 3D-карту, обобщающую эти сведения.
5. В файле **Top20.xlsx** можно найти данные о численности населения в 20 крупнейших городах США. Создайте пузырьковую карту Power Map, отображающую эти сведения.
6. Создайте лист, который послужит основой для создания карты Power Map, отражающей объема продаж товара в Блумингтоне, штат Индиана, Блумингтоне, штат Иллинойс и Блумингтоне, штат Миннесота.
7. В файле **Electoralvotes.xlsx** содержится количество членов палаты представителей и сената в каждом штате (плюс округ Колумбия). Создайте столбчатую диаграмму с накоплением, отражающую число членов палаты представителей и сената по каждому штату.

## ГЛАВА 49

# Спарклайны

### Обсуждаемые вопросы

- Как графически отобразить ежедневные расчеты с клиентами для каждого филиала банка в одной ячейке?
- Как изменять спарклайны?
- Как обобщить серию из побед и поражений команд НФЛ в одной ячейке?
- Обновляются ли спарклайны автоматически при добавлении новых данных?

Спарклайны — это графики, позволяющие отобразить строку или столбец данных в одной ячейке. Термин «спарклайн» был введен Эдвардом Тафти, известным специалистом в области представления и анализа данных. Он описал спарклайны как «насыщенные данными графики с простым дизайном и размером в слово». В Microsoft Excel 2019 можно быстро создать графики, хранящиеся в одной ячейке.

## Ответы на вопросы

### ❓ Как графически отобразить ежедневные расчеты с клиентами для каждого филиала банка в одной ячейке?

В файле **Sparklines.xlsx** вы найдете данные о ежедневных расчетах с клиентами в нескольких филиалах банка Нью-Йорка (рис. 49.1).

Пусть необходимо обобщить еженедельные расчеты с клиентами, построив график ежедневных расчетов для каждого филиала в одной ячейке. Выделите диапазон для размещения спарклайнов (я выбрал ячейки **I8:I14**) и затем на вкладке **Вставка (Insert)** в группе **Спарклайны (Sparklines)** выберите **График (Line)**, как показано на рис. 49.2.

Заполните диалоговое окно **Создание спарклайнов (Create Sparklines)** данными, на которых будут основаны спарклайны (я выбрал **D8:H14** для ввода в поле **Диапазон данных (Data Range Box)**), — рис. 49.3.

Вы увидите линейный график (рис. 49.4), обобщающий график клиентов для каждого филиала. Результаты находятся на листе **Line Sparkline** в файле **Sparklines.xlsx**.

Как видите, в каждом филиале понедельник и пятница являются самыми напряженными днями.

	C	D	E	F	G	H
7		Monday	Tuesday	Wednesday	Thursday	Friday
8	New York	1176	768	808	864	1235
9	Rochester	475	323	333	356	515
10	Utica	360	250	228	275	378
11	Syracuse	594	412	408	459	618
12	Buffalo	698	475	504	551	803
13	Ossining	306	208	204	234	322
14	Ithaca	437	288	294	299	450

Рис. 49.1. Данные для спарклайнов

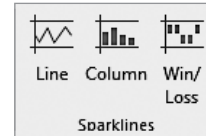


Рис. 49.2. Выбор спарклайнов

Create Sparklines ? X

Choose the data that you want

Data Range: D8:H14

Choose where you want the sparklines to be placed

Location Range: I8:I14

OK Cancel

Рис. 49.3. Диалоговое окно Создание спарклайнов

	C	D	E	F	G	H	I
7		Monday	Tuesday	Wednesday	Thursday	Friday	Summary
8	New York	1176	768	808	864	1235	
9	Rochester	475	323	333	356	515	
10	Utica	360	250	228	275	378	
11	Syracuse	594	412	408	459	618	
12	Buffalo	698	475	504	551	803	
13	Ossining	306	208	204	234	322	
14	Ithaca	437	288	294	299	450	

Рис. 49.4. Линейные спарклайны для отображения трафика филиалов

### ? Как изменять спарклайны?

Если вы щелкнете по любой ячейке, содержащей спарклайн, появится вкладка Конструктор для спарклайнов (Sparkline Tools Design). На вкладке Конструктор (Design) в спарклайны вы можете внести множество изменений. Например, как показано на рис. 49.5, можно отметить максимальную и минимальную точки.

Спарклайны представлены на рис. 49.6. См. лист High Low в файле Sparklines.xlsx.

Эти спарклайны показывают, что самым напряженным днем для каждого филиала является пятница, а самыми незагруженными днями — вторник и среда.

<input checked="" type="checkbox"/> High Point	<input type="checkbox"/> First Point
<input checked="" type="checkbox"/> Low Point	<input type="checkbox"/> Last Point
<input type="checkbox"/> Negative Points	<input type="checkbox"/> Markers
Show	

**Рис. 49.5.** Выбор максимальной и минимальной точек для отметки на спарклайнах

	C	D	E	F	G	H	I
7		Monday	Tuesday	Wednesday	Thursday	Friday	Summary
8	New York	1176	768	808	864	1235	
9	Rochester	475	323	333	356	515	
10	Utica	360	250	228	275	378	
11	Syracuse	594	412	408	459	618	
12	Buffalo	698	475	504	551	803	
13	Ossining	306	208	204	234	322	
14	Ithaca	437	288	294	299	450	

**Рис. 49.6.** На спарклайнах отмечены максимальные и минимальные точки

На вкладке **Конструктор (Design)** со спарклайнами можно выполнить следующие действия.

- Изменить тип спарклайна (график, гистограмма, выигрыш/проигрыш). Спарклайн-гистограмма и спарклайн выигрыша/проигрыша рассмотрены далее в этой главе.
- Выбрать **Изменить данные (Edit Data)** для изменения данных, используемых для создания спарклайнов. Кроме того, можно изменить настройки по умолчанию так, что в спарклайны будут включаться скрытые данные.
- Выбрать любую комбинацию из максимальной, минимальной, отрицательной, первой и последней точек.
- Изменить стиль или цвет спарклайнов и маркеров.
- Выбрать **Ось (Axis)** для изменения осей каждого спарклайна. Например, можно указать одинаковый масштаб оси *x* или *y* для каждого спарклайна. По умолчанию масштаб для каждой оси рассчитывается на основе данных каждого спарклайна. Такое масштабирование показано на рис. 49.4. **Пользовательское значение (Custom)** позволяет выбрать верхний и нижний пределы для каждой оси.
- Если точки данных соответствуют хаотично расположенным точкам на оси дат, в группе **Ось (Axis)** можно выбрать **Тип оси дат (Date Axis)**, и, таким образом, точки на графике будут находиться друг от друга на расстоянии, пропорциональном разнице во времени. Рисунок 49.7 показывает продажи компании через нерегулярные интервалы. (См. лист **Date Axis** в файле **Sparklines.xlsx**.) В ячейке F12 график построен так, будто между датами продаж имеются рав-

ные промежутки времени. В группе **Ось (Axis)** выберите **Тип оси дат (Date Axis)** и укажите диапазон **C10:C13** со значениями дат. Спарклайн в ячейке **F10** отображает нерегулярные периоды времени.

	C	D	E	F	G
9	<b>Date</b>	<b>Sales</b>			
10	1/1/2010	1000			Date axis
11	6/1/2010	1200			
12	9/1/2012	1400			No date axis
13	1/1/2015	1900			

**Рис. 49.7.** Спарклайны с осью дат

Спарклайн-график можно преобразовать в спарклайн-гистограмму. Для этого щелкните по любому спарклайну и на вкладке **Конструктор (Design)** в группе **Тип (Type)** выберите **Гистограмма (Column)**. См. рис. 49.8 и лист **Column Sparkline** в файле **Sparklines.xlsx**.

	C	D	E	F	G	H	I
7		<b>Monday</b>	<b>Tuesday</b>	<b>Wednesday</b>	<b>Thursday</b>	<b>Friday</b>	<b>Summary</b>
8	New York	1176	768	808	864	1235	■ _ _ _ ■
9	Rochester	475	323	333	356	515	■ _ _ _ ■
10	Utica	360	250	228	275	378	■ _ _ _ ■
11	Syracuse	594	412	408	459	618	■ _ _ _ ■
12	Buffalo	698	475	504	551	803	■ _ _ _ ■
13	Ossining	306	208	204	234	322	■ _ _ _ ■
14	Ithaca	437	288	294	299	450	■ _ _ _ ■


















**Рис. 49.8.** Спарклайн-гистограмма

### ❓ Как обобщить серию из побед и поражений команд НФЛ в одной ячейке?

Файл **Nflwinslosses.xlsx** содержит результаты игр команд НФЛ в регулярном сезоне 2009 г. Часть этих данных приведена на рис. 49.9. (См. лист **Win Loss**.)

1 (единица) обозначает победу, а -1 (минус единица) — поражение. Каждый спарклайн выигрыша/проигрыша отображает любое положительное число как столбик вверх, а любое отрицательное число как столбик вниз. 0 обозначается на графике как разрыв. Для создания спарклайнов выигрыша/проигрыша выделите диапазон, в котором должны быть размещены спарклайны (ячейки **C8:C39**), и на вкладке **Вставка (Insert)** в группе **Спарклайны (Sparklines)** выберите **Выигрыш/проигрыш (Win/Loss)**. Укажите диапазон данных **E8:T39**. Спарклайны очень оживляют сезон 2009 г. в НФЛ. Как видно из рисунка, команды «Индианаполис Колтс» и «Нью-Орлеан Сэйнтс» великолепно начали сезон. Команда «Теннеси Тайтанс»

стартовала неудачно, но потом выровнялась. Команда «Нью-Йорк Джайентс» стартовала хорошо, а затем у нее наступила черная полоса. Спарклайны выигрыша/проигрыша прекрасно подходят для отслеживания прогресса организации (см. задание 4 в конце главы).

	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
6			Game															
7		Team	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
8		Arizona Cardinals	-1	1	-1	1	1	1	-1	1	1	1	-1	1	-1	1	1	-1
18		Detroit Lions	-1	-1	1	-1	-1	-1	-1	-1	1	-1	-1	-1	-1	-1	-1	-1
19		Green Bay Packers	1	-1	1	-1	1	1	-1	1	1	1	1	1	1	-1	1	1
20		Houston Texans	-1	1	-1	1	-1	1	1	1	-1	-1	-1	1	1	1	1	1
21		Indianapolis Colts	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-1	-1
22		Jacksonville Jaguars	-1	-1	1	1	-1	1	-1	1	1	1	-1	1	-1	-1	-1	-1
23		Kansas City Chiefs	-1	-1	-1	-1	-1	1	-1	1	1	-1	-1	-1	-1	-1	-1	1
24		Miami Dolphins	-1	-1	-1	1	1	-1	1	-1	1	1	-1	1	1	-1	-1	-1
25		Minnesota Vikings	1	1	1	1	1	1	-1	1	1	1	1	-1	1	-1	-1	1
26		New England Patriots	1	-1	1	1	-1	1	1	1	-1	1	-1	-1	1	1	1	-1
27		New Orleans Saints	1	1	1	1	1	1	1	1	1	1	1	1	1	-1	-1	-1
28		New York Giants	1	1	1	1	1	-1	-1	-1	-1	1	-1	1	-1	1	-1	-1
29		New York Jets	1	1	1	-1	-1	-1	1	-1	-1	-1	1	1	1	-1	1	1
30		Oakland Raiders	-1	1	-1	-1	-1	1	-1	-1	1	1	-1	1	-1	1	1	-1
31		Philadelphia Eagles	1	-1	1	1	-1	1	1	-1	-1	1	1	1	1	1	1	-1
32		Pittsburgh Steelers	1	-1	-1	1	1	1	1	1	-1	-1	-1	-1	-1	1	1	1
33		San Diego Chargers	1	-1	1	-1	-1	1	1	1	1	1	1	1	1	1	1	1

**Рис. 49.9.** Спарклайны для побед и поражений команд НФЛ

**?** Обновляются ли спарклайны автоматически при добавлении новых данных?

Если вы хотите, чтобы при добавлении новых данных спарклайны обновлялись автоматически, преобразуйте данные в таблицу (см. главу 26 «Таблицы») или в динамический диапазон (см. главу 22 «Функция СМЕЩ»).

## Задания

- На основе данных о филиалах банка создайте спарклайны-графики с одним и тем же масштабом для каждого филиала.
- В файле **Dow.xlsx** содержатся значения индекса Доу — Джонса со 2 января по 10 августа 2010 г. Создайте спарклайн-график, отображающий взлеты и падения рынка.
- Зафиксируйте взлеты и падения рынка с помощью спарклайнов выигрыша/проигрыша. Воспользуйтесь файлом **Dow.xlsx**.
- В файле **Goals.xlsx** для каждого месяца содержатся процентные доли, описывающие то, насколько филиалы банка достигли своих целей в каждом месяце. Обработайте эти данные с помощью спарклайнов выигрыша/проигрыша.

## ГЛАВА 50

# Обработка данных с помощью статистических функций для баз данных

### Обсуждаемые вопросы

- Joolas, небольшая компания по продаже косметики, хранит данные о каждой сделке на листе Microsoft Excel. При этом ей часто требуется узнать ответы на следующие вопросы:
  - На какую сумму продала блеск для губ Джен?
  - Каково среднее количество проданной помады для каждой сделки Джен в регионе Восток?
  - На какую общую сумму было продано косметики либо продавцом Эмили, либо в регионе Восток?
  - На какую сумму продали помаду Коллин и Зарет в регионе Восток?
  - Сколько сделок с помадой было совершено не в регионе Восток?
  - На какую сумму продала помаду Джен в 2004 г.?
  - Сколько единиц косметики было продано по цене не менее \$3,20?
  - Какова общая сумма выручки для каждого косметического товара, реализованного каждым продавцом?
- Какие приемы применяются при создании диапазона условий?
- У меня есть база со следующими данными для каждой сделки: выручка, дата продажи и идентификационный код продукта. Как быстро извлечь сумму выручки для сделки по заданной дате продажи и идентификационному коду?

Как вы убедились в главе 45 «Сводные таблицы и срезы для описания данных», сводные таблицы Excel являются отличным инструментом для обобщения данных. Однако часто сводные таблицы содержат гораздо больше информации, чем необходимо. Специальные статистические функции для баз данных отвечают на любые вопросы отчетности без необходимости создания сводной таблицы.

Вы уже знакомы с такими функциями, как СУММ (SUM), СРЗНАЧ (AVERAGE), СЧЁТ (COUNT), МАКС (MAX) и МИН (MIN). Префикс БД/Б/Д (D) (расшифровывающийся как «для баз данных») к именам этих и других функций обозначает статистиче-

ские функции для баз данных. Но что такого делает функция **БДСУММ** (DSUM), чего, например, не может функция **СУММ**? Функция **СУММ** суммирует значения всех ячеек в диапазоне, а функция **БДСУММ** позволяет указать (с помощью условий) подмножество строк, суммируемых в диапазоне ячеек. Пусть в компании по продаже косметики имеется база данных со следующей информацией по каждой сделке:

- имя продавца;
- дата сделки;
- продаваемый продукт;
- количество единиц проданной продукции;
- сумма выручки в долларах для каждой сделки;
- регион, в котором произошла сделка.

Эти данные находятся в файле **Makeupdb.xlsx**, часть их показана на рис. 50.1.

	G	H	I	J	K	L	M
4	Trans Number	Name	Date	Product	Units	Dollars	Location
5	1	Betsy	4/1/2004	lip gloss	45	\$ 137.20	south
6	2	Hallagan	3/10/2004	foundation	50	\$ 152.01	midwest
7	3	Ashley	2/25/2005	lipstick	9	\$ 28.72	midwest
8	4	Hallagan	5/22/2006	lip gloss	55	\$ 167.08	west
9	5	Zaret	6/17/2004	lip gloss	43	\$ 130.60	midwest
10	6	Colleen	11/27/2005	eye liner	58	\$ 175.99	midwest
11	7	Cristina	3/21/2004	eye liner	8	\$ 25.80	midwest
12	8	Colleen	12/17/2006	lip gloss	72	\$ 217.84	midwest
13	9	Ashley	7/5/2006	eye liner	75	\$ 226.64	south
14	10	Betsy	8/7/2006	lip gloss	24	\$ 73.50	east
15	11	Ashley	11/29/2004	mascara	43	\$ 130.84	east
16	12	Ashley	11/18/2004	lip gloss	23	\$ 71.03	west
17	13	Emilee	8/31/2005	lip gloss	49	\$ 149.59	west
18	14	Hallagan	1/1/2005	eye liner	18	\$ 56.47	south
19	15	Zaret	9/20/2006	foundation	-8	\$ (21.99)	east
20	16	Emilee	4/12/2004	mascara	45	\$ 137.39	east
21	17	Colleen	4/30/2006	mascara	66	\$ 199.65	south
22	18	Jen	8/31/2005	lip gloss	88	\$ 265.19	midwest
23	19	Jen	10/27/2004	eye liner	78	\$ 236.15	south
24	20	Zaret	11/27/2005	lip gloss	57	\$ 173.12	midwest
25	21	Zaret	6/2/2006	mascara	12	\$ 38.08	west
26	22	Betsy	9/24/2004	eye liner	28	\$ 86.51	midwest
27	23	Colleen	2/1/2006	mascara	25	\$ 77.31	midwest
28	26	Emilee	12/6/2006	lip gloss	24	\$ 74.62	west

**Рис. 50.1.** Данные для работы со статистическими функциями для баз данных

Применив функцию **БДСУММ** с соответствующими критериями, можно, например, подсчитать выручку, полученную только от продажи блеска для губ в регионе Восток в 2004 г. Благодаря выставленным вами критериям будут отмечены строки, которые нужно включить для вычисления общей суммы. Для этих строк функция **БДСУММ** работает как обычная функция **СУММ**.

Синтаксис функции: **БДСУММ(база\_данных;поле;условия)**, где:



- **база\_данных** — диапазон ячеек, составляющих базу данных. Первая строка содержит заголовки столбцов;
- **поле** — столбец, содержащий значения, которые требуется суммировать. Его можно определить, заключив заголовок столбца в кавычки. (Например, для столбца, содержащего сумму в долларах, можно указать «Сумма».) Также поле можно определить, указав положение столбца в базе данных, при отсчете слева направо. Например, база данных сделок с косметикой занимает столбцы с N по M. (Столбец **Номер сделки** не является частью базы данных.) Столбец N можно указать как поле 1, а столбец M как поле 6;
- **условия** — ссылка на диапазон, определяющий строки, которыми будет оперировать функция. Первая строка диапазона условий должна включать один или несколько заголовков столбцов. (Единственным исключением из этого правила являются *вычисляемые условия*, которые обсуждаются в двух последних примерах этой главы.) Как видно из примера, при создании диапазона условий важно уяснить, что несколько условий в одной и той же строке объединяются с помощью функции И (AND), а условия в разных строках — с помощью функции ИЛИ (OR).

## Ответы на вопросы

### ❓ На какую сумму продала блеск для губ Джен?

Здесь необходимо применить функцию БДСУММ к столбцу 5 базы данных. Столбец 5 содержит сумму выручки в долларах для каждой сделки. (База данных Data состоит из диапазона H4:M1895.) Диапазон условий в O4:P5 охватывает все строки в базе данных, в которых Имя эквивалентно Jen, а Продукт эквивалентен lip gloss (блеск для губ). Таким образом, введя формулу =БДСУММ(data;5;O4:P5) в ячейку N5 (рис. 50.2), мы получим общую сумму в долларах, вырученную Джен за продажу

	N	O	P	Q	R	S	T	U	V
4	Jen lip gloss \$	Name	Product						
5	\$	5,461.61	Jen	lip gloss					
6		5461.61479				Name	Dollars		
7	Jen avg.units lipstick in east	Name	Product	Location		Betsy	>160	= ">"&T10	
8		42.25	Jen	lipstick	east				
9		42.25							
10	\$s Emilee or east	Name	Location			Cutoff	160		
11	\$	76,156.48	Emilee						
12			east				18212.09	=DSUM(data,5,S6:T7)	
13	\$s lipstick by Colleen or Zaret in east	Name	Location	Product					
14	\$	1,073.20	Colleen	east	lipstick				
15		1073.203709	Zaret	east	lipstick				
16	Number of lipstick transactions not in east	Product	Location						
17		164	lipstick	<=east	164				
18	total \$s jen lipstick in 2004	Name	Product	Date	Date				
19	\$	1,690.79	Jen	lipstick	>=1/1/2004	<1/1/2005			
20		1690.793115							
21	Units sold for >=\$3.20	Big price							
22		1127	FALSE	= (L5/K5)>=3.2					

Рис. 50.2. Примеры использования статистических функций для баз данных

блеска для губ. Также можно ввести формулу как `=БДСУММ(data;"Сумма";O4:P5)`. Джен продала блеск для губ на \$5461,61. В ячейке N6 тот же самый ответ получен при помощи функции `СУММЕСЛИМН (SUMIFS)`, описанной в главе 21 «Функции СУММЕСЛИ, СРЗНА ЧЕСЛИ, СУММЕСЛИМН, СРЗНА ЧЕСЛИМН, МАКСЕСЛИ и МИНЕСЛИ» по формуле `=СУММЕСЛИМН(Сумма;Имя;"Jen";Продукт;"lip gloss")`.

**❓ Каково среднее количество проданной помады для каждой сделки Джен в регионе Восток?**

Вы можете вычислить это количество, введя в ячейку N8 формулу `=ДСРЗНАЧ(data;4;O7:Q8)`. Значение 4 для аргумента поле определяет столбец Количество, а диапазон условий O7:Q8 помечает все строки в базе данных, в которых Имя эквивалентно Jen, Продукт эквивалентен lipstick, а Регион эквивалентен East. Функция `ДСРЗНАЧ` возвращает среднее количество проданных единиц продукции для помеченных строк. Как видите, в среднем Джен продала 42,25 шт помады в каждой сделке в регионе Восток. В ячейке N9 тот же самый результат вычислен по формуле `=СРЗНАЧЕСЛИМН(Количество; Имя; "Jen"; Продукт; "lipstick"; Регион; "east")`.

**❓ На какую общую сумму было продано косметики либо продавцом Эмили, либо в регионе Восток?**

В ячейке N11 вы можете вычислить общую сумму (\$76 156,48) для сделок Эмили и сделок в регионе Восток по формуле `=БДСУММ(data;5;O10:P12)`. Критерии в диапазоне O10:P12 помечают сделки в регионе Восток или сделки Эмили, поскольку критерии в разных строках обрабатываются при помощи операции ИЛИ. Программистам из Microsoft удалось добиться, чтобы сделки Эмили в регионе Восток не учитывались в этой формуле дважды. Здесь быстро получить ответ с помощью функции `СУММЕСЛИМН` невозможно.

**❓ На какую сумму продали помаду Коллин и Зарет в регионе Восток?**

В ячейке N14 формула `=БДСУММ(data;5;O13:Q15)` вычисляет общую выручку (\$1073,20) от сделок с помадой, заключенных Коллин и Зарет в регионе Восток. Обратите внимание, что в диапазоне O14:Q14 указаны критерии отбора сделок Коллин с помадой в регионе Восток, а в диапазоне O15:Q15 указаны те же условия отбора для сделок Зарет. Напомню, что условия в разных строках объединяются оператором ИЛИ. В ячейке N15 ответ на тот же вопрос получен по формуле: `=СУММЕСЛИМН(Сумма;Имя;"Colleen";Продукт;"lipstick";Регион;"east")+СУММЕСЛИМН(Сумма;Имя;"Zaret";Продукт;"lipstick";Регион;"east")`.

**❓ Сколько сделок с помадой было совершено не в регионе Восток?**

В ячейке N17 вы можете вычислить общее количество сделок с помадой (164) за пределами региона Восток по формуле `=БСЧЁТ(data;4;O16:P17)`. Мы используем

функцию **БСЧЁТ**, так как хотим указать критерии, по которым функция подсчитает число строк, включающих сделки по продаже помады и регионы, кроме Востока. В Excel выражение **<>East** обрабатывается как «не East». Функцию **СУММЕСЛИМН** в этом примере пришлось бы применять к каждому региону.

Так как функция **СЧЁТ** обрабатывает числа, ссылка должна указывать на столбец, содержащий числовые значения. Поскольку столбец 4, **Количество**, содержит числа, в формуле указан именно он. Формула **=БСЧЁТ(data;3;O16:P17)** возвратила бы 0, потому что в третьем столбце базы данных (столбец J на листе) нет чисел. При этом формула **=БСЧЁТА(data;3;O16:P17)** возвратила бы правильный результат, поскольку функция **БСЧЁТА** обрабатывает не только числа, но и текст.

### ? На какую сумму продала помаду Джен в 2004 г.?

Здесь необходимо знать, как пометить только сделки за 2004 г. Это можно сделать, включив в строку диапазона условий ссылку на поле **Дата** и выражения **>=1/1/2004** и **<1/1/2005**. Таким образом, формула **=БДСУММ(data;5;O18:R19)** в ячейке N19 вычисляет общую сумму выручки в долларах от сделок Джен с помадой (\$1690,79) с 1 января 2004 г. и до 1 января 2005 г. В ячейке N20 ответ вычислен по формуле **=СУММЕСЛИМН(Сумма; Дата; ">=1/1/2004"; Дата; "<=31/12/2004"; Продукт; "lipstick"; Имя; "Jen")**.

### ? Сколько единиц косметики было продано по цене не менее \$3,20?

В этом примере задействованы вычисляемые критерии. Как правило, вычисляемые критерии охватывают строки базы данных на основе истинности или ложности критерия для конкретной строки. В этом примере требуется пометить каждую строку, для которой **Сумма/Количество** **>=3,20**. При определении вычисляемого критерия (рис. 50.3) заголовок в первой строке над вычисляемыми условиями не должен быть заголовком столбца. Например, невозможно указать **Имя**, **Продукт** или другой заголовок из строки 4 этого листа. Вычисляемый критерий представляет собой формулу, которая возвращает значение **ИСТИНА** на основе первой строки информации в базе данных. Таким образом, чтобы отметить строки, в которых средняя цена больше или равна \$3,20, введем **=(L5/K5)>=3,2** в диапазон условий и дадим диапазону заголовок, не совпадающий с заголовком столбца. Если первая проверяемая строка данных не удовлетворяет критерию, то на листе появится значение **ЛОЖЬ**, но тем не менее все строки, в которых цена единицы продукции больше или равна \$3,20, будут помечены. Введенная в ячейку N22 формула **=БДСУММ(data;4;O21:O22)** вычислит общее количество единиц косметики (1127),

	N	O	P	Q
21	Units sold for >=\$3.20	Big price		
22	1127	FALSE	= (L5/K5)>=3.2	
23	=DSUM(data,4,O21:O22)			

Рис. 50.3. Вычисляемые критерии

проданных по цене больше или равной \$3,20. Обратите внимание, что в ячейке Q22 содержится формула  $= (L5/K5) >= 3,2$ .

### ❓ Какова общая сумма выручки для каждого косметического товара, реализованного каждым продавцом?

Для решения этой задачи я применил функцию БДСУММ с диапазоном критериев на основе столбцов Имя и Продукт. При помощи таблицы данных можно быстро перебрать все возможные комбинации имен и продуктов в диапазоне критериев и рассчитать общую выручку для каждой комбинации имени и продукта.

Сначала введите любое имя в ячейку X26 и любой продукт в ячейку Y26. Затем в ячейку Q25 введите формулу  $=\text{БДСУММ}(\text{data};5;\text{X25};\text{Y26})$ , вычисляющую общую сумму выручки (в данном случае) для Бетси и подводки для глаз. Введите имя каждого продавца в ячейки Q26:Q33 и каждого продукта в ячейки R25:V25. Теперь выделите таблицу данных (Q25:V33). На вкладке Данные (Data) в группе Работа с данными (Data Tools) выберите Анализ "что, если" (What-If Analysis) и затем выберите Таблица данных (Data Table). В поле Подставлять значения по строкам укажите ячейку X26 и в поле Подставлять значения по столбцам в ячейку Y26. Появится результат, показанный на рис. 50.4. В каждом элементе таблицы данных вычисляется сумма выручки для комбинации соответствующего имени и продукта, поскольку имена из таблицы данных помещаются в ячейку X26, а продукты — в ячейку Y26. Например, Эшли продала помаду на \$3245,45.

	O	P	Q	R	S	T	U	V	W	X	Y
25	dbase functions		6046.53428	lip gloss	foundation	lipstick	mascara	eye liner		Name	Product
26			Betsy	5675.65005	8043.486	3968.605	4827.254	6046.534		Betsy	eye liner
27			Hallagan	5603.11938	6985.734	3177.871	5703.347	6964.621			
28			Zaret	5670.32933	6451.65	2448.707	3879.95	8166.749			
29			Colleen	5573.32373	6834.768	2346.414	6746.525	3389.625			
30			Cristina	5297.97981	5290.99	2401.668	5461.647	5397.274			
31			Jen	5461.61479	5628.648	3953.3	6887.175	7010.441			
32			Ashley	6053.68456	4186.059	3245.443	6617.1	5844.949			
33			Emilee	5270.25031	5313.788	2189.138	4719.3	7587.389			
34											
35											
36	sumifs			lip gloss	foundation	lipstick	mascara	eye liner			
37			Betsy	5675.65005	8043.486	3968.605	4827.254	6046.534			
38			Hallagan	5603.11938	6985.734	3177.871	5703.347	6964.621			
39			Zaret	5670.32933	6451.65	2448.707	3879.95	8166.749			
40			Colleen	5573.32373	6834.768	2346.414	6746.525	3389.625			
41			Cristina	5297.97981	5290.99	2401.668	5461.647	5397.274			
42			Jen	5461.61479	5628.648	3953.3	6887.175	7010.441			
43			Ashley	6053.68456	4186.059	3245.443	6617.1	5844.949			
44			Emilee	5270.25031	5313.788	2189.138	4719.3	7587.389			

**Рис. 50.4.** Комбинирование таблицы данных с функцией БДСУММ

Этот пример показывает, как комбинирование таблиц данных со статистическими функциями для баз данных позволяет создать множество статистических данных. Обратите внимание, что эту задачу также можно решить, скопировав формулу  $=\text{СУММЕСЛИМН}(\text{Сумма}; \text{Имя}; \$Q\$37; \text{Продукт}; \$R\$36)$  из R37 в R37:V44.

### ? Какие приемы применяются при создании диапазона условий?

Вот несколько приемов создания диапазона условий. Допустим, заголовок столбца в первой строке диапазона критериев ссылается на столбец, содержащий текст (например, на столбец Н).

- **\*Allie\*** — помечаются записи, содержащие текстовую строку Allie в столбце Н.
- **A?X** — запись помечается, если значение в ячейке в столбце Н начинается с А и его третьим символом является Х. (Второй символ может быть любым!)
- **<>\*B\*** — запись помечается, если в ячейке в столбце Н не содержится символа В.

Если столбец (например, столбец I) содержит числовые значения, то:

- **>100** — запись помечается, если столбец I содержит значение больше 100;
- **<>100** — запись помечается, если столбец I содержит значение, не равное 100;
- **>=1000** — запись помечается, если значение в столбце I больше или равно 1000.
- С помощью знака & вы можете сослаться в условии на другую ячейку. Например, как показано на рис. 50.5 (см. файл **Makeupdb.xlsx**), использование формулы **= ">"&T10** позволяет задать условие попадания значения в сумму в зависимости от содержимого ячейки T10. Таким образом, формула **=БДСУМ(data,5,S6:T7)** вычислит общую выручку от продаж Бетси в транзакциях, в которых сумма была больше, чем некое значение в ячейке T10.

	S	T	U	V
6	Name	Dollars		
7	Betsy	>160	= ">"&T10	
8				
9				
10	Cutoff	160		
11				
12		18212.09	=ДСУМ(data,5,S6:T7)	

**Рис. 50.5.** Использование знака & при задании условия

### ? У меня есть база со следующими данными для каждой сделки: выручка, дата продажи и идентификационный код продукта. Как быстро извлечь сумму выручки для сделки по заданной дате продажи и идентификационному коду?

Файл **Dget.xlsx** (рис. 50.6) содержит базу данных с суммами выручки, датами продаж и идентификационными кодами продуктов для ряда сделок. Если известна дата сделки и идентификационный код продукта, каким образом можно найти сумму выручки? Проще всего воспользоваться функцией **БИЗВЛЕЧЬ (DGET)**. Синтаксис функции: **БИЗВЛЕЧЬ(база\_данных;поле;условия)**. При заданном диапазоне для базы данных и заданном значении для поля в базе данных (отсчет столбцов

выполняется слева направо в пределах диапазона) функция **БИЗВЛЕЧЬ** возвращает содержимое заданного столбца в записи базы данных, удовлетворяющей критериям. Если ни одна запись не удовлетворяет критериям, функция **БИЗВЛЕЧЬ** возвращает сообщение об ошибке **#ЗНАЧ!**. Если критериям удовлетворяет несколько записей, функция **БИЗВЛЕЧЬ** возвращает сообщение об ошибке **#ЧИСЛО!**.

Пусть вам нужно извлечь выручку для сделки с идентификационным кодом продукта 62426, имевшей место 9/1/2006 (9 января). Допустим, с этим продуктом в указанный день была проведена только одна сделка. Формула **=БИЗВЛЕЧЬ(B7:D28;1;G5:H6)**, введенная в ячейку G9, дает результат \$169. Обратите внимание, что для аргумента поле было указано значение 1, поскольку столбец **Выручка** является первым столбцом базы данных (занимающей диапазон B7:D28). Поиск сделки с кодом продукта 62426 и датой совершения 9/1/2006 обеспечивает диапазон условий G5:H6.

	A	B	C	D	E	F	G	H
2		<b>Example of DGET Function</b>						
3								
4								
5							<b>date</b>	<b>id code</b>
6							1/9/2006	62426
7		<b>revenue</b>	<b>date</b>	<b>id code</b>				
8		\$ 856.00	2/2/2000	89550			<b>revenue</b>	
9		\$ 461.00	4/12/2003	34506			169	
10		\$ 662.00	2/6/1999	57664				
11		\$ 622.00	9/25/2005	25449				
12		\$ 228.00	7/12/2006	26461				
13		\$ 997.00	6/9/2008	73945				
14		\$ 857.00	7/30/2006	78607				
15		\$ 454.00	12/5/1999	8605				
16		\$ 571.00	2/6/2001	33684				
17		\$ 690.00	5/17/2006	81984				
18		\$ 467.00	10/17/2008	4530				
19		\$ 281.00	8/6/2000	72489				
20		\$ 965.00	11/17/2003	66050				
21		\$ 169.00	1/9/2006	62426				
22		\$ 378.00	10/21/2001	34422				
23		\$ 273.00	2/28/2001	41064				
24		\$ 628.00	6/15/2006	29231				
25		\$ 382.00	12/11/2005	9625				
26		\$ 896.00	10/16/1999	66644				
27		\$ 260.00	3/8/2002	3346				
28		\$ 344.00	11/4/2002	38858				

**Рис. 50.6.** Применение функции **БИЗВЛЕЧЬ**

## Задания

1. Сколько блеска для губ продала Зарет в 2004 и 2005 г. (используйте данные из файла **Makeupdb.xlsx**)?
2. Создайте таблицу данных, содержащую общую выручку и количество проданных единиц продукции для каждого продавца.
3. Сколько блеска для губ продала Коллин не в регионе Запад?

4. На основе данных в файле **Makeupdb.xlsx** создайте таблицу данных, показывающую среднюю выручку за единицу продукции для каждого продавца для сделок, где средняя цена за единицу продукции превысила \$3,30.
5. На основе данных в файле **Sales.xlsx** определите следующее:
  - общий объем продаж в регионе Средний Запад;
  - общий объем продаж для Хизер в регионе Восток;
  - общий объем продаж для Хизер и для региона Восток;
  - общий объем продаж в регионе Восток для Хизер или для Джона;
  - число сделок в регионе Восток;
  - число сделок с продажами выше среднего;
  - общий объем продаж не в регионе Средний Запад.
6. Файл **Housepricedata.xlsx** содержит следующую информацию для ряда домов:
  - площадь в кв. футах;
  - цена;
  - количество ванных комнат;
  - количество спален.

На основе этой информации ответьте на следующие вопросы:

- Какова средняя цена для всех домов с общим количеством ванных комнат и спален не меньше шести?
  - Сколько домов с общим количеством спален и ванных комнат не более пяти продано более чем за \$300 000?
  - Сколько домов имеют как минимум три ванные комнаты, но общее количество спален и ванных комнат в них не превышает шести?
  - Какова максимальная цена дома площадью не более 3000 кв. футов и общим количеством спален и ванных комнат не более шести? (Подсказка: для решения задачи используйте функцию ДМАКС.)
7. Файл **Deciles.xlsx** содержит данные о невыплаченных остатках для 20 счетов. С помощью функций для баз данных вычислите невыплаченные остатки по счетам для каждого дециля.



## ГЛАВА 51

# Фильтрация данных и удаление дубликатов

### Обсуждаемые вопросы

Joolas — небольшая компания, производящая косметику. Менеджеры компании хранят данные обо всех сделках на листе Microsoft Excel. Иногда им необходимо извлечь или отфильтровать подмножество данных о продажах. Например, на основе этих данных требуется ответить на следующие вопросы:

- Как определить все сделки Джен по продаже помады в регионе Восток?
- Как определить все сделки Кики и Коллин по продаже помады и туши в регионе Восток или Юг?
- Как скопировать все сделки Кики и Коллин по продаже помады и туши в регионе Восток или Юг на другой лист?
- Как очистить фильтры в столбце или базе данных?
- Как определить все сделки по продаже более 90 единиц продукции и более чем на \$280?
- Как определить все сделки, заключенные в 2005 и 2006 г.?
- Как определить все сделки, заключенные в последние три месяца 2005 г. и в первые три месяца 2006 г.?
- Как определить все сделки, заключенные продавцами, имена которых начинаются с С?
- Как определить все сделки, для которых ячейка с названием продукта окрашена в красный цвет?
- Как определить все сделки среди 30 сделок с максимальной выручкой, в которых продавцами были Халлаган или Джен?
- Как быстро получить полный список продавцов?
- Как просмотреть все комбинации «продавец — продукт — регион», встречающиеся в базе данных?
- Если данные изменились, как повторно применить тот же самый фильтр?
- Как извлечь все сделки по продаже основы под макияж, заключенные Эмили и Джен за первые шесть месяцев 2005 г., в которых средняя цена единицы продукции составила больше \$3,20?



Microsoft Excel 2019 обладает возможностями фильтрации для определения любого подмножества данных. Кроме того, Excel позволяет быстро удалить дублирующиеся записи с листа. Данные для работы находятся в файле **Makeupfiltertemp.xlsx** в папке **Templates**. Для 1891 сделки в этом файле (часть сделок представлена на рис. 51.1) хранятся следующие данные:

- номер сделки;
- имя продавца;
- дата сделки;
- продаваемый продукт;
- количество единиц проданной продукции;
- сумма сделки в долларах;
- место сделки.

	C	D	E	F	G	H	I
3	Transaction number	Name	Date	Product	Units	Dollars	Location
4	1	Betsy	4/1/2004	lip gloss	45	\$ 137.20	south
5	2	Hallagan	3/10/2004	foundatio	50	\$ 152.01	midwest
6	3	Ashley	2/25/2005	lipstick	9	\$ 28.72	midwest
7	4	Hallagan	5/22/2006	lip gloss	55	\$ 167.08	west
8	5	Zaret	6/17/2004	lip gloss	43	\$ 130.60	midwest
9	6	Colleen	11/27/2005	eye liner	58	\$ 175.99	midwest
10	7	Cristina	3/21/2004	eye liner	8	\$ 25.80	midwest
11	8	Colleen	12/17/2006	lip gloss	72	\$ 217.84	midwest
12	9	Ashley	7/5/2006	eye liner	75	\$ 226.64	south
13	10	Betsy	8/7/2006	lip gloss	24	\$ 73.50	east
14	11	Ashley	11/29/2004	mascara	43	\$ 130.84	east
15	12	Ashley	11/18/2004	lip gloss	23	\$ 71.03	west
16	13	Emilee	8/31/2005	lip gloss	49	\$ 149.59	west
17	14	Hallagan	1/1/2005	eye liner	18	\$ 56.47	south
18	15	Zaret	9/20/2006	foundatio	-8	\$ (21.99)	east
19	16	Emilee	4/12/2004	mascara	45	\$ 137.39	east
20	17	Colleen	4/30/2006	mascara	66	\$ 199.65	south
21	18	Jen	8/31/2005	lip gloss	88	\$ 265.19	midwest
22	19	Jen	10/27/2004	eye liner	78	\$ 236.15	south

**Рис. 51.1.** Данные сделок по продаже косметики

Далее любой прямоугольный диапазон ячеек с заголовками в первой строке каждого столбца называется *базой данных*. Каждый столбец (с C по I) базы данных (диапазон ячеек C4:I1894) называется *полем*. Каждая строка базы данных, содержащая данные, называется *записью*. (Таким образом, записи базы данных хранятся в диапазоне C4:I1894.) Первая строка каждого поля должна содержать имя поля. Например, поле в столбце F имеет имя **Продукт**. Для идентификации подмножества записей можно обратиться с запросом к базе данных при помощи **автофильтра Excel (AutoFilter)** и **И-условий**. То есть вы можете использовать запросы

в форме «Найти все записи, в которых поле 1 удовлетворяет определенным условиям, и поле 2 удовлетворяет определенным условиям, и поле 3 удовлетворяет определенным условиям». Возможности автофильтра Excel демонстрируются на примерах в данной главе.

## Ответы на вопросы

### ❓ Как определить все сделки Джен по продаже помады в регионе Восток?

Для начала установите курсор в какую-нибудь ячейку базы данных и на вкладке Данные (Data) ленты Excel в группе Сортировка и фильтр (Sort & Filter) выберите Фильтр (Filter). Как показано на рис. 51.2, теперь у каждого столбца базы данных в строке заголовка появилась стрелка раскрывающегося списка (см. лист Data).

	C	D	E	F	G	H	I
3	Transaction number	Name	Date	Product	Units	Dollars	Location
4		1 Betsy	4/1/2004	lip gloss	45	\$ 137.20	south
5		2 Hallagan	3/10/2004	foundatior	50	\$ 152.01	midwest
6		3 Ashley	2/25/2005	lipstick	9	\$ 28.72	midwest
7		4 Hallagan	5/22/2006	lip gloss	55	\$ 167.08	west
8		5 Zaret	6/17/2004	lip gloss	43	\$ 130.60	midwest
9		6 Colleen	11/27/2005	eye liner	58	\$ 175.99	midwest
10		7 Cristina	3/21/2004	eye liner	8	\$ 25.80	midwest
11		8 Colleen	12/17/2006	lip gloss	72	\$ 217.84	midwest
12		9 Ashley	7/5/2006	eye liner	75	\$ 226.64	south
13		10 Betsy	8/7/2006	lip gloss	24	\$ 73.50	east
14		11 Ashley	11/29/2004	mascara	43	\$ 130.84	east
15		12 Ashley	11/18/2004	lip gloss	23	\$ 71.03	west
16		13 Emilee	8/31/2005	lip gloss	49	\$ 149.59	west
17		14 Hallagan	1/1/2005	eye liner	18	\$ 56.47	south
18		15 Zaret	9/20/2006	foundatior	-8	\$ (21.99)	east
19		16 Emilee	4/12/2004	mascara	45	\$ 137.39	east
20		17 Colleen	4/30/2006	mascara	66	\$ 199.65	south
21		18 Jen	8/31/2005	lip gloss	88	\$ 265.19	midwest
22		19 Jen	10/27/2004	eye liner	78	\$ 236.15	south

**Рис. 51.2.** Стрелки раскрывающегося списка фильтрации

Данные можно отфильтровать, используя раскрывающиеся списки.

Откройте список в столбце Имя (рис. 51.3). В нем можно было бы выбрать элемент Текстовые фильтры (Text Filters), позволяющий фильтровать данные на основе характеристик имен сотрудников (подробнее об этом далее). Но сейчас необходимо обработать данные для Джен (Jen), поэтому сначала снимите флажок (Выделить все) (Select All), установите флажок для Джен и нажмите кнопку ОК.

Теперь видны только те записи, где Джен была продавцом. Далее в столбце Продукт установите флажок для помады (lipstick). В столбце Регион установите флажок

для региона Восток (East). Теперь видны только сделки Джен по продаже помады в регионе Восток (см. рис. 51.4 и лист Lipstick Jen East в файле Makeupfilter.xlsx). Обратите внимание, что в столбцах, где были установлены условия фильтрации, стрелка заменилась на воронку.

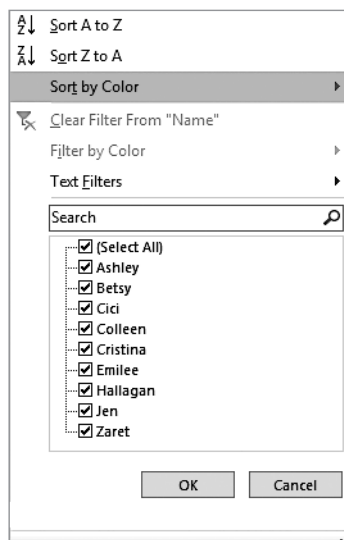


Рис. 51.3. Условия фильтрации для столбца Имя

	C	D	E	F	G	H	I
3	Transaction number	Name	Date	Product	Units	Dollars	Location
503	500	Jen	3/19/2005	lipstick	6	\$ 20.04	east
509	506	Jen	2/17/2004	lipstick	67	\$ 202.62	east
591	688	Jen	3/6/2006	lipstick	36	\$ 110.26	east
763	760	Jen	1/23/2005	lipstick	12	\$ 37.85	east
846	843	Jen	7/27/2006	lipstick	34	\$ 103.40	east
232	1229	Jen	9/24/2004	lipstick	92	\$ 277.63	east
781	1778	Jen	8/31/2005	lipstick	24	\$ 74.61	east
815	1812	Jen	5/22/2006	lipstick	67	\$ 203.18	east

Рис. 51.4. Сделки Джен по продаже помады в регионе Восток

### ❓ Как определить все сделки Кики и Коллин по продаже помады и туши в регионе Восток или Юг?

Просто выберите в столбце Имя Cici и Colleen, в столбце Продукт — lipstick и mascara и в столбце Регион — East и South. Записи, отвечающие этим условиям, показаны на рис. 51.5. (См. лист Cici Colleen Lipstick And Masc.)

	C	D	E	F	G	H	I
3	Transaction number	Name	Date	Product	Units	Dollars	Location
20	17	Colleen	4/30/2006	mascara	66	\$ 199.65	south
84	81	Cici	6/13/2006	lipstick	-3	\$ (7.62)	south
106	103	Cici	11/18/2004	mascara	38	\$ 115.86	south
139	136	Cici	9/24/2004	mascara	89	\$ 269.40	east
143	140	Cici	3/8/2005	lipstick	37	\$ 113.65	east
174	171	Colleen	5/26/2004	lipstick	60	\$ 182.02	east
191	188	Colleen	2/6/2004	mascara	-5	\$ (12.90)	east
201	198	Colleen	2/6/2004	mascara	60	\$ 181.94	east
237	234	Colleen	7/9/2004	lipstick	-3	\$ (7.40)	south
239	236	Cici	7/9/2004	mascara	5	\$ 17.20	east
244	241	Colleen	11/3/2006	mascara	64	\$ 194.25	east
253	250	Colleen	2/14/2005	lipstick	65	\$ 196.49	south
255	252	Cici	7/16/2006	mascara	-1	\$ (1.93)	south
317	314	Colleen	3/17/2006	lipstick	43	\$ 130.95	south
337	334	Colleen	2/14/2005	lipstick	15	\$ 46.64	east
339	336	Cici	7/18/2005	mascara	85	\$ 257.29	south
400	397	Colleen	2/17/2004	mascara	91	\$ 274.81	south
435	432	Cici	3/19/2005	mascara	27	\$ 83.00	south

**Рис. 51.5.** Сделки Кики и Коллин по продаже помады и туши в регионе Восток или Юг

**❓ Как скопировать все сделки Кики и Коллин по продаже помады и туши в регионе Восток или Юг на другой лист?**

Выделите все отфильтрованные ячейки на листе Cici Colleen lipstick and mascara и скопируйте их на пустой лист. (В новом листе в заголовках столбцов настройки фильтров исчезнут, и может потребоваться сделать шире столбец Е, содержащий даты.) Для создания пустого листа в книге щелкните правой кнопкой мыши по ярлычку любого листа, выберите **Вставить (Insert)** и нажмите **ОК** (или щелкните знак + справа от последнего листа). Лист **Visible Cells Copied** содержит записи о сделках Кики и Коллин по продаже помады и туши в восточном и южном регионах.

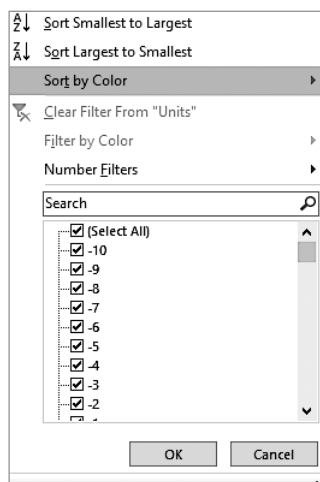
**❓ Как очистить фильтры в столбце или базе данных?**

Для удаления всех фильтров на вкладке **Данные (Data)** нажмите кнопку **Фильтр (Filter)**. Для удаления фильтра из столбца щелкните по значку воронки возле нужного столбца и выберите **Удалить фильтр из столбца** (при этом сбрасываются только результаты фильтрации, но не удаляется кнопка фильтра).

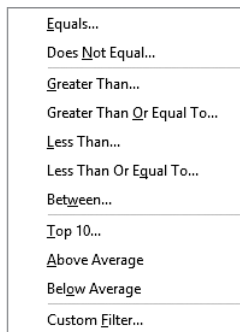
**❓ Как определить все сделки по продаже более 90 единиц продукции и более чем на \$280?**

Для включения фильтрации на вкладке **Данные (Data)** в группе **Сортировка и фильтр (Sort & Filter)** нажмите кнопку **Фильтр (Filter)**. Затем щелкните на стрелке раскрывающегося списка в столбце **Количество** (рис. 51.6).

В этом списке можно указать любое подмножество числовых значений (например, все сделки, в которых фигурировали –10 или –8 единиц продукции). Здесь выбран список **Числовые фильтры (Number Filters)**, который показан на рис. 51.7.

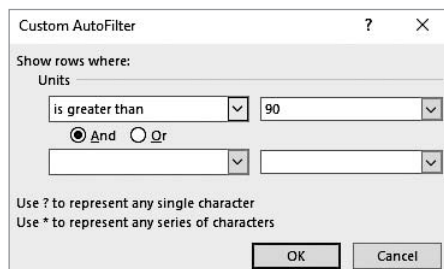


**Рис. 51.6.** Параметры фильтрации для столбца с числовыми значениями



**Рис. 51.7.** Список Числовые фильтры

Большая часть этого списка не требует разъяснения. Выберите **Больше (Greater Than)** и заполните диалоговое окно **Пользовательский автофильтр (Custom AutoFilter)**, как показано на рис. 51.8.



**Рис. 51.8.** Выбор всех записей с количеством проданных единиц продукции больше 90

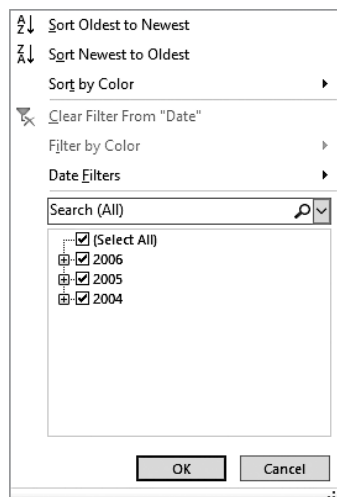
Затем в столбце **Сумма** аналогичным образом выберите только те записи, сумма в которых превышает \$280. Выберите **Числовые фильтры (Number Filters)**, больше (**Greater Than**), введите 280 и нажмите **OK**. На листе останутся записи, показанные на рис. 51.9. (См. лист **Sales>90 units dollars>\$280**.) Обратите внимание, что во всех записях проданных единиц продукции больше 90 и сумма превышает \$280.

	C	D	E	F	G	H	I
3	Transaction number	Name	Date	Product	Units	Dollars	Location
40	37	Ashley	8/9/2005	mascara	93	\$ 280.69	east
57	54	Cici	6/17/2004	eye liner	95	\$ 287.76	midwest
64	61	Emilee	12/21/2004	eye liner	95	\$ 287.05	midwest
65	62	Ashley	11/16/2005	lipstick	93	\$ 280.77	west
165	162	Betsy	10/14/2005	foundation	93	\$ 280.17	west
217	214	Betsy	9/13/2004	foundation	94	\$ 283.74	west
284	281	Emilee	11/16/2005	foundation	94	\$ 283.85	midwest
289	286	Betsy	7/27/2006	mascara	94	\$ 284.14	west
314	311	Jen	5/15/2004	eye liner	94	\$ 283.88	east
338	335	Ashley	5/26/2004	mascara	94	\$ 283.62	midwest
340	337	Colleen	11/18/2004	lip gloss	95	\$ 286.86	west
406	403	Jen	6/4/2005	mascara	93	\$ 281.17	south
450	447	Cici	10/1/2006	mascara	94	\$ 283.45	west
545	542	Ashley	4/19/2006	foundation	93	\$ 281.17	midwest
579	576	Emilee	4/10/2005	lipstick	93	\$ 280.80	midwest

**Рис. 51.9.** Сделки по продаже более 90 единиц продукции на сумму свыше \$280

### ❓ Как определить все сделки, заключенные в 2005 или 2006 г.?

После включения фильтрации кнопкой **Фильтр (Filter)** щелкните на стрелке раскрывающегося списка в столбце **Дата** для отображения списка, показанного на рис. 51.10.



**Рис. 51.10.** Параметры фильтрации для столбца **Дата**

После установки флажков для 2005 и 2006 г. останутся только записи, включающие продажи в 2005 и 2006 гг. (рис. 51.11, лист **Sales In 2005 And 2006** слева от предыдущих листов).

	C	D	E	F	G	H	I
3	Transaction number	Name	Date	Product	Units	Dollars	Location
6		3 Ashley	2/25/2005	lipstick	9	\$ 28.72	midwest
7		4 Hallagan	5/22/2006	lip gloss	55	\$ 167.08	west
9		6 Colleen	11/27/2005	eye liner	58	\$ 175.99	midwest
11		8 Colleen	12/17/2006	lip gloss	72	\$ 217.84	midwest
12		9 Ashley	7/5/2006	eye liner	75	\$ 226.64	south
13		10 Betsy	8/7/2006	lip gloss	24	\$ 73.50	east
16		13 Emilee	8/31/2005	lip gloss	49	\$ 149.59	west
17		14 Hallagan	1/1/2005	eye liner	18	\$ 56.47	south
18		15 Zaret	9/20/2006	foundatio	-8	\$ (21.99)	east
20		17 Colleen	4/30/2006	mascara	66	\$ 199.65	south
21		18 Jen	8/31/2005	lip gloss	88	\$ 265.19	midwest
23		20 Zaret	11/27/2005	lip gloss	57	\$ 173.12	midwest
24		21 Zaret	6/2/2006	mascara	12	\$ 38.08	west
26		23 Colleen	2/1/2006	mascara	25	\$ 77.31	midwest
27		24 Emilee	12/6/2006	lip gloss	24	\$ 74.62	west
29		26 Cristina	9/22/2005	foundatio	77	\$ 233.05	midwest
30		27 Cristina	3/28/2006	lip gloss	53	\$ 161.46	midwest
32		29 Zaret	9/9/2006	mascara	19	\$ 59.15	west

**Рис. 51.11.** Данные о продажах в 2005 и 2006 г.

Кроме того, в раскрывающемся списке можно выбрать элемент **Фильтры по дате (Date Filters)** и просмотреть список параметров (рис. 51.12).

Equals...
Before...
After...
Between...
Tomorrow
Today
Yesterday
Next Week
This Week
Last Week
Next Month
This Month
Last Month
Next Quarter
This Quarter
Last Quarter
Next Year
This Year
Last Year
Year to Date
All Dates in the Period ▶
Custom Filter...

**Рис. 51.12.** Параметры фильтрации по дате

Большинство этих параметров фильтрации также не требует пояснений. Параметр **Настраиваемый фильтр (Custom Filter)** позволяет выбрать в качестве условий фильтрации любой диапазон дат.

### ❓ Как определить все сделки, заключенные в последние три месяца 2005 г. и в первые три месяца 2006 г.?

Откройте в столбце **Дата** раскрывающийся список (см. рис. 51.10). Откройте список месяцев, щелкнув по знаку **+** слева от обозначения соответствующего года. Выберите месяцы с октября по декабрь 2005 г. и с января по март 2006 г. для отображения всех продаж за этот период. (См. рис. 51.13 и лист **Filter By Months**.)

	C	D	E	F	G	H	I
3	Transaction number	Name	Date	Product	Units	Dollars	Location
9	6	Colleen	11/27/2005	eye liner	58	\$ 175.99	midwest
23	20	Zaret	11/27/2005	lip gloss	57	\$ 173.12	midwest
26	23	Colleen	2/1/2006	mascara	25	\$ 77.31	midwest
30	27	Cristina	3/28/2006	lip gloss	53	\$ 161.46	midwest
33	30	Cici	2/23/2006	foundation	-9	\$ (24.63)	west
53	50	Cristina	12/8/2005	lip gloss	8	\$ 26.24	midwest
55	52	Colleen	11/16/2005	foundation	62	\$ 189.25	midwest
61	58	Hallagan	11/5/2005	foundation	63	\$ 191.37	south
65	62	Ashley	11/16/2005	lipstick	93	\$ 280.77	west
70	67	Colleen	3/6/2006	foundation	-2	\$ (3.94)	west
75	72	Jen	1/10/2006	lip gloss	69	\$ 208.69	east
82	79	Zaret	12/19/2005	eye liner	26	\$ 80.30	south
93	90	Colleen	1/21/2006	lip gloss	75	\$ 226.74	south
94	91	Betsy	10/3/2005	eye liner	74	\$ 224.23	west
98	95	Jen	3/17/2006	lipstick	-8	\$ (22.11)	west
109	106	Emilee	1/21/2006	foundation	39	\$ 118.63	south

**Рис. 51.13.** Продажи с октября по декабрь 2005 г. и с января по март 2006 г.

### ❓ Как определить все сделки, заключенные продавцами, имена которых начинаются с С?

Щелкните по стрелке раскрывающегося списка столбца **Имя** и выберите в нем элемент **Текстовые фильтры (Text Filters)**. В появившемся меню выберите **Начинается с**

Custom AutoFilter

Show rows where:

Name

begins with d

☒ And ☐ Or

Use ? to represent any single character  
Use \* to represent any series of characters

OK Cancel

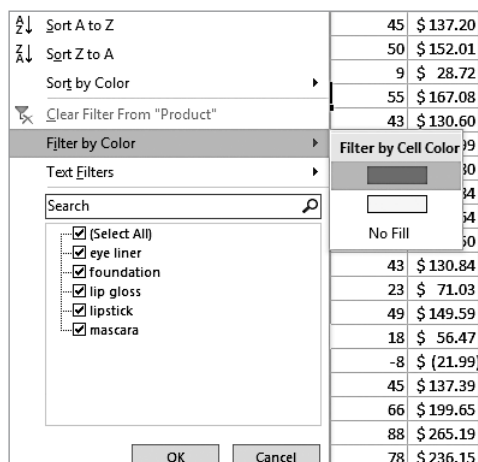
**Рис. 51.14.** Диалоговое окно Пользовательский автофильтр для выбора всех записей, в которых имя продавца начинается с С



(Begins With) и заполните диалоговое окно Пользовательский автофильтр (Custom AutoFilter), как показано на рис. 51.14.

### ? Как определить все сделки, для которых ячейка с названием продукта окрашена в красный цвет?

Щелкните по стрелке раскрывающегося списка столбца **Продукт** и выберите в нем элемент **Фильтр по цвету (Filter By Color)**. Затем выберите цвет, используемый как фильтр (рис. 51.15). На рис. 51.16 показаны только те записи, в которых ячейка с названием продукта окрашена в красный цвет (см. лист **Filter By Color**).



**Рис. 51.15.** Диалоговое окно для фильтрации по цвету ячейки

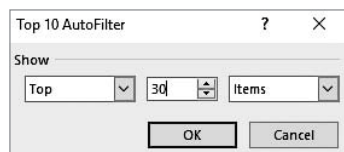
	C	D	E	F	G	H	I
3	Transaction number	Name	Date	Product	Units	Dollars	Location
11	8	Colleen	12/17/2006	lip gloss	72	\$ 217.84	midwest
247	244	Ashley	9/24/2004	foundation	84	\$ 253.29	east
292	289	Zaret	5/26/2004	lipstick	56	\$ 169.19	south

**Рис. 51.16.** Все записи, в которых ячейка с названием продукта окрашена в красный цвет

### ? Как определить все сделки среди 30 сделок с максимальной выручкой, в которых продавцами были Халлаган или Джен?

Нажмите кнопку **Фильтр (Filter)**, щелкните по стрелке раскрывающегося списка в столбце **Имя** и установите флажки для Халлаган (Hallagan) и Джен (Jen). Щелкните по стрелке раскрывающегося списка в столбце **Сумма** и выберите **Числовые фильтры (Number Filters)**. В списке выберите **Первые 10 (Top 10)** и заполните диалоговое окно (рис. 51.17), заменив значение 10 на 30. В результате все записи будут отфильтрованы и останутся только 30 записей для сделок с наибольшими

значениями выручки, в которых продавцами были Джен или Халлаган. (См. результаты на рис. 51.18 и на листе **Top 30 \$s With Hallagan Or Jen**). Следует отметить, что из первых по сумме 30 сделок только в пяти продавцами были Халлаган и Джен. Аналогичным образом можно выбрать показ наибольших 5% от количества элементов, наименьших 20% от количества элементов и т. д. для любых числовых столбцов.



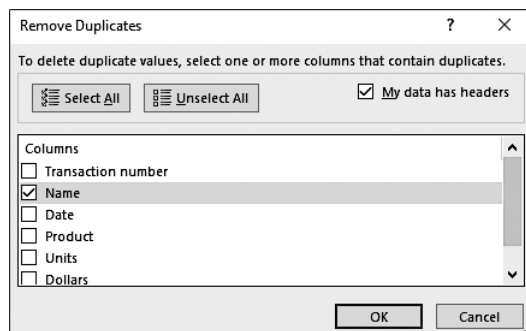
**Рис. 51.17.** Диалоговое окно для выбора 30 первых записей по значению суммы

	D	E	F	G	H	I
3	Name	Date	Product	Units	Dollars	Location
314	Jen	5/15/2004	eye liner	94	\$ 283.88	east
722	Hallagan	6/13/2006	lip gloss	95	\$ 286.68	midwest
797	Hallagan	8/18/2006	foundatio	95	\$ 287.80	midwest
1259	Hallagan	4/8/2006	lip gloss	95	\$ 286.76	midwest
1619	Hallagan	7/29/2005	eye liner	95	\$ 287.15	west

**Рис. 51.18.** Тридцать первых сделок по значению суммы, в которых продавцами были Халлаган или Джен

### ❓ Как быстро получить полный список продавцов?

Мы хотим получить список всех продавцов без повторения имен. На вкладке **Данные (Data)** в группе **Работа с данными** выберите **Удалить дубликаты (Remove Duplicates)**. Появится диалоговое окно **Удалить дубликаты (Remove Duplicates)**, показанное на рис. 51.19. Сначала нажмите кнопку **Снять выделение (Unselect All)**, затем установите флажок только для столбца **Имя** и нажмите **ОК**. После фильтрации останутся только первые записи, включающие имя каждого продавца. См. результаты на рис. 51.20 и на листе **Name Duplicates Removed**.



**Рис. 51.19.** Диалоговое окно Удалить дубликаты для столбца Имя

### ВАЖНО!

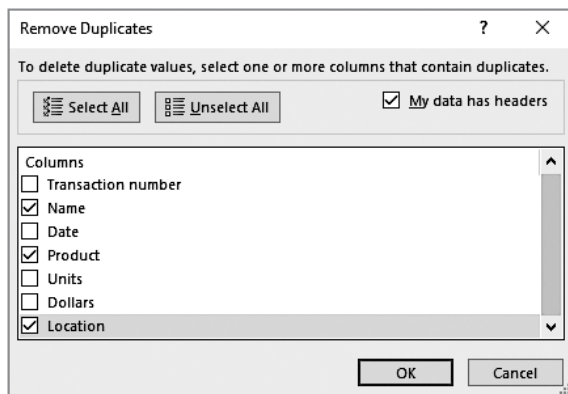
Поскольку при удалении дубликатов удаляется часть данных, рекомендуется перед этим сделать копию данных.

	C	D	E	F	G	H	I
3	Transaction number	Name	Date	Product	Units	Dollars	Location
4		1 Betsy	4/1/2004	lip gloss	45	\$ 137.20	south
5		2 Hallagan	3/10/2004	foundatio	50	\$ 152.01	midwest
6		3 Ashley	2/25/2005	lipstick	9	\$ 28.72	midwest
7		5 Zaret	6/17/2004	lip gloss	43	\$ 130.60	midwest
8		6 Colleen	11/27/2005	eye liner	58	\$ 175.99	midwest
9		7 Cristina	3/21/2004	eye liner	8	\$ 25.80	midwest
10		13 Emilee	8/31/2005	lip gloss	49	\$ 149.59	west
11		18 Jen	8/31/2005	lip gloss	88	\$ 265.19	midwest
12		28 Cici	6/17/2004	mascara	41	\$ 125.27	west

**Рис. 51.20.** Список имен продавцов

❓ Как просмотреть все комбинации «продавец — продукт — регион», встречающиеся в базе данных?

Здесь снова требуется начать с удаления дубликатов на вкладке Данные (Data). Заполните диалоговое окно Удалить дубликаты (Remove Duplicates), как показано на рис. 51.21 (нажмите кнопку Снять выделение (Unselect All), установите флажки для столбцов Имя, Продукт и Регион и нажмите ОК).



**Рис. 51.21.** Поиск уникальных комбинаций «продавец — продукт — регион»

На рис. 51.22 представлены первые записи для каждой комбинации «продавец — продукт — регион», встречающейся в базе данных. (См. лист Unique Name Product Location.) Всего получается 180 уникальных комбинаций. Как видите, двадцатая запись отсутствует, поскольку она относится к сделке Запет (Zaret) с блеском для губ (lip gloss) в регионе Средний Запад (Midwest), а пятая запись уже включает такую комбинацию.

	C	D	E	F	G	H	I
3	Transaction number	Name	Date	Product	Units	Dollars	Location
4	1	Betsy	4/1/2004	lip gloss	45	\$ 137.20	south
5	2	Hallagan	3/10/2004	foundation	50	\$ 152.01	midwest
6	3	Ashley	2/25/2005	lipstick	9	\$ 28.72	midwest
7	4	Hallagan	5/22/2006	lip gloss	55	\$ 167.08	west
8	5	Zaret	6/17/2004	lip gloss	43	\$ 130.60	midwest
9	6	Colleen	11/27/2005	eye liner	58	\$ 175.99	midwest
10	7	Cristina	3/21/2004	eye liner	8	\$ 25.80	midwest
11	8	Colleen	12/17/2006	lip gloss	72	\$ 217.84	midwest
12	9	Ashley	7/5/2006	eye liner	75	\$ 226.64	south
13	10	Betsy	8/7/2006	lip gloss	24	\$ 73.50	east
14	11	Ashley	11/29/2004	mascara	43	\$ 130.84	east
15	12	Ashley	11/18/2004	lip gloss	23	\$ 71.03	west
16	13	Emilee	8/31/2005	lip gloss	49	\$ 149.59	west
17	14	Hallagan	1/1/2005	eye liner	18	\$ 56.47	south
18	15	Zaret	9/20/2006	foundation	-8	\$ (21.99)	east
19	16	Emilee	4/12/2004	mascara	45	\$ 137.39	east
20	17	Colleen	4/30/2006	mascara	66	\$ 199.65	south
21	18	Jen	8/31/2005	lip gloss	88	\$ 265.19	midwest
22	19	Jen	10/27/2004	eye liner	78	\$ 236.15	south
23	21	Zaret	6/2/2006	mascara	12	\$ 38.08	west

**Рис. 51.22.** Список уникальных комбинаций «продавец — продукт — регион»

### ❓ Если данные изменились, как повторно применить тот же самый фильтр?

Щелкните правой кнопкой мыши на любой ячейке в результатах фильтрования и в группе Сортировка и фильтр (Sort & Filter) укажите на Фильтр (Filter) и щелкните по значку Повторить (Reapply). В отфильтрованных данных отразятся любые изменения данных.

### ❓ Как извлечь все сделки по продаже основы под макияж, заключенные Эмили и Джен за первые шесть месяцев 2005 г., в которых средняя цена единицы продукции составила больше \$3,20?

Автофильтр (даже настраиваемый) ограничен И-запросами по столбцам. Это означает, например, что поиск всех сделок Джен с помадой за 2005 г. и сделок Зарет с основой под макияж за 2004 г. невозможен. Для выполнения более сложных запросов, подобных этому, необходимо использовать Расширенный фильтр. Для него задают диапазон условий, определяющий записи, которые требуется извлечь. (Этот процесс подробно описан в главе 50 «Обработка данных с помощью статистических функций для баз данных».) После этого указывают диапазон, в который будут помещены извлеченные записи. Для определения всех сделок с основой под макияж (foundation) за первые шесть месяцев 2005 г., в которых продавцами были Эмили и Джен и в которых средняя цена единицы продукции составила больше \$3,20, можно задать диапазон условий, показанный в ячейках O4:S6 на рис. 51.23. (См. лист Jen+Emilee в файле Advancedfilter.xlsx.)

В ячейки R5 и R6 я ввел формулу  $= (L5/K5) > 3,2$ . Вспомните из главы 50, что такая формула является вычисляемым условием, по которому помечается каждая

	O	P	Q	R	S
4	Name	Date	Date	Price	Product
5	Jen	>=1/1/2005	<=6/30/2005	FALSE	Foundation
6	Emilee	>=1/1/2005	<=6/30/2005	FALSE	Foundation

**Рис. 51.23.** Диапазон условий для расширенного фильтра

строка со средней ценой за единицу продукции больше \$3,20. Также напомним, что заголовок для вычисляемого условия не должен быть именем поля. Я указал заголовок **Цена**. Согласно условиям, в диапазоне O5:S5 помечаются все записи для сделок, в которых продавцом является Джен, дата приходится на период между 1/1/2005 и 30/6/2005, проданным продуктом является основа под макияж, а средняя цена за единицу продукции составляет больше \$3,20. Таким образом, помечаются именно те записи, которые требуются. Напомним, что условия в разных строках объединяются оператором **ИЛИ**.

Теперь выберите любую ячейку в базе данных и на вкладке **Данные (Data)** в группе **Сортировка и фильтр (Sort & Filter)** на вкладке **Данные (Data)** щелкните по значку **Дополнительно (Advanced)**, значок воронки с карандашиком). Заполните диалоговое окно **Расширенный фильтр (Advanced Filter)** и нажмите **OK** — рис. 51.24.

Advanced Filter

Action

☒ Filter the list, in-place

☐ Copy to another location

List range: \$G\$4:\$M\$1895

Criteria range: \$O\$4:\$S\$6

Copy to: \$O\$14

☐ Unique records only

OK Cancel

**Рис. 51.24.** Диалоговое окно Расширенный фильтр

Согласно этим данным, из базы данных (в диапазоне G4:M1895) должны быть извлечены все записи, которые удовлетворяют условиям в O4:S6. Эти записи должны быть скопированы в диапазон, левым верхним углом которого является ячейка O14. Извлеченные записи представлены в диапазоне ячеек O14:U18. Указанным условиям соответствуют только четыре записи (см. рис. 51.25 и лист Jen+Emilee.)

	O	P	Q	R	S	T	U
14	Trans Number	Name	Date	Product	Units	Dollars	Location
15	392	Jen	2/25/2005	foundation	8	\$ 26.31	south
16	479	Emilee	5/24/2005	foundation	2	\$ 7.68	east
17	1035	Emilee	4/10/2005	foundation	8	\$ 26.40	east
18	1067	Jen	3/19/2005	foundation	1	\$ 4.86	east

**Рис. 51.25.** Результаты работы расширенного фильтра

Если в диалоговом окне **Расширенный фильтр (Advanced Filter)** установить флажок **Только уникальные записи (Unique records only)**, дублирующиеся записи не будут извлечены. Например, если бы Джен провела еще одну сделку с основой под макияж в регионе Восток 19 марта 2005 г. для одной единицы продукции на сумму \$4,88, была бы извлечена запись только по одной сделке.

## Задания

1. Используя файл **Makeupfiltertemp.xlsx** (для этого и всех следующих заданий), найдите все сделки **Халлаган (Hallagan)** по продаже подводки для глаз в регионе Запад.
2. Найдите все сделки, входящие в 5% наиболее крупных сделок по количеству проданных единиц продукции.
3. Найдите первые 20 наиболее доходных сделок, включающих продажи основы под макияж (**foundation**).
4. Найдите все сделки за 2004 г. по продаже не менее 60 единиц продукции, для которых цена за единицу продукции составила максимум \$3,10.
5. Найдите все сделки с основой под макияж (**foundation**) за первые три месяца 2004 г., в которых цена за единицу продукции была выше средней цены, полученной за основу под макияж, за весь период.
6. Найдите все сделки **Зарет (Zaret)** или **Бетси (Betsy)** по продаже помады (**lipstick**) или основы под макияж.
7. Найдите все уникальные комбинации «продукт — продавец».
8. Найдите 30 самых крупных сделок (по количеству единиц продукции) по продаже блеска для губ (**lip gloss**) или туши (**mascara**) в 2005 г.
9. Найдите все сделки **Джен (Jen)** с 10 августа по 15 сентября 2005 г.
10. Найдите все сделки **Коллин (Colleen)** по продаже помады (**lipstick**), в которых количество проданных единиц продукции больше, чем среднее количество единиц продукции в сделках по продаже помады.
11. Найдите все уникальные комбинации «имя — продукт — регион», встречающиеся в сделках за первые три месяца 2006 г.
12. Найдите все записи, в которых ячейка с названием продукта окрашена в желтый цвет.

## ГЛАВА 52

# Консолидация данных

### Обсуждаемый вопрос

- Компания продает продукцию в нескольких регионах США. В каждом регионе хранятся записи о количестве проданных единиц продукции в январе, феврале и марте. Существует ли простой способ создания мастер-книги, в которой постоянно объединяются продажи для всех регионов и проводится подсчет общей суммы выручки для каждого продукта, проданного в США в течение каждого месяца?

Бизнес-аналитики часто получают листы с одними и теми же обобщенными данными (таким как ежемесячные продажи продукции) из различных филиалов или регионов. Аналитику, как правило, приходится объединять или консолидировать эту информацию в одной книге Microsoft Excel для определения общей прибыльности компании. Для этой цели можно создать из нескольких диапазонов консолидации сводные таблицы, но существует другой, редко используемый способ — инструмент Консолидация (Consolidate) на вкладке Данные (Data) в группе Работа с данными (Data Tools). С его помощью можно обеспечить автоматическое отражение на объединенном листе изменений, произошедших на отдельных листах.

## Ответ на вопрос

**?** Компания продает продукцию в нескольких регионах США. В каждом регионе хранятся записи о количестве проданных единиц продукции в январе, феврале и марте. Существует ли простой способ создания мастер-книги, в которой постоянно объединяются продажи для всех регионов и проводится подсчет общей суммы выручки для каждого продукта, проданного в США в течение каждого месяца?

Файл East.xlsx (рис. 52.1) содержит данные о ежемесячных продажах продуктов А–Н в восточных штатах США за январь, февраль и март. Аналогично файл West.xlsx (рис. 52.2) содержит данные о ежемесячных продажах продуктов А–Н в западных штатах США за январь, февраль и март. Необходимо создать объединенный лист, в котором были бы сведены в таблицу общие объемы продаж для каждого продукта за месяц.

	A	B	C	D	E	F
1	Product	January	February	March		East Sales
2	A	205	263	20		
3	B	164	17	146		
4	C	278	177	179		
5	D	156	214	240		
6	D	72	134	48		
7	D	7	256	104		
8	A	141	87	148		
9	A	2	15	135		
10	A	44	47	72		
11	B	7	81	2		
12	E	25	120	171		
13	E	197	90	124		
14	E	221	121	48		
15	A	84	103	134		
16	G	13	250	51		
17	D	5	159	70		
18	E	136	152	28		

**Рис. 52.1.** Продажи в регионе Восток с января по март

	A	B	C	D
1	Product	January	February	March
2	A	173	1	256
3	A	208	201	224
4	B	176	33	350
5	B	190	249	215
6	D	162	74	156
7	D	90	150	170
8	D	112	284	141
9	G	154	217	113
10	G	152	200	275
11	G	277	183	372
12	H	131	71	266
13	F	294	211	249
14	F	146	125	5
15	A	115	214	141
16	F	157	241	73
17	A	125	227	135
18	A	314	189	180
19	C	189	154	101
20	C	313	182	68
21	C	389	247	257
22	B	353	151	99
23	C	62	162	238
24	D	173	153	270

**Рис. 52.2.** Продажи в регионе Запад с января по март

Перед применением инструмента Консолидация (Consolidate) полезно просмотреть оба листа вместе на экране. Для этого откройте обе книги и на вкладке Вид (View) в группе Окно (Window) щелкните по значку Рядом (View Side By Side). Экран должен выглядеть, как на рис. 52.3. Теперь откройте пустой лист, щелкните на Упорядочить все (Arrange All) и выберите Рядом (Tiled).

На пустом листе на вкладке Данные (Data) в группе Работа с данными (Data Tools) выберите инструмент Консолидация (Consolidate). Откроется диалоговое окно Консолидация (Consolidate), показанное на рис. 52.4.

Для объединения данных из регионов Восток и Запад на новом пустом листе в диалоговом окне Консолидация (Consolidate) в поле Ссылка (Reference) по очереди выберите диапазоны, которые требуется объединить. После выбора каждого диапазона нажмите Добавить (Add). Флажки Подписи верхней строки (Top row) и Значения левого столбца (Left column), установленные в разделе Использовать в качестве имени (Use labels in), позволяют убедиться в том, что объединяются выделенные диапазоны путем просмотра меток в верхней строке и левом столбце каждого диапазона. Флажок Создавать связи с исходными данными (Create links to source data) обеспечивает отображение изменений в выделенных диапазонах на объединенном листе. В раскрывающемся списке Функция (Function) выберите Сумма (Sum), поскольку требуется сложить объемы продаж каждого продукта за месяц. С помощью функции Количество (Count) можно посчитать количество сделок для каждого продукта в каждом месяце, а с помощью функции Максимум (Max) определить самую крупную сделку для каждого продукта в каждом месяце. Диалоговое окно Консолидация (Consolidate) должно быть заполнено так, как на рис. 52.5.



	A	B	C	D	E
1	Product	January	February	March	
2	A	173	1	256	
3	A	208	201	224	
4	B	176	33	350	
5	B	190	249	215	
6	D	162	74	156	
7	D	90	150	170	
8	D	112	284	141	
9	G	154	217	113	
10	G	152	200	275	
11	G	277	183	372	
12	H	131	71	266	
13	F	294	211	249	
14	F	146	125	5	
15	A	115	214	141	
16	F	157	241	73	
17	A	125	227	135	
18	A	314	189	180	
19	C	189	154	101	
20	C	313	182	68	
21	C	389	247	257	
22	B	353	151	99	
23	C	62	162	238	
24	D	173	153	270	

	A	B	C	D	E	V
1	Product	January	February	March		
2	A	173	1	256		
3	A	208	201	224		
4	B	176	33	350		
5	B	190	249	215		
6	D	162	74	156		
7	D	90	150	170		
8	D	112	284	141		
9	G	154	217	113		
10	G	152	200	275		
11	G	277	183	372		
12	H	131	71	266		
13	F	294	211	249		
14	F	146	125	5		
15	A	115	214	141		
16	F	157	241	73		
17	A	125	227	135		
18	A	314	189	180		
19	C	189	154	101		
20	C	313	182	68		
21	C	389	247	257		
22	B	353	151	99		
23	C	62	162	238		
24	D	173	153	270		

Рис. 52.3. Продажи в регионах Восток и Запад на одном экране

Consolidate

Function: Sum

Reference:

All references:

Use labels in

☒ Top row

☒ Left column

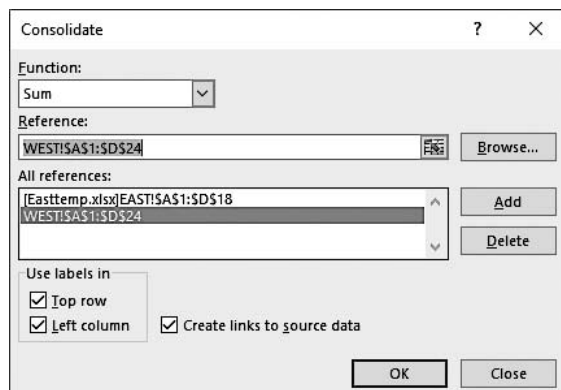
☐ Create links to source data

OK Close

Рис. 52.4. Диалоговое окно Консолидация

Когда вы нажмете ОК, новый лист будет выглядеть так, как на рис. 52.6. (См. файл Eastandwestconsolidated.xlsx.) Видно, что в феврале было продано 1317 единиц продукта А, в январе — 597 единиц продукта F и т. д.

Теперь в ячейке C2 файла East.xlsx измените объем продаж продукта А в феврале с 263 на 363. Обратите внимание, что на объединенном листе значение объема



**Рис. 52.5.** Диалоговое окно Консолидация, заполненное данными

1	2		A	B	C	D	E	F
		3				January	February	March
+		5		H		131	71	266
+		9		F		597	577	327
+		20		A		1323	1317	1445
+		26		B		890	335	812
+		32		C		1231	922	843
+		41		D		767	1424	1199
+		46		E		579	483	371
+		51		G		570	850	811

**Рис. 52.6.** Данные об общих объемах продаж после консолидации

продаж продукта **A** в феврале также увеличилось на 100 единиц (с 1317 до 1417). Это изменение произошло из-за того, что мы установили флажок **Создавать связи с исходными данными (Create links to source data)** в диалоговом окне **Консолидация (Consolidate)**. Кстати, если щелкнуть по цифре 2 в левом верхнем углу объединенного листа, можно увидеть, как в Excel для выполнения консолидации были сгруппированы данные. Результат содержится в файле **Eastandwestconsolidated.xlsx**. Разумеется, вы можете объединить данные с разных листов и при помощи сводных таблиц с несколькими диапазонами консолидации (см. главу 45 «Сводные таблицы и срезы для описания данных») или при помощи модели данных (глава 46 «Модель данных»).

Если вам приходится часто загружать в исходную книгу (в данном примере **East.xlsx** и **West.xlsx**) новые данные, рекомендуется диапазонам с данными присвоить имена, как таблицам. Тогда новые данные будут автоматически включаться в консолидацию. Или можно под текущим набором данных выделить несколько пустых строк. После заполнения пустых строк новыми данными они будут учтены при консолидации. Третий вариант — преобразовать диапазоны данных в динамические диапазоны (см. главу 22 «Функция СМЕЩ»).

## Задания

Эти задания по файлам `Jancon.xlsx` и `Febcon.xlsx`. Каждый файл содержит проданный продукт, показатели продаж в единицах продукции и выручку в долларах для сделок за месяц.

1. Создайте объединенный лист, на котором определен общий показатель продаж в единицах продукции и общая выручка в долларах для каждого продукта по регионам.
2. Создайте объединенный лист, на котором за первый квартал по регионам для каждого продукта определяется самая крупная сделка с точки зрения выручки и количества проданных единиц продукции.

## ГЛАВА 53

# Создание промежуточных итогов

### Обсуждаемые вопросы

- Существует ли быстрый способ настроить лист для вычисления общей выручки и количества проданных единиц продукции по регионам?
- Как можно при этом получить разбивку продаж по продавцам в каждом регионе?

Joolas — небольшая компания по производству косметики. Для каждой сделки она отслеживает имя продавца, место совершения сделки, проданный продукт, количество проданных единиц товара и выручку. Руководству компании необходимо получить ответы на вопросы, приведенные выше.

В Microsoft Excel для получения срезов данных можно использовать сводные таблицы. Однако зачастую требуется просто обработать список или базу данных в виде списка. Например, в базе данных по продажам необходимо суммировать выручку по регионам, по продуктам и по продавцам. Если отсортировать список по столбцу с нужными данными, с помощью инструмента **Промежуточный итог (Subtotal)** можно создать в списке подытог. Например, если базу данных косметических товаров отсортировать по регионам, то можно вычислить общую выручку и количество проданных единиц продукции для каждого региона и поместить промежуточные данные непосредственно под последней строкой соответствующего региона. Или еще один пример. После сортировки базы данных по продукту можно с помощью инструмента **Промежуточный итог (Subtotal)** вычислить общую выручку и количество проданных единиц продукции для каждого продукта и вывести промежуточные данные под строкой, в которой изменяется продукт. Переходим к рассмотрению конкретных примеров.

## Ответы на вопросы

- ❓ **Существует ли быстрый способ настроить лист для вычисления общей выручки и количества проданных единиц продукции по регионам?**

Данные к этому примеру находятся в файле `Makeupsubtotals.xlsx`. На рис. 53.1 показана часть данных после сортировки списка по столбцу **Регион (Location)**.

Для вычисления выручки и количества проданных единиц продукции по регионам установите курсор в пределах базы данных и затем на вкладке **Данные (Data)**

в группе Структура (Outline) выберите инструмент Промежуточный итог (Subtotal). Диалоговое окно Промежуточные итоги (Subtotal) заполните, как на рис. 53.2.

	A	B	C	D	E	F	G
4	Trans Number	Name	Date	Product	Units	Dollars	Location
5	10	Betsy	8/7/2006	lip gloss	24	\$ 73.50	east
6	11	Ashley	11/29/2004	mascara	43	\$ 130.84	east
7	15	Zaret	9/20/2006	foundatio	-8	\$ (21.99)	east
8	16	Emilee	4/12/2004	mascara	45	\$ 137.39	east
9	45	Emilee	9/20/2006	lip gloss	2	\$ 7.85	east
10	46	Ashley	8/9/2005	mascara	93	\$ 280.69	east
11	58	Cristina	4/12/2004	foundatio	34	\$ 104.09	east
12	60	Jen	10/27/2004	mascara	89	\$ 269.09	east
13	69	Cristina	1/23/2005	eye liner	73	\$ 221.41	east
14	77	Cristina	1/15/2004	mascara	27	\$ 83.29	east
15	81	Jen	1/10/2006	lip gloss	69	\$ 208.69	east
16	86	Jen	8/9/2005	eye liner	-2	\$ (4.24)	east
17	87	Emilee	8/31/2005	eye liner	5	\$ 17.03	east
18	98	Jen	4/12/2004	lip gloss	92	\$ 277.54	east
19	108	Cici	8/31/2005	lip gloss	-10	\$ (28.41)	east
20	114	Betsy	9/22/2005	lipstick	77	\$ 233.33	east
21	116	Zaret	6/24/2006	eye liner	22	\$ 68.07	east
22	119	Colleen	5/22/2006	eye liner	20	\$ 62.37	east
23	131	Jen	11/29/2004	mascara	56	\$ 168.87	east
24	133	Betsy	7/9/2004	mascara	11	\$ 34.42	east

**Рис. 53.1.** После сортировки списка по значениям в конкретном столбце можно быстро создать промежуточные итоги для соответствующих данных

**Рис. 53.2.** Диалоговое окно Промежуточные итоги

В списке При каждом изменении в (At each change in) выберите столбец Регион для создания промежуточных итогов в каждой точке, в которой значение в этом столбце изменяется, то есть происходит переход к другому региону. В списке Операция (Use function) выберите Сумма (Sum) для итогов по каждому региону на основе любого набора столбцов. Выберите в списке Добавить итоги по (Add subtotal to) столб-

цы Количество и Сумма для промежуточных итогов по этим столбцам. Флажок Заменить текущие итоги (Replace current subtotals) позволяет Excel удалять любые ранее вычисленные промежуточные итоги. Поскольку таковые в данном примере еще не созданы, для нас флажок не важен. Если установлен флажок Конец страницы между группами (Page break between groups), каждый раз после промежуточных итогов автоматически будет вставлен разрыв страницы. При включенном флажке Итоги под данными (Summary below data) промежуточные итоги располагаются под данными. Если этот флажок не активен, промежуточные итоги создаются над теми данными, которые использовались для вычислений. Кнопка Убрать все (Remove All) удаляет все промежуточные итоги из списка.

На рис. 53.3 приведен пример промежуточных итогов. Как видите, в регионе Восток (East) было продано 18 818 единиц продукции на сумму \$57 372,09.

1	2	3		A	B	C	D	E	F	G
	•	446		1849	Zaret	7/29/2005	lip gloss	12	\$ 37.85	east
	•	447		1853	Cici	4/23/2004	lip gloss	87	\$ 262.81	east
	•	448		1855	Emilee	7/18/2005	lip gloss	50	\$ 152.13	east
	•	449		1860	Emilee	8/11/2004	lip gloss	-9	\$ (24.76)	east
	•	450		1863	Zaret	3/6/2006	eye liner	73	\$ 220.67	east
	•	451		1867	Cristina	8/18/2006	mascara	73	\$ 220.77	east
	•	452		1869	Betsy	10/3/2005	lipstick	18	\$ 56.08	east
	•	453		1877	Cici	9/13/2004	eye liner	66	\$ 199.36	east
	•	454		1881	Emilee	10/3/2005	foundatio	0	\$ 2.66	east
	•	455		1883	Betsy	7/20/2004	lip gloss	-6	\$ (15.74)	east
	•	456		1890	Cici	6/15/2005	foundatio	16	\$ 49.75	east
	•	457		1891	Betsy	4/10/2005	foundatio	39	\$ 119.19	east
	•	458		1894	Colleen	5/15/2004	lip gloss	60	\$ 181.87	east
	•	459		1895	Emilee	11/27/2005	eye liner	15	\$ 47.16	east
	•	460		1896	Ashley	2/14/2005	foundatio	36	\$ 109.84	east
		461						18818	\$ 57,372.09	east Total

**Рис. 53.3.** Промежуточные итоги для региона Восток

Обратите внимание, что в левом верхнем углу окна (см. рис. 53.3) появились кнопки с числами 1, 2 и 3. Если нажать кнопку с наибольшим числом (в данном случае 3), то на листе останутся и данные, и итоги. Если нажать кнопку 2, то появятся промежуточные итоги по регионам, как на рис. 53.4. Если нажать кнопку 1, то появится общий итог, как на рис. 53.5. Таким образом, меньшее число соответствует меньшему уровню детализации.

	E	F	G
4	Units	Dollars	Location
461	18818	\$ 57,372.09	east Total
886	17985	\$ 54,805.41	midwest Total
1408	21083	\$ 64,296.35	south Total
1899	20821	\$ 63,438.82	west Total
1900	78707	\$ 239,912.67	Grand Total

**Рис. 53.4.** При создании промежуточных итогов добавляются кнопки для отображения различных уровней детализации

	E	F	G
4	Units	Dollars	Location
1900	78707	\$ 239,912.67	Grand Total

**Рис. 53.5.** Общий итог без детализации

### ❓ Как можно при этом получить разбивку продаж по продавцам в каждом регионе?

Если хотите, вы можете представить итоговые суммы во вложенном виде. Другими словами, можно получить разбивку продаж по каждому продавцу в каждом регионе или даже разбивку по количеству проданных продуктов для каждого продавца в каждом регионе. (См. файл `Nestedsubtotals.xlsx`.) Для демонстрации вложенных промежуточных итогов создадим разбивку продаж по каждому продавцу в каждом регионе.

Прежде всего, надо отсортировать данные по столбцу **Регион**, затем по столбцу **Имя**. Получится разбивка для каждого продавца по количеству проданных единиц товара и по выручке в каждом регионе. Если отсортировать сначала по **Имя**, а затем по **Регион**, получится разбивка по количеству проданных единиц товара и по выручке для каждого продавца по регионам. После сортировки данных выполните действия, аналогичные описанным ранее, и создайте промежуточные итоги по регионам. Снова выберите инструмент **Промежуточный итог (Subtotal)** и заполните диалоговое окно, как на рис. 53.6.

Subtotal ? X

At each change in:  
Name

Use function:  
Sum

Add subtotal to:  
☐ Name  
☐ Date  
☐ Product  
☒ Units  
☒ Dollars  
☐ Location

☒ Replace current subtotals  
☐ Page break between groups  
☒ Summary below data

Remove All OK Cancel

**Рис. 53.6.** Создание вложенных промежуточных итогов

Теперь нам нужна разбивка по столбцу **Имя**. Снимите флажок **Заменить текущие итоги (Replace current subtotals)** для гарантии того, что разбивка по регионам не будет замещена. Как видно из рис. 53.7, теперь появилась разбивка продаж по каждому продавцу в каждом регионе.

	A	B	C	D	E	F	G
4	Trans Number	Name	Date	Product	Units	Dollars	Location
56		Ashley Total			2558	\$ 7,772.70	
122		Betsy Total			2879	\$ 8,767.43	
173		Cici Total			1951	\$ 5,956.32	
220		Colleen Total			1874	\$ 5,713.07	
261		Cristina Total			1348	\$ 4,126.27	
312		Emilee Total			2064	\$ 6,295.47	
355		Hallagan Total			1626	\$ 4,965.62	
409		Jen Total			2282	\$ 6,949.21	
469		Zaret Total			2236	\$ 6,826.00	
470					18818	\$ 57,372.09	east Total
511		Ashley Total			1635	\$ 4,985.90	
555		Betsy Total			1598	\$ 4,878.09	
616		Cici Total			2671	\$ 8,129.62	
670		Colleen Total			2159	\$ 6,586.14	
721		Cristina Total			1923	\$ 5,870.03	
766		Emilee Total			1852	\$ 5,642.20	
809		Hallagan Total			2431	\$ 7,378.32	
862		Jen Total			2092	\$ 6,381.32	
903		Zaret Total			1624	\$ 4,953.80	
904					17985	\$ 54,805.41	midwest Total

**Рис. 53.7.** Вложенные промежуточные итоги

## Задания

Данные к заданиям находятся в файле **Makeupsubtotals.xlsx**. Используя инструмент Промежуточный итог (Subtotal), выполните следующие расчеты:

1. Найдите количество проданных единиц продукции и выручку для каждого продавца.
2. Найдите количество сделок по каждому продукту.
3. Найдите самую крупную сделку (с точки зрения выручки) для каждого продукта.
4. Найдите среднюю сумму в долларах для сделки по регионам.
5. Покажите разбивку проданных единиц продукции и выручки для каждого продавца, которая представляла бы результаты для каждого продукта по регионам.



# Приемы работы с диаграммами

### Обсуждаемые вопросы

- Как создать комбинированную диаграмму?
- Как создать вспомогательную ось?
- Как обработать недостающие данные?
- Как показать скрытые данные на диаграмме?
- Как с помощью картинок сделать гистограмму интереснее?
- Я отобразил данные о ежегодных продажах в виде гистограммы, но годы как метки столбиков не отображаются. Где ошибка?
- Как включить метки данных и таблицу данных в диаграмму?
- Как поместить в диаграмму метки данных на основе содержимого ячеек?
- Как отследить эффективность отдела сбыта по продажам за определенный период?
- Как создать ленточную диаграмму для проверки наличия запасов?
- Как сохранить диаграмму в качестве шаблона?
- Как с помощью диаграммы «термометр» отобразить прогресс?
- Как создать на диаграмме динамические метки?
- Как с помощью флажков указать, какие ряды данных требуется отобразить на диаграмме?
- Как с помощью списка обеспечить выбор ряда данных для отображения на диаграмме?
- Как создать диаграмму Ганта?
- Как создать диаграмму на основе отсортированных данных?
- Как создать гистограмму, которая при добавлении новых данных обновлялась бы автоматически?
- Как добавить условные цвета в диаграмму?
- Как с помощью диаграммы «водопад» отследить приближение объемов продаж к плану или разложить компоненты продажной цены?

- Как с помощью функции ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ и таблицы Excel создать динамические информационные панели?
- Как вставить в диаграмму вертикальную линию для разделения производительности до и после слияния?
- Как с помощью лепестковой диаграммы отобразить различия баскетболистов в силе, скорости и прыгучести?
- Известно, что изменение двух переменных можно представить с помощью точечной диаграммы. А как представить изменение трех переменных с помощью пузырьковой диаграммы?
- Как создать диаграмму «водопад»?
- Как используются для обобщения иерархических данных диаграммы «дерево» и «солнечные лучи»?
- Что такое диаграмма типа «воронка»?
- Как отразить в диаграмме свежую динамику по интересующим меня акциям?

Древняя китайская поговорка гласит: «Одна картина стоит тысячи слов». В Excel можно создавать великолепные диаграммы, и в данной главе я дам вам множество советов и рекомендаций. Только помните, что в этой сфере в Microsoft Excel произошли значительные изменения (в лучшую сторону!) по сравнению с версиями до 2013 года.

## Ответы на вопросы

### ❓ Как создать комбинированную диаграмму?

В файле `Combinationstemp.xlsx` содержатся данные о фактических и плановых продажах с января по июль. Вы хотите создать диаграмму, отображающую фактические и плановые продажи за каждый месяц. Сначала выделите диапазон `F5:H12` и на вкладке **Вставка (Insert)** в списке **Гистограмма или линейчатая диаграмма (Insert Column Or Bar Chart)** выберите первую 2D-гистограмму (с группировкой) — рис. 54.1.

Использование двух столбиков не позволяет сразу увидеть различия между фактическими и плановыми продажами, поэтому я предпочел создать комбинированную диаграмму, в которой один ряд данных представлен как линия, а другой — как столбики. Для создания такой диаграммы щелкните правой кнопкой мыши по любому ряду и выберите **Изменить тип диаграммы (Change Chart Type)**. Выберите в левой панели **Комбинированная (Combo)** и затем — первый вариант (Гистограмма с группировкой и график), как на рис. 54.2. Появится комбинированная диаграмма (рис. 54.3).

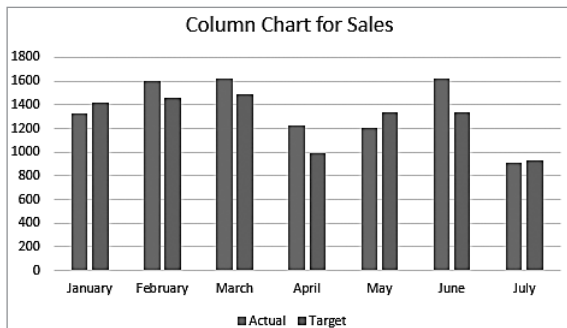


Рис. 54.1. Гистограмма фактических и плановых продаж

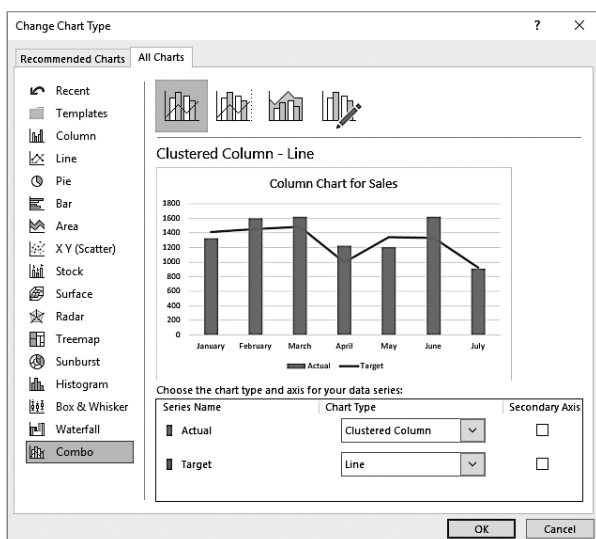


Рис. 54.2. Выбор комбинированной диаграммы

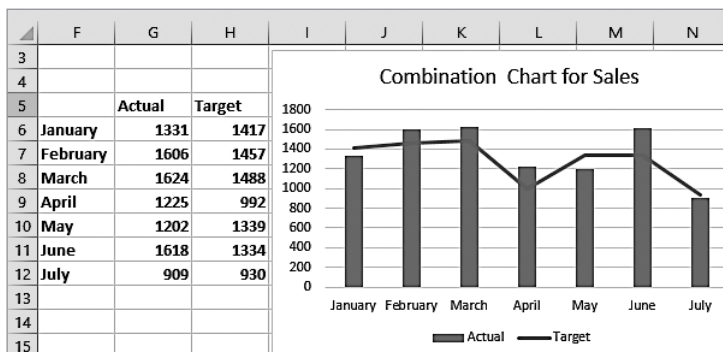


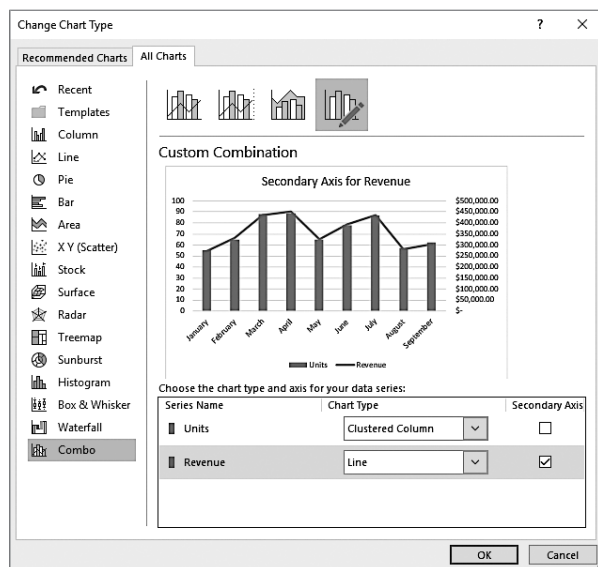
Рис. 54.3. Комбинированная диаграмма

## ? Как создать вспомогательную ось?

При составлении диаграммы для двух величин с различной амплитудой часто применяется вспомогательная ось, призванная повысить наглядность диаграммы. Для примера рассмотрим рис. 54.6 (см. файл `Secondaryaxis.xlsx`), на котором показаны ежемесячная выручка и количество проданных единиц продукции. Если представить эти данные на единственной оси  $y$ , значения ежемесячной выручки будут едва различимы. Для решения этой проблемы сначала выделите диапазон D7:F16, затем выберите на вкладке **Вставка (Insert)** **Вставить комбинированную диаграмму (Combo Chart)** (рис. 54.4) и затем первый в списке вариант выбора, **Гистограмма с группировкой и график (Clustered Column-Line)**. Заполните диалоговое окно, как на рис. 54.5: в дополнение параметрам по умолчанию установите флажок **Вспомогательная ось (Secondary Axis)** для ряда **Выручка (Revenue)** и нажмите **OK**.

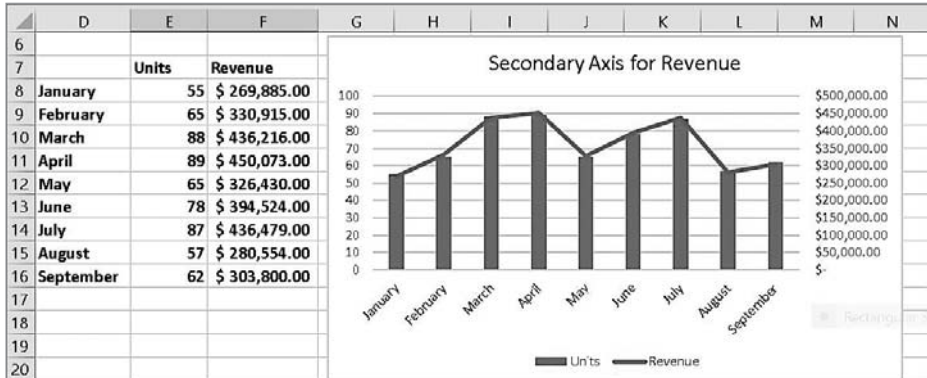


**Рис. 54.4.** Значок Вставить комбинированную диаграмму



**Рис. 54.5.** Создание комбинированной диаграммы со вспомогательной осью

В получившейся диаграмме (рис. 54.6) выручка суммируется на вспомогательной оси (на вертикальной оси справа) в виде графика. Как видите, ежемесячная выручка и количество проданных единиц продукции изменяются практически синхронно.



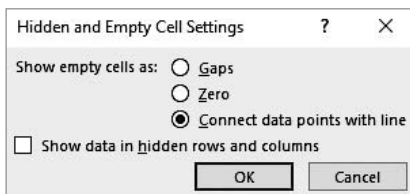
**Рис. 54.6.** Итоговая выручка на вспомогательной оси

### ❓ Как обработать недостающие данные?

Нередко случается так, что в некоторых строках таблицы не хватает данных. Excel предоставляет три способа составления диаграмм с недостающими данными:

- обработка данных как нулевых значений;
- обработка данных как пустых значений (пропуски);
- замена недостающей точки данных линией, соединяющей соседние точки данных.

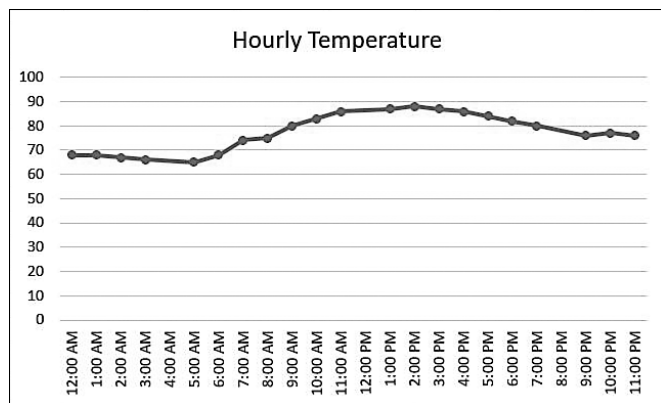
Для примера рассмотрим файл `Missingdata.xlsx`, содержащий почасовые значения температуры, но часть из них отсутствует. После представления данных в виде графика щелкните правой кнопкой мыши на графике и выберите команду **Выбрать данные (Select Data)**. В появившемся диалоговом окне **Выбрать источник данных (Select Data Source)** нажмите **Скрытые и пустые ячейки (Hidden And Empty Cells)**. Откроется диалоговое окно **Настройка скрытых и пустых ячеек (Hidden and Empty Cell Settings)**, показанное на рис. 54.7.



**Рис. 54.7.** Диалоговое окно Настройка скрытых и пустых ячеек

Установите переключатель **Показывать пустые ячейки как** в положение **линию (Connect data points with line)**. Затем с помощью контекстного меню измените тип диаграммы и выберите **График с маркерами (Line Graph with Dots and Lines)**. Полу-

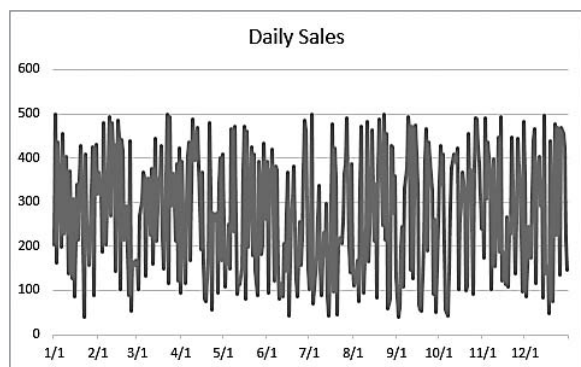
чившийся график представлен на рис. 54.8. На нем легко отличить недостающие данные, поскольку для них маркеры отсутствуют.



**Рис. 54.8.** Недостающие данные заменены линиями

### ? Как показать скрытые данные на диаграмме?

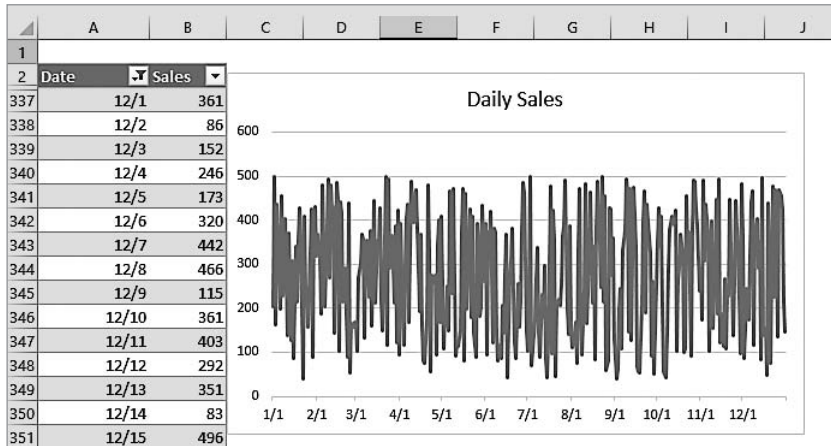
Довольно часто приходится создавать диаграммы для ежедневных продаж и фильтровать данные в таблице. В такой ситуации Excel может сделать две вещи: либо продолжать показывать все данные на диаграмме, либо показывать только отфильтрованные данные. В файле `Hidden.xlsx` находятся данные о ежедневных продажах продукта в течение года. Данные, представленные в виде графика, показаны на рис. 54.9.



**Рис. 54.9.** Диаграмма ежедневных продаж продукта

Если щелкнуть по графику правой кнопкой мыши, затем по **Выбрать данные (Select Data)** и в окне **Настройка скрытых и пустых ячеек (Hidden and Empty Cell Settings)** (см. рис. 54.7) включить флажок **Показывать данные в скрытых строках и столбцах (Show data in**

hidden rows and columns), на диаграмме будут представлены все точки данных даже после фильтрации данных. Например, на рис. 54.10 данные отфильтрованы для показа только декабрьских продаж, но на графике по-прежнему представлены ежедневные продажи за весь год. Если флажок **Показывать данные в скрытых строках и столбцах** не активен, на диаграмме будут видны только декабрьские продажи.



**Рис. 54.10.** Данные отфильтрованы, а точки на диаграмме — нет

### ❓ Как с помощью картинок сделать гистограмму интереснее?

Как правило, амплитуда продаж продукции обобщается как скучные столбцы или полосы, где высота столбцов или ширина полос пропорциональна продажам продукции. Разве не было бы интереснее представить итоги в виде картинки с продуктом, масштабированной пропорционально фактическим продажам? Для иллюстрации этой идеи рассмотрим данные в файле **Picturegraph.xlsx** и рис. 54.11. Допустим, вы продаете книги.

Сначала выделите диапазон **C5:D8**. На вкладке **Вставка (Insert)** щелкните **Гистограмма или линейчатая диаграмма (Insert Column Or Bar Chart)** и выберите первую 2D-гистограмму (с группировкой). Затем щелкните правой кнопкой мыши по любому столбику. Выберите в контекстном меню **Заливка (Fill)** и затем **Рисунок (Picture)**. Откроется диалоговое окно, представленное на рис. 54.12. Введите **книги (books)** в поле **Поиск изображений Bing (Bing Image Search)** и нажмите **Enter**. Появится множество картинок с книгами. Выбрав книгу, нажмите **Вставить (Insert)**. Картинка будет вставлена в диаграмму (см. рис. 54.11). Высота книг пропорциональна фактическим продажам.

Возможно, будет лучше сделать так, чтобы вместо одной книги продажи отображались стопками книг, где одна книга равна 1000 проданных экземпляров. На листе **stack and stretch** (рис. 54.14) создана как раз такая диаграмма. Чтобы изменить диаграмму, где одна книга отображает месячные продажи, щелкните правой кнопкой

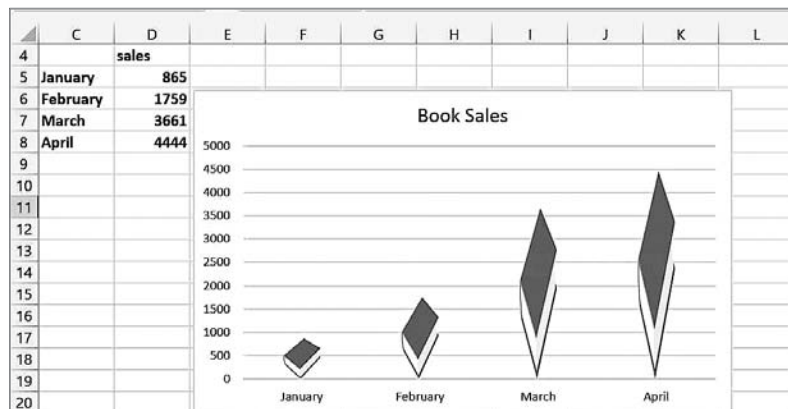


Рис. 54.11. Объемы продаж книг в виде изображений книг

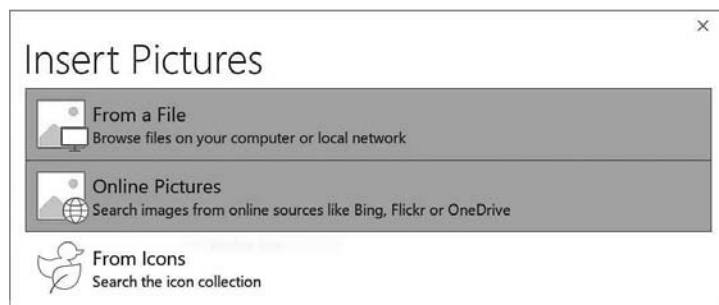


Рис. 54.12. Диалоговое окно Вставка картинок

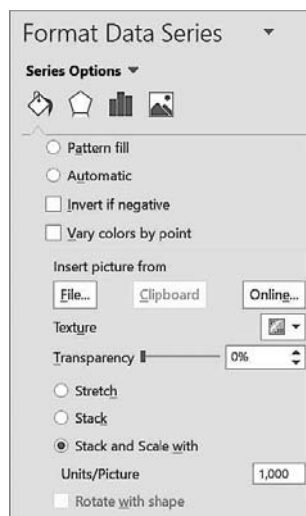
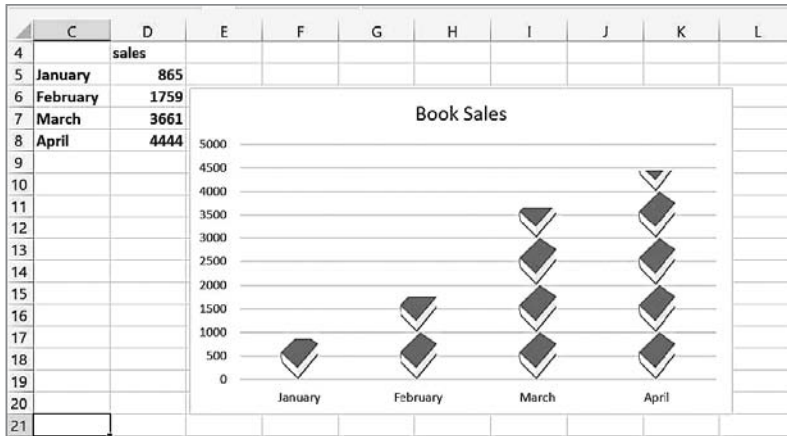


Рис. 54.13. Изменение диаграммы так, чтобы одна книга отражала 1000 экземпляров



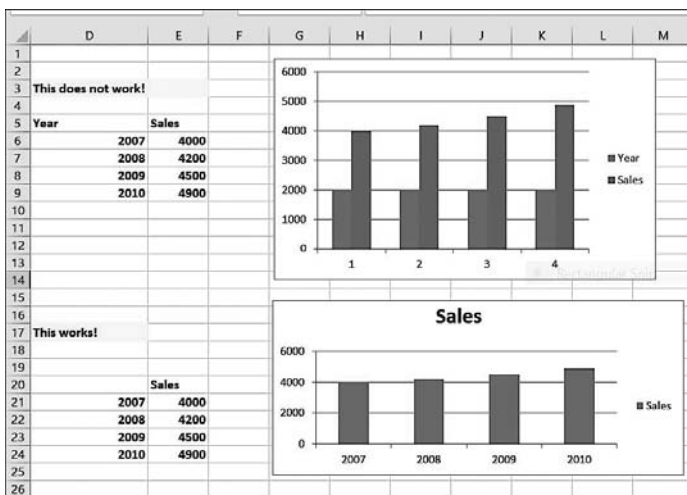
мышью по любой книге и выберите **Формат ряда данных (Format Data Series)**. Как показано на рис. 54.13, выбираем **Заливка (Fill)** (первый значок слева) и изменяем выбор с **Растянуть (Stretch)** на **Размножить (Stack)**, а точнее — **Размножить в масштабе (Stack and Scale with)**, и устанавливаем масштаб (**Units/Picture**) 1000 единиц = 1 книга.



**Рис. 54.14.** Теперь одна книга отражает 1000 экземпляров

❓ **Я отобразил данные о ежегодных продажах в виде гистограммы, но годы как метки столбиков не отображаются. Где ошибка?**

В файле *Categorylabel.xlsx* (рис. 54.15) содержатся данные о ежегодных продажах продукта за 2007–2010 гг. В гистограмме, созданной на основе диапазона D5:E9,



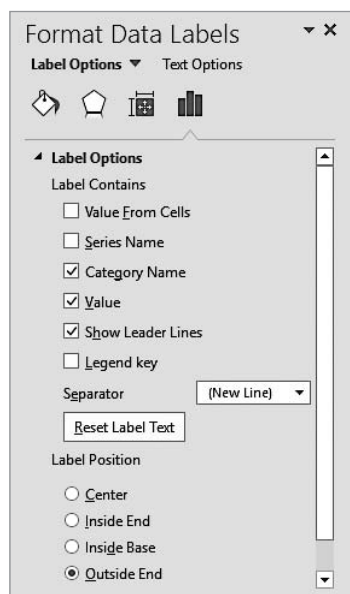
**Рис. 54.15.** Отсутствие метки позволяет придать гистограмме правильный вид

на оси  $x$  годы не указаны, поскольку Excel думает, будто я хочу, чтобы год обрабатывался как ряд данных. Если убрать метку категории **Год** в левом верхнем углу исходного диапазона и выбрать в качестве данных для гистограммы диапазон D20:E24, то на гистограмме появится ось  $x$  для годов.

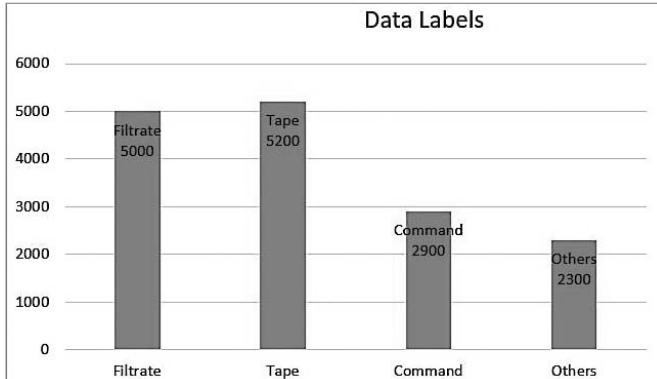
### ? Как включить метки данных и таблицу данных в диаграмму?

Часто нам нужно вставить метки данных рядом со столбиками или полосами или сделать красивую таблицу данных под диаграммой. Рассмотрим этот процесс на примере файла *Labelsandtables.xlsx*. В нем хранятся продажи за текущий месяц по четырем группам продуктов. Сначала создадим на основе этих данных гистограмму и поместим метки, содержащие имена продуктов и фактические продажи, над столбиками. Для этого создайте обычную гистограмму, затем выберите столбик с данными на гистограмме и выберите **Добавить элемент диаграммы (Add Chart Element)** на появившейся вкладке **Конструктор (Chart Tools Design)**. Щелкните стрелку справа от **Метки данных (Data Labels)**, затем **Дополнительные параметры подписей данных (More Data Label Options)**. В панели **Формат подписей данных (Format Data Labels)** заполните раздел **Параметры подписи (Label Options)**, как показано на рис. 54.16. Появится диаграмма, представленная на рис. 54.17.

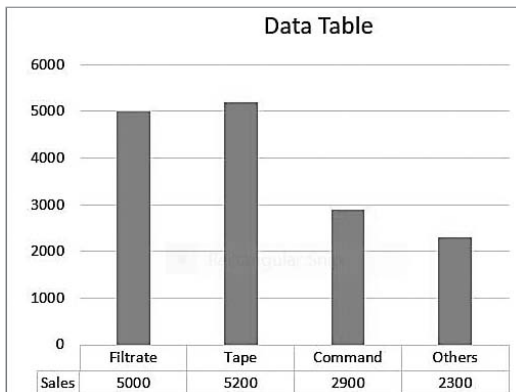
Теперь я покажу, как поместить таблицу данных, обобщающую продажи, под диаграмму. Просто выделите любую часть диаграммы и затем нажмите **Добавить элемент диаграммы** на вкладке **Конструктор**. Затем выберите **Таблица данных (Data Tables)** и вариант **Без легенды (No Legend Keys)**. Диаграмма примет вид, как на рис. 54.18.



**Рис. 54.16.** Параметры, необходимые для настройки показа имен групп и объемов продаж в отдельных строках



**Рис. 54.17.** В гистограмму включены имена групп и объемы продаж



**Рис. 54.18.** Продажи обобщены в таблице данных

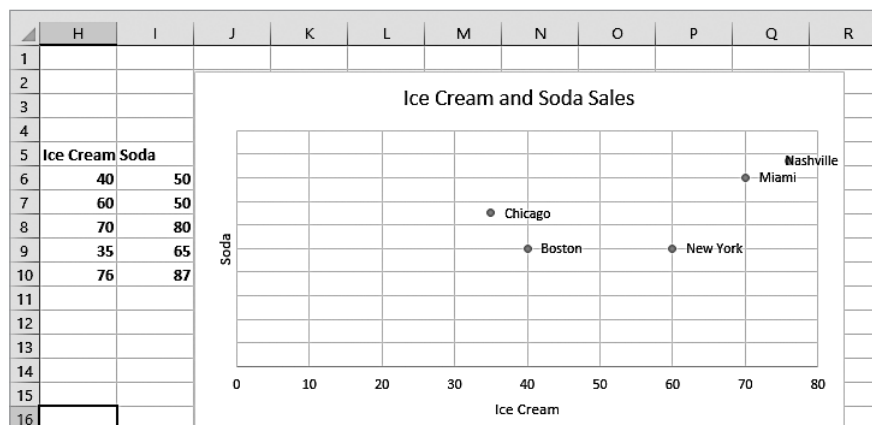
Для просмотра дополнительных возможностей щелкните на треугольнике раскрывающегося списка справа от флажка **Таблица данных (Data Table)**.



### Как поместить в диаграмму метки данных на основе содержимого ячеек?

Начиная с Excel 2013 можно помещать в диаграммы метки данных непосредственно из ячеек. В качестве примера рассмотрим файл `Labelsfromcells.xlsx`, представленный на рис. 54.19.

Сначала выделите ячейки H5:I10 и на вкладке **Вставка (Insert)** выберите первый вариант для точечной диаграммы. Значок **Точечная или пузырьковая диаграмма (Insert Scatter (X, Y) Or Bubble Chart)** расположен рядом со значком круговой диаграммы. Появится точечная диаграмма, как на рис. 54.19, но без меток городов. Для создания меток городов выделите диаграмму, на вкладке **Конструктор (Работа с диаграммами) (Design)** в раскрывающемся списке **Добавить элемент диаграммы (Add Chart Element)**



**Рис. 54.19.** Точечная диаграмма с метками из ячеек

выберите Метки данных (Data Labels) и затем Дополнительные параметры подписей данных (More Data Label Options). На панели Формат подписей данных (Format Data Labels) снимите флажок Значения Y (Y Value), установите флажок значения из ячеек (Value from Cells) и укажите диапазон G6:G10 для вставки меток городов в диаграмму, как показано на рис. 54.19.

### ❓ Как отследить эффективность отдела сбыта по продажам за определенный период?

В файле Salestracker.xlsx представлены ежемесячные продажи для каждого сотрудника отдела продаж с января по май (рис. 54.20).

	D	E	F	G	H	I
1	p	h	up			
2	q	i	down			
3	u	g	flat			
4						
5		January	February	March	April	May
6	Lebron	85	66	81	61	56
7	Wade	82	63	74	78	75
8	Dirk	45	100	115	127	150
9	Manning	75	88	89	76	83
10	Brady	96	90	98	76	93
11	Halliday	75	73	79	91	95
12	Britney	98	91	109	99	84
13	Lindsay	83	84	97	81	98
14	Paris	106	98	84	93	82
15	JLO	104	88	109	101	115
16	Emma	115	94	105	101	107
17	Melo	118	98	128	126	108
18	KD	100	114	104	116	131
19	Vick	112	122	102	124	107
20	Rodgers	127	114	116	139	108

**Рис. 54.20.** Данные о ежемесячных продажах

Вы хотите использовать значки (стрелку вверх, вниз, вправо) для отслеживания изменения ранга продавцов от месяца к месяцу (ранг повысился, понизился, остался прежним). Для этой цели можно было бы выбрать набор значков Excel (см. главу 24 «Условное форматирование»), но тогда пришлось бы вставлять набор значков для каждого месяца, что довольно трудоемко. Браздо эффективнее (но не так эстетично) создать эти значки путем ввода буквы **h** для стрелки вверх, **i** для стрелки вниз и **g** для стрелки вправо. Если затем изменить шрифт на Wingdings 3, появятся требуемые стрелки, поскольку буквы алфавита в Wingdings 3 соответствуют символам, показанным на рис. 54.21.

	J	K	L
22	Letter	Wingdings 3	
23	a	↗	
24	b	↖	
25	c	↔	
26	d	↘	
27	e	↙	
28	f	←	
29	g	→	
30	h	↑	
31	i	↓	
32	j	↖	
33	k	↗	
34	l	↙	
35	m	↘	
36	n	↗	
37	o	↖	
38	p	▲	
39	q	▼	
40	r	△	
41	s	▽	
42	t	◀	
43	u	▶	
44	v	◁	
45	w	▷	
46	x	◀	
47	y	▶	
48	z	◀	

**Рис. 54.21.** Соответствие между буквами и символами Wingdings 3

Для создания значков, показанных на рис. 54.22, предпримите следующее:

1. Скопируйте формулу `=РАНГ(E6;E$6:E$20;0)` из J6 в J6:N20 для вычисления ранга каждого продавца в каждом месяце.
2. Скопируйте формулу `=ЕСЛИ(K6<J6;"h";ЕСЛИ(K6>J6;"i";"g"))` из O6 в O7:R20. В ячейках появится буква **h**, если ранг продавца повысился, буква **i**, если понизился, и буква **g**, если остался прежним.
3. После изменения в диапазоне O6:R20 шрифта на Wingdings 3 появятся стрелки, как на рис. 54.22.

Преимущество такого подхода в том, что вы можете применить функцию **ЕСЛИ** (IF) для быстрого создания критериев выбора соответствующего значка.

	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	p	h	up												
2	q	i	down												
3	u	g	flat												
4							rank	rank	rank	rank	rank	trend	trend	trend	trend
5		January	February	March	April	May	January	February	March	April	May	Feb	March	April	May
6	Lebron	85	66	81	61	56	10	14	13	15	15	15	↓	↑	↓
7	Wade	82	63	74	78	75	12	15	15	12	14	↓	→	↑	↓
8	Dirk	45	100	115	127	150	15	4	3	2	1	↑	↑	↑	↑
9	Manning	75	88	89	76	83	13	10	11	13	12	↑	↓	↓	↑
10	Brady	96	90	98	76	93	9	9	9	13	10	→	→	↓	↑
11	Halliday	75	73	79	91	95	13	13	14	10	9	→	↓	↑	↑
12	Britney	98	91	109	99	84	8	8	4	8	11	→	↑	↓	↓
13	Lindsay	83	84	97	81	98	11	12	10	11	8	↓	↑	↓	↑
14	Paris	106	98	84	93	82	5	5	12	9	13	→	↓	↑	↓
15	JLO	104	88	109	101	115	6	10	4	6	3	↓	↑	↓	↑
16	Emma	115	94	105	101	107	3	7	6	6	6	↓	↑	→	→
17	Melo	118	98	128	126	108	2	5	1	3	4	↓	↑	↓	↓
18	KD	100	114	104	116	131	7	2	7	5	2	↑	↓	↑	↑
19	Vick	112	122	102	124	107	4	1	8	4	6	↑	↓	↑	↓
20	Rodgers	127	114	116	139	108	1	2	2	1	4	↓	→	↑	↓

Рис. 54.22. Значки, использованные для оценки эффективности продавцов

**?** Как создать ленточную диаграмму для проверки наличия запасов?

Довольно часто требуется отслеживать количественные величины (запасы, наличные денежные средства, количество аварий) и определять, остается ли отслеживаемая величина в границах между историческим максимумом и минимумом. Ленточная диаграмма — полезный инструмент для отслеживания изменений с течением времени. Пример ленточной диаграммы приведен на рис. 54.23. (См. файл Bandchart.xlsx.)

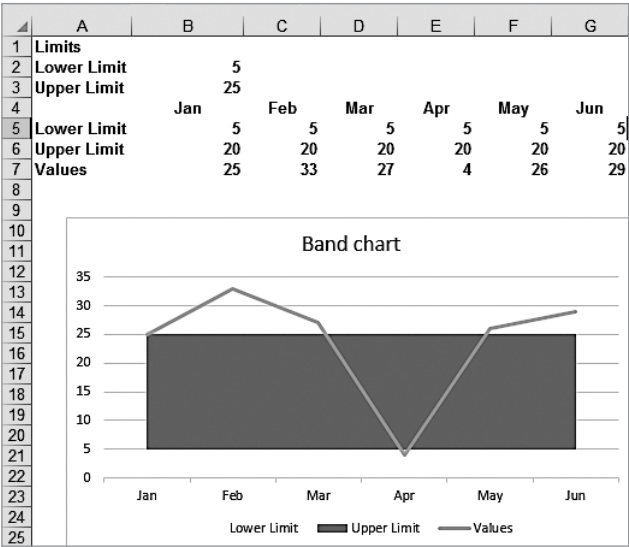


Рис. 54.23. Ленточная диаграмма для отслеживания запасов

Для создания ленточной диаграммы выполните следующее:

- Введите нижнюю границу запаса (5) в ячейку B2 и верхнюю границу (25) в ячейку B3.
- В строке 5 введите нижнюю границу для каждого месяца, скопировав формулу  $=B\$2$  из B5 в C5:G5.
- Скопируйте формулу  $=B\$3-B\$2$  из B6 в C6:G6 для вычисления разницы между верхней и нижней границами. Назовем эту строку *Верхняя граница*, поскольку при построении диаграммы именно эта строка будет использоваться для создания линии, представляющей верхнюю границу запаса в 25 единиц.
- Выделите диапазон A4:G7 и на вкладке *Вставка* (Insert) в группе *Диаграммы* (Chart) выберите *График или диаграмма с областями* (Insert Line Or Area Chart) и затем *С накоплением* (Stacked Column), это второй вариант среди 2D-диаграмм с областями. Измените название диаграммы на *Ленточная диаграмма*.
- Щелкните правой кнопкой мыши по ряду *Значения* (зеленая полоса), выберите в контекстном меню *Изменить тип диаграммы для ряда* (Change Series Chart Type). В открывшемся диалоговом окне под заголовком *Выберите тип диаграммы и ось для рядов данных: (Choose The Chart Type And Axis For Your Data Series)* для ряда *Верхняя граница* выберите *Заливка* (Area), первый вариант в соответствующем разделе. Для ряда *Значения* выберите *График*, первый вариант в разделе графиков. Нажмите ОК.
- Щелкните правой кнопкой мыши по ряду *Нижняя граница* на диаграмме и выберите *Заливка* (Fill). В меню *Заливка* (Fill) выберите *Нет заливки* (No Fill).

Полученная ленточная диаграмма показывает, что запас с трудом удается удерживать между требуемыми верхней и нижней границами.

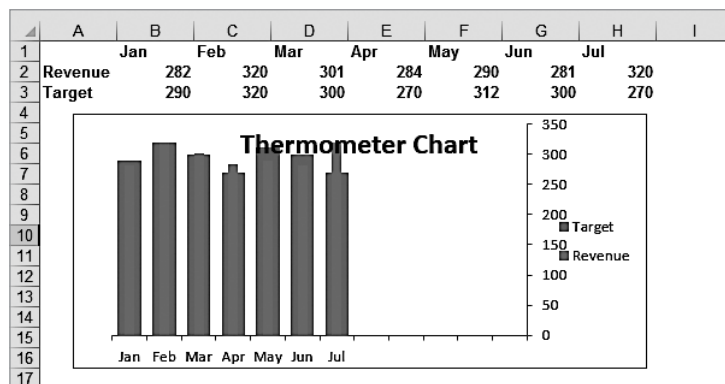
### ❓ Как сохранить диаграмму в качестве шаблона?

Я только что создал прекрасную ленточную диаграмму. Возможно, вы думаете, что теперь каждый раз придется повторять вышеописанные шаги для того же результата. Если так, то вы ошиблись. Можно сохранить ленточную (или любую другую) диаграмму как шаблон и менять настройки диаграммы в любое время. В качестве примера откройте файл *Bandchart.xlsx*, щелкните правой кнопкой мыши по диаграмме, выберите *Сохранить как шаблон* (Save As Template) и присвойте диаграмме любое имя (например, Band). Предположим, вам нужна ленточная диаграмма для месяцев с января по март. Выделите диапазон данных A4:D7 и на вкладке *Вставка* (Insert) в группе *Диаграммы* (Chart) щелкните по стрелке в правом нижнем углу. На вкладке *Все диаграммы* (All Charts) выберите *Шаблоны* (Templates). Найдите требуемый шаблон диаграммы, нажмите ОК, и готово!

### ❓ Как с помощью диаграммы «термометр» отобразить прогресс?

Диаграмма «термометр» показывает такие величины, как выручка, в виде столбика, отображающего плановое значение. Получившаяся диаграмма (на

рис. 54.24 и в файле *Thermometer.xlsx*) напоминает термометр, отсюда и идет ее название.



**Рис. 54.24.** Диаграмма «термометр»

Для создания диаграммы «термометр» выполните следующее:

- Выделите диапазон **A1:H3** и на вкладке **Вставка (Insert)** в группе **Диаграммы (Chart)** выберите 2D-диаграмму **Гистограмма с группировкой (Clustered Column)**.
- Щелкните правой кнопкой мыши по ряду **Выручка** на диаграмме (синий столбик) и выберите команду **Формат ряда данных (Format Data Series)**. Выберите **По вспомогательной оси (Secondary Axis)**.
- В той же панели **Формат ряда данных (Format Data Series)** (ряд **Выручка** выделен) установите в поле **Перекрытие рядов (Overlap)** значение **0%** и в поле **Боковой зазор (Gap Width)** значение **261%**.
- Щелкните правой кнопкой мыши по ряду **План** (красный столбик) и установите в поле **Перекрытие рядов (Overlap)** значение **0%** и в поле **Боковой зазор (Gap Width)** значение **48%**.

Теперь диаграмма выглядит, как на рис. 54.24. При необходимости ширину бокового зазора можно отрегулировать. Уменьшение бокового зазора для ряда **План**, например, сделает красные столбики шире; увеличение бокового зазора для ряда **Выручка** сделает синие столбики уже.

### ❓ Как создать на диаграмме динамические метки?

Вероятно, многим приходилось иметь дело с диаграммами Excel, в которых при изменении данных метки на диаграммах не менялись. Такая ситуация часто приносит путаницу. Давайте научимся связывать метки рядов и названия диаграммы с ячейками на листе. (См. файл *DynamicLabels.xlsx* и рис. 54.25.) Предположим, что вам нужно создать диаграмму для прогнозируемого ВВП в США и Китае. Название диаграммы должно содержать год, в который Китай догонит США по ВВП,



а названия рядов — ежегодные темпы роста для каждой страны. В ячейках C5 и C8 текущие расчетные значения темпов роста — 3% для ВВП США и 10% для ВВП Китая — можно изменить.

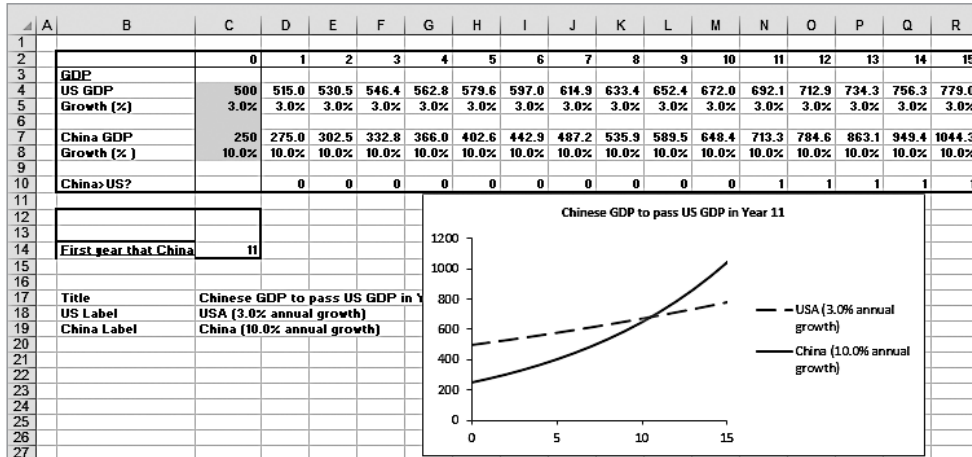


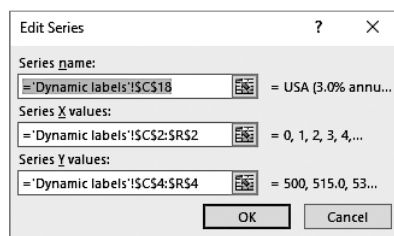
Рис. 54.25. Создание динамических меток

Основная идея в том, чтобы связать названия диаграммы и метки с ячейками, которые изменяются при изменении темпов роста. Для этого сделайте следующее:

- Скопируйте формулу `=ЕСЛИ(D7>=D4;1;0)` из D10 в E10:R10, которая даст в результате 1, если ВВП Китая окажется не меньше ВВП США.
- В ячейке C14 определите год, когда Китай догонит США по ВВП, по формуле `=ЕСЛИОШИБКА(ПОИСКПОЗ(1;D10:R10;0);"none")`. Обратите внимание, что если Китай никогда не догонит США по ВВП, формула возвратит «none».
- В ячейке C17 создайте желаемое название диаграммы по формуле `=ЕСЛИ(C14="none";"США опережает Китай по ВВП";"Китай догнал США по ВВП на "&ТЕКСТ(C14;"0")&" году")`. Обратите внимание, что если Китай никогда не догонит США по ВВП, то диаграмма будет называться "США опережает Китай по ВВП". В противном случае название диаграммы привязывается к ячейке C14 таким образом, что в него входит год, на который Китай догонит США по ВВП. "0" в функции ТЕКСТ гарантирует, что значение для года будет отформатировано как целое число.
- В ячейке C18 создайте название для ряда данных ВВП США по формуле `"США (темпы роста "&ТЕКСТ(C5;"0,0%")&")"`. Часть текстовой функции "0,0%" гарантирует, что значение для темпов роста будет отформатировано в процентах.
- В ячейке C19 создайте название для ряда данных ВВП Китая по формуле `"Китай (темпы роста "&ТЕКСТ(C8;"0,0%")&")"`.

Теперь все готово для создания диаграммы с динамическими метками.

- Удерживая Ctrl, выделите несмежные диапазоны ячеек C2:R2, C4:R4 и C7:R7 и создайте точечную диаграмму с гладкими кривыми: на вкладке Вставка (Insert) выберите Точечная или пузырьковая диаграмма (Insert Scatter (X, Y) Or Bubble Chart) и затем Точечная с гладкими кривыми (Scatter With Smooth Lines) (третий вариант).
- На вкладке Конструктор (Design) выберите Добавить элемент диаграммы (Add Chart Element) и Название диаграммы по центру (Centered Overlay Title). В строке формул введите знак равенства (=), затем выберите C17 и нажмите Enter. На диаграмме появится название с динамической меткой.
- Выделите ряд данных 1 (ряд ВВП США), щелкните правой кнопкой мыши и выберите Выбрать данные (Select Data). Нажмите Изменить (Edit) и заполните диалоговое окно Изменение ряда (Edit Series), как на рис. 54.26: введите в поле Имя ряда (Series Name) = 'Dynamic labels'!\$C\$18, оставьте остальные значения по умолчанию и нажмите OK в обоих диалоговых окнах.



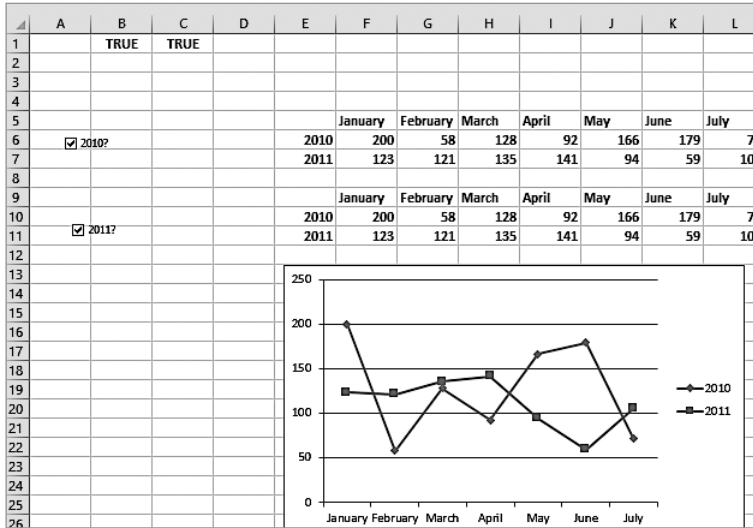
**Рис. 54.26.** Создание динамической метки ряда

Это связь метки ряда ВВП США с ячейкой C18, в которой содержится значение годовых темпов роста. Аналогичным образом создайте связь метки ряда ВВП Китая с ячейкой C19. Диаграмма с динамическими метками готова.

### **❓ Как с помощью флажков указать, какие ряды данных требуется отобразить на диаграмме?**

В главе 27 «Счетчики, полосы прокрутки, переключатели, флажки, группы и поля со списками» как вы, возможно, помните, мы научились переключать в ячейке с помощью флажка значения ЛОЖЬ и ИСТИНА. Оказалось, если Excel видит в ячейке ошибку #Н/Д, то точка на диаграмме не создается. Поэтому если требуется не включать в диаграмму какой-либо ряд данных, просто создайте формулу с функцией ЕСЛИ для записи в этот ряд данных ошибки #Н/Д после того, как с помощью флажка в ячейке установлено значение ЛОЖЬ. См. пример на рис. 54.27.

Откройте файл Checkbox.xlsx. Сначала методами, описанными в главе 27, создайте два флажка: один для управления рядом данных 2010, а другой для управления рядом данных 2011. Флажок 2010 управляет содержимым ячейки B1, а флажок 2011 — содержимым ячейки C1. Исходные данные находятся в диапазоне E6:L7. Формула =ЕСЛИ(\$B\$1;F6;НД()), скопированная из F10 в F10:L11, дает в результате либо исходные данные соответствующего года, если для этого года флажок



**Рис. 54.27.** Флажки для управления представлением данных на диаграмме

установлен, либо ошибку #Н/Д, если для этого года флажок года снят. Теперь выберите диапазон E9:L11 в качестве исходных данных для построения диаграммы **График с маркерами (Line With Markers)** — четвертый график в разделе 2D-графиков. Сброс флажка с номером года скроет на графике данные за этот год, а установка флажка с номером года представит на графике соответствующие данные.

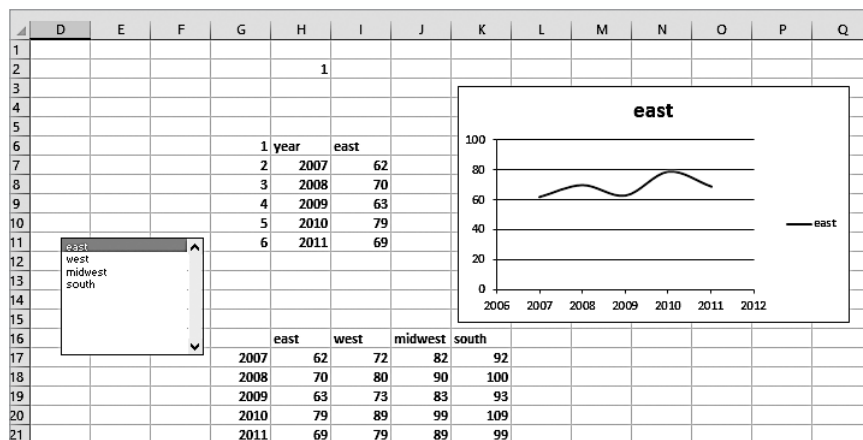


### Как с помощью списка обеспечить выбор ряда данных для отображения на диаграмме?

На листе (см. файл **Listbox.xlsx**) содержатся данные о продажах за 2007–2011 гг. в регионах Восток, Запад, Средний Запад и Юг. Вам нужен простой способ управления включением рядов данных в диаграмму. Такой способ предоставляет список (см. главу 27). Сначала на вкладке **Разработчик (Developer)** в группе **Элементы управления (Form Controls)** в списке **Вставить (Insert)** выберите элемент управления формы **Список (List Box)** — пятый значок в верхнем ряду. Поместите список на лист, щелкните правой кнопкой мыши и выберите **Формат объекта (Format Control)**. В поле **Формировать список по диапазону (Input Range)** выберите диапазон \$A\$14:\$A\$17 и в поле **Связь с ячейкой (Cell Link)** ячейку \$N\$2. Теперь, умно используя функцию **ИНДЕКС**, вы можете использовать ячейку H2 для управления рядами данных, которые нужно включить в диаграмму (рис. 54.28).

Скопируйте формулу **=ИНДЕКС(\$N\$16:\$K\$21;G7;\$N\$2)** из I7 в I8:I10 для извлечения правильного ряда данных. Например, в списке, показанном на рис. 54.28, выбран регион Восток. Это помещает в ячейку H2 значение 1. Функция **ИНДЕКС** возвращает первый столбец данных (столбец H), что и требовалось. Теперь выделите диапазон H7:Y11 и создайте точечную диаграмму с гладкими кривыми. Вы увидите, что

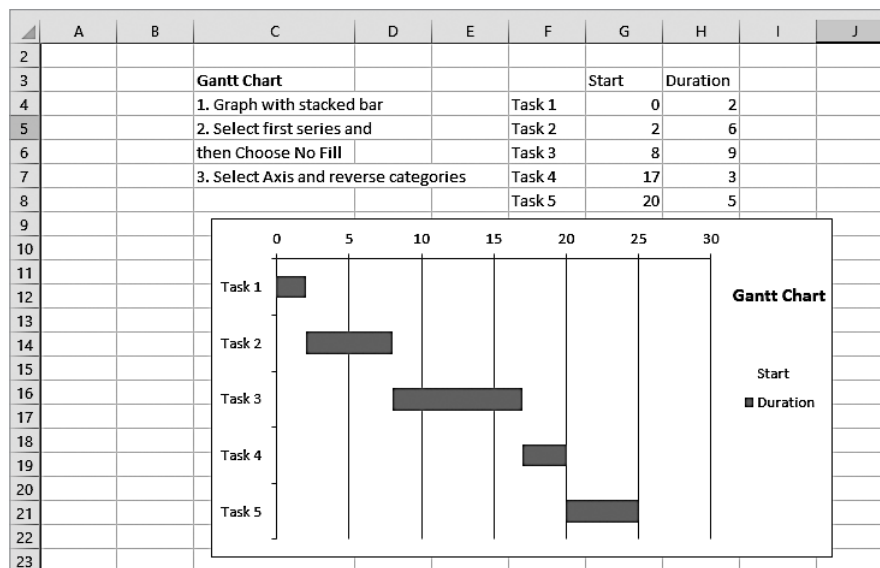
после щелчка на конкретном регионе из списка диаграмма построится на основе правильного ряда данных.



**Рис. 54.28.** Список для управления включением рядов данных в диаграмму

### ? Как создать диаграмму Ганта?

Часто для проекта требуется завершение ряда других проектов. На диаграмме Ганта показывается время начала каждого проекта. На рис. 54.29 (см. файл



**Рис. 54.29.** Диаграмма Ганта

Gantt.xlsx) представлена диаграмма Ганта для проекта, состоящего из пяти задач. Для создания такой диаграммы выполните следующее:

- Выберите диапазон F3:H8 и создайте 2D-гистограмму с накоплением (второй вариант).
- Щелкните правой кнопкой мыши по ряду Начало, выберите Заливка (Fill) и Нет заливки (No Fill), чтобы скрыть ряд Начало.
- Щелкните правой кнопкой мыши по вертикальной оси и выберите Формат оси (Format Axis). Затем установите флажок Обратный порядок категорий (Categories In Reverse Order), обеспечивающий Задаче 1 первое место в списке.

### ? Как создать диаграмму на основе отсортированных данных?

Пусть имеются объемы продаж в нескольких штатах (рис. 54.30 и файл Sortedgraphx.xlsx). Вы хотите построить диаграмму с перечислением штатов в порядке убывания объемов продаж (рис. 54.31). Перед созданием диаграммы необходимо реорганизовать данные в столбцах J–L и отсортировать данные продаж в порядке убывания.

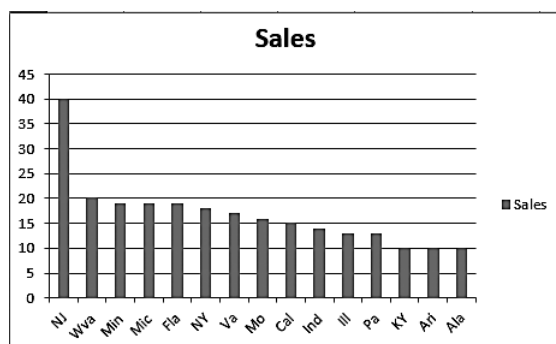
	E	F	G	H	I	J	K	L
8	State	Sales	Original rank	Revised		Rank	State	Sales
9	NJ	40	1	1		1	NJ	40
10	NY	18	6	6		2	Wva	20
11	Ind	14	10	10		3	Min	19
12	Cal	15	9	9		4	Mic	19
13	KY	10	13	13		5	Fla	19
14	Ari	10	13	14		6	NY	18
15	Ala	10	13	15		7	Va	17
16	Min	19	3	3		8	Mo	16
17	Ill	13	11	11		9	Cal	15
18	Mic	19	3	4		10	Ind	14
19	Mo	16	8	8		11	Ill	13
20	Fla	19	3	5		12	Pa	13
21	Pa	13	11	12		13	KY	10
22	Wva	20	2	2		14	Ari	10
23	Va	17	7	7		15	Ala	10

**Рис. 54.30.** Данные продаж для отсортированной диаграммы

Для этого сделайте следующее:

- Скопируйте формулу =РАНГ(F9;\$F\$9:\$F\$23;0) из G9 в G10:G23 для вычисления ранга продаж в каждом штате. Например, ранг Нью-Йорка ( NY) равен шести. Обратите внимание, что в Кентукки (KY), Алабаме (Ala) и Аризоне (Ari) объемы продаж одинаковые. Для сортировки штатов в порядке убывания продаж необходимо связать с каждым штатом уникальный ранг.
- Скопируйте формулу =G9+СЧЁТЕСЛИ(\$G\$8:G8;G9) из H9 в H10:H23 для создания уникального ранга каждого штата путем увеличения ранга штата, имеющего связанный ранг, каждый раз, когда встречается связанный ранг.

- Скопируйте формулу =ИНДЕКС(Е\$9:Е\$23;ПОИСКПОЗ(\$J9;\$N\$9:\$N\$23;0);1) из K9 в K10:L23 для сортировки штатов и продаж на основе объемов продаж по убыванию.
- Теперь создайте гистограмму на основе диапазона K9:L23, как на рис. 54.31 (выберите Гистограмма с группировкой (Clustered Column) в группе Диаграммы (Charts)).

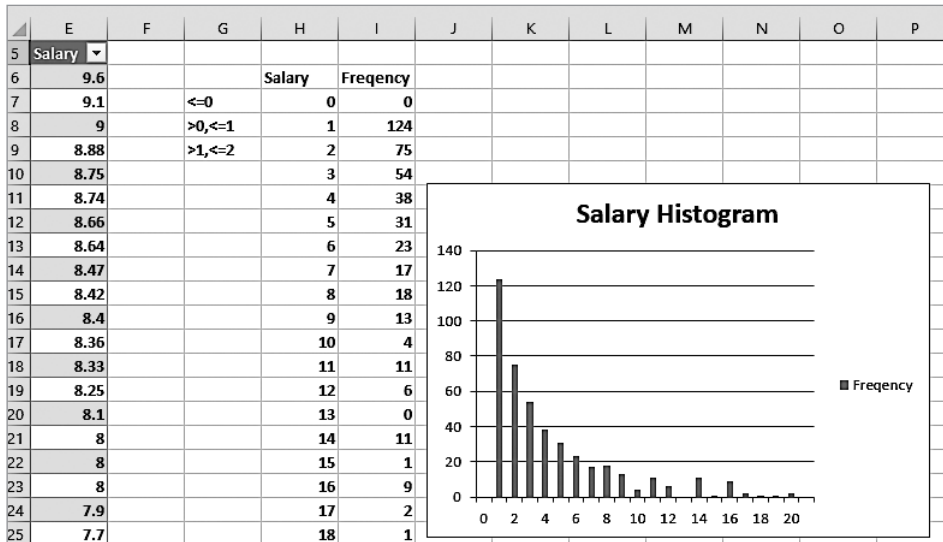


**Рис. 54.31.** Объемы продаж отсортированы по убыванию

### ❓ Как создать гистограмму, которая при добавлении новых данных обновлялась бы автоматически?

В главе 43 «Обобщение данных на гистограммах и диаграммах Парето» вы научились создавать гистограммы с помощью надстройки Анализ данных (Data Analysis). К сожалению, гистограммы, созданные с помощью пакета анализа, при изменении или добавлении данных не обновляются. Для того чтобы диаграммы автоматически подстраивались под изменения в исходных данных, нужно преобразовать данные в таблицу Excel. (См. пример в файле *Dynamichistograms.xlsx* и на рис. 54.32.)

Столбец E содержит зарплаты игроков НБА в миллионах долларов в сезоне 2003–2004 гг. Сначала выделите диапазон E5:E446 и преобразуйте его в таблицу. Введите бин-диапазоны для гистограммы в ячейки H7:H28. Обратите внимание, что в строке 28 вычисляется количество игроков с зарплатой свыше \$20 млн. Теперь воспользуемся функцией ЧАСТОТА (FREQUENCY) для вычисления того, сколько зарплат попадают в каждый бин-диапазон. Функция ЧАСТОТА (см. главу 91 «Формулы массива и функции, возвращающие массив») является функцией, возвращающей массив результатов. Для ее использования сначала необходимо выбрать диапазон, который заполнит функция (в данном примере I7:I28), затем ввести формулу с функцией ЧАСТОТА и нажать Ctrl+Shift+Enter. После выбора диапазона I7:I28 введите формулу =ЧАСТОТА(E6:E446;H7:H27) и нажмите Ctrl+Shift+Enter. В ячейку I7 Excel введет количество игроков (0) с зарплатой менее или равной 0, в ячейку I8 — количество игроков с зарплатой > 0 и <= 1 млн долларов (124), ..., в ячейку I27 — количество игроков с зарплатой > 19 и <= 20 млн долларов (2)



**Рис. 54.32.** Динамическая гистограмма

и в ячейку I28 — количество игроков с зарплатой > 20 млн долл. (0). Для создания динамической гистограммы выберите диапазон H7:I28 и вставьте гистограмму (выберите Гистограмма с группировкой (Clustered Column) в группе Диаграммы (Charts)). Мы получили динамическую гистограмму, которая будет обновляться автоматически после изменения входных данных. Для доказательства обновления гистограммы добавьте несколько зарплат в 30 млн долл. в столбце E (ряд 447 и больше). В правой части диаграммы появится новый столбец.

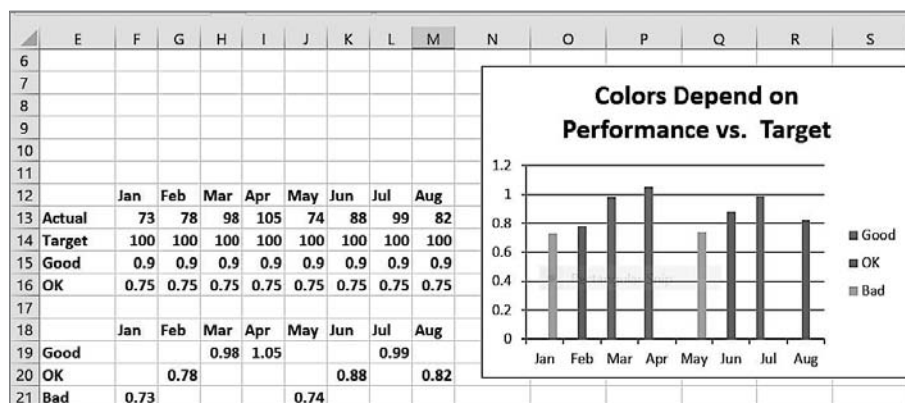
### ❓ Как добавить условные цвета в диаграмму?

Допустим, вы создаете диаграмму фактических и плановых продаж для каждого месяца. При этом вы хотите выделить месяцы, в которых фактические продажи составили 90% и больше от плановых, синим цветом; месяцы, в которых фактические продажи составили 75% и меньше — зеленым цветом; а остальные месяцы — красным цветом. В файле Condcolors.xlsx (рис. 54.33) показано, как это делается. Хитрость создания такой диаграммы заключается в размещении данных для каждого цвета в отдельной строке (см. строки 19–21).

1. Сначала поместите продажи, которые требуется выделить синим, в строку 19, скопировав формулу =ЕСЛИ(F13/F14>F15;F13/F14;" ") из F19 в G19:M19.
2. Затем поместите продажи, которые требуется выделить красным цветом, в строку 20, скопировав формулу =ЕСЛИ(И(G13/G14>G16;G13/G14<=G15);G13/G14;" ") из F20 в G20:M20.
3. Наконец, поместите продажи, которые требуется выделить зеленым цветом, в строку 21, скопировав формулу =ЕСЛИ(И(G13/G14>G16; G13/G14<=G15);

G13/G14; " ") из F21 в G21:M21. Теперь выделите диапазон E18:M21 и создайте гистограмму (выберите Гистограмма с группировкой (Clustered Column) в группе Диаграммы (Charts)).

4. Если вы хотите назначить цвет для какого-либо ряда данных, выделите этот ряд и щелкните на значке кисти справа от диаграммы; щелкните на параметре Цвет (Color) и затем выберите в палитре желаемый цвет.



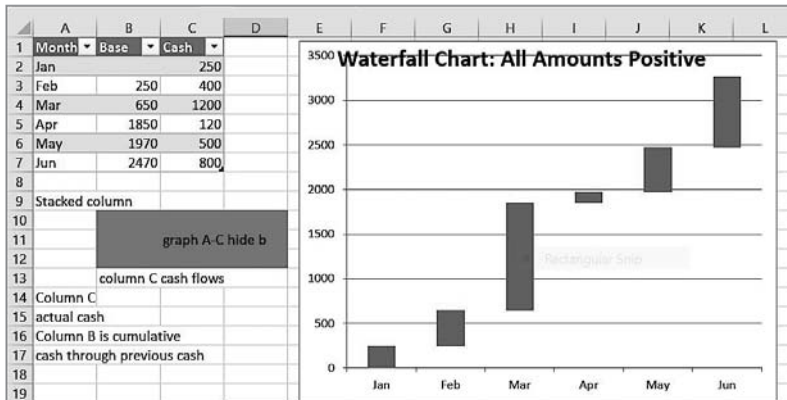
**Рис. 54.33.** Синий цвет для удачных месяцев, зеленый для неудачных и красный для средних

### ❓ Как с помощью диаграммы «водопад» отследить приближение объемов продаж к плану или разложить компоненты продажной цены?

С помощью диаграммы «водопад», первоначально разработанной консалтинговой компанией McKinsey, часто представляют прогресс в достижении конечной кассовой позиции или разбивку общего дохода компании на составляющие затраты и прибыль. В этом разделе мы рассмотрим, как создается диаграмма «водопад» в версиях Excel 2013 и ранее. Далее в главе будет показано, насколько проще создавать эти диаграммы в Excel 2016 и Excel 2019. В файле **Waterfallcharts.xlsx** содержится несколько примеров диаграммы «водопад».

На листе **All Positive** (рис. 54.34) создана диаграмма «водопад» для ситуации, в которой все денежные потоки являются положительными. Диаграмма должна показывать ежемесячный прогресс на пути к цели — к денежному потоку в размере \$3270. После ввода денежных потоков в столбец С необходимо отследить суммарный денежный поток в столбце В. Скопируйте формулу  $=B2+C2$  из В3 в В4:В7 для вычисления суммарного денежного потока в конце каждого месяца. Затем выделите диапазон ячеек А1:С7 и создайте гистограмму с накоплением. Измените название диаграммы: *Диаграмма «водопад»: все денежные потоки положительные*. Если выделить ряд **База** (синие столбики) и выбрать **Нет заливки (No Fill)**, получится диаграмма, показанная на рис. 54.34.

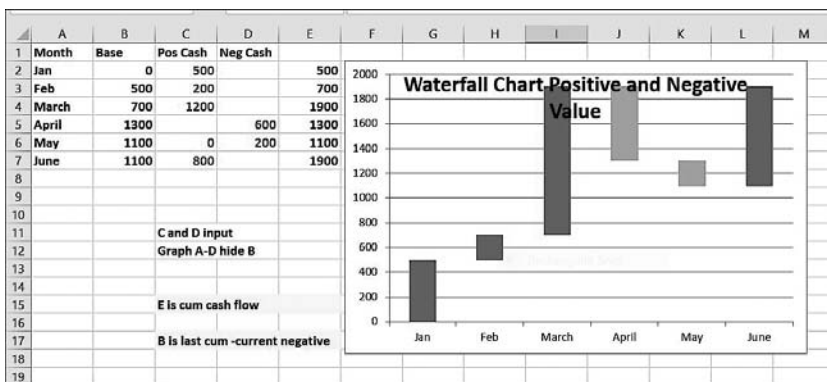




**Рис. 54.34.** На этой диаграмме все денежные потоки положительные

На рис. 52.35 (см. лист **Positive and Negative**) показана диаграмма «водопад» при наличии нескольких отрицательных денежных потоков.

После ввода положительных денежных потоков в столбец **C** и отрицательных денежных потоков в столбец **D** вычислите общий денежный поток в конце каждого месяца посредством копирования формулы  $=E1+C2-D2$  из **E2** в **E3:E7**. Затем в столбце **B**, скопировав формулу  $=E1-D2$  из **B2** в **B3:B7**, вычислите суммарный денежный поток за последний месяц, скорректированный отрицательным денежным потоком за текущий месяц. После выбора диапазона **A1:D7** и создания гистограммы с накоплением остается только скрыть столбец **B** на диаграмме. Полученная в результате диаграмма представлена на рис. 54.35. (Чтобы скрыть столбец, щелкните правой кнопкой мыши по синему столбику и выберите **Нет заливки (No Fill)** в меню **Заливка (Fill)**.) Кроме того, можно поменять название диаграммы на **Диаграмма «водопад» с положительными и отрицательными денежными потоками**.



**Рис. 54.35.** Диаграмма «водопад» с положительными и отрицательными денежными потоками

Следуя методике, примененной на листе **Positive and Negative**, можно быстро заполнить разбивку дохода компании на составляющие затрат (включая прибыль). См. лист **Profitability Waterfall** и рис. 54.36.

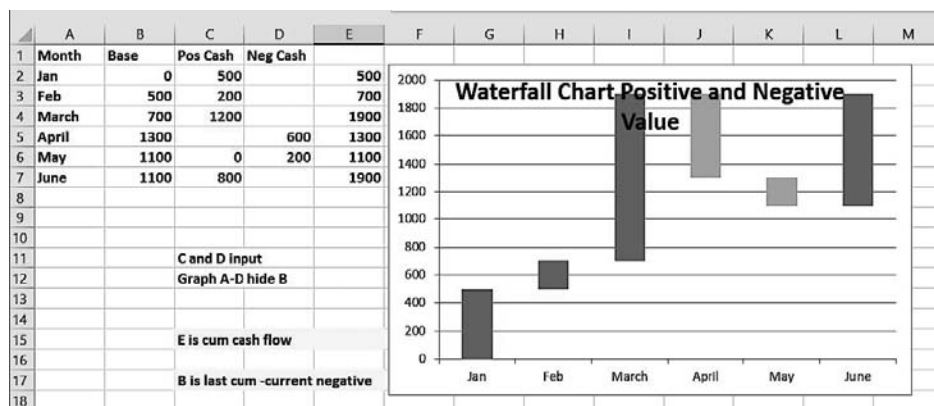


Рис. 54.36. Диаграмма рентабельности

### ❓ Как с помощью функции **ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ** и таблицы Excel создать динамические информационные панели?

Часто нам нужно загрузить ежемесячные, ежеквартальные или еженедельные данные о продажах и на их основе создать диаграммы, автоматически обновляющиеся при добавлении новых данных. Зная функцию Excel **ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ** (**GETPIVOTDATA**) и возможности таблиц Excel (см. главы 45 «Сводные таблицы и срезы для описания данных» и 26 «Таблицы» соответственно), это можно сделать относительно просто (см. файл **Randy.xlsx**). Допустим, вы загрузили еженедельные данные о продажах четырех категорий продуктов: **Safety, Tape, Command** и **Abrasives**, в четырех магазинах: **Menards, Target, Loews** и **Home Depot**. Цель — создать информационную панель, позволяющую быстро составить диаграмму на основе еженедельных данных о продажах в указанном магазине с возможностью включения в диаграмму желаемых категорий продуктов. Как показано на рис. 54.37, ваши данные загружаются в столбцы **D–G**. Преобразуйте диапазон **D4:G243** в таблицу. Это обеспечит автоматический учет в сводной таблице новых данных при обновлении таблицы.

Для создания сводной таблицы выберите вкладку **ВСТАВКА (Insert)** и затем в группе **Таблицы (Tables)** — **Сводная таблица (PivotTable)**. В открывшейся панели перетащите поле **Неделя** в область **СТРОКИ (ROWS)**, поля **Магазин** и **Категория** в область **СТОЛБЦЫ (COLUMNS)** и поле **Выручка** в область **ЗНАЧЕНИЯ (VALUES)**. В сводной таблице (рис. 54.38) обобщены еженедельные данные о продажах по каждой категории продуктов в каждом магазине.

Теперь мы готовы применить функцию **ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ** для извлечения данных, необходимых для создания диаграммы. Сначала создайте

	D	E	F	G
4	WEEK	Category	Store	Revenue
5	3	Abrasives	Lowes	2043
6	12	Safety	Menards	2343
7	3	Tape	Home Depot	1414
8	12	Command	Target	1820
9	9	Tape	Home Depot	943
10	7	Tape	Target	1219
11	7	Command	Menards	1156
12	11	Abrasives	Lowes	2127
13	12	Safety	Menards	1315
14	3	Tape	Target	1580
15	10	Abrasives	Home Depot	1598
16	4	Command	Lowes	1000
17	7	Tape	Menards	1087
18	7	Abrasives	Menards	1728
19	1	Abrasives	Target	1911
20	7	Abrasives	Menards	1563
21	2	Tape	Target	2482
22	7	Safety	Lowes	1534
23	12	Safety	Menards	1471
24	2	Abrasives	Lowes	990
25	11	Tape	Lowes	1580
26	1	Safety	Target	2389
27	2	Safety	Menards	2263

Рис. 54.37. Данные о продажах

	I	J	K	L	M	N	O	P	Q	R	S	T
9												
10												
11	Sum of Revenue	Column Labels										
12		Lowes				Lowes Total	Menards			Menards Total	Home Depot	
13	Row Labels	Abrasives	Safety	Tape	Command		Abrasives	Safety	Tape	Command		Abrasives
14	1		1065		1247	2312	3772	526	3466		7764	3272
15	2	2349	4510			6859	6220	3296	2159		11675	1921
16	3	2797	1233	2046	6335	12411	3472		959	2369	6800	
17	4	1501	3822	1180	1000	7503	6018	2212			8230	
18	5	1940		833	2945	5718	708		501	597	1806	1758
19	6			1918	6494	8412	4145	6205	4653		15003	
20	7	3049	3228			6277	3291	2303	6472	1156	13222	3580
21	8	4208	2890	1873		8971	910		1494	3776	6180	2445
22	9			2052	2825	4877	6265		1888		8153	
23	10	6117		1875		7992			2138		2138	4807
24	11	2127	1971	4461	2394	10953	3596	1169			4765	
25	12		1542	4786	850	7178	1124	7565	7264	2431	18384	2471
26	13				1196	1196				3400	3400	
27	14	1455				1455						
28	Grand Total	25543	20261	21024	25286	92114	39521	23276	####	13729	107520	20254

Рис. 54.38. Сводная таблица еженедельных данных о продажах по категориям продуктов для каждого магазина

раскрывающийся список (см. главу 42 «Проверка достоверности данных») в ячейке AG8 (на листе Model) для выбора магазина. Для каждой категории продуктов создайте флажки (см. главу 27), с помощью которых можно управлять диапазоном AH9:AK9. Эти ячейки определяют категории продуктов, появляющиеся в диаграмме. Скопируйте формулу =ЕСЛИ(АН\$9=ЛОЖЬ; НД(); ЕСЛИОШИБКА(ПОЛУЧИТЬ.ДАННЫЕ.СВОДНОЙ.ТАБЛИЦЫ("Выручка"; \$I\$11; "Неделя"; \$AG11; "Категория"; АН\$10;

"Магазин"; \$AG\$8; " ") из AH11 в AH11:AK24. Если для какой-либо категории продуктов установлен флажок, формула извлекает еженедельные продажи для этой категории; если флажок не установлен, то в ячейку записывается сообщение об ошибке #Н/Д. Кроме того, если объем продаж равен 0, часть формулы с функцией ЕСЛИОШИБКА (IFERROR) обеспечивает запись нулевых продаж. Преобразуйте ячейки AG10:AK24 в таблицу для автоматического обновления любой диаграммы, построенной на основе этих данных. На рис. 54.39 показаны исходные данные для диаграммы, на которой требуется обобщить данные по продажам продуктов Abrasives, Safety и Tape в магазине Loew's.

А теперь долгожданная награда! Выделите диапазон AG10:AK24 и создайте точечную диаграмму с гладкими кривыми, выбрав Точечная или пузырьковая диаграмма

	AE	AF	AG	AH	AI	AJ	AK
3							
4		<input checked="" type="checkbox"/> Abrasives	<input type="checkbox"/> Command				
5							
6		<input checked="" type="checkbox"/> Tape	<input checked="" type="checkbox"/> Safety				
7			store				
8			Lowes				
9				TRUE	FALSE	TRUE	TRUE
10			Week	Abrasives	Command	Tape	Safety
11			1	0	#N/A	0	1065
12			2	2349	#N/A	0	4510
13			3	2797	#N/A	2046	1233
14			4	1501	#N/A	1180	3822
15			5	1940	#N/A	833	0
16			6	0	#N/A	1918	0
17			7	3049	#N/A	0	3228
18			8	4208	#N/A	1873	2890
19			9	0	#N/A	2052	0
20			10	6117	#N/A	1875	0
21			11	2127	#N/A	4461	1971
22			12	0	#N/A	4786	1542
23			13	0	#N/A	0	0
24			14	1455	#N/A	0	0

Рис. 54.39. Исходные данные для диаграммы

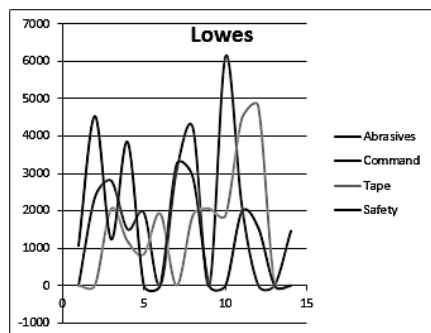
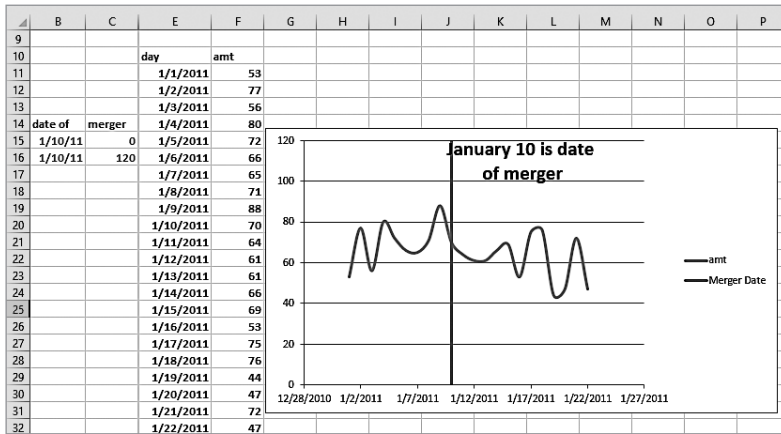


Рис. 54.40. Динамическая информационная панель по продажам в магазине Loew's продуктов Abrasives, Tape и Safety

в группе **Диаграммы (Charts)** и затем **Точечная диаграмма с гладкими кривыми** в разделе **Точечные (Scatter)**. Готовая диаграмма представлена на рис. 54.40. Она может показать продажи в любом магазине по любой комбинации категорий продуктов и обновляется при загрузке новых данных.

### ? Как вставить в диаграмму вертикальную линию для разделения производительности до и после слияния?

Допустим, 10 января 2011 г. произошло слияние вашей компании с другой компанией (см. файл *Verticalline.xlsx* и рис. 54.41), и вы суммируете ежедневные продажи в диаграмме. Для обозначения даты слияния в диаграмму необходимо вставить вертикальную линию. Если нарисовать вертикальную линию из набора фигур, то при перемещении диаграммы линия окажется в неправильном месте. Для устранения этой проблемы сначала выберите диапазон **E10:F32** и создайте точечную диаграмму с гладкими кривыми (третий вариант). В диапазоне **B15:C16** введите дату слияния, а также верхнюю и нижнюю границы для вертикальной линии по оси *y*. (В данном случае нижняя граница равна 0, а верхняя граница равна 120.) Скопируйте диапазон **B15:C16** и щелкните правой кнопкой мыши в области диаграммы. Щелкните по значку **Вставить (Paste)** для вставки вертикальной линии на дату 10 января.



**Рис. 54.41.** Вертикальная линия указывает дату слияния 10 января 2011 г.

### ? Как с помощью лепестковой диаграммы отобразить различия баскетболистов в силе, скорости и прыгучести?

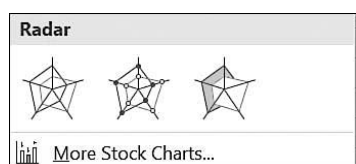
В файле *Radar.xlsx* находятся данные о силе, скорости и прыгучести четырех спортсменов. Выделите диапазон **C2:F6** и на вкладке **Вставка (Insert)** щелкните по значку лепестковой диаграммы (рис. 54.42), он также служит и для диаграмм «водопад», «воронка», «биржевая» и «поверхностная». В раскрывающемся меню выберите **Лепестковая с маркерами (Radar Chart with Markers, второй вариант в раз-**

деле Лепестковая, рис. 54.43) для получения лепестковой диаграммы, показанной на рис. 54.44.

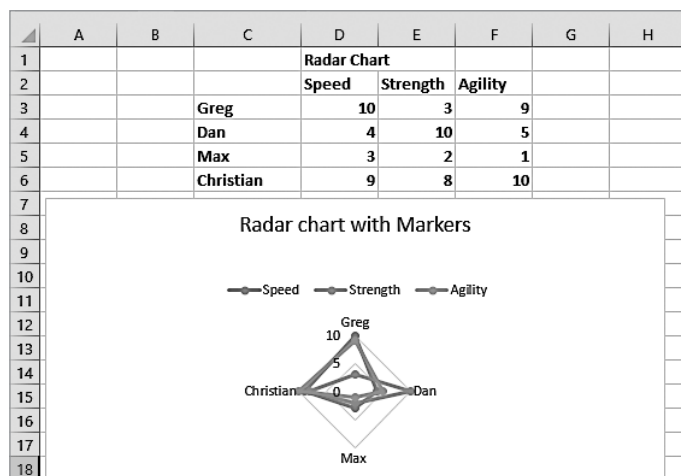


**Рис. 54.42.** Значок диаграмм «водопад», «воронка», «биржевая» и «поверхностная» или лепестковых

Центр лепестковой диаграммы соответствует показателю 0, и чем дальше маркер от центра, тем лучше показатель. По диаграмме сразу видно, что Макс, например, плохо забивает, с какой стороны ни подходить к этому вопросу, а вот Кристиан хорош по всем показателям.



**Рис. 54.43.** Варианты лепестковых диаграмм

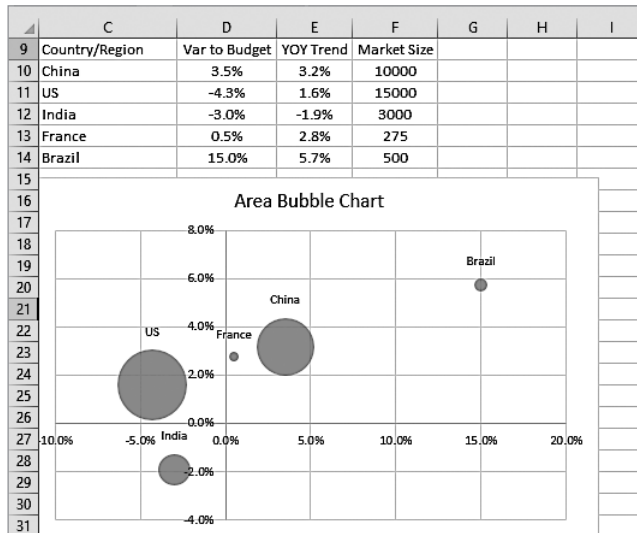


**Рис. 54.44.** Лепестковая диаграмма с маркерами, обобщающая эффективность спортсменов по нескольким параметрам

**?** Известно, что изменение двух переменных можно представить с помощью точечной диаграммы. А как представить изменение трех переменных с помощью пузырьковой диаграммы?

Точечная диаграмма показывает изменение двух переменных, тогда как пузырьковая диаграмма позволяет отобразить три переменные. Файл **Bubble.xlsx**

(рис. 54.45) содержит для нескольких стран/регионов процентную разницу объемов продаж относительно бюджета, а также годовой рост продаж и размер рынка каждой страны. Для отображения этих данных в виде пузырьковой диаграммы выделите диапазон **D9:F14**, на вкладке **Вставка (Insert)** щелкните на стрелке раскрывающегося списка значка **Точечная диаграмма (Scatter Plot)** и выберите **Пузырьковая (Bubble Chart)**, первый вариант. Поменяйте заголовок на **Пропорциональная пузырьковая диаграмма**. После добавления подписей данных (как было описано ранее в этой главе) для каждого пузырька на основе страны (диапазон ячеек **C10:C14**) и размещения подписей над каждым пузырьком получится диаграмма, как на рис. 54.45. (Для добавления подписей в панели **Формат подписей данных (Format Data Labels)**, ниже заголовка **Параметры подписей (Label Options)**, установите первый флажок, **Значения из ячеек (Value From Cells)**, и снимите флажок **Значение Y (Y Value)**). После этого выберите диапазон **C10:C14** для вставки названий стран в диаграмму. В разделе **Положение подписи (Label Position)** выберите **выше (Above)**.) Площади пузырьков пропорциональны размеру рынка каждой страны. Например, площадь пузырька для США на 50% больше, чем площадь пузырька для Китая.

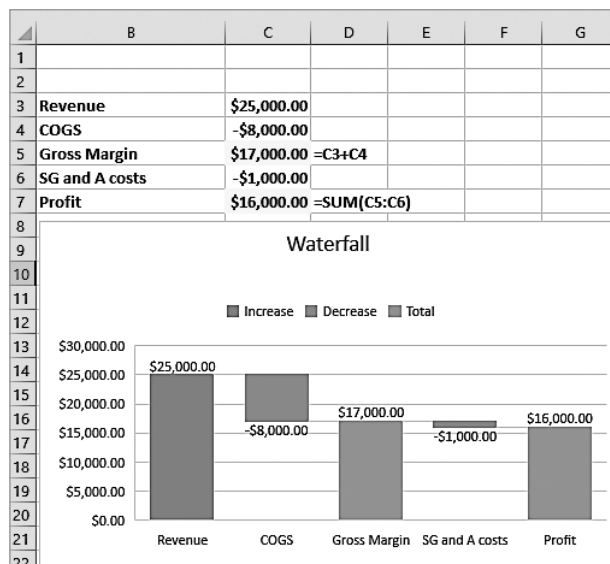


**Рис. 54.45.** Пузырьковая диаграмма с пропорциональными размерами

### ❓ Как создать диаграмму «водопад» в Excel 2019?

Выше было показано, насколько замысловатым могло оказаться создание диаграмм «водопад» (или «каскадная») в версиях Excel, предшествовавших Excel 2016. Здесь мы продемонстрируем, насколько просто это стало в Excel 2016 или 2019. В файле **Waterfallpandltemp.xlsx** (см. рис. 54.46) приведены следующие данные: доход компании; затраты на реализацию (COGS); коммерческие, общие

и административные издержки (SG&A) и прибыль. Требуется создать диаграмму «водопад», отражающую динамику превращения дохода компании, равного \$25 000, в прибыль, составляющую \$16 000. Валовая прибыль и прибыль — итоговые значения, которые не будут изменяться, поэтому их требуется обрабатывать отдельно.



**Рис. 54.46.** Диаграмма «водопад», увязывающая доход, затраты и прибыль

Сначала выделите диапазон B3:C7. На вкладке **Вставка (Insert)** выберите опцию "Водопад" или биржевая диаграмма (Insert Waterfall Or Stock Chart), показанную на рис. 54.42. Затем, как показано на рис. 54.47, выберите опцию диаграмма "водопад" ("каскадная").

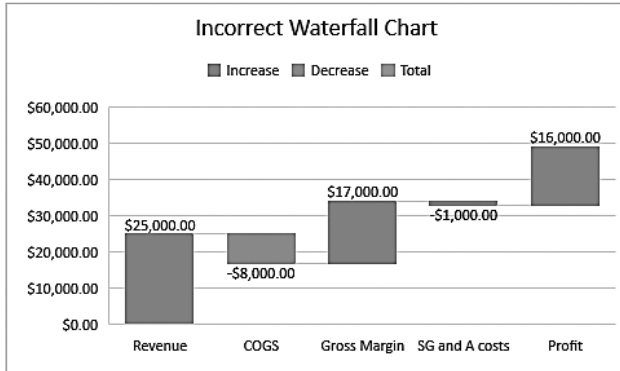


**Рис. 54.47.** Выбор диаграммы «водопад»

Вы получите диаграмму «водопад», представленную на рис. 54.48.

Ошибка этой диаграммы в том, что столбики **валовая прибыль (Gross Margin)** и **прибыль (Profit)** должны стоять на оси *x*. Чтобы добиться этого, щелкните дважды на столбике **валовая прибыль**, щелкните правой кнопкой мыши и выберите **Формат элемента данных (Format Data Point)**; в открывшейся панели выберите **Установить в качестве итога (Set As Total)**, а затем щелкните столбик **прибыль** и для него





**Рис. 54.48.** Неправильная диаграмма «водопад»

также выберите **Установить в качестве итога (Set As Total)**. Теперь вы получите правильную диаграмму «водопад», приведенную выше на рис. 54.46, на которой итоговые столбики выделены серым цветом. Теперь понятно представлено, как доход в \$25 000 преобразуется в прибыль, составляющую \$16 000.

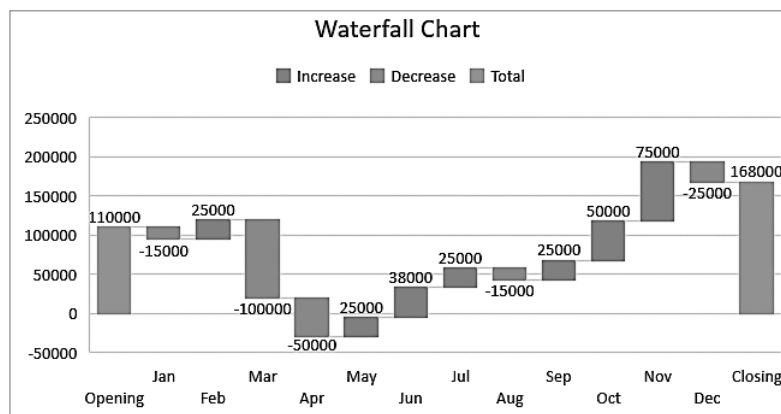
В качестве другого примера диаграммы «водопад» рассмотрим данные из файла **Waterfallcashtemp.xlsx** в папке **Templates** (см. рис. 54.49), которые отражают остаток денежных средств на начало года, ежемесячные изменения кассовой наличности и остаток кассовых средств на конец года.

	A	B	C	D
1	Opening	110000		
2	Jan	-15000		
3	Feb	25000		
4	Mar	-100000		
5	Apr	-50000		
6	May	25000		
7	Jun	38000		
8	Jul	25000		
9	Aug	-15000		
10	Sep	25000		
11	Oct	50000		
12	Nov	75000		
13	Dec	-25000		
14	Closing	168000	=B1+SUM(B2:B13)	

**Рис. 54.49.** Денежные потоки в течение года

Выбрав диапазон **A1:B14** и создав диаграмму «водопад», мы снова видим, что первую и последнюю точки необходимо зафиксировать как итоговые (щелкните дважды столбик **Opening**, щелкните правой кнопкой мыши и выберите **Формат**

элемента данных (Format Data Point); в открывшейся панели выберите **Установить в качестве итога (Set As Total)**; затем сделайте то же самое для столбика (Closing). Так вы получите диаграмму «водопад», представленную на рис. 54.50.



**Рис. 54.50.** Диаграмма «водопад» для денежных потоков

### ❓ Как в Excel используются для обобщения иерархических данных диаграммы «дерево» и «солнечные лучи»?

Диаграммы «дерево» (также известная как *мозаичная* диаграмма) и «солнечные лучи» позволяют визуализировать иерархические данные. Иерархические данные — это способ организации данных с множественными отношениями один-к-многим. Структура подчиняется правилу, что один родитель может иметь много потомков, но потомок может иметь только одного родителя. Например, если смотреть на продажи продукта по *кварталам года*, по *месяцам года* и по *неделям месяца*, то *квартал* является родителем для трех потомков (месяцы квартала), а каждый месяц имеет четыре или пять потомков в соответствии с неделями, входящими в этот месяц.

В качестве другого примера рассмотрим небольшой книжный магазин. Родителями здесь могут быть разделы книг (детские книги, книги по искусству и фотоальбомы и т. п.). Потомками в категории «Искусство и фотография» могли бы оказаться книги по рукоделию и иллюстрированные подарочные издания. В детских книгах потомками могут быть книги в категориях по возрасту: 3–5 и 6–8 лет. В категории 3–5 лет могут быть подкатегории: «Первая книжка», «Азбука», «Толстой малышам».

Далее, основываясь на публикации в блоге Microsoft Office «Breaking down hierarchical data with treemap and sunburst charts» (<https://blogs.office.com/2015/08/11/breaking-down-hierarchical-data-with-treemap-and-sunburst-charts/>), мы создадим диаграммы «дерево» и «солнечные лучи». В файле *Treemapbookstoretemp.xlsx* в папке *Templates* (см. рис. 54.51) представлены данные о доходах по различным категориям книг в небольшом книжном магазине.

	A	B	C	D
1	GENRE	SUB-GENRE	TOPIC	REVENUE
2	Arts & Photogra	How-to Crafts		\$ 2,711
3		Coffee-table	Photography	\$ 2,309
4	Children's Book	Baby Books		\$ 16,092
5		Age 3-5	1st Readers	\$ 24,514
6			ABCs	\$ 17,771
7			Tolstoy for Tot	\$ 13,295
8		Age 6-8		\$ 14,046
9		Pre-Teen & Teen		\$ 18,046
10	Computers & In	Troubleshooting		\$ 4,527
11	Mystery	Crime	Fiction	\$ 11,186
12			True Crime	\$ 8,790
13		Spy		\$ 6,516
14			True Spy	\$ 3,809
15	Nonfiction	Health	Diet	\$ 3,293
16			Fitness	\$ 6,891
17		History		\$ 1,131
18	Magazine	Fashion	Women's	\$ 7,315
19			Men's	\$ 2,222
20		Home		\$ 2,612
21		Other		\$ 3,140
22		Sports	Sport's Illustra	\$ 8,009
23			MMA	\$ 4,257
24	Romance	Break up	Teen	\$ 6,205
25			Young Adult	\$ 25,193
26			Audiobooks	\$ 3,045
27		Make Up		\$ 15,050
28	Science Fiction	Apocalyptic		\$ 10,200
29		Comics		\$ 3,456

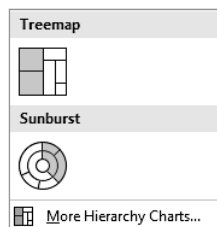
**Рис. 54.51.** Данные для диаграмм «дерево» и «солнечные лучи»

Создание древовидной диаграммы начните с выделения диапазона A1:D29. Затем на вкладке **Вставка (Insert)** выберите опцию **Иерархическая диаграмма (Insert Hierarchy Chart)**, представленную на рис. 54.52.

Из набора опций, представленных на рис. 54.53, выберите **Древовидная (Treemap)**, и вы получите диаграмму, приведенную на рис. 54.54.



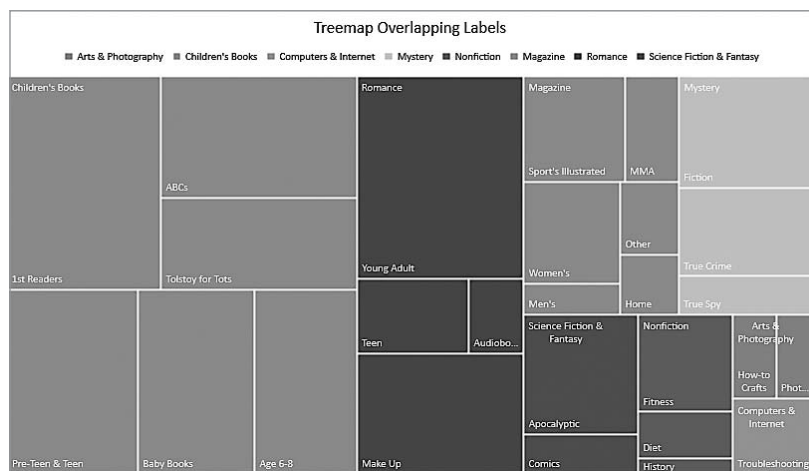
**Рис. 54.52.** Значок иерархических диаграмм



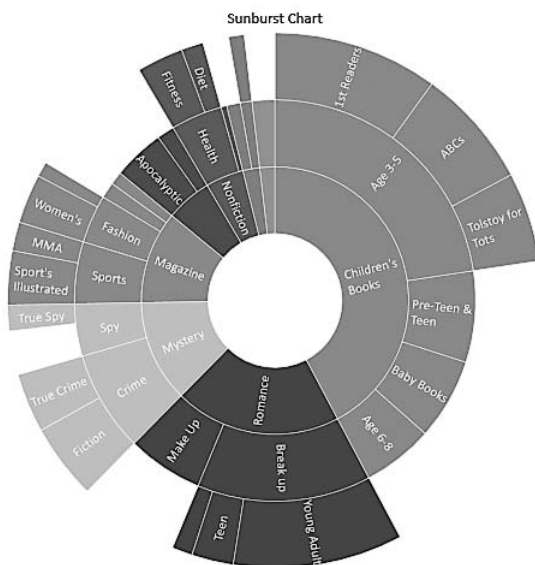
**Рис. 54.53.** Выбор диаграммы «дерево» или «солнечные лучи»

Эта диаграмма отображает долю каждой категории книг в общем объеме продаж. Обратите внимание, что типы книг, относящиеся к одному родителю, представлены одним и тем же цветом. Диаграмма показывает, например, что чуть меньше половины всех продаж (42%) составляют детские книги. Чтобы получить конкретные данные, нужно навести курсор на какой-либо прямоугольник. Если щелкнуть

правой кнопкой мыши по прямоугольнику, можно выбрать **Формат ряда данных (Format Data Series)** и установить в параметрах подписи **Баннер**, после чего подписи к родительским прямоугольникам будут помещены отдельно, каждая в свой собственный прямоугольник. При желании вы можете добавить цифры продаж к диаграмме, выбрав **Значения (Values)** из **Дополнительные параметры подписей данных (More Data Labels Options)**.



**Рис. 54.54.** Перекрывающиеся надписи на древовидной диаграмме



**Рис. 54.55.** Диаграмма «солнечные лучи» для книжного магазина

Обратите внимание, что из-за недостатка места на диаграмме не отображается подпись 3–5 лет (Age 3–5). Использование диаграммы «солнечные лучи» позволяет решить эту проблему. Если выделить диапазон A1:D29 и выбрать опцию Солнечные лучи (Sunburst), вы получите диаграмму, показанную на рис. 54.55. (см. файл Sunburstfinal.xlsx).

Теперь видна разбивка книг категории 3–5 лет (Age 3–5) на три подкатегории. Конечно, следует признать, что по диаграмме «солнечные лучи» труднее на вид определить долю каждого типа книг в общем объеме продаж.

### ? Что такое диаграмма типа «воронка»?

Диаграмма типа «воронка» показывает параметры процесса на разных его стадиях. Возможность создавать такие диаграммы была добавлена в Excel 2019. В файле Funnel.xlsx продемонстрировано применение диаграммы-«воронки». В Faber University изначально было 5000 потенциальных доноров, а в конечном итоге оказалось 400 доноров, как показано на рис. 54.56. Чтобы создать диаграмму-«воронку» этого процесса, выберите диапазон C5:D9. Затем, выбрав значок, показанный на рис. 54.42, выберите вариант Воронка (Funnel Chart).

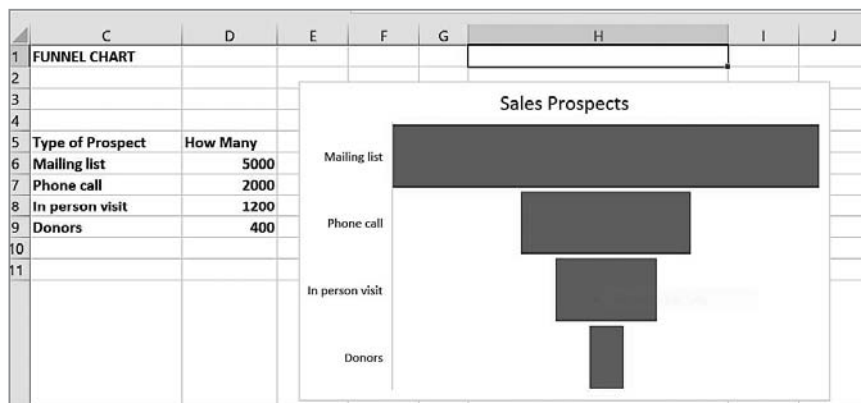
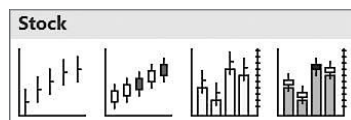


Рис. 54.56. Диаграмма типа «воронка»

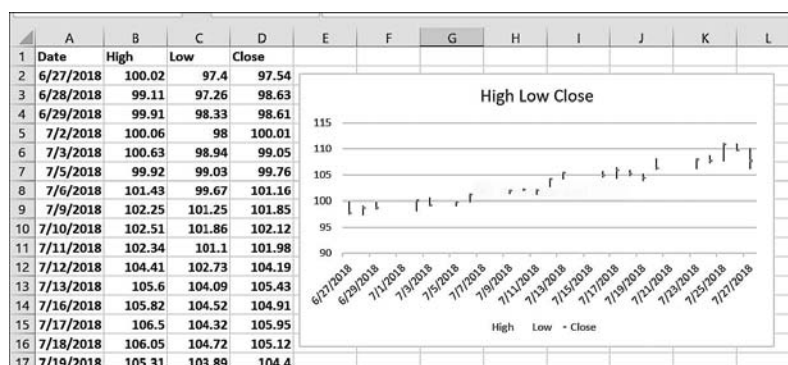
### ? Как отразить в диаграмме свежую динамику по интересующим меня акциям?

В Excel есть множество возможностей обобщать цены и тенденции объемов продаж на бирже. Здесь мы рассмотрим четыре типа биржевых диаграмм (см. файл Stockcharts.xlsx), содержащих ежедневную информацию о цене и объемах продаж акций компании Microsoft в течение июня и июля 2018 г. Нажав на значок диаграммы, показанный на рис. 54.42, вы увидите значки для четырех вариантов биржевых диаграмм, как показано на рис. 54.57.

Самый высокий курс, самый низкий курс, курс закрытия (High Low Close Chart). Такая диаграмма показывает ежедневную цену на момент закрытия дня торгов (рис. 54.58) как точку, наивысшую цену за день — вверх линии, а самую низкую — вниз линии. Вам нужно просто перечислить данные в столбик, как показано на рис. 54.58, и выбрать первый значок из показанных на рис. 54.57.



**Рис. 54.57.** Варианты биржевых диаграмм

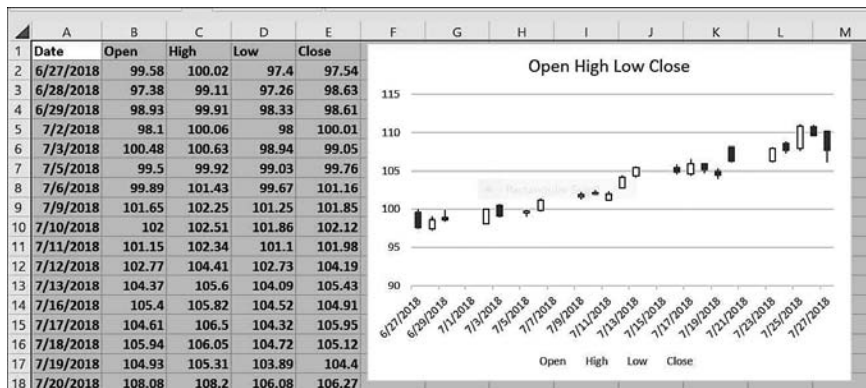


**Рис. 54.58.** Диаграмма «Самый высокий курс, самый низкий курс, курс закрытия»

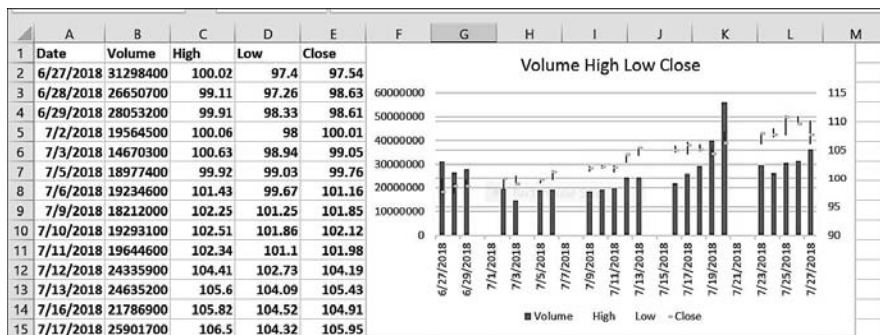
Курс открытия, самый высокий курс, самый низкий курс, курс закрытия (Open High Low Close Chart). Такая диаграмма показывает ежедневную наивысшую цену (наверху верхней линии), самую низкую цену (внизу нижней линии) и цены открытия и закрытия дня торгов (рис. 54.59). Если закрашивается ящик, цена открытия находится наверху ящика, а цена закрытия — внизу. Если ящик не закрашивается, цена открытия находится внизу ящика, а цена закрытия — наверху. Данные должны быть перечислены в столбик, как показано на рис. 54.59.

Объем, самый высокий курс, самый низкий курс, курс закрытия (Volume High Low Close Chart). Как показано на рис. 54.60, эта диаграмма точно такая же, как и первая (High Low Close Chart): ежедневный объем продаж представлен столбиком и при помощи вспомогательной оси.

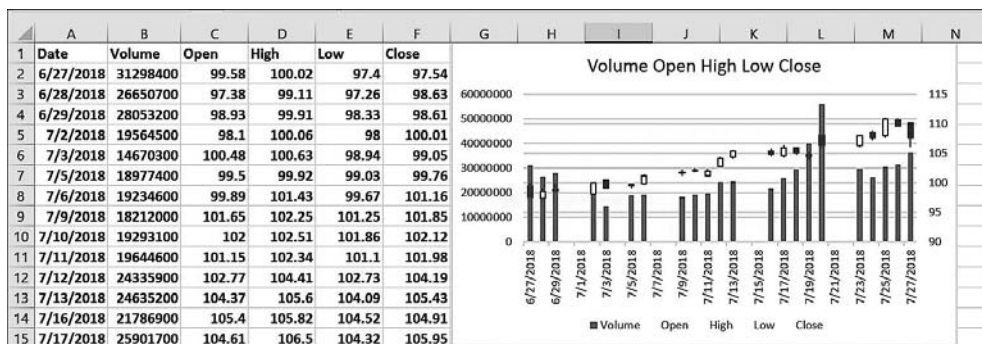
Объем, курс открытия, самый высокий курс, самый низкий курс, курс закрытия (Volume Open High Low Close Chart). Как показано на рис. 54.61, эта диаграмма идентична второй (Open, High, Low, Close Chart) с ежедневным объемом продаж в столбик и вспомогательной осью.



**Рис. 54.59.** Диаграмма «Курс открытия, самый высокий курс, самый низкий курс, курс закрытия»



**Рис. 54.60.** Диаграмма «Объем, самый высокий курс, самый низкий курс, курс закрытия»



**Рис. 54.61.** Диаграмма «Объем, курс открытия, самый высокий курс, самый низкий курс, курс закрытия»

## Задания

В файле **Cakes.xlsx** содержится информация о количестве продавцов и выручке пекарни за каждый месяц. На основе этих данных выполните задания 1–4.

1. Создайте комбинированную диаграмму со вспомогательной осью для количества продавцов и выручки за каждый месяц.
2. Постройте диаграмму ежемесячной выручки и вставьте в диаграмму подписи данных.
3. Постройте диаграмму для количества продавцов и вставьте под диаграмму таблицу данных.
4. Нанесите на диаграмму ежемесячных продаж в качестве столбиков изображение торта вместо стандартной гистограммы.

В файле **Hiddenpivot.xlsx** содержатся данные о продажах конфет в 2009–2017 гг. На основе этих данных выполните задания 5 и 6.

5. Отфильтруйте данные так, чтобы на листе остались только продажи за 2013 г, а на диаграмме были представлены все данные.
6. Поместите продажи за каждый год в сводную таблицу и покажите процентный рост продаж с каждым годом.
7. На основе данных в файле **Salestracker.xlsx** выделите двух лучших продавцов каждого месяца стрелками вверх, двух худших — стрелками вниз, а остальных — стрелками вправо.
8. В баскетбольной команде Университета Индианы хорошей эффективностью считается не менее 50 блокировок мяча за игру, а низкой эффективностью — 30 и менее блокировок. В последних шести играх этот показатель у Индианы составил 25, 55, 45, 43, 59 и 39 единиц. Представьте эти данные на ленточной диаграмме.
9. Представьте данные из задания 8 на гистограмме, в которой удачные игры в защите выделены зеленым цветом, неудачные игры — красным, а остальные игры — оранжевым.
10. В течение первого года работы доходы фирмы составили \$5 млн, а расходы — \$6 млн. Для различных предполагаемых темпов роста доходов и расходов составьте диаграмму доходов и расходов на 10 ближайших лет и поместите в название диаграммы год, в который доходы впервые превысят расходы. Пусть заголовок графика показывает этот год.

В файле **Crimedata.xlsx** содержатся ежегодные данные о насильственных и имущественных преступлениях и об убийствах в США. На основе этих данных выполните задания 11 и 12.



11. Создайте диаграмму, отображающую количество преступлений для каждого года. Для управления показом рядов данных примените флажки.
12. Для данных об убийствах создайте гистограмму обновляющуюся автоматически при добавлении новых данных.
13. В таблице ниже содержатся отчеты о работе пяти сотрудников. Представьте эти данные на лепестковой диаграмме.

Сотрудник	Трудолюбие	Работа в команде	Выполнение заданий в срок	Пунктуальность
Уэйн	1	2	3	4
Вивиан	5	6	7	8
Грег	10	9,5	9	8,5
Джен	9	2	9	4
Ванда	1	1,5	2	2,5

14. Проект состоит из пяти задач. Ниже даны время начала и продолжительность выполнения каждой задачи. Представьте эти данные на диаграмме Ганта.

Задача	Время начала	Продолжительность
A	0	4
B	3	6
C	5	7
D	6	8
E	3	6

15. Ниже приведена прибыль компании, а также доходы и расходы в разбивке по годам. Представьте эти данные на диаграмме «водопад».

Статья	Сумма
Доходы	\$300 000
Прибыль	\$65 000
Трудозатраты	\$100 000
Затраты на материал	\$80 000
Накладные расходы	\$55 000

16. В файле **Problem16data.xlsx** находятся данные о продажах в ресторане Mel's Diner в утренние и обеденные часы. Создайте древовидную диаграмму для обобщения этих данных. Подсказка: вам нужно будет агрегировать данные.
17. Используя данные из файла **Problem16data.xlsx**, создайте диаграмму «солнечные лучи», обобщающую продажи ресторана.
18. Создайте четыре типа биржевых диаграмм на основе данных из файла **Dowdata.xlsx**.
19. Каждый год 250 000 старшеклассников играют в футбол за свою школу, 16 000 старшекурсников играют за свой колледж, а 256 старшекурсников рекрутируются в НФЛ. Обобщите эти данные при помощи подходящей диаграммы.

## ГЛАВА 55

# Оценка линейных зависимостей

### Обсуждаемые вопросы

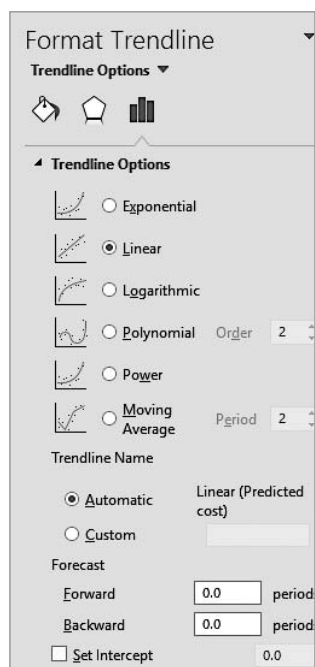
- Как определить зависимость между ежемесячным производством и ежемесячными эксплуатационными расходами?
- Насколько точно эта зависимость объясняет ежемесячные колебания эксплуатационных расходов предприятия?
- Насколько точны мои прогнозы?
- Какие функции при оценке линейной зависимости позволяют получить угловой коэффициент и начальную ординату прямой, точнее всего соответствующие данным?

Вы — директор завода, производящего небольшие холодильники. Головной офис спускает сверху данные о том, сколько холодильников необходимо производить каждый месяц. Для составления бюджета требуется спрогнозировать ежемесячные эксплуатационные расходы и получить ответы на вопросы, обсуждаемые в данной главе.

Любой бизнес-аналитик должен иметь возможность оценить взаимозависимость важных коммерческих показателей (переменных). В Excel определить зависимость между двумя переменными часто позволяет кривая тренда, рассматриваемая в данной главе, а также в главах 56 «Моделирование экспоненциального роста» и 57 «Степенная кривая». Переменная, значения которой пытаются предсказать аналитики, называется *зависимой переменной*. Переменная, используемая для построения прогноза, называется *независимой переменной*. В следующей таблице приведено несколько примеров зависимостей, которые можно оценить.

Независимая переменная	Зависимая переменная
Количество продукции, произведенное заводом в месяц	Ежемесячные эксплуатационные расходы завода
Сумма, затраченная на рекламу в месяц	Продажи за месяц
Количество сотрудников	Ежегодные командировочные расходы
Доход компании	Количество сотрудников (в штате)
Ежемесячный доход на фондовом рынке	Доход от акций (например, Dell)
Площадь дома	Стоимость дома

Первым шагом в определении зависимости между двумя переменными является составление графика для точек данных (точечной диаграммы) с независимой переменной по оси  $x$  и зависимой переменной по оси  $y$ . Затем следует выделить график, щелкнуть по точке данных (все точки данных отображаются синим цветом), на вкладке Конструктор (Design) в группе Макеты диаграмм (Chart Layouts) выбрать Добавить элемент диаграммы (Add Chart Element), выбрать в списке Линия тренда (Trendline) и затем Дополнительные параметры линии тренда (More Trendline Options) или щелкнуть правой кнопкой мыши и выбрать Добавить линию тренда (Add Trendline). Появится диалоговое окно Формат линии тренда (Format Trendline) — рис. 55.1.



**Рис. 55.1.** Диалоговое окно Формат линии тренда

Если на диаграмме видно, что точки данных расположены примерно вдоль прямой линии, необходимо установить переключатель в положение **Линейная (Linear)**. Если диаграмма показывает, что зависимая переменная возрастает более быстрыми темпами, возможно, переключатель стоит установить в положение **Экспоненциальная (Exponential)** или **Степенная (Power)**. Если из диаграммы следует, что зависимая переменная возрастает (или убывает) с уменьшающейся скоростью, вероятно, опция **Степенная (Power)** окажется наиболее подходящей.

В этой главе рассматривается **Линейная зависимость**. В главе 56 обсуждается **Экспоненциальная кривая**, а в главе 57 — **Степенная**. В главе 64 «Скользящие средние

для временных рядов» проанализирована Линейная фильтрация (Moving Average), а в главе 89 «Ценообразование продуктов с помощью субъективно определяемого спроса» описывается Полиномиальная кривая. (Логарифмическая кривая не представляет интереса для описываемых задач, а потому здесь не рассматривается.)

## Ответы на вопросы

### ❓ Как определить зависимость между ежемесячным производством и ежемесячными эксплуатационными расходами?

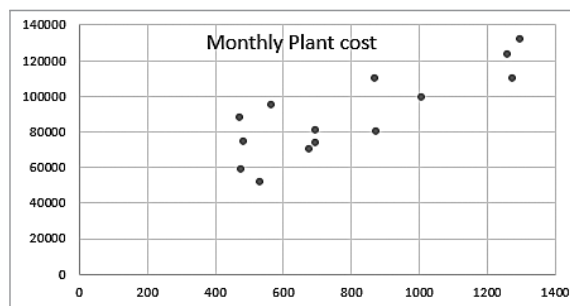
В файле Costestimate.xlsx (рис. 55.2) содержатся данные о количестве произведенной продукции и ежемесячных эксплуатационных расходах завода для периода в 14 месяцев. Вы хотите спрогнозировать ежемесячные эксплуатационные расходы на основе количества произведенного товара, что позволит вам как директору завода составить текущий бюджет и лучше разобраться в том, какие затраты нужны на производство холодильников.

	B	C	D	E	F
1				sum errors	-0.0304
2	Month	Units Produced	Monthly Plant cost	Predicted cost	Error
3	1	1260	123118	118872.66	4245.342
4	2	1007	99601	102612.68	-3011.68
5	3	1296	132000	121186.33	10813.67
6	4	873	80000	94000.671	-14000.7
7	5	532	52000	72085.044	-20085
8	6	476	58625	68485.997	-9861
9	7	482	74624	68871.609	5752.391
10	8	1273	110000	119708.15	-9708.15
11	9	692	81000	82368.036	-1368.04
12	10	690	73507	82239.499	-8732.5
13	11	564	95024	74141.642	20882.36
14	12	470	88004	68100.385	19903.62
15	13	675	70000	81275.468	-11275.5
16	14	870	110253	93807.865	16445.14
17	15	1100		108589.67	

**Рис. 55.2.** Эксплуатационные данные завода

Начнем строить диаграмму XY (или точечную диаграмму), отображающую независимую переменную (количество произведенной продукции) по оси  $x$  и зависимую переменную (ежемесячные эксплуатационные расходы) по оси  $y$ . Столбец данных по оси  $x$  должен располагаться слева от столбца данных, показываемых на оси  $y$ . Для создания диаграммы выделите данные в диапазоне C2:D16 (включая метки в ячейках C2 и D2). На вкладке Вставка (Insert) в группе Диаграммы (Charts) выберите Точечная или пузырьковая диаграмма (Insert Scatter (X, Y) Or Bubble Chart) и далее первый вариант — Точечная (с маркерами) (Scatter With Only Markers). Появится диаграмма, представленная на рис. 55.3.

Глядя на эту диаграмму, можно предположить, что зависимость между количеством произведенной продукции и ежемесячными эксплуатационными расходами



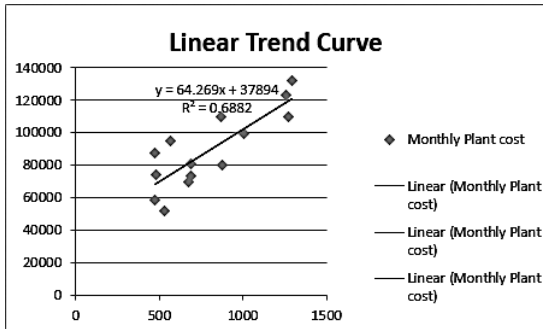
**Рис. 55.3.** Точечная диаграмма эксплуатационных расходов в противопоставлении количеству произведенной продукции

можно представить в виде прямой линии (линейная зависимость). Щелкните в пределах диаграммы и выберите какую-либо точку данных. Все точки данных отображаются синим цветом. Щелкните по точке правой кнопкой мыши и выберите **Добавить линию тренда** (Add Trendline). В диалоговом окне **Формат линии тренда** (Format Trendline) установите переключатель в положение **Линейная** (Linear) и затем установите флажки **Показывать уравнение на диаграмме** (Display Equation on chart) и **Поместить на диаграмму величину достоверности аппроксимации ( $R^2$ )** (Display R-squared value on chart) — рис. 55.4.

**Рис. 55.4.** Выбор параметров линии тренда

Закройте диалоговое окно и добавьте те такие элементы диаграммы, как названия осей и название диаграммы. Результат представлен на рис. 55.5. Названия диаграммы и осей можно добавить, если на вкладке **Конструктор (Работа с диаграммами)** (Chart Tools Design) выбрать **Добавить элемент диаграммы** (Add Chart Element) и затем **Названия осей** (Axis Titles) и **Название диаграммы** (Chart Title).

Если вы хотите добавить больше цифр после запятой в уравнении, выберите уравнение линии тренда и на панели **Формат подписи линии тренда** (вы попадаете туда из контекстного меню) выберите в списке категорий числовой формат и укажите для него параметр **Число десятичных знаков**.



**Рис. 55.5.** Завершенная линия тренда

Каким образом Excel определяет наиболее подходящую линию? Excel выбирает линию (из множества линий, которые могут быть проведены) с минимальной суммой квадратов расстояний по вертикали от каждой точки до линии. Расстояние по вертикали от каждой точки до линии называется *ошибкой*, или *отклонением*. Линия, создаваемая Excel, называется *линией наименьших квадратов*. Минимизируется, как правило, сумма квадратов отклонений, а не сумма отклонений, поскольку при простом суммировании отклонений положительные и отрицательные отклонения компенсируют друг друга. Например, при сложении отклонений точка на 100 единиц выше линии и точка на 100 единиц ниже линии компенсируют друг друга. Но если используются квадратичные отклонения, то при поиске наиболее подходящей линии в Excel будут учтены отклонения для обеих этих точек.

Таким образом, Excel вычисляет наиболее подходящую прямую линию для прогнозирования ежемесячных эксплуатационных расходов на основе ежемесячного количества производимой продукции по формуле:

прогнозируемые ежемесячные эксплуатационные расходы =  $37\,894,0956 + 64,2687 \times$   
 $\times$  (количество произведенной продукции).

Скопировав формулу  $=64.2687 \times C3 + 37894.0956$  из E3 в E4:E17, вы вычислите прогнозируемые расходы для каждой точки наблюдения данных. Например, для 1260 единиц произведенной продукции прогнозируемые эксплуатационные расходы составят \$118 872,66 (см. рис. 55.2).

Вам не следует использовать линию наименьших квадратов для прогнозирования значений зависимой переменной за пределами диапазона имеющихся данных. В этом примере можно предсказывать ежемесячные эксплуатационные расходы завода только для месяцев, в которых производилось приблизительно от 450 до 1300 единиц продукции.

Начальную ординату этой прямой (равную \$37 894,0956) можно интерпретировать как ежемесячные постоянные расходы. Таким образом, даже если завод не произведет ни одного холодильника за месяц, из диаграммы следует, что он по-

несет расходы в размере \$37 894,10. Угловой коэффициент (64,2687) этой прямой показывает, что каждый дополнительно произведенный холодильник увеличивает ежемесячные расходы на \$64,27. Следовательно, переменные расходы на производство холодильника оцениваются в \$64,27.

В ячейках F3:F16 я вычислил ошибки (или отклонения) для каждой точки данных. Ошибку для каждой точки данных я определил как величину, на которую эта точка отстоит от линии наименьших квадратов. Для каждого месяца ошибка равна разности фактических и прогнозируемых расходов. Ошибку для каждой точки данных я вычислил, скопировав формулу  $=D3-E3$  из F3 в F4:F16. Положительное значение ошибки указывает, что точка находится над линией наименьших квадратов, а отрицательное — что она находится под этой линией. Сумма ошибок вычислена в ячейке F1 и равна -0,0304. В действительности для любой линии наименьших квадратов сумма ошибок должна быть равна 0. (-0,0304 было получено, поскольку величины в уравнении были округлены до четырех знаков после запятой.) Факт равенства суммы ошибок нулю означает, что линия наименьших квадратов обладает интуитивно удовлетворительным свойством разделения точек на два подмножества.

### ❓ Насколько точно эта зависимость объясняет ежемесячные колебания эксплуатационных расходов предприятия?

Очевидно, что каждый месяц и эксплуатационные расходы, и количество произведенной продукции меняются. Встает вопрос: какой процент ежемесячных изменений эксплуатационных расходов объясняется ежемесячными изменениями количества производимой продукции? Ответ на него дает значение  $R^2$  (0,69 на рис. 55.5). Можно утверждать, что линейная зависимость объясняет 69% изменений ежемесячных эксплуатационных расходов. Это означает, что 31% изменений ежемесячных эксплуатационных расходов объясняется другими факторами. Эти другие факторы, влияющие на эксплуатационные расходы, можно попытаться определить с помощью множественной регрессии (см. главы 59–61).

Меня часто спрашивают, какое значение  $R^2$  является хорошим. Точного ответа на этот вопрос не существует. Конечно, для одной независимой переменной большее значение  $R^2$  означает наилучшее приближение данных, чем меньшее значение  $R^2$ . Лучшей мерой точности прогнозов является *стандартная ошибка регрессии*, которую я опишу в следующем разделе.

### ❓ Насколько точны мои прогнозы?

Когда вы подбираете линию к точкам, то получаете стандартную ошибку регрессии, которой измеряется *разброс* точек вокруг линии наименьших квадратов. Стандартная ошибка, связанная с линией наименьших квадратов, может быть вычислена посредством функции СТОШУХ (STEYX). Синтаксис функции — СТОШУХ(известные\_значения\_y; известные\_значения\_x), где известные\_значения\_y содержит значения зависимой переменной, а известные\_значения\_x — значения независимой переменной. Стандартная ошибка линии оценки расходов вычислена



в файле **Costestimate.xlsx** в ячейке K1 по формуле **=СТОШУХ(D3:D16;C3:C16)**. Результат представлен на рис. 55.6.

Приблизительно 68% точек должны находиться в пределах одной стандартной ошибки регрессии от линии наименьших квадратов и примерно 95% точек — в пределах двух стандартных ошибок регрессии. Эти меры напоминают правило описательной статистики из главы 43 «Обобщение данных на гистограммах и диаграммах Парето». В данном примере абсолютное значение примерно 68% ошибок не должно превышать \$13 772, а абсолютное значение примерно 95% ошибок не должно превышать \$27 544 (или  $2 \times 13\,772$ ). В столбце F, как видите, 10 из 14, или 71% точек находится в пределах одной стандартной ошибки регрессии от линии наименьших квадратов и все (100%) точки — в пределах двух стандартных ошибок регрессии. Любая точка, находящаяся более чем в двух стандартных ошибках от линии наименьших квадратов, называется *выбросом*. Поиск причин выбросов поможет вам повысить эффективность управления компаниями. Например, месяц, в котором фактические эксплуатационные расходы на \$30 000 превысили плановые, будет считаться выбросом вверх. Если удастся установить причину такого выброса и не допустить его повторения, эффективность производства определенно повысится. Все то же самое касается месяца, в котором фактические расходы оказались на \$30 000 ниже плановых. Если найти причину такого стремительного уменьшения расходов и обеспечить его частое повторение, эффективность производства возрастет.

	H	I	J	K	L
1	slope	64.2687	std err	13771.85	=STEYX(D3:D16,C3:C16)
2	intercept	37894.1	=INTERCEPT(D3:D16,C3:C16)		
3	RSq	0.688203	=RSQ(D3:D16,C3:C16)		
4	l1 formula				
5	=SLOPE(D3:D16,C3:C16)				

**Рис. 55.6.** Вычисление значений углового коэффициента, начальной ординаты,  $R^2$  и стандартной ошибки регрессии

**?** Какие функции при оценке линейной зависимости позволяют получить угловым коэффициент и начальную ординату прямой, точнее всего соответствующие данным?

Угловой коэффициент и начальную ординату линии наименьших квадратов возвращают соответственно функции Excel **НАКЛОН (SLOPE)** и **ОТРЕЗОК (INTERCEPT)** с аргументами **известные\_значения\_у** и **известные\_значения\_х**. Таким образом, по формуле **=НАКЛОН(D3:D16;C3:C16)** в ячейке I1 (см. рис. 55.6) вычисляется угловой коэффициент (64,27) линии наименьших квадратов. А по формуле **=ОТРЕЗОК(D3:D16;C3:C16)** в ячейке I2 вычисляется начальная ордината (37 894,1) линии наименьших квадратов. Кстати, функция **КВПИРСОН (RSQ)** с аргументами **известные\_значения\_у** и **известные\_значения\_х** возвращает значе-

ние  $R^2$ , связанное с линией наименьших квадратов. Таким образом, по формуле `=КВПИРСОН(D3:D16;C3:C16)` в ячейке I3 вычисляется значение  $R^2$  для линии наименьших квадратов, равное 0,6882.

## Задания

Файл `Delldata.xlsx` содержит ежемесячные доходы для фондового индекса Standard&Poore и акций Dell. Бета-коэффициент акций определяется как угловой коэффициент линии наименьших квадратов, используемой для прогнозирования ежемесячного дохода от акции по ежемесячному доходу от всех операций на рынке. Используйте этот файл для решения задач 1–4.

1. Оцените бета-коэффициент акций Dell.
2. Интерпретируйте смысл бета-коэффициента акций Dell.
3. Вы уверены, что грянет рецессия. В какие акции будете инвестировать, с высоким или низким бета-коэффициентом?
4. Для месяца, в котором рынок растет на 5%, определите диапазон, в котором цена акций Dell вырастет с вероятностью 95%.

Файл `Housedata.xlsx` содержит данные о площади в квадратных футах и продажной цене нескольких домов в Белвью, штат Вашингтон. Используйте этот файл для решения задач 5–7.

5. Вы собираетесь возвести пристройку к дому площадью 500 кв. футов. На сколько дом вырастет в цене?
6. Какой процент изменения цены дома можно объяснить изменением его размера?
7. Дом площадью 3000 кв. футов продается за \$500 000. Соответствует ли такая цена стандартным ценам на недвижимость в Белвью? В чем причина расхождения?
8. Известно, что  $32^\circ$  по Фаренгейту равны  $0^\circ$  по Цельсию и что  $212^\circ$  по Фаренгейту равны  $100^\circ$  по Цельсию. С помощью кривой тренда определите зависимость между градусами Фаренгейта и Цельсия. При создании исходной диаграммы перед нажатием кнопки **Готово (Finish)** необходимо указать, что данные находятся в столбцах, а не в строках, поскольку для двух точек данных в Excel считается, что переменные находятся в разных строках.
9. Файл `Betadata.xlsx` содержит ежемесячный доход для индекса Standard&Poore, а также ежемесячные доходы для акций Cinergy, Dell, Intel, Microsoft, Nortel и Pfizer. Вычислите бета-коэффициент каждой акции.
10. В файле `Electiondata.xlsx` содержатся данные по нескольким выборам: голоса, набранные республиканцами при обычном голосовании (подсчитаны в день выборов), и голоса, набранные республиканцами по бюллетеням для заочного

голосования (подсчитаны после дня выборов). Допустим, на выборах республиканцы набрали 49% голосов в день выборов и 62% голосов по бюллетеням для заочного голосования. Кандидат от демократов крикнул: «Результаты фальсифицированы!» Как вы думаете, правда это или нет?

11. В файле **Oldfaithful.xlsx** содержатся данные о частоте извержений гейзера Старый Служака (Олд Фейтфул) и продолжительности последующего извержения. Допустим, прошло 4 минуты после последнего извержения. Вы на 95% уверены, что продолжительность следующего извержения составит от \_\_\_\_\_ до \_\_\_\_\_ минут.
12. В файле **Dailydow.xlsx** приведены ежедневные значения индекса Доу — Джонса с 1996 по 2010 г. Каково значение  $R^2$  для завтрашнего значения индекса Доу — Джонса, прогнозируемого по сегодняшнему значению? Означает ли такое высокое значение  $R^2$ , что движение рынка легко предсказать?
13. В файле **Problem13data.xlsx** приведены размеры (в квадратных футах) и оценочная стоимость нескольких домов в Ривер Оукс, Хьюстон. Оцените связь между размером дома и его ценой. Какой процент изменения цены дома можно объяснить изменением его размера? Как вы можете оценить, на сколько увеличение размеров дома увеличивает его стоимость?

## ГЛАВА 56

# Моделирование экспоненциального роста

### Обсуждаемый вопрос

- Как моделировать рост доходов компании с течением времени?

Если вы хотите оценить компанию, вам нужно иметь некоторое представление о ее будущих доходах. Несмотря на то что будущее может быть не таким, как прошлое, довольно часто оценочный анализ корпорации начинается с изучения доходов компании в последнее время. Аналитики часто строят кривую тренда, основываясь на предшествовавшей динамике доходов. При построении кривой тренда годы откладываются на оси  $x$  (например, первый год данных Год 1, второй год данных Год 2 и т. д.), а по оси  $y$  откладывается доход компании.

Как правило, зависимость между временем и доходом не является прямолинейной. Напомню, что прямая линия на всем протяжении имеет одинаковый угловой коэффициент. Это означает, что при увеличении независимой переменной (в данном случае года) на 1 прогноз для зависимой переменной (дохода) увеличивается на ту же величину. В большинстве компаний доход каждый год возрастает на определенный постоянный процент. В таком случае по мере роста дохода годовое увеличение дохода также увеличивается. Так, рост дохода на 10% для 1 млн долларов означает рост дохода на \$100 000. Рост дохода на 10% для 100 млн долларов означает рост дохода на 10 млн долларов. Анализ такой ситуации предполагает, что кривая тренда для прогноза дохода должна расти более резко и иметь увеличивающийся угловой коэффициент. Экспоненциальная функция обладает тем свойством, что при увеличении независимой переменной на 1 зависимая переменная увеличивается на такую же величину в процентном отношении. Именно такая зависимость требуется для моделирования роста дохода.

Уравнение экспоненциальной функции:  $y = ae^{bx}$ . Здесь  $x$  — значение независимой переменной (в данном примере это год),  $y$  — значение зависимой переменной (в данном примере это годовой доход). Значение  $e$  (приблизительно 2,7182) — основание натурального логарифма. Если в диалоговом окне **Формат линии тренда (Format Trendline)** установить переключатель в положение **Экспоненциальная (Exponential)**, значения  $a$  и  $b$ , наиболее точно соответствующие данным, будут вычислены автоматически. Обратимся к примеру.

## Ответ на вопрос

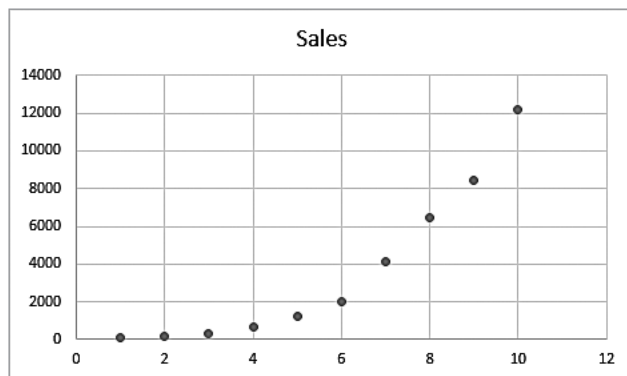
### ? Как моделировать рост доходов компании с течением времени?

Файл Ciscoexp.xls (рис. 56.1) содержит доход для Cisco с 1990 по 1999 г. Доход указан в миллионах долларов. Например, в 1990 г. доход Cisco составил 103,47 млн долларов.

	A	B	C	D
1		Year 1=1990		
2				
3	Year	Sales	Prediction	Ratio
4	1	70	103.471229	
5	2	183	182.848984	1.767148
6	3	340	323.121233	1.767148
7	4	649	571.003071	1.767148
8	5	1243	1009.04699	1.767148
9	6	1979	1783.13546	1.767148
10	7	4096	3151.06443	1.767148
11	8	6440	5568.39749	1.767148
12	9	8459	9840.18301	1.767148
13	10	12154	17389.0606	1.767148
14				
15	16		529558.325	

**Рис. 56.1.** Годовой доход Cisco с 1990 по 1999 г.

Для приближения этих данных к экспоненциальной кривой сначала выделите диапазон A3:B13. Затем на вкладке Вставка (Insert) в группе Диаграммы (Charts) выберите точечную диаграмму, первый вариант Точечная (с маркерами) (Scatter With Only Markers) — рис. 56.2.



**Рис. 56.2.** Точечная диаграмма для кривой тренда Cisco

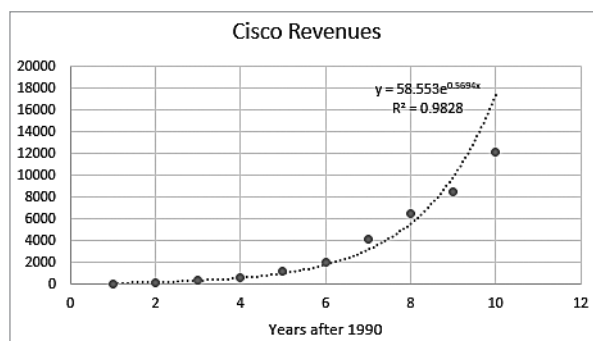
Поскольку угловой коэффициент прямой линии является константой, подбирать прямую линию к таким данным было бы странно. Для стремительно растущего на-клона графика, как в этом случае, больше всего подойдет экспоненциальный рост.

Чтобы получить экспоненциальную кривую, больше всего приближающуюся к данным, щелкните правой кнопкой мыши по точке данных (все точки отображаются синим цветом) и затем выберите **Добавить линию тренда (Add Trendline)**. В диалоговом окне **Формат линии тренда (Format Trendline)** установите переключатель в положение **Экспоненциальная (Exponential)** и отметьте флажки **Показывать уравнение на диаграмме (Display Equation On Chart)** и **Поместить на диаграмму величину достоверности аппроксимации ( $R^2$ ) (Display R-Squared Value On Chart)**. Получившаяся диаграмма представлена на рис. 56.3.

Предполагаемый доход Cisco за год  $x$  (напомним, что для  $x = 1$  это 1990 г.) вычисляется по следующей формуле:

предполагаемый доход =  $58,552664e^{0,569367x}$ .

Я вычислил предполагаемый доход в ячейках **C4:C13**, скопировав формулы  $=58,552664*EXP(0,569367*A4)$  из **C4** в **C5:C13**. Например, прогноз доходов Cisco в 1999 г. (на 10-й год) составляет 17,389 млрд долларов.



**Рис. 56.3.** Экспоненциальная кривая доходов Cisco

Обратите внимание, что большая часть точек расположена очень близко к подобранной экспоненциальной кривой. Это означает, что на этом наборе данных экспоненциальная кривая хорошо отображает рост доходов Cisco в 1990-х гг. Тот факт, что значение  $R^2$  (0,98) близко к 1, также согласуется с визуальным доказательством хорошей подгонки.

Следует помнить, что каждый раз при увеличении  $x$  на 1 значение оценки по экспоненциальной кривой увеличивается на столько же в процентном отношении. Это можно проверить, вычислив отношение предполагаемого дохода в каком-либо году к предполагаемому доходу в предыдущем году. Для этого скопируйте формулу  $=C5/C4$  из **D5** в **D6:D13**. Предполагаемые темпы роста компании Cisco составляют 76,7% в год, что является наилучшей оценкой годовых темпов роста Cisco с 1990 по 1999 г. Такая оценка годовых темпов роста компании в течение некоторого периода называется среднегодовым темпом роста в сложных процентах (Compound Annual Growth Rate, CAGR).

Разумеется, перед использованием этих предполагаемых годовых темпов роста дохода в оценочном анализе вы должны узнать, какова вероятность того, что такой рост удержится какое-то время, ведь экспоненциальный рост не может длиться вечно. Например, если бы вы воспользовались экспоненциальной кривой тренда для прогнозирования доходов в 2005 г. (год 16), то предполагаемый доход Cisco в 2005 г. составил бы 530 млрд долларов. Если бы такая оценка была верна, доход Cisco в 2005 г. в три раза превысил бы доход крупнейшей в мире компании (Walmart). Такое представляется крайне маловероятным. Вывод следующий: в первые годы доход технологической компании растет экспоненциально. Через некоторое время темпы роста падают. Если бы аналитики Уолл-стрит в конце 1990-х гг. осознали этот факт, пузыри доткомов можно было бы избежать. Обратите внимание, что в 1999 г. фактический доход Cisco был ниже дохода, рассчитанного по кривой тренда. Возможно, эти цифры указывали на начало технологического спада, начавшегося в конце 2000 г.

Часто кривая Гомпертца, она же логистическая кривая (также часто называемая *S-кривой*), для новых компаний гораздо больше походит на кривую ежегодного роста выручки, чем на экспоненциальный рост. Это потому, что S-кривые отражают существенное замедление роста выручки. Пространное обсуждение S-кривых можно найти в моей книге «Маркетинговый анализ: техники на основе обработки данных с помощью *Microsoft Excel*» (Wiley, 2014). Кстати, почему необходимо использовать  $x = 1$ , а не  $x = 1990$ ? Если указать  $x = 1990$ , Excel придется обрабатывать числа, подобные  $e^{1990}$ , а это Excel плохо умеет делать.

## Задания

Файл *Exponentialdata.xlsx* содержит годовой доход от продаж компаний Staples, Walmart и Intel. На основе этих данных выполните первые пять заданий:

1. Для каждой компании подберите экспоненциальную кривую тренда к данным продаж.
2. Для какой компании экспоненциальный рост наиболее соответствует росту дохода?
3. Для какой компании экспоненциальный рост менее всего соответствует росту дохода?
4. Для каждой компании вычислите годовые темпы роста дохода в процентах.
5. Для каждой компании спрогнозируйте по кривой тренда доход в 2003 г.
6. В файле *Impalas.xls* приведены цены Impalas 2009, 2008, 2007 и 2006 годов выпуска в 2010 году. Исходя из этих данных, оцените снижение стоимости автомобиля по мере его старения.
7. Файл *Problem7data.xlsx* содержит данные о годовых доходах компании Google в 2005–2013 гг. Используя эти данные, оцените среднегодовой темп роста в сложных процентах для Google и предскажите доход Google за 2014 г.

## ГЛАВА 57

# Степенная кривая

### Обсуждаемый вопрос

- Моя компания наращивает производство. Можно ли смоделировать зависимость между количеством произведенной продукции и временем, необходимым для производства одной единицы продукции?

Степенная кривая вычисляется посредством уравнения  $y = ax^b$ . В этом уравнении  $a$  и  $b$  являются константами. Используя кривую тренда, можно определить значения  $a$  и  $b$  для степенной кривой, наиболее приближенной к точечной диаграмме. В большинстве случаев  $a > 0$ . При этом наклон степенной кривой зависит от значения  $b$  следующим образом:

- для  $b > 1$   $y$  возрастает при увеличении  $x$ , и наклон степенной кривой увеличивается при увеличении  $x$ ;
- для  $0 < b < 1$   $y$  возрастает при увеличении  $x$ , и наклон степенной кривой уменьшается при увеличении  $x$ ;
- для  $b = 1$  степенная кривая представляет собой прямую линию;
- для  $b < 0$   $y$  убывает при увеличении  $x$ , и степенная кривая выпрямляется при увеличении  $x$ .

Далее я приведу примеры различных зависимостей, которые можно смоделировать с помощью степенной кривой. Эти примеры находятся в файле **Powerexamples.xlsx**.

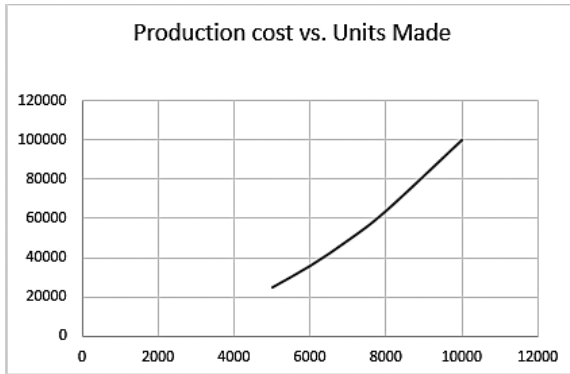
Диаграмма для прогнозируемой общей себестоимости продукции как функции количества произведенной продукции показана на рис. 57.1. Обратите внимание, что  $b = 2$ . Как я уже говорил, для этого значения  $b$  себестоимость продукции растет с количеством произведенной продукции. Наклон становится круче, что указывает на рост себестоимости каждой дополнительно произведенной единицы продукции. Такая зависимость может возникнуть из-за сверхурочных, оплата которых выше, чем оплата обычного рабочего времени.

Если вы пытаетесь понять, как объемы продаж зависят от расходов на рекламу, то можете обнаружить кривую, представленную на рис. 57.2.

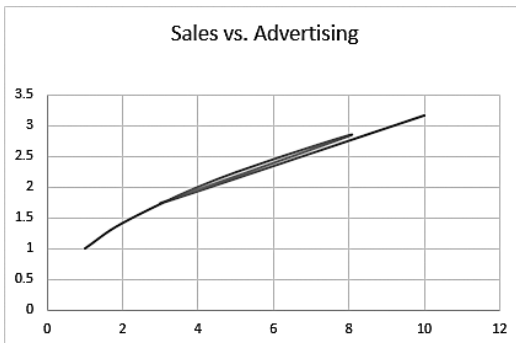
Здесь  $b = 0,5$  (то есть между 0 и 1). Если значение  $b$  находится в этом интервале, по мере увеличения расходов на рекламу объем продаж увеличивается, но с уменьшающейся скоростью. Таким образом, степенная кривая позволяет смодели-



лизовать закон убывающей доходности: каждый дополнительно потраченный на рекламу доллар приносит меньше прибыли.



**Рис. 57.1.** Прогнозируемая себестоимость как функция количества произведенной продукции

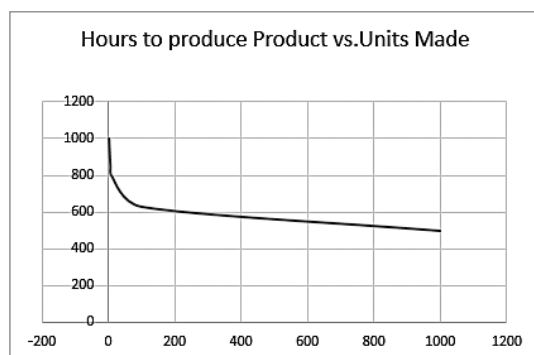


**Рис. 57.2.** Объем продаж как функция расходов на рекламу

Когда вы пытаетесь спрогнозировать время, необходимое для производства последней единицы продукции, на основании количества продукции, произведенной на сегодняшний день, то полезной часто оказывается точечная диаграмма, показана на рис. 57.3.

Здесь  $b = -0,1$ . Поскольку  $b$  меньше 0, время, необходимое для производства каждой единицы продукции, уменьшается, но скорость уменьшения, то есть скорость «обучения», снижается. Эта зависимость означает, что сначала при производстве продукции идет значительная экономия рабочего времени по мере накопления рабочих навыков. Однако при последующем производстве этой продукции рабочее время экономится все меньше. Зависимость между общим количеством произведенной продукции и временем, необходимым для производства единицы про-

дукции, называется *кривой обучения*, или *кривой опыта*. Степенная кривая имеет следующие свойства.



**Рис. 57.3.** Время, необходимое для производства единицы продукции, как функция совокупного количества продукции

- **Свойство 1.** Если  $x$  увеличивается на 1%,  $y$  увеличивается примерно на  $b\%$ .
- **Свойство 2.** Каждый раз, когда  $x$  удваивается,  $y$  увеличивается на тот же процент.

Допустим, спрос на продукт в зависимости от цены можно смоделировать как  $1000 \cdot (\text{Цена})^{-2}$ . Тогда свойство 1 подразумевает, что 1%-ное увеличение цены снизит спрос (независимо от величины цены) на 2%. В этом случае показатель степени  $b$  (без минуса) называется *коэффициентом эластичности*. Коэффициент эластичности рассматривается в главе 87 «Оценка кривой спроса». Теперь, исходя из вышеизложенного, посмотрим, как подобрать степенную кривую к данным.

## Ответ на вопрос

- ❓ **Моя компания наращивает производство. Можно ли смоделировать зависимость между количеством произведенной продукции и временем, необходимым для производства одной единицы продукции?**

В файле *Fax.xlsx* содержатся данные о количестве произведенных факсов и себестоимости (в долларах 1982 г.) производства «последнего» факса для каждого года. Например, в 1983 г. было произведено 70 000 факсов, и себестоимость производства «последнего» факса составила \$3416. Данные представлены на рис. 57.4.

Поскольку кривая обучения прогнозирует либо себестоимость, либо время, затрачиваемое на производство единицы продукции, на основе данных об об-

щем количестве продукции, в столбце С вычисляется общее количество факсов, произведенных к концу каждого года. В ячейке С4 ссылка на ячейку В4 показывает количество факсов, произведенных в 1982 г. Скопировав формулу =СУММ(\$B\$4:B4) из С5 в С6:С10, я вычислил общий объем производства к концу каждого года.

	A	B	C	D	E	F	G
1	Learning curve FAX data						
2							
3	Year	Production	Cumulative Production	Unit Cost	Forecast		
4	1982	64000	64000	\$ 3,700.00	3955.81		
5	1983	70000	134000	\$ 3,416.00	3280.54		
6	1984	100000	234000	\$ 3,125.00	2848.51		
7	1985	150000	384000	\$ 2,583.00	2512.63		
8	1986	175000	559000	\$ 2,166.00	2284.66		
9	1987	400000	959000	\$ 1,833.00	1992.72		
10	1988	785000	1744000	\$ 1,788.00	1712.60		
11	1989	1000000	2744000		1526.85		learning
12		double	3488000		1436.83		percentage
13		cumulative production					0.838975
14							

**Рис. 57.4.** Данные, на основе которых строится кривая обучения для производства факсов

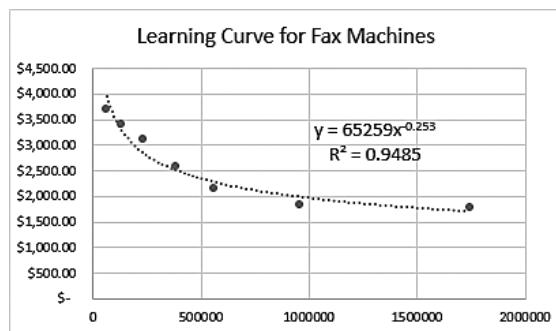
Теперь можно создать точечную диаграмму с совокупным количеством произведенной продукции по оси *x* и себестоимостью единицы продукции по оси *y*. После создания диаграммы щелкните на одной из точек данных (точки данных отображаются синим цветом), а затем щелкните правой кнопкой мыши и выберите **Добавить линию тренда (Add Trendline)**. В диалоговом окне **Формат линии тренда (Format Trendline)** установите переключатель в положение **Степенная (Power)** и отметьте флажки **Показывать уравнение на диаграмме (Display Equation on chart)** и **Поместить на диаграмму величину достоверности аппроксимации (R<sup>2</sup>) (Display R-Squared value on chart)**. Диаграмма с такими параметрами показана на рис. 57.5. Нарисованная кривая представляет собой степенную кривую, больше всего соответствующую данным.

Степенная кривая позволяет прогнозировать себестоимость производства факса следующим образом:

$$\text{себестоимость производства факса} = 65\,259 * (\text{совокупное количество})^{-0,2533}.$$

Обратите внимание, что большая часть точек данных находится возле подобранной степенной кривой и что значение  $R^2$  близко к 1, то есть степенная кривая подогнана должным образом.

Скопируйте формулу =65259\*C4^-0,2533 из ячейки Е4 в Е5:Е10 и вычислите прогнозируемую себестоимость последнего факса, произведенного в конце каждого года. (Знак вставки «^», который находится на клавише «6», означает возведение в степень.)



**Рис. 57.5.** Кривая обучения для производства факсов

Если вы установили, что в 1989 г. было произведено 1 000 000 факсов, то, сначала подсчитав в ячейке C11 общий объем производства к концу 1989 г. (2 744 000 шт.), вы можете скопировать уравнение прогноза в ячейку E11 и получить себестоимость последнего произведенного в 1989 г. факса — \$1526,85.

Напомню 2-е свойство степенной кривой: каждый раз, когда  $x$  удваивается,  $y$  увеличивается на тот же процент. Введите удвоенный объем производства в 1988 г. в ячейку C12 и скопируйте формулу из E10 в E12. Оказывается, удвоение объема производства снижает прогнозируемую себестоимость до 83,9% от предыдущего значения (1436,83/1712,60). Поэтому текущая кривая обучения называется 84%-ной кривой обучения. Каждый раз при удвоении объема производства трудовые затраты на производство факса уменьшаются на 16,1%.

Если кривая становится круче, данным может соответствовать не только степенная кривая, но и экспоненциальная. И какая же из них больше подходит к данным? В большинстве случаев ответ получить просто: сравните на глаз кривые и выберите ту, которая кажется более подходящей. Для более точной оценки можно вычислить *сумму квадратичных ошибок* для каждой кривой (сложив квадраты разности значений на кривой и фактических значений для каждой точки данных) и выбрать кривую с меньшей суммой квадратичных ошибок.

Кривая обучения была изобретена в 1936 г. на базе военно-воздушных сил Райт-Паттерсон (*Wright-Patterson Air Force Base*) в Дейтоне, штат Огайо, когда было установлено, что каждый раз при удвоении совокупного количества произведенных самолетов время, затраченное на производство одного самолета, уменьшалось примерно на 15%.

В «Википедии» можно найти оценки кривой обучения для различных отраслей промышленности:

- авиационно-космическая промышленность: 85%;
- судостроение: 80–85%;
- сложные станки, для новых моделей: 75–85%;

- повторяющееся производство электроники: 90–95%;
- повторяющиеся производственные операции или штамповочные операции: 90–95%;
- повторяющиеся электромонтажные работы: 75–85%;
- повторяющиеся сварочные работы: 90%;
- сырье: 93–96%;
- покупные детали: 85–88%.

## Задания

- На основе примера по производству факсов смоделируйте зависимость между общим объемом производства и общей себестоимостью произведенной продукции.
- На основе примера по производству факсов смоделируйте зависимость между общим объемом производства и средней себестоимостью одного факса.
- Коммерческий директор считает, что общий объем продаж продукта зависит от цены так, как показано в таблице ниже. Определите зависимость между ценой и спросом, а также прогнозируемым спросом для цены \$46. На сколько процентов уменьшается спрос при увеличении цены на 1%?

Цена	Спрос
\$30,00	300
\$40,00	200
\$50,00	110
\$60,00	60

- Бренд-менеджер нового лекарственного препарата считает, что годовой объем продаж препарата является функцией количества рекламных звонков (см. таблицу). Определите объем продаж препарата, если врачам сделано 80 000 рекламных звонков.

Коммерческие звонки	Количество проданного препарата
50 000	25 000
100 000	52 000
150 000	68 000
200 000	77 000

5. Время, затрачиваемое на производство каждого из первых десяти самолетов, приведено в следующей таблице.

Самолет	Часы
1	1000
2	800
3	730
4	630
5	600
6	560
7	560
8	500
9	510
10	510

Определите общее количество часов, необходимое для производства следующих десяти самолетов.

# Представление зависимостей с помощью корреляции

Обсуждаемый вопрос

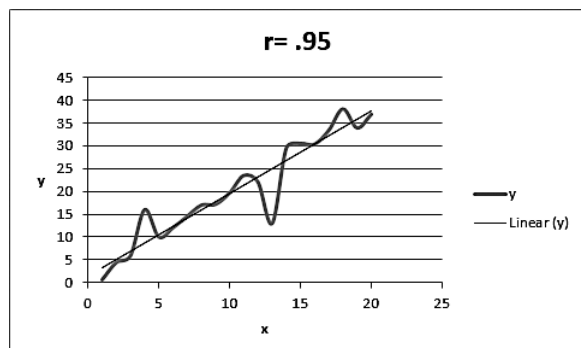
- Каким образом связаны ежемесячные доходы от акций Microsoft, GE, Intel, GM и Cisco?

Кривые тренда помогают понять, как две переменные связаны между собой. Однако мы часто имеем дело более чем с двумя переменными. Вычисление корреляции между любыми парами переменных позволяет увидеть совместное движение вверх и вниз значений нескольких переменных.

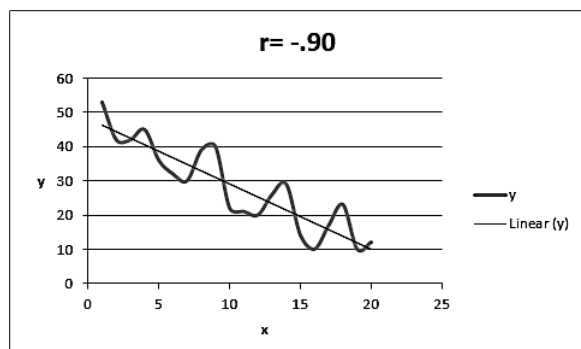
Корреляция (обычно обозначаемая  $r$ ) между двумя переменными (назовем их  $x$  и  $y$ ) — это мера степени линейной зависимости между  $x$  и  $y$ . Коэффициент корреляции между любыми двумя переменными всегда лежит в интервале от  $-1$  до  $+1$ . При этом важна не столько точная формула вычисления коэффициента корреляции между двумя переменными, сколько возможность интерпретации корреляции переменных.

Коэффициент корреляции, близкий  $+1$ , означает, что между  $x$  и  $y$  имеется сильная положительная линейная зависимость. То есть если  $x$  больше среднего, то  $y$  также имеет тенденцию превышать среднее, и если  $x$  меньше среднего,  $y$  также, как правило, будет меньше среднего. Если для точек данных подбирается прямая линия, то это будет прямая линия с положительным угловым коэффициентом, хорошо соответствующая точкам данных. Например, для данных, представленных на рис. 58.1 ( $x$  — количество произведенной продукции, а  $y$  — ежемесячная себестоимость продукции), коэффициент корреляции между  $x$  и  $y$  равен  $+0,9$ . (Для рис. 58.1–58.3 см. файл `Correlationexamples.xlsx`.)

Коэффициент корреляции, близкий  $-1$ , означает наличие между  $x$  и  $y$  сильной отрицательной линейной зависимости, то есть если  $x$  больше среднего,  $y$ , скорее всего, меньше среднего, и если  $x$  меньше среднего,  $y$  имеет тенденцию превышать среднее. Если для данных подбирается прямая линия, то она имеет отрицательный угловой коэффициент и хорошо соответствует точкам данных. Например, для данных, представленных на рис. 58.2, коэффициент корреляции между  $x$  и  $y$  равен  $-0,9$ .



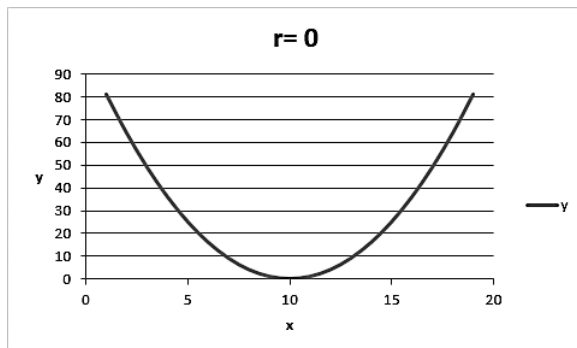
**Рис. 58.1.** Коэффициент корреляции, близкий  $+1$ , указывает на сильную положительную линейную зависимость двух переменных



**Рис. 58.2.** Коэффициент корреляции, близкий  $-1$ , указывает на сильную отрицательную линейную зависимость двух переменных

Коэффициент корреляции, приблизительно равный 0, свидетельствует о слабой линейной зависимости между  $x$  и  $y$ . Другими словами, на основании информации о том, больше  $x$  или меньше среднего, невозможно предположить, больше  $y$  своего среднего или меньше. На рис. 58.3 показан график зависимости объемов продаж в штуках ( $y$ ) от опыта продаж в годах ( $x$ ). Опыт продаж в годах и объем продаж в штуках имеют коэффициент корреляции 0,003. Для этого набора данных средний опыт равен 10 годам. Как видите, у продавцов со стажем более 10 лет объемы продаж могут быть низкими или высокими. Также можно отметить, что у продавцов со стажем менее 10 лет объемы продаж также могут быть низкими либо высокими. Несмотря на то что между годами опыта и объемами продаж существует незначительная линейная зависимость или линейная зависимость вовсе отсутствует, между годами опыта и объемами продаж имеется сильная нелинейная зависимость (см. подобранную кривую). Коэффициент корреляции не является мерой степени нелинейной зависимости.





**Рис. 58.3.** Коэффициент корреляции, близкий 0, указывает на слабую линейную зависимость между двумя переменными

## Ответ на вопрос

❓ **Каким образом связаны ежемесячные доходы от акций Microsoft, GE, Intel, GM и Cisco?**

В файле Stockcorrel.xlsx (рис. 58.4) содержатся ежемесячные доходы от акций Microsoft, GE, Intel, GM и Cisco за 1990-е гг. Вы можете использовать корреляции, чтобы попытаться понять, как связаны между собой движения этих акций.

Для нахождения корреляции (лист Initial Correlation Matrix) между каждой парой акций на вкладке Данные (Data) в группе Анализ (Analysis) сначала выберите Анализ данных (Data Analysis), а затем в списке — инструмент Корреляция (Correlation). (Перед использованием этого инструмента необходимо установить пакет анализа. Для этого выберите вкладку ФАЙЛ (File), нажмите Параметры (Options) и затем выберите Надстройки (Add-Ins) в левой панели. В списке Управление (Manage) выберите Надстройки Excel (Excel Add-Ins) и нажмите Перейти (Go). В открывшемся диалоговом окне установите флажок Пакет анализа (Analysis ToolPak) и нажмите ОК. Примеры и подробные инструкции см. в главах 43 «Обобщение данных на гистограммах и диаграммах Парето» и 44 «Обобщение данных с помощью описательной статистики».) Нажмите ОК и заполните диалоговое окно Корреляция (Correlation), как на рис. 58.5.

Самый простой способ задать исходный диапазон состоит в следующем: выберите левый верхний угол диапазона (B51), нажмите Ctrl+Shift+→, а затем Ctrl+Shift+↓. Первая строка диапазона исходных данных содержит метки, поэтому установите флажок Метки в первой строке (Labels in first row). Введите ячейку H52 как левую верхнюю ячейку диапазона выходных данных и нажмите ОК. Полученный результат представлен на рис. 58.6.

Например, коэффициент корреляции между акциями Cisco и Microsoft составляет 0,513, а коэффициент корреляции между акциями GM и Microsoft — 0,069. Анализ

	A	B	C	D	E	F
51	Date	MSFT	GE	INTC	GM	CSCO
52	3/30/1990	0.122	0.040	0.037	0.022	0.011
53	4/30/1990	0.047	-0.004	-0.054	-0.035	0.011
54	5/31/1990	0.259	0.084	0.222	0.116	0.042
55	6/29/1990	0.041	0.005	-0.026	-0.021	0.071
56	7/31/1990	-0.125	0.034	-0.053	-0.021	-0.038
57	8/31/1990	-0.075	-0.134	-0.250	-0.131	-0.029
58	9/28/1990	0.024	-0.113	-0.004	-0.088	-0.091
59	10/31/1990	0.012	-0.046	0.008	0.014	0.311
60	11/30/1990	0.133	0.053	0.119	0.014	0.339
61	12/31/1990	0.042	0.057	0.027	-0.058	0.136
62	1/31/1991	0.304	0.115	0.188	0.055	0.304
63	2/28/1991	0.057	0.070	0.044	0.101	-0.043
64	3/28/1991	0.023	0.024	-0.021	-0.044	-0.129
65	4/30/1991	-0.067	0.016	0.053	-0.053	0.221
66	5/31/1991	0.109	0.099	0.132	0.217	0.084
67	6/28/1991	-0.069	-0.042	-0.166	-0.055	-0.054
68	7/31/1991	0.079	-0.010	0.011	-0.025	0.287
69	8/30/1991	0.160	0.022	0.053	-0.034	0.156
70	9/30/1991	0.044	-0.067	-0.146	-0.016	-0.096
71	10/31/1991	0.055	-0.005	-0.038	-0.060	0.189
72	11/29/1991	0.036	-0.062	0.009	-0.113	0.015
73	12/31/1991	0.144	0.190	0.195	-0.061	0.338

Рис. 58.4. Ежемесячный доход от акций за 1990-е гг.

Correlation ? X

Input

Input Range:

Grouped By: ☒ Columns ☐ Rows

☒ Labels in first row

Output options

☒ Output Range:

☐ New Worksheet Ply:

☐ New Workbook

OK Cancel Help

Рис. 58.5. Диалоговое окно Корреляция

	H	I	J	K	L	M	N
51	Full correlation matrix						
52	Date	MSFT	GE	INTC	GM	CSCO	
53	Date	1.000	-0.101	0.073	-0.019	0.011	-0.126
54	MSFT	-0.101	1.000	0.445	0.517	0.069	0.513
55	GE	0.073	0.445	1.000	0.324	0.380	0.376
56	INTC	-0.019	0.517	0.324	1.000	0.317	0.488
57	GM	0.011	0.069	0.380	0.317	1.000	0.159
58	CSCO	-0.126	0.513	0.376	0.488	0.159	1.000

Рис. 58.6. Корреляционные связи между доходами от акций

показывает, что доходы от акций Cisco, Intel и Microsoft наиболее тесно связаны между собой. Поскольку коэффициент корреляции между каждой парой этих акций приблизительно равен 0,5, эти акции демонстрируют умеренную положительную зависимость. Другими словами, если одни акции приносят доход выше средне-

го, вполне вероятно (но необязательно), что другие акции также принесут доход выше среднего. Поскольку доходы от акций Cisco, Intel и Microsoft тесно связаны с капиталовложениями в технологии, их довольно сильная корреляция не является неожиданной. Из рисунка видно, что ежемесячные доходы от акций Microsoft и GM практически не коррелируют. То есть если акции Microsoft поднимутся выше обычного, это не поможет нам узнать, «выступят» ли акции GM лучше или хуже. Опять же это неудивительно, поскольку GM не является высокотехнологичной компанией и поэтому более восприимчива к колебания делового цикла.

## Заполнение корреляционной матрицы

Как вы видите в этом примере, Excel оставил некоторые ячейки корреляционной матрицы пустыми. Например, коэффициент корреляции между акциями Microsoft и GE (равный коэффициенту корреляции между акциями GE и Microsoft) опущен. Если вы хотите заполнить всю корреляционную матрицу, выделите матрицу, щелкните правой кнопкой мыши и выберите команду **Копировать (Copy)**. Щелкните правой кнопкой мыши по свободной части листа и выберите **Специальная вставка (Paste Special)**. В диалоговом окне **Специальная вставка (Paste Special)** установите флажок **Транспонировать (Transpose)**. Строки и столбцы данных поменяются местами. Теперь выделите транспонированную матрицу, щелкните правой кнопкой мыши и выберите команду **Копировать (Copy)**. Выделите исходную корреляционную матрицу, щелкните правой кнопкой мыши и снова выберите **Специальная вставка (Paste Special)**. В диалоговом окне **Специальная вставка (Paste Special)** установите флажок **Пропускать пустые ячейки (Skip Blanks)** и нажмите **ОК**. Транспонированные данные будут скопированы в исходную матрицу, за исключением пустых ячеек. Полная корреляционная матрица показана на рис. 58.7.

	H	I	J	K	L	M	N
48		correl function					
49		0.159					
50							
51		Full correlation matrix					
52		Date	MSFT	GE	INTC	GM	CSCO
53	Date	1.000	-0.101	0.073	-0.019	0.011	-0.126
54	MSFT	-0.101	1.000	0.445	0.517	0.069	0.513
55	GE	0.073	0.445	1.000	0.324	0.380	0.376
56	INTC	-0.019	0.517	0.324	1.000	0.317	0.488
57	GM	0.011	0.069	0.380	0.317	1.000	0.159
58	CSCO	-0.126	0.513	0.376	0.488	0.159	1.000
59							
60							
61		1.000	-0.101	0.073	-0.019	0.011	-0.126
62			1.000	0.445	0.517	0.069	0.513
63				1.000	0.324	0.380	0.376
64					1.000	0.317	0.488
65						1.000	0.159
66							1.000
67							
68		Transposed correlations					

Рис. 58.7. Полная корреляционная матрица

## Функция КОРРЕЛ

В качестве альтернативы инструменту Корреляция (Correlation) из пакета анализа можно использовать функцию КОРРЕЛ (CORREL). Например, введя в ячейку I49 формулу =КОРРЕЛ(E52:E181;F52:F181) (на листе Complete Matrix), мы также получим коэффициент корреляции ежемесячного дохода от акций Cisco (показанного в столбце F) и дохода от акций GM (показанного в столбце E), равный 0,159.

## Связь между коэффициентом корреляции и $R^2$

В главе 55 «Оценка линейных зависимостей» вы вычислили значение  $R^2$  для количества произведенной продукции и ежемесячных эксплуатационных расходов — 0,688. Как это значение связано с коэффициентом корреляции количества произведенной продукции и ежемесячных эксплуатационных расходов? Корреляционные связи между двумя наборами данных — это просто следующая формула для линии тренда:

$$\sqrt{\text{значение } R^2}.$$

В этой формуле вы выбираете знак перед квадратным корнем такой же, как знак наклона линии тренда. Таким образом, коэффициент корреляции количества произведенной продукции и ежемесячных эксплуатационных расходов для данных из главы 53 равен  $\sqrt{+0,688} = +0,829$ .

## Корреляция и регрессия к среднему

Возможно, вам знакомо выражение «регрессия к среднему». По сути оно означает, что прогнозируемое значение зависимой переменной будет в некотором смысле ближе к среднему значению этой переменной, чем независимая переменная. Пусть вам требуется спрогнозировать зависимую переменную  $y$  по независимой переменной  $x$ . Если  $x$  равно  $k$  стандартных отклонений от среднего, прогнозом для  $y$  будет  $rk$  стандартных отклонений от среднего. (Здесь  $r$  — коэффициент корреляции между  $x$  и  $y$ .) Поскольку значение  $r$  находится в пределах между  $-1$  и  $+1$ , это означает, что  $y$  имеет меньшее стандартное отклонение от среднего, чем  $x$ . Это истинное определение *регрессии к среднему*. Интересное применение регрессии к среднему см. в задании 5.

## Задания

Данные к заданиям 1–3 находятся в файле Ch56data.xlsx.

1. На листе Problem 1 указано количество автомобилей, каждый день запаркованных как на открытой, так и закрытой автостоянке возле школы Университе-

та Индианы. Вычислите коэффициент корреляции и объясните корреляцию между количеством автомобилей, запаркованных на открытой и закрытой стоянках.

2. На листе **Problem 2** указаны ежедневные объемы продаж (в долларах) для лазерных принтеров, картриджей и канцелярских товаров. Вычислите коэффициенты корреляции и объясните корреляцию между этими величинами.
3. На листе **Problem 3** указаны годовые доходы по акциям, казначейским векселям и облигациям. Вычислите коэффициенты корреляции и объясните корреляцию между годовыми доходами по этим трем классам инвестиций.

И еще два задания.

4. Файл **Dow.xlsx** содержит ежемесячные данные о доходах по 30 акциям, входящим в индекс Доу — Джонса. Найдите коэффициенты корреляции для всех акций. Затем для каждого пакета акций с помощью условного форматирования выделите три пакета, наиболее сильно коррелирующих с ним. (Разумеется, корреляция пакета с самим собой не должна быть выделена.)
5. Команды НФЛ проводят 16 игр за регулярный сезон. Допустим, стандартное отклонение от количества выигранных игр для всех команд равно 2, и коэффициент корреляции между количеством выигранных командой игр в двух подряд идущих сезонах равен 0,5. Если команда ведет в регулярном сезоне со счетом 12:4, каков наилучший прогноз для количества игр, которые команда выиграет в следующем сезоне?

## ГЛАВА 59

# Введение во множественную регрессию

### Обсуждаемые вопросы

- Завод выпускает три продукта. Как спрогнозировать эксплуатационные расходы завода на основе количества произведенной продукции?
- Насколько точны мои прогнозы ежемесячных расходов на основе количества произведенной продукции?
- Множественную регрессию можно запустить командой Анализ данных. Существует ли способ запустить регрессионный анализ без этой команды и поместить результаты регрессии на тот же лист, где находятся данные?

### Ответы на вопросы

- ② Завод выпускает три продукта. Как спрогнозировать эксплуатационные расходы завода на основе количества произведенной продукции?

В главах 56–58 я описал, как в Microsoft Excel с помощью кривой тренда спрогнозировать одну переменную (*y*, или *зависимую* переменную) по другой переменной (*x*, или *независимой* переменной). Однако довольно часто бывает так, что значение зависимой переменной нужно предсказать на основе нескольких независимых переменных ( $x_1, x_2, \dots, x_n$ ). В таких случаях оценить зависимость можно либо с помощью инструмента Регрессия в пакете Excel Анализ данных (Data Analysis), либо с помощью функции ЛИНЕЙН (LINEST).

Множественная регрессия предполагает, что зависимость между *y* и  $x_1, x_2, \dots, x_n$  имеет следующий вид:

$$y = \text{константа} + B_1x_1 + B_2x_2 + \dots + B_nx_n.$$

Excel вычисляет значения *константа*,  $B_1$ ,  $B_2$ , ...,  $B_n$  для составления наиболее точного (в смысле минимизации суммы квадратичных ошибок) прогноза по этому уравнению. Работа множественной регрессии продемонстрирована в следующем примере.

Лист **Data** в файле **Mrcostest.xlsx** (рис. 59.1) содержит эксплуатационные расходы завода за 19 месяцев, а также количество произведенного продукта А, продукта В и продукта С в каждом месяце.

	A	B	C	D	E
3	Month	Cost	A Made	B Made	C Made
4	1	44439	515	541	928
5	2	43936	929	692	711
6	3	44464	800	710	824
7	4	41533	979	675	758
8	5	46343	1165	1147	635
9	6	44922	651	939	901
10	7	43203	847	755	580
11	8	43000	942	908	589
12	9	40967	630	738	682
13	10	48582	1113	1175	1050
14	11	45003	1086	1075	984
15	12	44303	843	640	828
16	13	42070	500	752	708
17	14	44353	813	989	804
18	15	45968	1190	823	904
19	16	47781	1200	1108	1120
20	17	43202	731	590	1065
21	18	44074	1089	607	1132
22	19	44610	786	513	839

**Рис. 59.1.** Данные для прогноза по ежемесячным эксплуатационным расходам

Требуется найти лучший прогноз для ежемесячных эксплуатационных расходов в следующем виде (называемом *Вид 1*):

ежемесячные эксплуатационные расходы = константа +  $B_1$ \*(количество A) +  $B_2$ \*(количество B) +  $B_3$ \*(количество C).

Анализ данных (Data Analysis) в Excel позволяет найти уравнение вида, наиболее точно соответствующее данным. На вкладке **Данные (Data)** в группе **Анализ (Analysis)** щелкните на **Анализ данных (Data Analysis)** и затем выберите инструмент **Регрессия (Regression)**. Заполните диалоговое окно **Регрессия (Regression)**, как на рис. 59.2.

Regression

Input

Input Y Range:

Input X Range:

☒ Labels ☐ Constant is Zero

☐ Confidence Level:  %

Output options

☐ Output Range:

☒ New Worksheet Ply:

☐ New Workbook

Residuals

☒ Residuals ☐ Residual Plots

☐ Standardized Residuals ☐ Line Fit Plots

Normal Probability

☐ Normal Probability Plots

OK Cancel Help

**Рис. 59.2.** Диалоговое окно Регрессия

- Входной интервал  $Y$  (B3:B22 в этом примере) содержит зависимую переменную или данные (включая метку Расходы), которые требуется предсказать.
- Входной интервал  $X$  (C3:E22) содержит данные или независимые переменные (включая метки Количество А, Количество В и Количество С), которые требуется использовать в прогнозе. В Excel установлено ограничение — 15 независимых переменных, которые должны находиться в смежных столбцах.
- Поскольку оба входных интервала  $X$  и  $Y$  включают метки, я установил флажок Метки (Labels).
- Я выбрал поместить выходные данные на новый лист с именем Регрессия.
- При установленном флажке Остатки (Residuals) для каждого наблюдения вычисляется прогноз (по уравнению *Вуда 1*) и остатки, равные разности между фактическими и прогнозируемыми расходами.

### ПРИМЕЧАНИЕ

Если Пакет анализа (Analysis ToolPak) еще не установлен, на вкладке Файл (File) выберите команду Параметры (Options) и затем Надстройки (Add-Ins). В списке Управление (Manage) выберите Надстройки Excel (Excel Add-ins) и нажмите Перейти (Go). Установите флажок Пакет анализа (Analysis ToolPak) и нажмите OK.

Нажав OK в диалоговом окне Регрессия (Regression), на листе Регрессия появятся выходные данные, показанные на рис. 59.3 и 59.4.

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	Regression Statistics								
4	Multiple R	0.803398744							
5	R Square	0.645449542							
6	Adjusted R Square	0.57453945							
7	Standard Error	1252.763898							
8	Observations	19							
9									
10	ANOVA								
11		df	SS	MS	F	Significance F			
12	Regression	3	42856229.89	1.4E+07	9.1024	0.001126532			
13	Residual	15	23541260.74	1569417					
14	Total	18	66397490.63						
15									
16		Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
17	Intercept	35102.90045	1837.226911	19.1065	6E-12	31186.94158	39018.8593	31186.94158	39018.8593
18	A Made	2.065953296	1.664981779	1.24083	0.2337	-1.482873542	5.61478013	-1.48287354	5.61478013
19	B Made	4.176355531	1.681252566	2.48407	0.0253	0.592848311	7.75986275	0.592848311	7.75986275
20	C Made	4.790641037	1.789316107	2.67736	0.0172	0.97680169	8.60448038	0.97680169	8.60448038

**Рис. 59.3.** Первоначальный вывод итогов множественной регрессии

Какое уравнение прогноза самое лучшее? В столбце Коэффициенты (ячейки B16:B20) показано, что наилучшим уравнением *Вуда 1*, по которому можно прогнозировать ежемесячные эксплуатационные расходы, является следующее:



прогнозируемые расходы =  $35102,90 + 2,07 \cdot (\text{количество A}) + 4,18 \cdot (\text{количество B}) + 4,79 \cdot (\text{количество C})$ .

	A	B	C
24	RESIDUAL OUTPUT		
25			
26	Observation	Predicted Cost	Residuals
27	1	42871.99	1567.01
28	2	43318.35	617.65
29	3	43668.36	795.64
30	4	43575.81	-2042.81
31	5	45342.07	1000.93
32	6	44685.80	236.20
33	7	42784.48	418.52
34	8	43662.85	-662.85
35	9	42753.82	-1786.82
36	10	47339.70	1242.30
37	11	46550.10	-1547.10
38	12	43484.02	818.98
39	13	42668.27	-598.27
40	14	44764.61	-411.61
41	15	45329.26	638.74
42	16	47574.96	206.04
43	17	44179.19	-977.19
44	18	45310.78	-1236.78
45	19	42888.56	1721.44

**Рис. 59.4.** Первоначальный вывод остатков множественной регрессии

Возникает вопрос: какие независимые переменные полезны для прогноза ежемесячных эксплуатационных расходов? В конце концов, если в качестве независимой переменной выбрать количество игр, выигранных командой «Сиэтл Маринерс» за месяц, можно ожидать, что эта переменная окажет небольшое влияние на прогноз ежемесячных эксплуатационных расходов. При запуске регрессионного анализа каждая независимая переменная имеет  $p$ -значение между 0 и 1. Любая независимая переменная с  $p$ -значением (см. ячейки E16:E20), меньшим или равным 0,05, считается приемлемой для прогноза зависимой переменной. Таким образом, чем меньше  $p$ -значение, тем выше прогностическая способность независимой переменной. Здесь три независимых переменных имеют следующие  $p$ -значения: 0,23 (для количества произведенного продукта A), 0,025 (для продукта B) и 0,017 (для продукта C). Эти  $p$ -значения можно интерпретировать следующим образом.

- Если вы используете показатели B и C для предсказания ежемесячных эксплуатационных расходов, то получаете 77%-ную вероятность ( $1 - 0,23$ ), что показатель A увеличит прогностическую силу.
- Если вы используете показатели A и C для предсказания ежемесячных эксплуатационных расходов, то получаете 97,5%-ную вероятность ( $1 - 0,025$ ), что показатель B увеличит прогностическую силу.
- Если вы используете показатели A и B для предсказания ежемесячных эксплуатационных расходов, то получаете 98,3%-ную вероятность ( $1 - 0,017$ ), что показатель C увеличит прогностическую силу.

$p$ -значения показывают, что показатель А незначительно увеличивает прогностическую силу для продуктов В и С. То есть если количество продуктов В и С известно, ежемесячные эксплуатационные расходы можно предсказать почти с той же точностью, что и в случае включения А в качестве независимой переменной. Следовательно, показатель А как независимую переменную можно удалить и использовать для прогноза только показатели В и С. Скопируйте данные на лист Столбец А удален и удалите столбец Количество А (столбец С). Измените параметр Входной интервал Х (С3:D22) — см. рис. 59.2. На новом листе, например с именем Без столбца А, появятся результаты регрессии (рис. 59.5 и 59.6).

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	Regression Statistics								
4	Multiple R	0.780421232							
5	R Square	0.609057299							
6	Adjusted R Square	0.560189461							
7	Standard Error	1273.715391							
8	Observations	19							
9									
10	ANOVA								
11		df	SS	MS	F	Significance F			
12	Regression	2	40439876.29	2E+07	12.463	0.0005456			
13	Residual	16	25957614.34	2E+06					
14	Total	18	66397490.63						
15									
16		Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
17	Intercept	35475.30255	1842.860853	19.25	2E-12	31568.613	39381.992	31568.613	39381.992
18	B Made	5.320968077	1.429095476	3.723	0.0018	2.2914217	8.3505145	2.2914217	8.3505145
19	C Made	5.417137848	1.745311646	3.104	0.0068	1.7172433	9.1170324	1.7172433	9.1170324

**Рис. 59.5.** Вывод итогов множественной регрессии без показателя А как независимой переменной

	A	B	C
25	Observation	Predicted Cost	Residuals
26	1	43381.05021	1057.949794
27	2	43008.99747	927.002527
28	3	43716.91148	747.0885248
29	4	43173.14649	-1640.146495
30	5	45018.33547	1324.664528
31	6	45352.53278	-430.5327792
32	7	42634.5734	568.4265963
33	8	43497.43576	-497.4357601
34	9	43096.66501	-2129.665007
35	10	47415.43478	1166.565215
36	11	46525.80688	-1522.806879
37	12	43366.11226	936.8877388
38	13	43312.00414	-1242.004144
39	14	45093.11881	-740.1188117
40	15	44751.5519	1216.448104
41	16	47438.12957	342.8704271
42	17	44383.92553	-1181.925527
43	18	44837.33022	-763.3302206
44	19	42749.93783	1860.062168

**Рис. 59.6.** Вывод остатка без показателя А как независимой переменной

Как видите, показатели В и С имеют очень низкие  $p$ -значения (0,002 и 0,007 соответственно). Эти значения показывают, что обе независимые переменные об-

ладают значительной прогностической силой. Новые коэффициенты в столбце В позволяют прогнозировать ежемесячные эксплуатационные расходы с помощью следующего уравнения:

прогнозируемые ежемесячные эксплуатационные расходы =  $35475,3 + 5,32 \cdot (\text{показатель В}) + 5,42 \cdot (\text{показатель С})$ .

**❓ Насколько точны мои прогнозы ежемесячных расходов на основе количества произведенной продукции?**

В выходных данных регрессии в ячейке В5 листа Без столбца А (см. рис. 59.5) значение  $R^2$  равно 0,61. Такое  $R^2$  означает, что вместе показатели В и С объясняют 61% отклонений в ежемесячных эксплуатационных расходах. Обратите внимание, что в первоначальной регрессии, включавшей показатель А в качестве независимой переменной,  $R^2$  было равно 0,65. То есть добавление показателя А как независимой переменной объясняет только 4% отклонений в ежемесячных эксплуатационных расходах. Такая незначительная разница согласуется с решением удалить показатель А из списка независимых переменных.

В выходных данных регрессии в ячейке В7 листа Без столбца А стандартная ошибка для регрессии с показателями В и С в качестве независимых переменных равна 1274. Можно ожидать, что примерно 68% множественных регрессионных прогнозов будут точными в пределах одной стандартной ошибки и 95% прогнозов будет точным в пределах двух стандартных ошибок. Любой прогноз, отличающийся от фактического значения более чем на две стандартные ошибки, считается выбросом. Таким образом, если прогнозируемые эксплуатационные расходы по ошибке превышают \$2548 ( $2 \times 1274$ ), такое наблюдение следует считать выбросом.

В части выходных данных с остатками, показанной ранее на рис. 59.6, для каждого наблюдения приведены предсказанные расходы и остатки, эквивалентные разности фактических и предсказанных расходов. Например, для первого наблюдения предсказанные расходы составляют \$43 381,10. Остаток \$1057,95 означает, что прогноз был ниже фактических расходов на \$1057,95.

**❓ Множественную регрессию можно запустить командой Анализ данных. Существует ли способ запустить регрессионный анализ без этой команды и поместить результаты регрессии на тот же лист, где находятся данные?**

Для вставки результатов регрессионного анализа прямо в книгу можно применить функцию Excel ЛИНЕЙН (LINEST). При наличии  $m$  независимых переменных начните с выделения пустого диапазона ячеек, состоящего из пяти строк и  $m + 1$  столбцов, где будут храниться результаты работы функции ЛИНЕЙН. Для этой цели на листе Столбец А удален я использовал диапазон F5:H9.

Синтаксис функции: ЛИНЕЙН(известные\_значения\_У; известные\_значения\_Х; ИСТИНА; ИСТИНА). Если указать для третьего аргумента ЛОЖЬ, Excel решит уравнение без свободного члена. Если указать ЛОЖЬ для четвертого аргумента, многие вы-

числения регрессионного анализа будут опущены, и функция **ЛИНЕЙН** возвратит только уравнение множественной регрессии.

Выделите на листе **Столбец А** удален диапазон необходимого размера (в данном примере F5:H9) и введите формулу **=ЛИНЕЙН(B4:B22; C4:D22; ИСТИНА; ИСТИНА)**. В этот момент не нажимайте **Enter**! Поскольку функция **ЛИНЕЙН** является функцией, возвращающей массив, для правильной работы функции нажмите **Enter**, удерживая **Ctrl+Shift**. В выделенном диапазоне появятся результаты, показанные на рис. 59.7. (См. лист **Столбец А** удален.)

	F	G	H	I	J
4				LINEST OUTPUT	
5	5.417137848	5.320968077	35475.30255		
6	1.745311646	1.429095476	1842.860853		
7	0.609057299	1273.715391	#N/A		
8	12.46335684	16	#N/A		
9	40439876.29	25957614.34	#N/A		
10					
11	Cmadecoef	Bmadecoef	Const		
12	Std err C	Std Err B	Std err const		
13	rsq	std err est			
14	F	df			
15	ssreg	ssresid			

**Рис. 59.7.** Множественная регрессия с функцией **ЛИНЕЙН**

В строке 5 вы найдете уравнение прогноза (коэффициенты читаются справа налево, начиная с начальной ординаты):

прогнозируемые ежемесячные расходы = 35 475,3 + 5,32\*(показатель В) + 5,43\*(показатель С).

Строка 6 содержит стандартные ошибки для каждого вычисленного коэффициента, но они не слишком важны. Ячейка F7 содержит значение  $R^2$  — 0,61, а в ячейке G7 находится стандартная ошибка регрессии — 1274. Строки 8 и 9 содержат менее важную информацию (F-статистику, степени свободы, сумму квадратов регрессии и сумму квадратов остатков).

## ПРИМЕЧАНИЕ

Задания по множественной регрессии находятся в конце главы 61 «Моделирование нелинейных характеристик и взаимосвязей».

## ГЛАВА 60

# Включение качественных факторов во множественную регрессию

### Обсуждаемые вопросы

- Как предсказать квартальные продажи автомобилей в США?
- Как предсказать президентские выборы в США?
- Какая функция Excel позволяет составить прогноз по уравнению множественной регрессии?

В первом примере с использованием множественной регрессии (глава 59) ежемесячные эксплуатационные расходы завода были спрогнозированы на основе объемов производства каждого продукта. Поскольку объемы производства можно подсчитать точно, объемы А, В и С возможно использовать как *независимые количественные переменные*. Однако во многих ситуациях независимые переменные быстро определить количественно невозможно. В данной главе мы посмотрим, как можно включить качественные факторы (например, сезонность, половую принадлежность, политическую партию) во множественный регрессионный анализ.

## Ответы на вопросы

### ❓ Как предсказать квартальные продажи автомобилей в США?

Допустим, вы хотите спрогнозировать квартальные продажи автомобилей в США с целью определить влияние квартала года на объемы продаж. Мы будем пользоваться данными из файла *Autotemp.xlsx* на листе *Data* (рис. 60.1). Объемы продаж указаны в тысячах штук, а ВВП — в миллиардах долларов.

Можно было бы попытаться определить независимую переменную как 1 для первого квартала, 2 для второго квартала и т. д. К сожалению, при таком подходе четвертый квартал превосходил бы по эффективности первый квартал в четыре раза, что не может быть правдой. Квартал года — это *независимая качественная переменная*. При моделировании независимой качественной переменной вы создаете независимую переменную, называемую *фиктивной переменной*, для всех возможных значений качественной переменной, кроме одного. (Невключаемое значение

выбирается произвольным образом. В данном примере я предпочел опустить четвертый квартал.) Фиктивные переменные показывают, какое значение качественной переменной имеет место.

	A	B	C	D	E	F
9	Historical data					
10	Year	Quarter	Sales	GNP	Unemp	Int
11	79	1	2541	5.9	9.4	
12	79	2	2910	2640	5.7	9.4
13	79	3	2562	2695	5.9	9.7
14	79	4	2385	2701	6	12
15	80	1	2520	2785	6.2	13
16	80	2	2142	2509	7.3	9.6
17	80	3	2130	2570	7.7	9.2
18	80	4	2190	2667	7.4	14
19	81	1	2370	2878	7.4	14
20	81	2	2208	2835	7.4	15
21	81	3	2196	2897	7.4	15
22	81	4	1758	2744	8.3	12
23	82	1	1944	2582	8.8	13
24	82	2	2094	2613	9.4	12
25	82	3	1911	2529	10	9.3
26	82	4	2031	2544	10.7	7.9
27	83	1	2046	2633	10.4	7.8
28	83	2	2502	2878	10.1	8.4
29	83	3	2238	3051	9.4	9.1
30	83	4	2394	3274	8.5	8.8
31	84	1	2586	3594	7.9	9.2
32	84	2	2898	3774	7.5	9.8
33	84	3	2448	3861	7.5	10
34	84	4	2460	3919	7.2	8.8
35	85	1	2646	4040	7.4	8.2
36	85	2	2988	4133	7.3	7.5
37	85	3	2967	4303	7.1	7.1
38	85	4	2439	4393	7	7.2
39	86	1	2598	4560	7.1	8.9
40	86	2	3045	4587	7.1	7.7
41	86	3	3213	4716	6.9	7.4
42	86	4	2685	4796	6.8	7.4

**Рис. 60.1.** Данные о продажах автомобилей

Таким образом, фиктивные переменные для первого, второго и третьего кварталов обладают следующими свойствами:

- фиктивная переменная для первого квартала равна 1, если квартал является первым кварталом, и равна 0 в противном случае;
- фиктивная переменная для второго квартала равна 1, если квартал является вторым кварталом, и равна 0 в противном случае;
- фиктивная переменная для третьего квартала равна 1, если квартал является третьим кварталом, и равна 0 в противном случае.

Наблюдения для четвертого квартала будут идентифицироваться по тому факту, что фиктивные переменные для кварталов с первого по третий равны 0. Теперь понятно, почему не нужна фиктивная переменная для четвертого квартала. Если включить в регрессию фиктивную переменную для четвертого квартала как независимую, Excel возвратит сообщение об ошибке, поскольку если между любым набором независимых переменных существует точная линейная зависимость, то в ходе множественной регрессии Excel должен выполнить математический эквивалент деления на 0 (что невозможно). Если в этом примере включить фиктивную

переменную для четвертого квартала, каждая точка данных будет удовлетворять следующей точной линейной зависимости:

(фиктивная переменная 1) + (фиктивная переменная 2) + (фиктивная переменная 3) + (фиктивная переменная 4) = 1.

### ПРИМЕЧАНИЕ

Точная линейная зависимость возникает в случае, если существуют константы  $c_0, c_1, \dots, c_N$ , такие, что для каждой точки данных  $c_0 + c_1x_1 + c_2x_2 + \dots + c_Nx_N = 0$ . Здесь  $x_1, \dots, x_N$  — значения независимых переменных.

Чтобы создать фиктивную переменную для первого квартала, я скопировал формулу =ЕСЛИ(B12=1;1;0) из G12 в G13:G42. Эта формула помещает в столбец G значение 1, если квартал является первым кварталом, и 0 в противоположном случае. Аналогичным образом я создал фиктивные переменные для второго квартала (в H12:H42) и третьего (в I12:I42). Результаты показаны на рис. 60.2.

	A	B	C	D	E	F	G	H	I	J	K	L
10	Year	Quarter	Sales	GNP	Unemp	Int	Q1	Q2	Q3	LagGNP	LagUnemp	LagInt
11	79	1	2541	2541	5.9	9.4	1	0	0	2541	5.9	9.4
12	79	2	2910	2640	5.7	9.4	0	1	0	2541	5.9	9.4
13	79	3	2562	2595	5.9	9.7	0	0	1	2640	5.7	9.4
14	79	4	2385	2701	6	12	0	0	0	2595	5.9	9.7
15	80	1	2520	2785	6.2	13	1	0	0	2701	6	11.9
16	80	2	2142	2509	7.3	9.6	0	1	0	2785	6.2	13.4
17	80	3	2130	2570	7.7	9.2	0	0	1	2509	7.3	9.6
18	80	4	2190	2667	7.4	14	0	0	0	2570	7.7	9.2
19	81	1	2370	2878	7.4	14	1	0	0	2667	7.4	13.6
20	81	2	2208	2835	7.4	15	0	1	0	2878	7.4	14.4
21	81	3	2196	2897	7.4	15	0	0	1	2835	7.4	15.3
22	81	4	1758	2744	8.3	12	0	0	0	2897	7.4	15.1
23	82	1	1944	2582	8.8	13	1	0	0	2744	8.3	11.8
24	82	2	2094	2613	9.4	12	0	1	0	2582	8.8	12.8
25	82	3	1911	2529	10	9.3	0	0	1	2613	9.4	12.4
26	82	4	2031	2544	10.7	7.9	0	0	0	2529	10	9.3
27	83	1	2046	2633	10.4	7.8	1	0	0	2544	10.7	7.9
28	83	2	2502	2878	10.1	8.4	0	1	0	2633	10.4	7.8
29	83	3	2238	3051	9.4	9.1	0	0	1	2878	10.1	8.4
30	83	4	2394	3274	8.5	8.8	0	0	0	3051	9.4	9.1
31	84	1	2586	3594	7.9	9.2	1	0	0	3274	8.5	8.8
32	84	2	2898	3774	7.5	9.8	0	1	0	3594	7.9	9.2
33	84	3	2448	3861	7.5	10	0	0	1	3774	7.5	9.8
34	84	4	2460	3919	7.2	8.8	0	0	0	3861	7.5	10.3
35	85	1	2646	4040	7.4	8.2	1	0	0	3919	7.2	8.8
36	85	2	2988	4133	7.3	7.5	0	1	0	4040	7.4	8.2
37	85	3	2967	4303	7.1	7.1	0	0	1	4133	7.3	7.5
38	85	4	2439	4393	7	7.2	0	0	0	4303	7.1	7.1
39	86	1	2598	4560	7.1	8.9	1	0	0	4393	7	7.2
40	86	2	3045	4587	7.1	7.7	0	1	0	4560	7.1	8.9
41	86	3	3213	4716	6.9	7.4	0	0	1	4587	7.1	7.7
42	86	4	2685	4796	6.8	7.4	0	0	0	4716	6.9	7.4
43												

**Рис. 60.2.** Фиктивные переменные для определения квартала, в котором произошла продажа

В дополнение к сезонности мы хотим использовать макроэкономические переменные, такие как ВВП (в миллиардах долларов 1986 года), процентные ставки и уровень безработицы, для прогнозирования продаж автомобилей. Допустим, вы пытаетесь оценить продажи за второй квартал 1979 г. . Поскольку значения



для ВВП, процентной ставки и уровня безработицы на начало второго квартала в 1979 г. неизвестны, прогнозировать объемы продаж второго квартала 1979 г. на основе этих величин невозможно. Вместо них для прогнозирования продаж следует использовать значения ВВП, процентной ставки и уровня безработицы с отставанием на один квартал. Например, скопируйте формулу =D11 из J12 в J12:L42 для создания значения с отставанием для ВВП, первой независимой макроэкономической переменной. Диапазон J12:L12 содержит ВВП, уровень безработицы и процентную ставку для первого квартала 1979 г. (Когда для прогнозирования зависимой переменной используются значения независимых переменных (или переменных) с отставанием, независимые переменные называются *лаговыми независимыми переменными*.)

Теперь можно запустить множественную регрессию. На вкладке **Данные (Data)** выберите **Анализ данных (Data Analysis)**. (Необходимо установить пакет **Анализ данных**. Инструкции см. в главе 42.) В диалоговом окне **Анализ данных (Data Analysis)** выберите **Регрессия (Regression)**. Укажите C11:C42 как **Входной интервал Y (Input Y Range)** и G11:L42 как **Входной интервал X (Input X Range)**, установите флажок **Метки (Labels)**, так как строка 11 содержит метки, и флажок **Остатки (Residuals)**. Введите имя нового листа и нажмите **ОК** для получения выходных данных (см. лист **Regression** и рис. 60.3–60.5).

	A	B	C	D	E	F	G	H	I
1	<b>SUMMARY OUTPUT</b>								
2									
3	<i>Regression Statistics</i>								
4	Multiple R	0.884139126							
5	R Square	0.781701994							
6	Adjusted R Square	0.727127492							
7	Standard Error	190.5240756							
8	Observations	31							
9									
10	<b>ANOVA</b>								
11		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
12	Regression	6	3119625.193	519938	14.324	6.7975E-07			
13	Residual	24	871186.1616	36299					
14	Total	30	3990811.355						
15									
16		<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
17	Intercept	3154.700285	462.6530922	6.8187	5E-07	2199.83143	4109.5691	2199.83143	4109.56914
18	Q1	156.833091	98.87110703	1.5862	0.1258	-47.2268026	360.89298	-47.2268026	360.8929846
19	Q2	379.7835116	96.08921514	3.9524	0.0006	181.46516	578.10186	181.46516	578.1018637
20	Q3	203.035501	95.40891864	2.1281	0.0438	6.12121161	399.94979	6.12121161	399.9497905
21	LagGNP	0.174166906	0.05842	2.9811	0.0065	0.05358398	0.2947298	0.05358398	0.294729835
22	LagUnemp	-93.83233214	28.32328716	-3.3129	0.0029	-152.288712	-35.375953	-152.288712	-35.3759525
23	Lagint	-73.9167147	17.78861573	-4.1553	0.0004	-110.630399	-37.20303	-110.630399	-37.2030302

**Рис. 60.3.** Итоги регрессии и таблица дисперсионного анализа для продаж автомобилей

На рис. 60.4 показано, что уравнение (уравнение 1) для прогноза квартальных продаж автомобилей должно иметь вид:

прогнозируемые квартальные продажи = 3154,7 + 156,833\*Q1 + 379,784\*Q2 + + 203,036\*Q3 + 0,174\*(ВВП с отставанием в млрд долл.) – 93,83\*(уровень безработицы с отставанием) – 73,91\*(процентная ставка с отставанием).



Кроме того, из рис. 60.4 видно, что  $p$ -значение каждой независимой переменной не превышает 0,15. На основании этого можно сделать вывод, что все независимые переменные имеют значимое влияние на квартальные продажи автомобилей. Все коэффициенты в уравнении регрессии интерпретируются *ceteris paribus* — при прочих равных условиях (то есть каждый коэффициент независимой переменной оказывает влияние после устранения влияния всех других переменных уравнения регрессии).

	A	B	C	D	E	F	G	H	I
16		Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
17	Intercept	3154.700285	462.6530922	6.8187	5E-07	2199.83143	4109.5691	2199.83143	4109.56914
18	Q1	156.833091	98.87110703	1.5862	0.1258	-47.2268026	360.89298	-47.2268026	360.8929846
19	Q2	379.7835116	96.08921514	3.9524	0.0006	181.46516	578.10186	181.46516	578.1018637
20	Q3	203.035501	95.40891864	2.1281	0.0438	6.12121161	399.94979	6.12121161	399.9497905
21	LagGNP	0.174166906	0.05842	2.9811	0.0065	0.05358398	0.2947298	0.05358398	0.294729835
22	LagUnemp	-93.83233214	28.32328716	-3.3129	0.0029	-152.288712	-35.375953	-152.288712	-35.3759525
23	LagInt	-73.9167147	17.78851573	-4.1553	0.0004	-110.630399	-37.20303	-110.630399	-37.2030302

**Рис. 60.4.** Информация о коэффициентах регрессии для продаж автомобилей

Вот пояснения к каждому коэффициенту.

- Увеличение в предыдущем квартале ВВП на 1 млрд долларов повышает квартальные продажи автомобилей на 174 шт.
- Увеличение в предыдущем квартале уровня безработицы на 1% уменьшает квартальные продажи автомобилей на 93 832 шт.
- Увеличение в предыдущем квартале процентной ставки на 1% уменьшает квартальные продажи на 73 917 шт.

При интерпретации коэффициентов фиктивных переменных следует осознавать, что они говорят вам о влиянии сезонности на значения, но не говорят о самих качественных переменных. Данные, приведенные на рис. 60.4, можно интерпретировать следующим образом:

- продажи первого квартала превышают продажи четвертого квартала на 156 833 шт.;
- продажи второго квартала превышают продажи четвертого квартала на 379 784 шт.;
- продажи третьего квартала превышают продажи четвертого квартала на 203 036 шт.

Оказалось, что продажи автомобилей выше всего во втором квартале (апрель — июнь; возврат налогов и наступающее лето) и ниже всего в четвертом квартале (октябрь — декабрь; зачем покупать новую машину, если посыпание улиц солью повредит ее?).

На основании рис. 60.3 можно сделать несколько выводов.

- Изменение независимых переменных (макроэкономических факторов и сезонности) объясняет 78% изменений зависимой переменной (квартальных продаж).

- Стандартная ошибка регрессии составляет 190 524 автомобиля. Можно ожидать, что приблизительно 68% прогнозов дано с точностью до 190 524 шт. и примерно 95% прогнозов — с точностью до 381 048 шт. ( $2 \times 190\,524$ ).
- Для подбора коэффициентов уравнения регрессии использовано 31 наблюдение.

Единственной величиной, представляющей интерес в таблице дисперсионного анализа на рис. 60.3, является **Значимость F (0,00000068)**. Эта мера подразумевает, что только в 6,8 случая на 10 000 000 все вместе взятые независимые переменные непригодны для прогнозирования продаж автомобилей. Таким образом, вы можете быть уверены в том, что независимые переменные пригодны для прогнозирования квартальных продаж. (Подробнее модель дисперсионного анализа (ANOVA; analysis of variance) представлена в главах 62 и 63.)

Рисунок 60.5 показывает предсказанные продажи и остатки для каждого наблюдения. Например, для второго квартала 1979 г. (первое наблюдение) предсказанные по уравнению 1 продажи составляют 2728,6 тысячи, а остаток равен 181 411 шт. (2910 – 2728,6). Обратите внимание, что остатки не превышают 381 000 по абсолютной величине, то есть выбросы отсутствуют.

	A	B	C
27	RESIDUAL OUTPUT		
28			
29	Observation	Predicted Sales	Residuals
30	1	2728.588616	181.4113836
31	2	2587.848606	-25.84860587
32	3	2336.034563	48.96543676
33	4	2339.328281	180.6717193
34	5	2447.266343	-305.2663429
35	6	2400.118977	-270.1189769
36	7	2199.7408	-9.740800106
37	8	2076.383266	293.6167341
38	9	2276.947422	-68.94742189
39	10	2026.185621	169.8143789
40	11	1848.731191	-80.73119119
41	12	2138.394335	-194.3943352
42	13	2212.298466	-118.2984663
43	14	2014.216596	-103.2165964
44	15	1969.394332	61.60666841
45	16	2166.640544	-120.6405443
46	17	2440.632301	61.3676994
47	18	2290.362403	-62.36240272
48	19	2131.386979	262.6130214
49	20	2433.681173	152.3188271
50	21	2739.094517	168.9064833
51	22	2586.877653	-138.8776532
52	23	2362.035446	97.96455439
53	24	2667.994409	-21.99440886
54	25	2937.601377	50.39862257
55	26	2838.174893	128.8251074
56	27	2713.079218	-274.0792178
57	28	2887.577992	-289.5779921
58	29	3004.570968	40.42903226
59	30	2921.225251	291.7747488
60	31	2781.597472	-96.59747188

**Рис. 60.5.** Остатки для данных продаж автомобилей

### ❓ Как предсказать президентские выборы в США?

Когда Джеймса Карвилла, советника президента, спросили о том, какие факторы влияют на президентские выборы, он ответил: «Это экономика, дурачок».

Экономист Йельского университета Рой Фэйр показал, что Карвилл был прав в своем убеждении, что состояние экономики сильно влияет на результаты президентских выборов. В качестве зависимой переменной (см. в файле President16.xlsx лист Data, показанный на рис. 60.6) Фэйр для каждого выбора (с 1916 по 2012 г.) использовал процент голосов, набранных партией, находящейся у власти, с учетом голосования за две партии (без голосов за кандидатов других партий). Для выборов 2016 г. были взяты данные на 2 июля 2016 г., представленные Фэйром на его веб-сайте. Фэйр попытался предсказать процент голосов правящей партии при голосовании за две партии с помощью следующих независимых переменных.

- **Партия у власти.** В исходных данных 1 обозначает республиканскую партию у власти, а 0 — демократическую.
- **Рост ВВП в процентах в течение первых девяти месяцев года выборов.**
- **Абсолютное значение уровня инфляции в течение первых девяти месяцев года выборов.** Абсолютное значение используется потому, что плох как положительный, так и отрицательный уровень инфляции.
- **Количество кварталов за последние четыре года, в течение которых наблюдался сильный экономический рост.** Сильный экономический рост определяется как ежегодный рост на уровне не менее 3,2%.

	A	B	C	D	E	F	G	H	I	J	K
	Candidates	Year	Incumbent Share	Party in Power	Growth rate in election year	Abs. Value Inflation rate in election year	Quarters Growth>3.2 %	Time Incumbent in Office	Incumbent Party 1 = Republican	War	President Running?
6											
7	Wilson-Hughes	1916	51.7	D	2.2	4.3	3	0	0	0	1
8	Harding-Cox	1920	36.1	D	-11.5	0	0	1	0	1	0
9	Coolidge-Davis	1924	58.3	R	-3.9	5.2	10	0	1	0	1
10	Hoover-Smith	1928	58.8	R	4.6	0.2	7	1	1	0	0
11	Hoover-FDR	1932	40.9	R	-14.9	7.1	4	1.25	1	0	1
12	FDR-Landon	1936	62.2	D	11.9	2.5	9	0	0	0	1
13	FDR-Wilkie	1940	55	D	3.7	0	8	1	0	0	1
14	FDR-Dewey	1944	53.8	D	4.1	0	0	1.25	0	1	1
15	Truman-Dewey	1948	52.3	D	1.8	0	0	1.5	0	1	1
16	Ike-Stevenson	1952	44.7	D	0.6	2.3	7	1.75	0	0	0
17	Ike-Stevenson	1956	57.1	R	-1.5	1.9	5	0	1	0	1
18	Kennedy-Nixon	1960	49.9	R	0.1	1.9	5	1	1	0	0
19	Johnson-Goldwater	1964	61.2	D	5.1	1.3	9	0	0	0	1
20	Nixon-Humphrey	1968	49.4	D	4.8	3.1	7	1	0	0	0
21	Nixon-McCgovern	1972	61.8	R	6.3	4.8	4	0	1	0	1
22	Ford-Carter	1976	49	R	3.7	7.6	5	1	1	0	0
23	Carter-Reagan	1980	44.8	D	-3.8	7.9	5	0	0	0	1
24	Reagan-Mondale	1984	59.1	R	5.4	5.2	8	0	1	0	1
25	Bush-Dukakis	1988	53.8	R	2.1	3	5	1	1	0	0
26	Bush-Clinton	1992	46.4	R	2.3	3.3	3	1.25	1	0	1
27	Clinton-Dole	1996	54.7	D	2.9	2	4	0	0	0	1
28	Gore-Bush	2000	50.3	D	2.2	1.6	7	0	0	0	0
29	Bush-Kerry	2004	51.2	R	2	2.2	2	0	1	0	1
30	Obama-McCain	2008	46.3	R	-2.3	3.1	2	1	1	0	0
31	Obama-Romney	2012	51.9	D	1.57	1.03	1	0	0	0	1
32	Clinton-Trump	2016	51.1	D	2.16	1.28	3	1	0	0	0

Рис. 60.6. Данные президентских выборов

- **Время пребывания правящей партии у власти.** Фэйр обозначает один срок пребывания у власти как 0, два срока как 1, три срока как 1,25, четыре срока как 1,5 и не менее пяти сроков как 1,75. Это определение подразумевает, что каждый срок после первого срока пребывания у власти имеет меньше влияния на результаты выборов, чем первый срок.
- **Выборы в военное время.** Выборы в 1920 г. (Первая мировая война), в 1944 г. (Вторая мировая война) и в 1948 г. (Вторая мировая война еще не закончилась в 1945 г.) были определены как выборы в военное время. (Выборы, состоявшиеся во время вьетнамской войны, не считались выборами в военное время.) В годы войны переменные, связанные с темпами экономического роста и инфляции, признаны неуместными, и для них установлено значение 0.
- **Действующий президент баллотируется на переизбрание.** В этом случае переменная устанавливается в 1; в противном случае она устанавливается в 0. В 1976 г. Джеральд Форд не считался президентом, баллотирующимся на переизбрание, поскольку он не был избран ни президентом, ни вице-президентом.

Теперь попытаемся на основе данных о выборах с 1916 по 2004 г. составить уравнение множественной регрессии для предсказания результатов будущих президентских выборов. Данные о выборах в 2008–2016 гг. были сохранены в качестве точек проверки достоверности. При подборе уравнения регрессии к данным всегда желательно иметь какие-либо данные для проверки уравнения регрессии. Так мы определим, может ли уравнение предсказать непредставленные данные. Любой инструмент прогнозирования, недостаточно хорошо предсказывающий данные, которые были ему недоступны, не следует использовать для прогнозирования будущего.

Для запуска регрессии на вкладке **Данные (Data)** в группе **Анализ (Analysis)** выберите **Анализ данных (Data Analysis)** и затем инструмент **Регрессия (Regression)**. Здесь как **Входной интервал Y (Input Y Range)** указаны ячейки C6:C29 и как **Входной интервал X (Input X Range)** — ячейки E6:L29. Установлен флажок **Метки (Labels)**, так как строка 6 содержит метки, и флажок **Остатки (Residuals)**. Результаты представлены справа от данных на листе **Data** и на рис. 60.7 (начиная с ячейки N14).

Как видно из рис. 60.7, все независимые переменные (если округлить коэффициент для военного времени до 0,10) имеют *p*-значения не больше 0,10, и все переменные должны быть применены в прогнозах. По данным из раздела коэффициентов на рис. 60.7 можно определить, что наилучшим уравнением для прогнозирования результатов выборов является следующее уравнение (уравнение 2):

прогнозируемый процент в президентских выборах =  $43,36 + 0,69 \cdot \text{рост\_ВВП} - 0,57 \cdot \text{абс.уровень\_инфляции} + 1,1 \cdot \text{кварталы\_с\_ростом} - 2,91 \cdot \text{время\_у\_власти} + 5,16 \cdot \text{республиканцы} + 6,34 \cdot \text{военное\_время} + 3,93 \cdot \text{президент\_баллотируется}$ .

Коэффициенты независимых переменных можно интерпретировать следующим образом (после устранения влияния всех остальных независимых переменных):

- 1%-ное увеличение годовых темпов роста ВВП в год выборов означает прибавку 0,69% голосов для правящей партии;

	N	O	P	Q	R	S	T	U	V
14	SUMMARY OUTPUT								
15									
16	Regression Statistics								
17	Multiple R	0.943094							
18	R Square	0.889425							
19	Adjusted R Square	0.837824							
20	Standard Error	2.721034							
21	Observations	23							
22									
23	ANOVA								
24		df	SS	MS	F	Significance F			
25	Regression	7	893.3326	127.6189	17.23643	4.05E-06			
26	Residual	16	111.0604	7.404026					
27	Total	22	1004.393						
28									
29		Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
30	Intercept	43.36238	2.798878	15.49388	1.23E-10	37.39714	49.32762	37.39714085	49.32762062
31	Growth rate in election year	0.687368	0.118927	6.784699	3.6E-06	0.434094	0.940642	0.43409368	0.940642066
32	Abs. Value Inflation rate in election year	-0.571465	0.308857	-1.850257	0.084073	-1.229779	0.086848	-1.229778965	0.086848378
33	Quarters Growth>3.2%	1.102468	0.298606	3.693286	0.002169	0.466218	1.738719	0.466217746	1.738719149
34	Time Incumbent in Office	-2.913594	1.194638	-2.438892	0.027638	-5.459906	-0.367283	-5.459905978	-0.3672828
35	Incumbent Party 1 = Republican	5.162856	1.321898	3.906638	0.001405	2.346296	7.980416	2.346296371	7.980416041
36	War	6.344981	2.998314	2.116183	0.051462	-0.045774	12.73574	-0.045773784	12.73573823
37	President Running?	3.926576	1.38939	2.826116	0.012768	0.965162	6.887991	0.965162059	6.887990806

**Рис. 60.7.** Выходные данные регрессии для прогнозирования результатов президентских выборов

- 1%-ное отклонение от идеала (нулевой инфляции) означает потерю 0,57% голосов для правящей партии;
- каждый квартал с высоким экономическим ростом во время президентского срока увеличивает долю его голосов на 1,1%;
- по отношению к первому сроку пребывания у власти второй срок уменьшает долю голосов за правящую партию на 2,91%, и каждый последующий срок уменьшает долю голосов еще на  $0,25 \times (2,91\%) = 0,73\%$ ;
- республиканцы имеют преимущество в 5,16% над демократами;
- если США находится в состоянии войны, к прогнозу доли голосов за правящую партию добавляется 6,34%;
- если президент баллотируется на переизбрание, необходимо добавить 3,93% к прогнозу доли голосов за действующего президента.

Остатки согласно уравнению для президентских выборов можно увидеть на рис. 60.8.

Значение в ячейке O18 показывает, что регрессия объясняет 88,9% отклонений в доле голосов за партию у власти. Стандартная ошибка 2,72% показывает, что приблизительно 95% прогноза должно находиться в пределах  $2 \times 2,72 = 5,44\%$ . Поскольку допустимый предел погрешности при опросе в день президентских выборов составляет приблизительно 3%, просто удивительно, насколько хорошо работает модель, даже не учитывающая качества кандидатов! Как видно из рис. 60.8, единственный выброс был в 1992 г. на выборах Буш — Клинтон, в которых Клинтон набрал на 5,41% голосов больше, чем ожидалось. Это может быть связано с тем, что Клинтон умел хорошо подать себя.

	M	N	O	P
41		RESIDUAL OUTPUT		
42				
43	Year	Observation	Predicted Incumbent Share	Residuals
44	1916	1	49.65	2.05
45	1920	2	38.89	-2.79
46	1924	3	57.82	0.48
47	1928	4	56.38	2.42
48	1932	5	38.92	1.98
49	1936	6	63.96	-1.76
50	1940	7	55.74	-0.74
51	1944	8	52.81	0.99
52	1948	9	50.50	1.80
53	1952	10	45.08	-0.38
54	1956	11	55.85	1.25
55	1960	12	50.11	-0.21
56	1964	13	59.97	1.23
57	1968	14	49.69	-0.29
58	1972	15	58.45	3.35
59	1976	16	49.32	-0.32
60	1980	17	45.67	-0.87
61	1984	18	62.01	-2.91
62	1988	19	50.85	2.95
63	1992	20	51.81	-5.41
64	1996	21	52.55	2.15
65	2000	22	51.68	-1.38
66	2004	23	54.77	-3.57

**Рис. 60.8.** Остатки согласно уравнению для президентских выборов

### ❓ Какая функция Excel позволяет составить прогноз по уравнению множественной регрессии?

Довольно утомительно делать прогнозы по уравнению 2, но можно упростить получение прогнозов из множественной регрессии с помощью функции Excel **ТЕНДЕНЦИЯ** (TREND). Для этого даже не требуется запускать регрессию с помощью команды **Анализ данных** (Data Analysis).

Для иллюстрации работы функции **ТЕНДЕНЦИЯ** прогнозирование выборов с 1916 по 2016 г. описано с использованием только данных для выборов с 1916 по 2004 г. См. лист **Trend Function** и рис. 60.9. Сначала выделите диапазон ячеек (в этом примере L7:L32), куда должны быть помещены прогнозы. Установите курсор в первую ячейку этого диапазона (ячейку L7) и введите формулу **=ТЕНДЕНЦИЯ(С7:С29; Е7:К29; Е7:К32; ИСТИНА)**. Затем нажмите **Ctrl+Shift+Enter**. Выбор строк 7–29 в первых двух аргументах функции **ТЕНДЕНЦИЯ** обеспечивает построение уравнения регрессии на основе данных по 1916–2004 гг.

На рис. 60.9 в ячейках L7:L32 представлен сгенерированный прогноз результатов выборов. Обратите внимание, что в 2008 г. партия у власти (республиканцы) получила на 1,8% меньше, чем по прогнозу; в 2012 г. правящие демократы получили на 3,0% больше, чем по прогнозу. Хотя результаты опросов и оценки экспертов предсказывали Х. Клинтон безоговорочную победу в 2016 г, наша модель говорила, что Клинтон получит на голосовании только 45%. В реальности она получила 51% на двухэтапном народном голосовании на прямых выборах!



	A	B	L	M
6	Candidates	Year	Prediction	Abs error
7	Wilson-Hughes	1916	49.651271	2.048729
8	Harding-Cox	1920	38.889038	2.789038
9	Coolidge-Davis	1924	57.824144	0.475856
10	Hoover-Smith	1928	56.376521	2.423479
11	Hoover-FDR	1932	38.92051	1.97949
12	FDR-Landon	1936	63.962187	1.762187
13	FDR-Wilkie	1940	55.738371	0.738371
14	FDR-Dewey	1944	52.810153	0.989847
15	Truman-Dewey	1948	50.500809	1.799191
16	Ike-Stevenson	1952	45.07892	0.37892
17	Ike-Stevenson	1956	55.84732	1.25268
18	Kennedy-Nixon	1960	50.106938	0.206938
19	Johnson-Goldwater	1964	59.973844	1.226156
20	Nixon-Humphrey	1968	49.693889	0.293889
21	Nixon-McGovern	1972	58.449071	3.350929
22	Ford-Carter	1976	49.32411	0.32411
23	Carter-Reagan	1980	45.674726	0.874726
24	Reagan-Mondale	1984	62.011728	2.911728
25	Bush-Dukakis	1988	50.853061	2.946939
26	Bush-Clinton	1992	51.812336	5.412336
27	Clinton-Dole	1996	52.549267	2.150733
28	Gore Bush	2000	51.677525	1.377525
29	Bush Kerry	2004	54.774262	3.574262
30	Obama McCain	2008	44.464091	1.835909
31	Obama Romney	2012	48.881984	3.018016
32	Clinton Trump	2016	44.509431	6.590569

Рис. 60.9. Прогноз для всех прошедших выборов

Функция **ТЕНДЕНЦИЯ** является функцией, возвращающей массив. Более подробное обсуждение этих функций см. в главе 91 «Формулы массива и функции, возвращающие массив». А здесь приведена некоторая базовая информация о функциях, возвращающих массив.

- Перед определением функции, возвращающей массив, необходимо определить диапазон ячеек, в который будет размещен результат работы этой функции.
- Для выполнения вычислений необходимо нажать не **Enter**, а сочетание клавиш **Ctrl+Shift+Enter**, завершающее ввод функции, возвращающей массив.
- После ввода функции, возвращающей массив, при выборе какой-либо ячейки из тех, в которых размещены результаты этой функции, вы увидите в строке формул формулу в фигурных скобках. Эти скобки указывают на то, что результат в этой ячейке получен в результате работы функции, возвращающей массив.
- Никакие данные из диапазона, созданные функцией, возвращающей массив, не могут быть изменены.

Задания по множественной регрессии приведены в конце главы 61.

## ГЛАВА 61

# Моделирование нелинейных характеристик и взаимосвязей

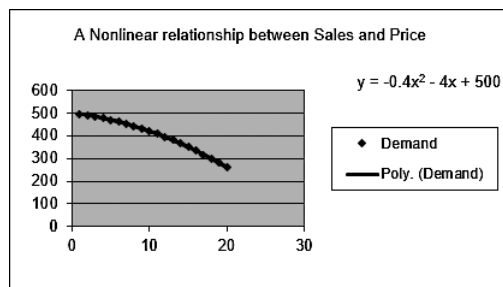
### Обсуждаемые вопросы

- Когда говорят, что независимая переменная оказывает нелинейное воздействие на зависимую переменную, что это означает?
- Когда говорят, что между воздействиями двух независимых переменных на зависимую переменную существует взаимосвязь, что это означает?
- Как выполнить проверку на наличие нелинейности и взаимосвязи в регрессии?

## Ответы на вопросы

❓ Когда говорят, что независимая переменная оказывает нелинейное воздействие на зависимую переменную, что это означает?

Зачастую независимая переменная воздействует на зависимую переменную нелинейным образом. Например, при прогнозе продаж продукта по уравнению вида  $\text{Продажи} = 500 - 10 \times \text{Цена}$  цена влияет на объем продаж *линейно*. Это уравнение показывает, что при повышении цены на единицу (для любого уровня цен) объем продаж уменьшается на 10 единиц. Если соотношение между объемом продаж и ценой определяется уравнением  $\text{Продажи} = 500 + 4 \times \text{Цена} - 0,40 \times \text{Цена}^2$ , цена и объем продаж связаны нелинейно. Как показано на рис. 61.1, чем больше повышается цена, тем сильнее падает спрос (см. лист Nonlinearity в файле Interactions.xlsx).



**Рис. 61.1.** Нелинейная зависимость между спросом и ценой



Коротко говоря, если изменение зависимой переменной, вызванное изменением на единицу независимой переменной, не является константой, между независимой и зависимой переменными имеется нелинейная зависимость.

**❓ Когда говорят, что между воздействиями двух независимых переменных на зависимую переменную существует взаимосвязь, что это означает?**

Если влияние одной независимой переменной на зависимую переменную зависит от значения другой независимой переменной, можно сказать, что две независимые переменные проявляют *взаимосвязь*. Рассмотрим для примера прогноз продаж в зависимости от цены и суммы, затраченной на рекламу. Если эффект от изменения расходов на рекламу значим при низкой цене и незначим при высокой цене, цена и расходы на рекламу проявляют взаимосвязь. Если эффект от изменения расходов на рекламу одинаков при любой цене, затраты на рекламу и цена не связаны.

**❓ Как выполнить проверку на наличие нелинейности и взаимосвязи в регрессии?**

Чтобы проверить, имеет ли независимая переменная нелинейный эффект на зависимую переменную, просто добавьте в уравнение регрессии член, равный квадрату независимой переменной. Если квадратный член имеет небольшое *p*-значение (меньше 0,15), это доказывает наличие нелинейной зависимости.

Для проверки наличия взаимосвязи между двумя независимыми переменными добавьте в уравнение регрессии член, равный произведению независимых переменных. Если член имеет небольшое *p*-значение (меньше 0,15), это является доказательством взаимосвязи.

Чтобы вы лучше поняли, я попытаюсь определить, как половая принадлежность и опыт влияют на размер оклада в небольшой производственной компании. Для каждого сотрудника в базе хранятся приведенные ниже данные. Информацию см. на листе **Data** в файле **Interactions.xlsx** (рис. 61.2).

- Годовая зарплата (в тысячах долларов).
- Опыт работы на производстве (в годах).
- Пол (1 — женский, 0 — мужской).

Используем эти данные, чтобы спрогнозировать зарплату (зависимую переменную) в зависимости от опыта (**Exp**) и от пола (**Gender**). Для проверки нелинейности влияния опыта на зарплату я добавил член **Опыт в квадрате**, скопировав формулу  $=B2^2$  из D2 в D3:D98. На листе **Data** замените D2 формулой  $=B2^2$ . Затем скопируйте D2 в D3:D98. Для проверки наличия значимого взаимодействия между опытом и полом я добавил член **Опыт\*Пол**, скопировав формулу  $=B2*C2$  из E2 в E3:E98. Замените E2 на формулу  $=B2*C2$ . В диалоговом окне **Регрессия (Regression)** я указал A1:A98 как **Входной интервал Y (Input Y Range)** и B1:E98 как **Входной интервал X (Input X Range)**. Установив флажок **Метки (Labels)** и нажав **ОК**, я запустил регрессионный анализ и получил результаты, как на рис. 61.3.

	A	B	C	D	E
1	Salary	Exp	Gender	Exp^2	Exp*Gender
2	55.92199	5	1	25	5
3	62.39869	15	0	225	0
4	67.84753	12	0	144	0
5	64.80815	19	0	361	0
6	58.29632	9	0	81	0
7	51.93969	3	0	9	0
8	36.34266	12	1	144	12
9	28.38828	15	1	225	15
10	73.57633	8	0	64	0
11	44.16733	12	1	144	12
12	57.03224	1	1	1	1
13	61.58797	5	0	25	0
14	63.21908	14	0	196	0
15	58.15548	1	1	1	1
16	47.4811	3	1	9	3
17	44.43084	6	1	36	6
18	59.10713	4	0	16	0
19	64.55505	9	0	81	0
20	64.65747	18	0	324	0
21	49.09593	7	1	49	7
22	69.01037	14	0	196	0
23	64.99559	16	0	256	0

Рис. 61.2. Данные для прогнозирования зарплаты в зависимости от пола и опыта

	H	I	J	K	L	M
1						
2	SUMMARY OUTPUT					
3						
4	<b>Regression Statistics</b>					
5	Multiple R	0.930645916				
6	R Square	0.866101822				
7	Adjusted R Square	0.860280162				
8	Standard Error	4.530277908				
9	Observations	97				
10						
11	<b>ANOVA</b>					
12		<b>df</b>	<b>SS</b>	<b>MS</b>	<b>F</b>	<b>Significance F</b>
13	Regression	4	12213.26555	3053.316	148.7723146	2.77158E-39
14	Residual	92	1888.154449	20.52342		
15	Total	96	14101.42			
16						
17		<b>Coefficients</b>	<b>Standard Error</b>	<b>t Stat</b>	<b>P-value</b>	<b>Lower 95%</b>
18	Intercept	58.30311924	1.960660525	29.73647	5.6775E-49	54.40907819
19	Exp	0.860472831	0.381574828	2.255057	0.02650141	0.102632286
20	Gender	1.119399728	1.94092596	0.576735	0.565527049	-2.735446764
21	Exp^2	-0.035148705	0.017761027	-1.978979	0.050808362	-0.070423637
22	Exp*Gender	-2.164501713	0.180872223	-11.96702	1.86636E-20	-2.523729559

Рис. 61.3. Результаты регрессии для теста на нелинейность и взаимодействие

Как видно из рисунка, пол не является значимым ( $p$ -значение для него больше 0,15). Все остальные независимые переменные — значимые (т. е. имеют  $p$ -значение не больше 0,15). Не оказывающую влияние переменную пола можно удалить из независимых переменных. Для этого я скопировал данные на новый лист с именем FinalRegression, щелкнул правой кнопкой мыши на вкладке листа, выбрал Переместить или скопировать (Move Or Copy) и в диалоговом окне установил флажок

Создать копию (Create A Copy). После удаления столбца Пол можно получить результаты регрессии, как на листе FinalRegression и на рис. 61.4.

	H	I	J	K	L	M
1						
2	SUMMARY OUTPUT					
3						
4	Regression Statistics					
5	Multiple R	0.9303858				
6	R Square	0.8656177				
7	Adjusted R Square	0.8612828				
8	Standard Error	4.5139938				
9	Observations	97				
10						
11	ANOVA					
12		df	SS	MS	F	Significance F
13	Regression	3	12206.43899	4068.813	199.6852	2.12292E-40
14	Residual	93	1894.981012	20.37614		
15	Total	96	14101.42			
16						
17		Coefficients	Standard Error	t Stat	P-value	Lower 95%
18	Intercept	59.057442	1.455412135	40.57781	6.06E-61	56.16728518
19	Exp	0.7811172	0.354623849	2.202664	0.03009	0.076905191
20	Exp^2	-0.0335942	0.017492209	-1.92052	0.057856	-0.06833021
21	Exp*Gender	-2.0733945	0.087774433	-23.6219	4.49E-41	-2.247696986

Рис. 61.4. Результаты регрессии после удаления незначимой переменной Пол

Теперь все независимые переменные являются значимыми (имеют  $p$ -значение не больше 0,15). Следовательно, можно предсказать зарплату (в тысячах долларов) по следующему уравнению (уравнение 3):

прогнозируемая зарплата =  $59,06 + 0,78 * (\text{опыт}) - 0,033 * (\text{опыт}^2) - 2,07 * (\text{опыт} * \text{пол})$ .

Отрицательный член ( $\text{опыт}^2$ ) показывает, что каждый дополнительный год опыта все меньше влияет на зарплату. То есть опыт оказывает нелинейное воздействие на зарплату. Фактически эта модель говорит о том, что после 13 лет работы каждый дополнительный год опыта на самом деле уменьшает зарплату.

Напомню, что переменная Пол равна 1 для женщин и 0 для мужчин. После подстановки 1 в уравнение 3 зарплату для женщин можно спрогнозировать следующим образом:

прогнозируемая зарплата =  $59,06 + 0,78 * (\text{опыт}) - 0,033 * (\text{опыт}^2) - 2,07 * (\text{опыт} * 1) = 59,06 - 0,033 * (\text{опыт}^2) - 1,29 * (\text{опыт})$ .

Для мужчин (после подстановки  $\text{пол} = 0$ ) получится следующее уравнение:

прогнозируемая зарплата =  $59,06 + 0,78 * (\text{опыт}) - 0,033 * (\text{опыт}^2) - 2,07 * (\text{опыт} * 0) = 59,06 + 0,78 * (\text{опыт}) - 0,033 * (\text{опыт}^2)$ .

Таким образом, взаимосвязь между полом и опытом показывает, что каждый дополнительный год опыта работы приносит женщинам выгоды в среднем на  $0,78 - (-1,29) = \$2070$  меньше, чем мужчинам. Это указывает на несправедливое отношение к женщинам.

## Задания к главам 59–61

Компания Fizzy Drugs должна оптимизировать размер выработки от важного химического процесса. Компания считает, что количество произведенной продукции в фунтах при каждом запуске процесса зависит от размера используемого резервуара, давления и температуры. Привлеченные к решению задачи ученые полагают, что влияние одной переменной может зависеть от значений других переменных. Резервуар должен иметь объем 1,3–1,5 м<sup>3</sup>, кроме того, необходимо поддерживать давление между 4 и 4,5 мм рт. ст. и температуру 22–30 °С. Ученые провели эксперименты для нижних и верхних уровней трех контрольных переменных, результаты которых представлены в файле *Fizzy.xlsx*. Используйте данные из этого файла для ответа на задания 1–3.

1. Определите связь между количеством продукции, объемом резервуара, температурой и давлением.
2. Рассмотрите взаимосвязь между давлением, объемом резервуара и температурой.
3. Какие значения температуры и давления и какой объем резервуара вы бы рекомендовали?

Еще несколько заданий по множественной регрессии.

4. В супермаркете Mr. D's в течение 12 недель подряд проводились наблюдения за продажами (количество покупок) консервированных томатов. (См. файл *Grocery.xlsx*.) Каждую неделю отслеживались следующие данные:
  - Во все ли тележки была положена реклама консервированных томатов?
  - Каждому ли покупателю был вручен скидочный купон на консервированные томаты?
  - Была ли снижена цена (нет, на 1 или 2 цента)?

На основе этих данных определите, как перечисленные факторы влияют на продажи. Предскажите продажи консервированных томатов для недели, на которой в тележки были положены рекламные листовки, покупателям предлагался купон и цена на томаты была снижена на 1 цент.

5. В файле *Countryregion.xlsx* для нескольких слаборазвитых стран/регионов содержатся следующие данные:
  - детская смертность;
  - уровень грамотности взрослого населения;
  - процент учеников, закончивших начальную школу;
  - ВВП на душу населения.

Исходя из этих данных, создайте уравнение для прогнозирования детской смертности. Имеются ли в этом наборе данных выбросы? Проинтерпретируйте коэффициенты полученного уравнения. В пределах каких значений 95% прогнозов по детской смертности должны быть точными?

6. В файле **Baseball96.xlsx** содержатся следующие данные: засчитанные пробежки, одинарные, двойные, тройные удары, хоум-раны, кражи базы и уолки для каждой команды Главной лиги бейсбола в сезоне 1996 г. На основе этих данных определите влияние одиночных и двойных ударов, а также других действий на пробежки.
7. В файле **Cardata.xlsx** приведена следующая информация для 392 моделей автомобилей:
  - количество цилиндров;
  - объем двигателя;
  - мощность в лошадиных силах;
  - вес;
  - разгон;
  - количество миль на галлон горючего.

Определите уравнение для прогнозирования количества миль на галлон горючего. Почему все независимые переменные не являются значимыми?

8. В файле **Priceads.xlsx** находятся данные о еженедельных объемах продаж продукта, а также о цене продукта и расходах на рекламу. Объясните, как реклама и цена продукта влияют на объемы продаж.
9. Файл **Teams.xlsx** содержит следующие данные: засчитанные пробежки, хит-бай-питчи, уолки, одиночные, двойные, тройные удары и хоум-раны для команд Главной лиги бейсбола в сезонах 2000–2006 гг.
  - Исходя из этих данных, постройте модель, предсказывающую на основе статистики ударов, сколько очков получит команда за пробежки.
  - Насколько точно эта модель предсказывает количество засчитанных пробежек для команды за сезон?
  - Предположим, что команда имеет положительный остаток, превышающий в три раза стандартную ошибку регрессии. Что могло привести к такому результату?
10. Процент занятия базы (ОВР) вычисляется как (хит-бай-питч + уолк + хит) / (выход на биту + хит-бай-питч + уолк + сакрифайс флай). Процент мощности отбивания (SLG) равен количеству баз (хоум-ран дает очки за четыре базы, тройной удар — за три, двойной удар — за две и одинарный — за одну), деленному на количество выходов на биту. В файле **OPSslug.xlsx** содержится информация об ударах для команд Главной лиги в сезонах 2000–2006 гг.

На основе этих данных создайте модель для прогнозирования засчитанных пробежек по ОВР и SLG команды.

- Какую из моделей, эту или модель из задания 9, можно порекомендовать для прогнозирования засчитанных пробежек?
- Как с помощью этой модели можно определить способность к отбиванию у отдельных игроков?

- Почему можно отнестись скептически к вычислению с помощью этой модели способности к отбиванию у такого действительно великого хиттера, как Альберт Пухолс (Albert Pujols) из «Сент-Луис Кардиналс»?

11. Файл **NFLinfo.xlsx** содержит следующие данные для команд НФЛ в сезонах 2003–2006 гг.:

- Margin (преимущество) — количество очков, на которое команда опережает своего соперника;
- NYP/A — набранные ярды относительно попыток выполнить пас;
- YR/A — набранные ярды относительно попыток выполнить вынос;
- DNYR/A — набранные противником ярды относительно попыток выполнить пас;
- DYR/A — набранные противником ярды относительно попыток выполнить вынос;
- TO — перехваты, выполненные командой;
- DTO — перехваты, выполненные командой противника.

Показывают ли эти данные, что больше влияет на успех команды: эффективный пас или вынос?

В какое количество очков приблизительно обходится команде перехват мяча?

Многие футбольные болельщики считают, что для организации паса необходим эффективный вынос. Соответствуют ли данные такой точке зрения?

12. В файле **Qbinfo.xlsx** перечислены рейтинги знаменитых квотербеков НФЛ в сезоне 2009 г. Рейтинг квотербеков основан на проценте завершенных пасов; проценте пасов, приведших к тачдаунам; процентах набранных ярдов относительно попыток выполнить пас и проценте перехваченных пасов.

- Создайте формулу прогнозирования рейтинга квотербека на основе его статистики за сезон.
- Насколько точны прогнозы?

13. В файле **Problem13data.xlsx** приведены еженедельные данные о продаже печенья в супермаркете. Кроме того, приведены еженедельные данные для каждого продукта о наличии скидки и/или о его рекламной выкладке. Объясните, как цена и рекламная выкладка влияют на продажи печенья. Между какими значениями, \_\_\_\_\_ и \_\_\_\_\_, должны находиться объемы продаж с точностью 95% в неделю, когда продукт находится на рекламной выкладке и цена на него снижена?

14. Фармкомпаниям требуется определить, как давление, температура и размер резервуара влияют на объем выхода нового лекарства. Даны значения выхода для нескольких партий этого нового лекарства, а также давление, температура и размер резервуара для каждой партии. Определите уравнение, объясняющее, как эти факторы влияют на выход. Если процесс проходит при высоком

давлении, низкой температуре и в большом резервуаре, можно с вероятностью 95% утверждать, что выход будет между \_\_\_\_\_ и \_\_\_\_\_. Можете игнорировать нелинейности и взаимодействия.

15. В файле **Problem15data.xls** приведены данные о количестве открыток, проданных аптекой в течение нескольких кварталов, и средняя цена за открытки в течение квартала. Определите, как цена и сезонность влияют на продажу открыток. Если в четвертом квартале вы продаете открытки по цене 4 доллара за штуку, то с вероятностью в 95% будет продано от \_\_\_\_\_ до \_\_\_\_\_ открыток.
16. В файле **Problem16data.xlsx** приведены данные о количестве круассанов с миндальным кремом, продаваемых в кондитерской каждый день. Для каждого дня приведены данные о том, был ли день солнечный или пасмурный. Создайте уравнение, прогнозирующее ежедневные продажи круассанов. Объясните, как погода и день недели влияют на продажи. В пасмурный выходной день можно с вероятностью в 95% утверждать, что продажи будут в интервале между \_\_\_\_\_ и \_\_\_\_\_.

## ГЛАВА 62

# Однофакторный дисперсионный анализ

### Обсуждаемые вопросы

- Владелец издательства, публикующего компьютерную литературу, хочет знать, влияет ли на продажи местоположение стенда с книгами в отделе компьютерной литературы в книжных магазинах. В частности, имеет ли значение, в какой зоне отдела компьютерной литературы (передней, задней или средней) находится стенд с книгами?
- Если я определяю, имеют ли популяции значимо различные средние значения, почему метод называется дисперсионным анализом?
- Как применить результаты однофакторного дисперсионного анализа для прогнозирования?

Аналитики данных часто имеют дело с данными о нескольких группах людей или предметов и стараются определить, различаются ли данные о группах на значимую величину. Вот несколько примеров:

- Есть ли значимое различие в длительности, на которую четыре врача оставляют своих пациенток в больнице после родов?
- Зависит ли количество производимого нового препарата от размера резервуара, в котором он производится (большой, маленький или средний)?
- Было ли вызвано падение давления после приема одного из четырех препаратов?

Если вы пытаетесь определить, являются ли средние значения для нескольких наборов данных, зависящих от одного фактора, значимо различными, следует воспользоваться однофакторным дисперсионным анализом. В перечисленных примерах к таким факторам относятся врачи, размер резервуара и препараты соответственно. При анализе данных можно выбрать одну из двух гипотез.

- Нулевую гипотезу, согласно которой средние значения всех групп идентичны.
- Альтернативную гипотезу, согласно которой между средними значениями групп имеется статистически значимое различие.

Для проверки этих гипотез в Microsoft Excel есть инструмент **Однофакторный дисперсионный анализ (Anova: Single Factor)**, который выбирается в диалоговом окне



**Анализ данных (Data Analysis).** Если вычисленное  $p$ -значение окажется небольшим (как правило, не больше 0,15), можно сделать вывод, что верна альтернативная гипотеза (средние значения значимо различны). Если  $p$ -значение больше 0,15, верна нулевая гипотеза (генеральные совокупности имеют одинаковые средние значения). Давайте рассмотрим пример.

## Ответы на вопросы

**?** Владелец издательства, публикующего компьютерную литературу, хочет знать, влияет ли на продажи местоположение стенда с книгами в отделе компьютерной литературы в книжных магазинах. В частности, имеет ли значение, в какой зоне отдела компьютерной литературы (передней, задней или средней) находится стенд с книгами?

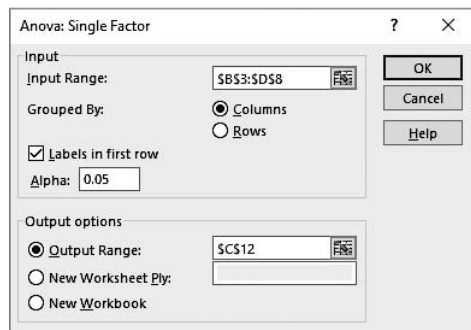
Издательству необходимо знать, лучше ли продаются книги в зависимости от того, выставлены ли они на стенде в передней, задней или средней зоне отдела компьютерной литературы. В 12 магазинах были отслежены еженедельные продажи (в сотнях штук). В пяти магазинах стенды были установлены в передней зоне отдела, в четырех магазинах — в задней зоне и в трех магазинах — в средней. Полученные объемы продаж представлены на листе **Signif** в файле **Onewayanovatemp.xlsx** (рис. 62.1). Указывают ли данные на значимость местоположения стенда или нет?

	A	B	C	D
1	<b>One-Way ANOVA</b>			
2				
3		<b>Front</b>	<b>Back</b>	<b>Middle</b>
4		7	12	10
5		10	13	11
6		8	15	12
7		9	16	
8		11		

**Рис. 62.1.** Данные о продажах книг

Допустим, во всех 12 магазинах одинаковая структура продаж и они приблизительно одинаковые по площади. Это предположение позволяет воспользоваться однофакторным дисперсионным анализом, поскольку в таком случае можно считать, что на продажи влияет не более одного фактора (а именно положение стенда с книгами в отделе компьютерной литературы). (Если бы магазины были разных размеров, данные пришлось бы анализировать с помощью двухфакторного дисперсионного анализа, рассматриваемого в главе 63 «Рандомизированные блоки и двухфакторный дисперсионный анализ».)

Для анализа данных на вкладке **Данные (Data)** щелкните на кнопке **Анализ данных (Data Analysis)** и выберите инструмент **Однофакторный дисперсионный анализ (Anova: Single Factor)**. Заполните диалоговое окно, как на рис. 62.2.



**Рис. 62.2.** Диалоговое окно Однофакторный дисперсионный анализ

Я использовал следующие конфигурации:

- данные для входного интервала, включая метки, находятся в ячейках B3:D8;
- выбран флажок **Метки в первой строке** (Labels in first row), поскольку первая строка входного интервала содержит метки;
- я установил переключатель в положение **по столбцам** (Columns), поскольку данные объединены в столбцы;
- я указал ячейку C12 как левую верхнюю ячейку выходного интервала ( Output Range);
- выбор значения **Альфа** (Alpha) несуществен. Можно оставить значение по умолчанию.

После нажатия ОК появятся результаты, показанные на рис. 62.3.

	C	D	E	F	G	H	I
11							
12	Anova: Single Factor						
13							
14	SUMMARY						
15	Groups	Count	Sum	Average	Variance		
16	Front	5	45	9	2.5		
17	Back	4	56	14	3.333333		
18	Middle	3	33	11	1		
19							
20							
21	ANOVA						
22	Source of Variation	SS	df	MS	F	P-value	F crit
23	Between Groups	55.66667	2	27.83333	11.38636	0.00343	4.25649
24	Within Groups	22	9	2.444444			
25							
26	Total	77.66667	11				
27							
28	est std error	1.563472					

**Рис. 62.3.** Результаты однофакторного дисперсионного анализа

В ячейках F16:F18 вы видите средний объем продаж в зависимости от местоположения стенда. Когда стенд располагается в передней зоне отдела компьютерной

литературы, средний объем продаж составляет 900 шт.; когда он располагается в задней зоне отдела, то средний объем продаж равен 1400 шт.; когда он располагается в средней зоне, средний объем продаж составляет 1100 шт. Поскольку  $p$ -значение, равное 0,003 (в ячейке H23), меньше 0,15, можно сделать вывод, что эти средние значения значимо различны.

**?** Если я определяю, имеют ли популяции значимо различные средние значения, почему метод называется дисперсионным анализом?

Предположим, что данные о продаже книг — это данные на листе **Insig** в файле **Onewayanovatemp.xlsx** (рис. 62.4). Если вы запустите однофакторный дисперсионный анализ для этих данных, то получите результаты, представленные на рис. 62.5.

	B	C	D
3	Front	Back	Middle
4	7	2	3
5	20	16	19
6	8	25	11
7	8	13	
8	2		

**Рис. 62.4.** Данные о продажах, для которых принята нулевая гипотеза

	E	F	G	H	I	J	K
12	Anova: Single Factor		Overall mean				
13			11.1666667				
14	SUMMARY						
15	Groups	Count	Sum	Average	Variance		
16	Front	5	45	9	44		
17	Back	4	56	14	90		
18	Middle	3	33	11	64		
19							
20							
21	ANOVA						
22	Source of Variation	SS	df	MS	F	P-value	F crit
23	Between Groups	55.66667	2	27.83333	0.436411	0.659334	4.25649
24	Within Groups	574	9	63.77778			
25							
26	Total	629.6667	11				
27							
28	est std err	7.986099					

**Рис. 62.5.** Результаты дисперсионного анализа, для которых принята нулевая гипотеза

Обратите внимание, что средний объем продаж для каждой зоны отдела точно такой же, как раньше. Тем не менее  $p$ -значение, равное 0,66, показывает, что следует принять нулевую гипотезу и сделать вывод об отсутствии влияния размещения стенда в отделе компьютерной литературы на объемы продаж. Причина такого странного результата в том, что во втором наборе данных вариация в объемах продаж для каждого положения стенда в секции более значительна. В первом наборе данных, например, вариация в объемах продаж, когда стенд находится в передней

зоне отдела, составляет от 700 до 1100 книг, в то время как во втором наборе вариация в объемах продаж для стенда в той же зоне составляет от 200 до 2000 книг. Вариация продаж для каждого местоположения стенда измеряется суммой квадратов отклонений от среднего для данных внутри групп. Эта мера приведена в ячейке D24 в первом наборе данных и в ячейке F24 во втором. В первом наборе данных сумма квадратов отклонений от среднего для данных внутри групп равна 22, в то время как во втором наборе данных сумма квадратов внутри групп равна 574! Такая большая вариация точек данных для каждого местоположения стенда маскирует вариацию между самими группами (положениями стенда), и для второго набора данных нет возможности заключить, что разница между объемами продаж для разных положений стенда значима.

### **❓ Как применить результаты однофакторного дисперсионного анализа для прогнозирования?**

Если между средними значениями по каждой группе есть значимая вариация, наилучшим прогнозом для каждой группы будет среднее значение по этой группе. Следовательно, для первого набора данных можно предсказать следующее:

- объем продаж для стенда в передней зоне отдела компьютерной литературы составит 900 книг в неделю;
- объем продаж для стенда в задней зоне отдела компьютерной литературы составит 1400 книг в неделю;
- объем продаж для стенда в средней зоне отдела компьютерной литературы составит 1100 книг в неделю.

Если значимая вариация отсутствует, наилучшим прогнозом для каждого наблюдения будет просто общее среднее. Таким образом, для второго набора данных можно предсказать, что еженедельные продажи составят 1117 книг независимо от того, где расположен стенд.

Так же можно оценить точность прогнозов. Средним квадратичным отклонением прогноза по результатам однофакторного дисперсионного анализа является корень квадратный из среднего квадратичного (MS) внутри групп. Как показано на рис. 62.3, среднее квадратичное отклонение прогноза для первого набора данных составляет 1,56 (см. лист Signif). Это означает, что следует ожидать, например, следующих результатов:

- для 68% недель, в течение которых книги будут находиться в передней зоне отдела компьютерной литературы, объем продаж составит от  $900 - 156 = 744$  до  $900 + 156 = 1056$  книг;
- для 95% недель, в течение которых книги будут находиться в передней зоне отдела компьютерной литературы, объем продаж составит от  $900 - 2 \times (156) = 588$  до  $900 + 2 \times (156) = 1212$  книг.

## Задания

Данные к этим заданиям находятся в файле `Chapter62data.xlsx`.

1. Для пациентов четырех кардиологов известно количество дней, в течение которых пациенты оставались в больнице после операции на открытом сердце.
  - Существуют ли доказательства того, что врачи проводят различную политику выписки из лечебного учреждения?
  - В течение скольких дней пациент первого кардиолога будет оставаться в больнице с вероятностью 95%?
2. Лекарственный препарат может быть изготовлен в термостате с рабочей температурой 400, 300 или 200 градусов. Дано количество полученного препарата в фунтах для различных партий, изготовленных при разных температурах.
  - Влияет ли температура на объем выработки препарата?
  - Каков интервал в фунтах, в котором с вероятностью 95% окажется выработка продукта из термостата с рабочей температурой 200 градусов?
  - Если считать, что давление внутри корпуса также влияет на результативность процесса, остается ли анализ надежным?

## ГЛАВА 63

# Рандомизированные блоки и двухфакторный дисперсионный анализ

### Обсуждаемые вопросы

- Я пытаюсь проанализировать эффективность отдела продаж. Проблема в том, что кроме эффективности самого торгового представителя сумма его продаж зависит от региона, в который он назначен. Как можно включить в анализ распределение продавцов по регионам?
- Как предсказать объем продаж на основе информации о торговых представителях и регионах? Насколько точными окажутся прогнозы?
- Как определить, влияют ли на продажи видеоигр изменение цены и расходов на рекламу? Как определить, значимо ли взаимодействие цены и рекламы?
- Как интерпретировать влияние цены и рекламы на продажи при отсутствии значимого взаимодействия между ценой и рекламой?

Во многих наборах данных на зависимую переменную могут оказывать влияние два фактора. В следующей таблице приведено несколько подобных примеров.

Факторы	Зависимая переменная
Продавец и назначенный регион	Объем продаж
Цена продукта и расходы на рекламу	Объем продаж
Температура и давление	Выход продукции
Хирург и марка используемого стента	Здоровье пациента после операции на открытом сердце

Если на зависимую переменную могут повлиять два фактора, то для определения фактора, оказывающего значительное влияние на зависимую переменную (если это имеет место), вы можете использовать либо анализ с рандомизированными блоками, либо двухфакторный дисперсионный анализ с повторениями. Кроме того, двухфакторный дисперсионный анализ позволяет определить, проявляют ли два фактора значимое взаимодействие. Допустим, вам необходимо спрогнозировать объем

продаж в зависимости от цены продукта и рекламного бюджета. Цена и реклама значимо влияют друг на друга, если эффект от рекламы зависит от цены продукта.

В модели с рандомизированными блоками каждая возможная комбинация факторов наблюдается только один раз. Вы не можете проверить взаимодействие в рандомизированном блочном плане. В модели двухфакторного дисперсионного анализа каждая комбинация факторов наблюдается одинаковое число раз (назовем его  $k$ ). Для проверки взаимодействия  $k$  должно быть больше 1. В модели двухфакторного дисперсионного анализа проверить взаимодействие не составляет труда.

## Ответы на вопросы

**?** Я пытаюсь проанализировать эффективность отдела продаж. Проблема в том, что кроме эффективности самого торгового представителя сумма его продаж зависит от региона, в который он назначен. Как можно включить в анализ распределение продавцов по регионам?

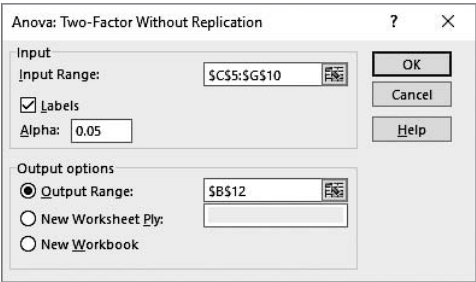
Пусть необходимо определить, как продавец и регион продаж, в который он назначен, влияют на объемы продаж продукта. Допустим, каждый из четырех торговых представителей провел в каждом из пяти регионов один месяц. Полученные в результате объемы продаж приведены на листе **Randomized Blocks** в файле **Twowayanova.xlsx** (рис. 63.1). Например, первый представитель продал за месяц 20 единиц продукции, когда работал в регионе 4.

	C	D	E	F	G
5		Rep 1	Rep 2	Rep 3	Rep 4
6	Dist 1	1	3	10	12
7	Dist 2	17	12	16	14
8	Dist 3	17	21	22	25
9	Dist 4	20	10	17	23
10	Dist 5	22	21	37	32

**Рис. 63.1.** Данные для примера с рандомизированными блоками

Эта модель называется двухфакторным дисперсионным анализом без повторений, так как на объем продаж могут оказывать влияние два фактора (регион и торговый представитель) при наличии единственного наблюдения для каждой комбинации «представитель — регион». Также эта модель называется *рандомизированным блочным планом*, поскольку в ней можно рандомизировать (в хронологическом порядке) распределение представителей по регионам. Другими словами, необходимо убедиться, что месяц, в течение которого представитель 1 работал в регионе 1, с одинаковой вероятностью может быть первым, вторым, третьим, четвертым или пятым месяцем. Есть надежда, что такая рандомизация уменьшает влияние временного периода на анализ (предположительно, со временем квалификация торгового представителя повышается); и в каком-то смысле «блокируется» влияние выбора региона на результат при попытке сравнения торговых представителей.

Для анализа этих данных в Microsoft Excel на вкладке **Данные (Data)** выберите **Анализ данных (Data Analysis)** и затем инструмент **Двухфакторный дисперсионный анализ без повторений (Anova: Two-Factor Without Replication)**. (Необходимо установить пакет **Анализ данных**. Инструкции см. в главе 44 «Обобщение данных с помощью описательной статистики».) Заполните диалоговое окно, как на рис. 63.2.



**Рис. 63.2.** Диалоговое окно Двухфакторный дисперсионный анализ без повторений для настройки модели с рандомизированными блоками

Я использовал следующую информацию:

- входной интервал находится в ячейках C5:G10;
- я выбрал флажок **Метки (Labels)**, поскольку первая строка входного интервала содержит метки;
- я ввел B12 как левую верхнюю ячейку выходного интервала;
- выбор значения **Альфа (Alpha)** несущественен. Можно оставить значение по умолчанию.

	B	C	D	E	F	G	H
12	Anova: Two-Factor Without Replication					17.6	
13							
14	SUMMARY	Count	Sum	Average	Variance		
15	Dist 1	4	26	6.5	28.3333	-11.1	
16	Dist 2	4	59	14.75	4.91667	-2.85	
17	Dist 3	4	85	21.25	10.9167	3.65	
18	Dist 4	4	70	17.5	31	-0.1	
19	Dist 5	4	112	28	60.6667	10.4	
20						-17.6	
21	Rep 1	5	77	15.4	69.3	-2.2	
22	Rep 2	5	67	13.4	59.3	-4.2	
23	Rep 3	5	102	20.4	104.3	2.8	
24	Rep 4	5	106	21.2	67.7	3.6	
25							
26							
27	ANOVA						
28	Source of Variation	SS	df	MS	F	P-value	F crit
29	Rows	1011	4	252.825	15.876	9.7E-05	3.2592
30	Columns	216.4	3	72.1333	4.52957	0.0241	3.4903
31	Error	191.1	12	15.925			
32			stdev	3.99061			
33	Total	1419	19				

**Рис. 63.3.** Выходные данные рандомизированного блочного плана



Результат анализа представлен на рис. 63.3. (Результаты в ячейках G12:G24 созданы не с помощью инструмента Excel Анализ данных. Как я ввел формулы в эти ячейки, объясню далее в этой главе.)

Оказывает ли фактор в строках (регионы) или фактор в столбцах (торговые представители) значимый эффект на продажи, можно определить *post-hoc*-значению. Если *p*-значение для фактора небольшое (меньше 0,15), фактор оказывает значимый эффект на продажи. Обратные значения, для строк (0,0000974) и для столбцов (0,024), меньше 0,15, поэтому и регион, и представитель оказывают существенное влияние на продажи.

### ❓ Как предсказать объем продаж на основе информации о торговых представителях и регионах? Насколько точными окажутся прогнозы?

Как предсказать объемы продаж продукта? Продажи за месяц можно предсказать по следующему уравнению 1:

прогнозируемый объем продаж = общее\_среднее + (влияние\_представителя) + (влияние\_региона).

В этом уравнении (влияние\_представителя) равно 0, если фактор «торговый представитель» не значимый. Если он значимый, (влияние\_представителя) равно среднему значению для данного представителя минус общее среднее. Аналогично, (влияние\_региона) равно 0, если фактор «регион» не значимый. Если он значимый, (влияние\_региона) равно среднему значению для данного региона минус общее среднее.

Я вычислил общее среднее (17,6) в ячейке G12 по формуле =CPЗНАЧ(D6:G10). Влияние представителей и регионов вычисляется путем копирования формулы =E15-\$G\$12 из ячейки G15 в G16:G24. В качестве примера вы можете вычислить прогнозируемые продажи представителя 4 в регионе 2 как  $17,6 - 2,85 + 3,6 = 18,35$ . Это значение вычислено в ячейке D38 (рис. 63.4) по формуле =G12+G16+G24. Если бы влияние региона было значимо, а влияние представителя незначимо, прогнозируемые продажи для представителя 4 в регионе 2 составили бы  $17,6 - 2,85 = 14,75$ .

	C	D
36	District 2	
37	Rep 4 Forecast	
38	Mean	18.35
39	Lower	10.3688
40	Upper	26.3312

Рис. 63.4. Прогноз объема продаж для представителя 4 в регионе 2

Как и в однофакторном дисперсионном анализе, стандартное отклонение ошибок прогноза равно корню квадратному из среднеквадратичной погрешности, показанной в ячейке E31. Стандартное отклонение вычислено в ячейке E32 по формуле =КОРЕНЬ(E31). Таким образом, можно быть на 95% уверенным в том, что если пред-

ставитель 4 будет назначен в регион 2, ежемесячные объемы продаж составят от  $18,35 - 2 \times 3,99 = 10,37$  до  $18,35 + 2 \times 3,99 = 26,33$ . Эти границы вычислены в ячейках D39 и D40 по формулам  $=D\$38-2*\$E\$32$  и  $=D\$38+2*\$E\$32$ , соответственно.

**?** Как определить, влияют ли на продажи видеоигр изменение цены и расходов на рекламу? Как определить, значимо ли взаимодействие цены и рекламы?

Если для каждой комбинации факторов в строках и столбцах имеется более одного наблюдения, воспользуйтесь двухфакторным дисперсионным анализом с повторениями. Для выполнения такого анализа Excel требует, чтобы число наблюдений для каждой комбинации факторов в строках и столбцах было одинаковым.

В дополнение к проверке значимости факторов в строках и столбцах можно также проверить значимость их взаимодействия. Например, если необходимо определить, как цена и расходы на рекламу влияют на объемы продаж, взаимосвязь между ценой и расходами на рекламу означает, что эффект от изменения стоимости рекламы будет зависеть от уровня цены (или, что то же самое, эффект от изменения цены будет зависеть от стоимости рекламы). Отсутствие взаимодействия между ценой и рекламой означает, что эффект от изменения цены не будет зависеть от стоимости рекламы.

В качестве примера для двухфакторного дисперсионного анализа с повторениями определим, как цена и стоимость рекламы влияют на ежемесячные продажи видеоигры. В файле *Twowayanova.xlsx* на листе *Two Way ANOVA No Interaction* находятся данные, представленные на рис. 63.5. Например, за три месяца, в течение которых расходы на рекламу были низкими, а цена — средней, было продано 21, 20 и 16 единиц продукции.

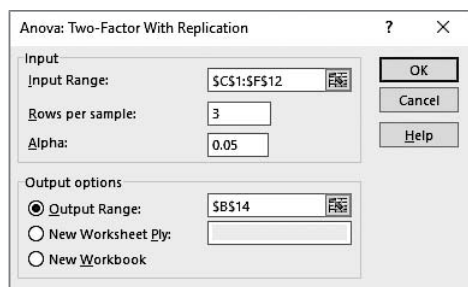
	B	C	D	E	F	G
1		Average	25.04			
2			Price			
3			Low	Medium	High	Effect
4		Low	41	21	10	-5.5926
5	Adv		25	20	11	
6			23	16	8	
7		Medium	28	28	11	-1.8148
8			30	22	22	
9			32	18	18	
10		High	35	26	21	7.4074
11			45	40	26	
12			47	32	20	
13		Effect	8.963	-0.2593	-8.7	

**Рис. 63.5.** Данные о продажах видеоигры; нет взаимодействия между факторами

Обратите внимание, что для каждой комбинации «цена — реклама» имеется ровно по три наблюдения. В ячейке D1 по формуле  $=CPЗНАЧ(D4:F12)$  я вычислил общее среднее (25,037) для всех наблюдений. В ячейках G4, G7 и G10 я вычислил эффекты для каждого уровня расходов. Например, эффект от низкого уровня расходов на

рекламу равен среднему значению для низкого уровня расходов на рекламу минус общее среднее. Эффект от низкого уровня расходов на рекламу (5,59) я вычислил в ячейке G4 по формуле  $=CP3HAЧ(D4:F6)-D\$1$ . Аналогичным образом, я вычислил эффекты для каждого уровня цены, скопировав формулу  $=CP3HAЧ(D4:D12)-D\$1$  из D13 в E13:F13.

Для анализа этих данных на вкладке **Данные (Data)** выберите **Анализ данных (Data Analysis)** и затем в диалоговом окне **Анализ данных (Data Analysis)** — инструмент **Двухфакторный дисперсионный анализ с повторениями (Anova: Two-Factor With Replication)**. Заполните диалоговое окно, как на рис. 63.6.



**Рис. 63.6.** Диалоговое окно Двухфакторный дисперсионный анализ с повторениями для запуска анализа с повторениями

Для анализа я указал следующую информацию.

- Входные данные, включая метки, находятся в диапазоне C3:F12. В Excel в двухфакторном дисперсионном анализе с повторениями для каждого уровня эффекта требуются метки в первой строке каждого столбца входного интервала. Таким образом, в ячейки D3:F3 я ввел Низкая, Средняя и Высокая (цена). В Excel также требуются метки для каждого уровня эффекта в первом столбце входного интервала. Эти метки должны находиться в строках, отмечающих начало данных для каждого уровня. Таким образом, метки, соответствующие низкому, среднему и высокому уровню расходов на рекламу, следует разместить в ячейках C4, C7 и C10.
- В поле Число строк для выборки (Rows Per Sample) я ввел 3, поскольку для каждой комбинации «цена — реклама» имеется три повторения.
- Для левой верхней ячейки выходного интервала я указал B14.

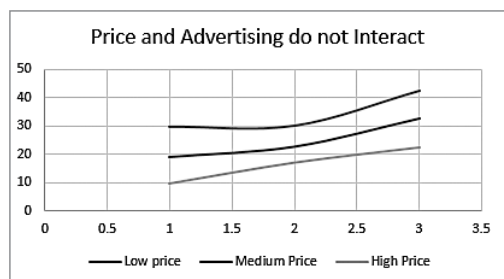
Затем я нажал ОК. Единственно важной частью выходных данных является таблица дисперсионного анализа, показанная на рис. 63.7.

Как и в случае с рандомизированными блоками, влияние (включая взаимодействие) значимо, если  $p$ -значение меньше 0,15. Здесь Выборка (строка с данными о воздействии рекламы, Sample) и Цена (показанная в строке с меткой Столбцы, Columns) являются высокосignификантными, а значимое взаимодействие отсутствует.

	B	C	D	E	F	G	H	I
42	ANOVA							
43	Source of Variation	SS	df	MS	F	P-value	F crit	F crit
44	Sample	804.96296	2	402.481	13.52	0.0003	3.5546	3.55456109
45	Columns	1405.4074	2	702.704	23.6	9E-06	3.5546	3.55456109
46	Interaction	50.592593	4	12.6481	0.425	0.7888	2.9277	2.92774871
47	Within	536	18	29.7778				
48								
49	Total	2796.963	26					

**Рис. 63.7.** Выходные данные двухфакторного дисперсионного анализа с повторениями; взаимодействия между факторами нет

( $p$ -значение для взаимодействия, Interaction, равно 0,79!). Таким образом, можно заключить, что цена и реклама влияют на объем продаж, и воздействие рекламы на объем продаж не зависит от уровня цены. На рис. 63.8 приведена диаграмма среднего объема продаж для каждой комбинации «цена — реклама». Как видите, цена и реклама не проявляют значимого взаимодействия. (См. лист Graph No Interaction.)



**Рис. 63.8.** Взаимодействие между ценой и рекламой в этом наборе данных отсутствует

Обратите внимание, что по мере повышения расходов на рекламу средние объемы продаж увеличиваются примерно с одинаковой скоростью, независимо от низкого, среднего или высокого уровня цены. В принципе, можно распознать отсутствие взаимодействия по тому факту, что линии для всех уровней цены почти параллельны.

### ❓ Как интерпретировать влияние цены и рекламы на продажи при отсутствии значимого взаимодействия между ценой и рекламой?

В отсутствие значимого взаимодействия можно прогнозировать объемы продаж в двухфакторном дисперсионном анализе с повторениями так же, как в двухфакторном дисперсионном анализе без повторений по следующему уравнению (уравнение 2):

прогнозируемый объем продаж = общее среднее + [строка или эффект рекламы (если значимо)] + [столбец или эффект цены (если значимо)].

В этом анализе предполагается, что цена и реклама являются единственными факторами, влияющими на объем продаж. Если продажи имеют ярко выраженный сезонный характер, в анализ необходимо включить сезонность. (Сезонность рассматривается в главах 65 «Метод Винтерса и лист прогноза» и 67 «Прогноз для особых случаев».) Например, если цена высокая, а уровень расходов на рекламу средний, прогнозируемый объем продаж составит  $25,037 + (-1,814) + (-8,704) = 14,52$ . (См. ячейку E54 на рис. 63.9, на котором снова представлен лист Two Way ANOVA No Interaction.) На рис. 63.5 было показано, что общее среднее равно 25,037, эффект от среднего уровня расходов на рекламу равен  $-1,814$  и эффект от высокой цены равен  $-8,704$ .

	B	C	D	E	F	G	H	I
42	ANOVA							
43	Source of Variation	SS	df	MS	F	P-value	F crit	F crit
44	Sample	804.96296	2	402.481	13.52	0.0003	3.5546	3.55456109
45	Columns	1405.4074	2	702.704	23.6	9E-06	3.5546	3.55456109
46	Interaction	50.592593	4	12.6481	0.425	0.7888	2.9277	2.92774871
47	Within	536	18	29.7778				
48								
49	Total	2796.963	26					
50								
51			Stdev	5.4569				
52			Medium Ad					
53			High Price					
54			Mean	14.5185				
55			Lower	3.60471				
56			Upper	25.4323				
57								

**Рис. 63.9.** Прогноз для объема продаж при высокой цене и среднем уровне рекламы

Стандартное отклонение ошибок прогноза равно корню квадратному из средне-квадратичной погрешности:

$$\sqrt{29,78} = 5,46.$$

С вероятностью 95% прогноз будет точен в пределах 10,92 единицы. Другими словами, можно быть на 95% уверенным, что объем продаж в течение месяца при высокой цене и среднем уровне расходов на рекламу составит от 3,60 до 25,43 единицы.

На листе Two Way ANOVA with Interaction я заменил данные из предыдущего примера данными, показанными на рис. 63.10. После выполнения двухфакторного дисперсионного анализа с повторениями я получил результаты, как на рис. 63.11.

В этом наборе данных  $p$ -значение для взаимодействия равно 0,001. Если  $p$ -значение для взаимодействия небольшое (меньше 0,15), можно даже не проверять  $p$ -значения для факторов в строке и столбце. Прогнозируемый объем продаж для любой комбинации цены и уровня расходов на рекламу равен среднему трех наблюдений, включающих эту комбинацию цены и уровня расходов на рекламу. Например, наилучший прогноз для объема продаж в течение месяца при высоком уровне расходов на рекламу и средней цене равен:

$$\frac{34 + 40 + 32}{4} = \frac{106}{3} = 35,33 \text{ единицы.}$$

	D	E	F	G
2			Price	
3				
4		Low	Medium	High
5	Low	41	21	15
6		25	20	14
7		23	16	13
8	Medium	28	28	14
9		30	22	13
10		32	18	12
11	High	50	34	13
12		51	40	13
13		52	32	13

**Рис. 63.10.** Данные продаж при взаимодействии цены и рекламы

	C	D	E	F	G	H	I
43	ANOVA						
44	Source of Variation	SS	df	MS	F	P-value	F crit
45	Sample	828.963	2	414.481	24.22	7.86E-06	3.55456
46	Columns	2498.74	2	1249.37	73.02	2.31E-09	3.55456
47	Interaction	509.926	4	127.481	7.45	0.001006	2.92775
48	Within	308	18	17.1111			
49							
50	Total	4145.63	26				
51		Std dev	4.137				

**Рис. 63.11.** Выходные данные для двухфакторного дисперсионного анализа с повторениями

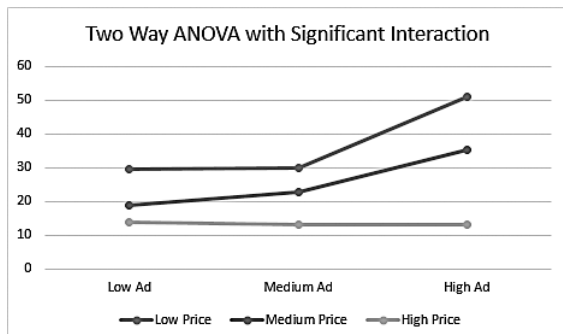
Стандартное отклонение ошибок прогноза также равно корню квадратному из среднеквадратичной погрешности:

$$\sqrt{17,11} = 4,14.$$

Таким образом, прогноз для объема продаж с вероятностью 95% окажется точным в пределах 8,26 единицы.

На рис. 63.12 показано, почему эти данные проявляют значимое взаимодействие цены и расходов на рекламу. Для низкой и средней цены повышение расходов на рекламу приводит к повышению среднего объема продаж, но если цена высокая, повышение расходов на рекламу не оказывает влияния на средний объем продаж. Это объясняет, почему нельзя использовать уравнение 2 для предсказания продаж при наличии значимого взаимодействия. В конце концов, как можно говорить о влиянии рекламы, когда влияние рекламы зависит от цены?

Главное, что можно сказать об этом графике, — серии данных не параллельны. Это говорит о том, что для различных уровней цен изменение расходов на рекламу имеет различное воздействие на объем продаж.



**Рис. 63.12.** В этом наборе данных цена и уровень рекламы проявляют значимое взаимодействие

## Задания

Данные к заданиям 1–3 находятся в файле Ch63.xlsx.

1. Считается, что давление (высокое, среднее или низкое) и температура (высокая, средняя или низкая) влияют на объем выработки. Приняв это за основу, решите следующие задачи.
  - Исходя из данных на листе **Problem 1**, определите, как температура и/или давление влияют на объем выработки.
  - В каком диапазоне с вероятностью 95% находится объем выработки при высоком давлении и низкой температуре?
2. Вам нужно определить, как конкретный торговый представитель и количество рекламных звонков врачу (один, три и пять) влияют на сумму (в тысячах долларов), на которую каждый врач выписывает ваш лекарственный препарат. Исходя из данных на листе **Problem 2**, найдите ответы на следующие вопросы:
  - Как представитель и количество звонков влияют на объем продаж?
  - Если представитель 3 сделал пять звонков врачу, то вы на 95% уверены, что врач выпишет рецепты вашего препарата всего на какую сумму?
3. Ответьте на вопросы к заданию 2, используя данные на листе **Problem 2**.
4. Файл **Coupondata.xlsx** содержит информацию о продажах арахисового масла в течение нескольких недель, когда покупателям выдавался (или не выдавался) купон и когда реклама была (или не была) напечатана в воскресной газете. Опишите, как купон и реклама влияют на продажи арахисового масла.

## ГЛАВА 64

# Скользящие средние для временных рядов

### Обсуждаемый вопрос

- Я пытаюсь проанализировать тенденцию роста ежеквартальных доходов Amazon.com с 1996 г. Объем продаж в четвертом квартале в США обычно больше (из-за Рождества), чем объем продаж в первом квартале следующего года. Эта закономерность затемняет тенденцию к росту продаж. Как отобразить тенденцию к росту доходов графически?

## Ответ на вопрос

**?** Я пытаюсь проанализировать тенденцию роста ежеквартальных доходов Amazon.com с 1996 г. Объем продаж в четвертом квартале в США обычно больше (из-за Рождества), чем объем продаж в первом квартале следующего года. Эта закономерность затемняет тенденцию к росту продаж. Как отобразить тенденцию к росту доходов графически?

Данные временных рядов просто отображают одну и ту же величину, измеренную в разные моменты времени. Например, данные в файле Amazon.com.xlsx, часть которых показана на рис. 64.1, отображают временной ряд для ежеквартальных доходов в миллионах долларов для Amazon.com. Данные охватывают временной интервал от четвертого квартала 1995 г. до третьего квартала 2017 г.

Для построения диаграммы этого временного ряда выделите диапазон E4:F93, который содержит номер квартала (от 1 до 89) и ежеквартальные доходы Amazon.com (в миллионах долларов). Затем на вкладке Вставка (Insert) в группе Диаграммы (Chart) выберите Точечная или пузырьковая диаграмма (Insert Scatter (X, Y) Or Bubble Chart) и затем второй тип точечной диаграммы — Точечная с гладкими кривыми и маркерами (Scatter With Smooth Lines And Markers). Диаграмма временного ряда показана на рис. 64.2.

Вы видите, что наблюдается тенденция к росту доходов, но тот факт, что из-за доходов в четвертом квартале доходы в первых трех кварталах каждого года кажутся небольшими, затрудняет ее отслеживание. Поскольку в году четыре квартала, было бы неплохо отобразить на диаграмме средние доходы за предыдущие четыре квартала. Это называется *скользящей средней для четырех периодов*. Использование



	C	D	E	F
4	FiscalYear	FiscalQuarter	Quarter Number	Revenue
5	1995	4	1	0.511
6	1996	1	2	0.875
7	1996	2	3	2.23
8	1996	3	4	4.173
9	1996	4	5	8.468
10	1997	1	6	16.005
11	1997	2	7	27.855
12	1997	3	8	37.887
13	1997	4	9	66.04
14	1998	1	10	87.361
15	1998	2	11	115.982
16	1998	3	12	153.649
17	1998	4	13	252.893
18	1999	1	14	293.643
19	1999	2	15	314.376
20	1999	3	16	355.778
21	1999	4	17	676.042
22	2000	1	18	573.889
23	2000	2	19	577.876
24	2000	3	20	637.858
25	2000	4	21	972.36

Рис. 64.1. Ежеквартальные доходы от продаж Amazon.com

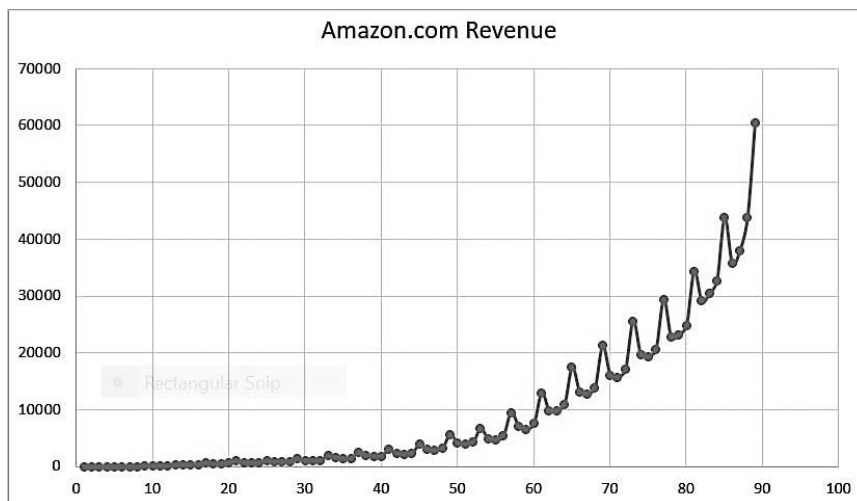
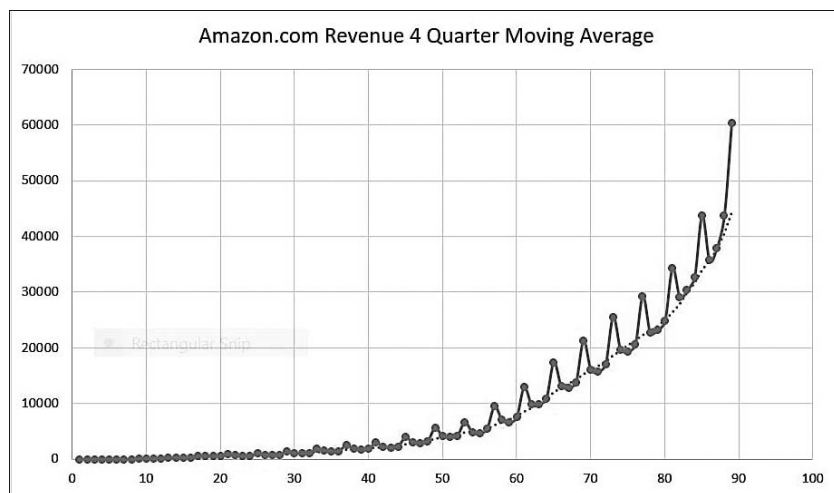


Рис. 64.2. Временной ряд ежеквартальных доходов Amazon.com

скользящей средней за четыре квартала сглаживает влияние сезонных колебаний, поскольку в каждое среднее будет входить по одной точке данных из каждого квартала. Такая диаграмма называется *графиком скользящей средней*, так как нанесенные на диаграмму средние значения «скользят» во времени. Графики скользящей средней также сглаживают случайные изменения, что позволяет получить более полное представление о том, что происходит с данными.

Для построения графика скользящей средней ежеквартальных доходов необходимо изменить диаграмму. Выделите диаграмму и щелкните на какой-либо точке данных, чтобы все точки стали синего цвета. Щелкните на точке правой кнопкой мыши и выберите **Добавить линию тренда (Add Trendline)**, затем установите переключатель в положение **Линейная фильтрация (Moving Average)** и выберите в списке **Период (Period)** значение 4. На диаграмму добавится кривая скользящей средней за четыре квартала (рис. 64.3, файл *Amazon.com.xlsx*).



**Рис. 64.3.** Кривая скользящей средней за четыре квартала

Для каждого квартала ставится точка со значением среднего для текущего и трех предыдущих кварталов. Очевидно, что для скользящей средней за четыре квартала соответствующая кривая начинается с четвертой точки данных. Кривая скользящей средней ясно показывает, что доходы Amazon.com имеют устойчивую тенденцию к повышению. Кроме того, наклон скользящей средней за четыре квартала увеличивается. По всей вероятности, наклон этого графика скользящей средней со временем выровняется, в результате чего график будет выглядеть как S-образная кривая. Инструмент Excel **Кривая тренда (Trend Curve)** не подбирает S-образные кривые, но для подбора к данным S-образных кривых в Microsoft Excel 2016 можно использовать инструмент **Поиск решения (Solver)**. Подбор S-образных кривых к данным см. в книге Уэйна Винстона «Анализ рынков» (Wayne L. Winston. *Marketing Analytics*, Wiley, 2014).

## Задание

Файл *Ch62data.xlsx* содержит квартальные доходы GM, Ford и GE. Постройте кривую скользящей средней за четыре квартала для доходов каждой компании. Что вы узнали по каждой кривой тренда?

# Метод Винтерса и Лист прогноза

Часто требуется предсказать будущие значения временного ряда, такие как ежемесячные расходы или ежемесячная выручка от продукта. Как правило, это не просто, поскольку характеристики любого временного ряда постоянно меняются. В большинстве случаев для прогнозирования будущих значений временного ряда лучше всего подходят методы сглаживания или адаптивные методы. В этой главе я опишу самый эффективный метод сглаживания — метод Винтерса (предложенный Петером Винтерсом (Peter Winters) и Чарльзом Хольтом (Charles Holt)). Я также покажу вам, как пользоваться чудесным инструментом Excel Лист прогноза (добавленным в Excel 2016).

Чтобы вы поняли, как работает метод Винтерса, я использую его для прогнозирования ежемесячного объема строительства нового жилья в США. Объем строительства нового жилья — это просто количество новых домов, которые начали строиться в течение месяца. Начнем с описания трех ключевых характеристик временных рядов.

## Характеристики временных рядов

Поведение большинства временных рядов можно объяснить тремя характеристиками: *база*, *тренд* и *сезонность*.

- *База* ряда описывает текущий уровень ряда в отсутствие какой-либо сезонности. Например, предположим, что базовый уровень объема строительства нового жилья в США равен 160 000 домов. В этом случае, если текущий месяц был средним месяцем относительно других месяцев года, можно считать, что объем нового строительства составит 160 000 домов.
- *Тренд* временного ряда — это процентное увеличение базы за период. Таким образом, тренд 1,02 означает, что, согласно оценке, объем нового строительства увеличивается на 2% каждый месяц. Тренд 0,97 означает, что, согласно оценке, объем нового строительства уменьшается на 3% каждый месяц.
- *Сезонность* (индекс сезонности) для периода говорит нам, насколько больше или меньше типичного ожидаемый ежемесячный объем нового строительства. Например, если индекс сезонности декабря составляет 0,8, то декабрьский объем нового строительства на 20% меньше типичного ежемесячного объема. Если индекс сезонности июня составляет 1,3, то июньский объем нового строительства на 30% больше типичного ежемесячного объема.

## Определение параметров

На основе всех данных, полученных в результате наблюдения в течение месяца  $t$ , вычисляются следующие представляющие интерес величины:

- $L_t$  — уровень ряда;
- $T_t$  — тренд ряда;
- $S_t$  — индекс сезонности для текущего месяца.

Ключом к методу Винтерса являются следующие три уравнения, с помощью которых обновляются величины  $L_t$ ,  $T_t$  и  $S_t$ . В следующих формулах  $alp$ ,  $bet$  и  $gam$  — это *сглаживающие постоянные*. Значения этих постоянных выбираются с целью оптимизации прогнозов. В следующих формулах с равно числу периодов в сезонном цикле (например,  $c = 12$  месяцев), а  $x_t$  равно наблюдаемому значению временного ряда в момент  $t$ .

- Формула 1:  $L_t = alp(x_t/s_{t-c}) + (1 - alp)(L_{t-1} \times T_{t-1})$ .
- Формула 2:  $T_t = bet(L_t/L_{t-1}) + (1 - bet)T_{t-1}$ .
- Формула 3:  $S_t = gam(x_t/L_t) + (1 - gam)S_{t-c}$ .

Формула 1 показывает, что новая оценка базы равна взвешенному среднему текущего наблюдения (с исключенной сезонной составляющей) и база предыдущего периода обновлена в соответствии с последней оценкой тренда. Формула 2 показывает, что новая оценка тренда равна взвешенному среднему отношения текущей базы к базе предыдущего периода (это текущая оценка тренда) плюс тренд предыдущего периода. Формула 3 показывает, что обновленная оценка индекса сезонности вычисляется как взвешенное среднее оценки индекса сезонности, основанного на текущем периоде, плюс предыдущая оценка. Обратите внимание, что бóльшие значения параметров сглаживания придают больший вес текущему наблюдению.

$F_{t,k}$  можно определить как прогноз  $F$  после периода  $t$  для периода  $t + k$ . Таким образом,  $F_{t,k} = L_t(T_t)^k S_{t+k-c}$  (формула 4).

По этой формуле сначала с помощью текущей оценки тренда обновляется база на  $k$  периодов вперед. Затем получившаяся в результате оценка базы для периода  $t + k$  корректируется соответствующим индексом сезонности.

## Определение начальных параметров для метода Винтерса

Для запуска метода Винтерса необходимы начальные оценки базы, тренда и индексов сезонности ряда. В качестве начальных данных для метода Винтерса я использовал ежемесячные объемы строительства нового жилья за 1986 и 1987 гг. Затем я выбрал параметры сглаживания для оптимизации прогнозов на месяц вперед для 1988–1996 гг. (рис. 65.1 и файл House2.xlsx).

	A	B	C	D	E	F	G	H	I	J
1	DATE	HS							1987 mean	150.45
2	Jan-86	105.4							1986 mean	145.1417
3	Feb-86	95.4							Trend	1.002998
4	Mar-86	145								
5	Apr-86	175.8								
6	May-86	170.2								
7	Jun-86	163.2								
8	Jul-86	160.7								
9	Aug-86	160.7								
10	Sep-86	147.7								
11	Oct-86	173				alp	bet	gam		
12	Nov-86	124.1				0.493579	0.014807	0.27241618		
13		120.5								
14	Jan-87	115.6								
15	Feb-87	107.2								
16	Mar-87	151								
17	Apr-87	188.2								
18	May-87	186.6								
19	Jun-87	183.6								
20	Jul-87	172								
21	Aug-87	163.8			MAPE	0.0731				
22	Sep-87	154								
23	Oct-87	154.8								
24	Nov-87	115.6	Base	Trend	Forecast	APE				
25	Dec-87	113	143.05	1.003						
26	Jan-88	105.1	142.04	1.003	107.2714	0.02066				
27	Feb-88	102.8	146.17	1.003	97.63509	0.05024				
28	Mar-88	141.2	143.86	1.003	146.8439	0.03997				
29	Apr-88	159.3	136.92	1.002	177.6762	0.11536				
30	May-88	158	134.1	1.002	165.6342	0.04832				

Рис. 65.1. Определение начальных параметров для метода Винтерса

Далее я выполнил следующие шаги.

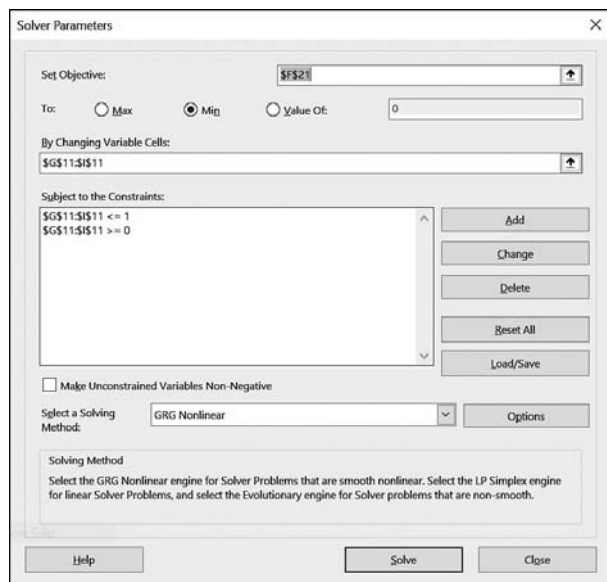
- **Шаг 1.** Я вычислил, например, январский индекс сезонности как среднее январских объемов строительства нового жилья для 1986 и 1987 г., деленное на средние ежемесячные объемы для 1986 и 1987 г. Таким образом, скопировав формулу =CPЗНАЧ(B2;B14)/CPЗНАЧ(\$B\$2:\$B\$25) из G14 в G15:G25, я получил расчетные индексы сезонности. Например, январский индекс равен 0,75, а июньский — 1,17.
- **Шаг 2.** Для вычисления среднего ежемесячного тренда я разделил средний объем для 1987 г. на средний объем для 1986 г. и возвел в одну двенадцатую. Я вычислил тренд в ячейке J3 (и скопировал в D25) по формуле =(J1/J2)^(1/12).
- **Шаг 3.** База ряда вычисляется как декабрьское значение для 1987 г. с исключенной сезонной составляющей. Я вычислил базу в ячейке C25 по формуле =(B25/G25).

## Вычисление сглаживающих постоянных

Теперь мы готовы вычислить сглаживающие постоянные. В столбце C обновляется база ряда, в столбце D — тренд ряда и в столбце G — индексы сезонности. В столбце E я вычислил прогноз для следующего месяца и в столбце F — абсолютную процентную ошибку для каждого месяца. Наконец, я использовал Поиск

решения (Solver) для выбора значений сглаживающих постоянных, минимизирующих сумму абсолютных процентных ошибок. Вот как я это сделал.

- **Шаг 1.** В ячейки G11:I11 я ввел пробные значения (от 0 до 1) для сглаживающих постоянных.
- **Шаг 2.** В ячейках C26:C119 я вычислил по формуле 1 обновленные уровни ряда, скопировав формулу  $=alp*(B26/G14)+(1-alp)*(C25*D25)$  из C26 в C27:C119.
- **Шаг 3.** В ячейках D26:D119 я обновил по формуле 2 тренд ряда, скопировав формулу  $=bet*(C26/C25)+(1-bet)*D25$  из D26 в D27:D119.
- **Шаг 4.** В ячейках G26:G119 я обновил по формуле 3 индексы сезонности, скопировав формулу  $=gam*(B26/C26)+(1-gam)*G14$  из G26 в G27:G119.
- **Шаг 5.** В ячейках E26:E119 я вычислил по формуле 4 прогноз для текущего месяца, скопировав формулу  $=(C25*D25)*G14$  из E26 в E27:E119.
- **Шаг 6.** В ячейках F26:F119 я вычислил абсолютную процентную ошибку для каждого месяца, скопировав формулу  $=ABS(B26-E26)/B26$  из F26 в F27:F119.
- **Шаг 7.** В ячейке F21 я вычислил среднюю абсолютную процентную ошибку для 1988–1996 гг. по формуле  $=CPЗНАЧ(F26:F119)$ .
- **Шаг 8.** Наконец, я щелкнул на Поиск решения (Solver) на вкладке Данные (Data), чтобы определить значения сглаживающих постоянных, минимизирующие среднюю абсолютную процентную ошибку. Диалоговое окно Параметры поиска решения (Solver Parameters) показано на рис. 65.2. Поиск решения (Solver) рассмотрен в главах 29–38.



**Рис. 65.2.** Диалоговое окно Параметры поиска решения для модели Винтерса

## ПРИМЕЧАНИЕ

Для активации надстройки Поиск решения (Solver) на вкладке Файл (File) выберите команду Параметры (Options), затем среди параметров Excel выберите раздел Надстройки (Add-Ins). В нижней части диалогового окна в раскрывающемся списке Управление (Manage) выберите Надстройки Excel (Excel Add-Ins) и нажмите Перейти (Go). В диалоговом окне Надстройки (Add-Ins) установите флажок Поиск решения (Solver Add-In) и нажмите ОК. На вкладке Данные (Data) в группе Анализ (Analysis) появится кнопка Поиск решения (Solver). Подробную информацию см. в главе 29 «Введение в оптимизацию с надстройкой Поиск решения».

Я использовал сглаживающие постоянные ( G11:I11) для минимизации средней абсолютной процентной ошибки (ячейка F21). Инструмент Поиск решения (Solver) обеспечивает поиск оптимальной комбинации сглаживающих постоянных. Сглаживающие постоянные должны находиться в пределах от 0 до 1. Здесь средняя абсолютная процентная ошибка минимизируется при  $\alpha = 0,50$ ,  $\beta = 0,01$  и  $\gamma = 0,27$ . Возможно, значения сглаживающих постоянных будут слегка различаться, но значение средней абсолютной процентной ошибки должно быть близко к 7,3%. В этом примере есть много комбинаций сглаживающих постоянных, которые дают прогнозы практически с одинаковой средней абсолютной процентной ошибкой. Мой прогноз на месяц вперед ошибочен в среднем на 7,3%.

## Замечания

- Вместо выбора сглаживающих постоянных для оптимизации ошибок прогноза вперед на один период можно было бы, например, выбрать оптимизацию средней абсолютной процентной ошибки прогнозирования общих объемов строительства для следующих шести месяцев.
- Если в конце месяца  $t$  требуется предсказать объемы продаж для следующих четырех кварталов, можно просто сложить  $F_{t,1} + F_{t,2} + F_{t,3} + F_{t,4}$ . При необходимости можно выбрать сглаживающие постоянные для минимизации абсолютной процентной ошибки прогнозирования объемов продаж для следующих нескольких лет.

## Инструмент Excel Лист прогноза

В инструменте Лист прогноза, впервые добавленном в Excel 2016, применяется обобщение метода Винтерса. Он позволяет просто выдавать прогнозы для данных типа временного ряда. Вы также получаете доверительные интервалы для каждого прогноза. Чтобы воспользоваться Листом прогноза, вы должны поместить данные в книгу с расширением \*.xlsx (НЕ \*.xls). В файле Airlinetemp.xlsx содержатся сведения о количестве тысяч пассажиро-миль налета на линиях США в 2001–2017 гг. . Воспользуемся данными за 2001–2016 гг. (рис. 65.3), чтобы спрогнозировать, каким будет месячный налет на пассажирских линиях в 2017 г. . Сначала поместите

курсор в любое место в таблице данных (столбцы A и B), с помощью которых будет делаться прогноз на 2017 г. Затем на вкладке **Данные (Data)** щелкните значок **Лист прогноза (Forecast Sheet)** в группе **Прогноз (Forecast)**. На диаграмме предпросмотра в диалоговом окне **Создание листа прогноза (Create Forecast Sheet)** (рис. 65.4) выбираем 1 декабря 2017 г. как **Конец периода прогноза (Forecast End)**. Эта диаграмма отражает имеющиеся данные, а красным показан период прогноза для каждого месяца. В данном случае мы видим текущий прогноз и верхний и нижний пределы 95-процентного доверительного интервала для каждого прогноза. Например, для января 2017 г. прогнозируется 70 249 426,22 мили, и мы уверены на 95%, что реальные пассажиро-мили в январе 2017 г. будут лежать в диапазоне между 66 497 179,856 и 74 001 672,59 мили. Чтобы создать **Лист прогноза**, показывающий прогнозы и доверительный интервал для требуемого периода, просто нажмите на кнопку **Создать (Create)** в диалоговом окне **Создать лист прогноза (Create Forecast Sheet)**. На отдельном листе появятся результаты, представленные на рис. 65.5.

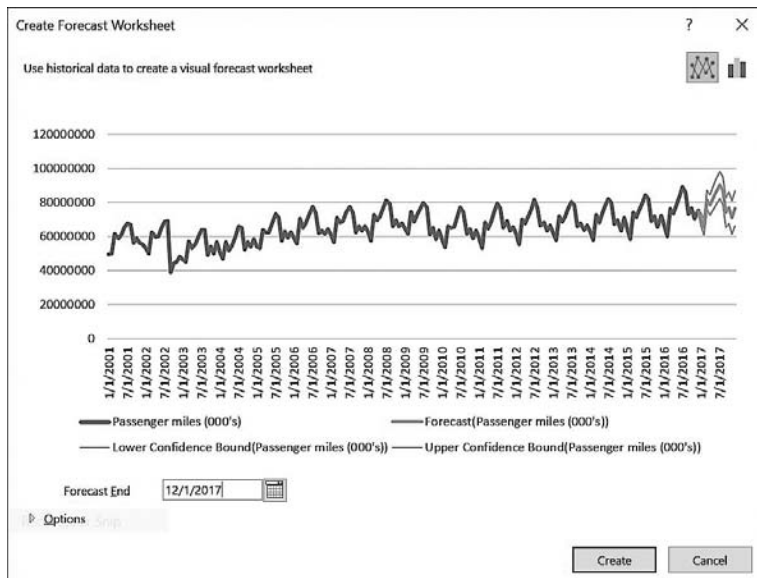
	A	B	C	D	E
4	Date	Passenger miles (000's)		Date	Passenger miles (000's)
5	1/1/2001	49843099		1/1/2017	69778658
6	2/1/2001	49931931		2/1/2017	65026219
7	3/1/2001	61478163		3/1/2017	79121758
8	4/1/2001	58981617		4/1/2017	75617434
9	5/1/2001	61223861		5/1/2017	81226986
10	6/1/2001	65601574		6/1/2017	87042101
11	7/1/2001	67898320		7/1/2017	91256833
12	8/1/2001	67028338		8/1/2017	86708521
13	9/1/2001	56441629		9/1/2017	75735172
14	10/1/2001	58834210		10/1/2017	77790770
15	11/1/2001	56283261		11/1/2017	72625059
16	12/1/2001	55380280		12/1/2017	77310983
17	1/1/2002	53129922			
18	2/1/2002	49992995			
19	3/1/2002	62323049			
20	4/1/2002	59801567			
21	5/1/2002	60246478			
22	6/1/2002	64987598			
23	7/1/2002	68573410			
24	8/1/2002	69003617			
25	9/1/2002	39106905			

**Рис. 65.3.** Данные о пассажиро-милях

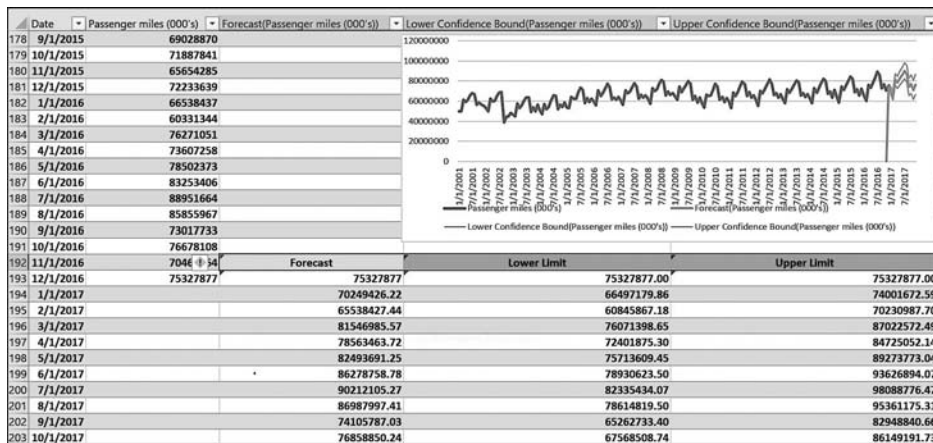
Для получения прогноза и 95%-ных доверительных интервалов на каждый месяц также возможно применить функции Excel (см. файл *Airlinemiles.xlsx* и рис. 65.6).

Скопировав формулу `=ПРЕДСКАЗ.ETS(A197;$B$5:$B$196;$A$5:$A$196;1;1)` из D197 в D198:D208, вы получите прогноз на каждый месяц 2017 г. на основе данных за 2016-й. Первая «1» позволит Excel автоматически определить интервал сезонности (12 месяцев в данном случае), а вторая «1» означает, что нужно применить линейную интерполяцию между точками имеющихся данных для замещения любых недостающих данных.





**Рис. 65.4.** В диалоговом окне Создать лист прогноза на диаграмме предпросмотра отражены как имеющиеся данные, так и прогноз, а также верхняя и нижняя границы 95%-ного доверительного интервала для данного прогноза



**Рис. 65.5.** Доверительные интервалы и прогнозы пассажиро-миль в 2017 г.

Скопировав формулу =D197-ПРЕДСКАЗ.ETS.ДОВИНТЕРВАЛ (A197;\$B\$5:\$B\$196;\$A\$5:\$A\$196;0.95;12) из E197 в E198:E208, получаем нижнюю границу 95%-ного доверительного интервала для прогноза каждого месяца. Аргументы функции ПРЕДСКАЗ.ETS.ДОВИНТЕРВАЛ (FORECAST.ETX.CONFINT) говорят программе о том, что нужно использовать 12-месячный интервал сезонности и данные 2016 г. для расчета половины ширины 95%-ного доверительного интервала для прогноза каждого месяца.

ца. В столбце F применена точно такая же формула, как и в столбце E, только знак «минус» заменен на «плюс», чтобы получить верхнюю границу 95%-ного доверительного интервала для пассажиро-миль каждого месяца. Как показано в столбце G на рис. 65.6, средняя абсолютная процентная ошибка для наших прогнозов на 2017 г. составляет всего лишь 1,5%.

	A	B	C	D	E	F	G
86	2/1/2016	60331344		Lower Limit January 2017 95% Confidence Interval			
87	3/1/2016	76271051		=D197-FORECAST.ETS.CONFINT(A197,\$B\$5:\$B\$196,\$A\$5:\$A\$196,0.95,1)			
88	4/1/2016	73607258					
89	5/1/2016	78502373					
90	6/1/2016	83253406		January 2017 Forecast			
91	7/1/2016	88951664		=FORECAST.ETS(A197,\$B\$5:\$B\$196,\$A\$5:\$A\$196,1,1)			
92	8/1/2016	85855967					
93	9/1/2016	73017733					
94	10/1/2016	76678108					MAPE
95	11/1/2016	70460064					0.015424
96	12/1/2016	75327877	Forecast	Forecast function	Lower Limit	Upper limit	APE
97	1/1/2017	69778658	70249426.22	70249426.22	66497179.86	74001672.59	0.006701
98	2/1/2017	65026219	65538427.44	65538427.44	60845867.18	70230987.70	0.007815
99	3/1/2017	79121758	81546985.57	81546985.57	76071398.65	87022572.49	0.02974
100	4/1/2017	75617434	78563463.72	78563463.72	72401875.30	84725052.14	0.037499
101	5/1/2017	81226986	82493691.25	82493691.25	75713609.45	89273773.04	0.015355
102	6/1/2017	87042101	86278758.78	86278758.78	78930623.50	93626894.07	0.008847
103	7/1/2017	91256833	90212105.27	90212105.27	82335434.07	98088776.47	0.011581
104	8/1/2017	86708521	86987997.41	86987997.41	78614819.50	95361175.31	0.003213
105	9/1/2017	75735172	74105787.03	74105787.03	65262733.40	82948840.66	0.021987
106	10/1/2017	77790770	76858850.24	76858850.24	67568508.74	86149191.73	0.012125
107	11/1/2017	72625059	71159787.61	71159787.61	61441625.27	80877949.95	0.020591
108	12/1/2017	77310983	76572990.12	76572990.12	66444005.56	86701974.67	0.009638

**Рис. 65.6.** Применение функций Excel для создания прогнозов и границ доверительного интервала в 2017 г.

## Задания

Данные для всех заданий находятся в файле Quarterly.xlsx.

1. С помощью метода Винтерса спрогнозируйте доходы Apple на один квартал вперед.
2. С помощью метода Винтерса спрогнозируйте доходы Amazon.com на один квартал вперед.
3. С помощью метода Винтерса спрогнозируйте доходы Home Depot на один квартал вперед.
4. С помощью метода Винтерса спрогнозируйте суммарный доход Home Depot для следующих двух кварталов.
5. Спрогнозируйте пассажиро-мили за 2016 г. на основе данных 2015 г. с помощью Листа прогноза. Рассчитайте среднюю абсолютную процентную ошибку для ваших прогнозов.

## ГЛАВА 66

# Метод прогнозирования «по отношению к скользящему среднему»

### Обсуждаемые вопросы

- Что такое тренд временного ряда?
- Как определить индексы сезонности для временного ряда?
- Существует ли простой способ включить тренд и сезонность в прогнозирование будущих продаж продукта?

Часто для прогнозирования будущих ежеквартальных доходов корпорации или ежемесячных объемов продаж продукта требуется простой точный метод. Для таких ситуаций подходит точный и простой в использовании метод прогнозирования «по отношению к скользящему среднему».

В файле *Ratioma.xlsx* вы найдете данные о продажах продукта за 20 кварталов (см. на рис. 66.1 строки 5–24). Требуется предсказать объемы продаж в следующих четырех кварталах (кварталы 21–24). Этот временной ряд характеризуется трендом и индексом сезонности.

## Ответы на вопросы

### ❓ Что такое тренд временного ряда?

Например, тренд 10 единиц продукции за квартал означает, что объем продаж увеличивается на 10 единиц продукции за квартал, а тренд –5 единиц продукции за квартал означает, что объем продаж имеет тенденцию к снижению на 5 единиц в квартал. Мультипликативный тренд 1,04 означает, что продажи возрастают на 4 % за квартал; мультипликативный тренд 0,94 означает, что продажи снижаются на 6% за квартал. В главе 65 «Метод Винтерса и Лист прогноза» мы как раз имели дело с мультипликативным трендом; в настоящей главе мы рассматриваем аддитивный тренд.

### ? Как определить индексы сезонности для временного ряда?

Известно, что компания Walmart ожидает большого роста продаж в четвертом квартале года (благодаря праздникам). Если это не учесть, трудно составить точный прогноз квартальных доходов компании Walmart. Более полно представить модель продаж компании позволяют индексы сезонности. Квартальные индексы сезонности для доходов Walmart:

- квартал 1 (январь — март): 0,90;
- квартал 2 (апрель — июнь): 0,98;
- квартал 3 (июль — сентябрь): 0,96;
- квартал 4 (октябрь — декабрь): 1,16.

Эти индексы означают, например, что объем продаж в четвертом квартале, как правило, на 16% выше, чем объем продаж для квартала в среднем. *Среднее для индексов сезонности должно быть равно 1.*

Проверьте свое понимание материала, ответив на следующий вопрос: если объем продаж компании Walmart в четвертом квартале 2013 г. составил 200 млрд долларов, а в первом квартале 2014 г. — 180 млрд долларов, идут ли дела в компании Walmart лучше или хуже? Ключом к ответу на вопрос является *устранение влияния сезонных изменений* на объемы продаж и выражение объема продаж для каждого квартала относительно «среднего» квартала. Например, объем продаж в четвертом квартале 2013 г. эквивалентен продажам на сумму  $200/1,16 = 172,4$  млрд долларов в «среднем» квартале, а объем продаж в первом квартале 2014 г. эквивалентен продажам на сумму  $180/0,9 = 200$  млрд долларов в «среднем» квартале. Таким образом, несмотря на то что фактический объем продаж в компании Walmart снизился на 10%, по всей видимости, объем продаж увеличивается на  $(200/172,4) - 1 = 16\%$  за квартал. Этот простой пример показывает, насколько важны индексы сезонности компании или продукта.

### ? Существует ли простой способ включить тренд и сезонность в прогнозирование будущих продаж продукта?

Теперь давайте прибегнем к простому методу прогнозирования «по отношению к скользящему среднему». Этот метод позволяет быстро вычислить тренд и индексы сезонности временного ряда и упрощает прогнозирование будущих значений временного ряда. Результаты моего ответа на вопрос — на рис. 66.1 и в файле Ratioma.xlsx.

Начнем с оценки уровня ряда с исключенной сезонной составляющей для каждого периода (с использованием центрированных скользящих средних). Затем к оценкам с исключенной сезонной составляющей (в столбцеG) подберем линию тренда. Далее определим индекс сезонности для каждого квартала. Наконец, вычислим будущий уровень ряда путем экстраполяции линии тренда и затем спрогнозируем объемы продаж путем включения сезонной составляющей в оценку линии тренда. Ниже описаны подробные шаги.

	B	C	D	E	F	G	H	I	J	K	L	M
1				slope	6.93878676	=SLOPE(G7:G22,B7:B22)						
2				intercept	30.1661765	=INTERCEPT(G7:G22,B7:B22)			quarter	seasonal index	normalized	
3										1	0.8185469	0.813736785
4	Quarter#	Year	Quarter	Sales	4 period MA	Centered MA	Actual/CMA	Forecast		2	0.9393395	0.933819599
5	1	1	1	24						3	1.0673637	1.06109143
6	2	1	2	44	52					4	1.1983944	1.191352187
7	3	1	3	61	58	55.00	1.11					
8	4	1	4	79	63.5	60.75	1.30					
9	5	2	1	48	71	67.25	0.71					
10	6	2	2	66	77.5	74.25	0.89					
11	7	2	3	91	82.5	80.00	1.14					
12	8	2	4	105	87.25	84.88	1.24					
13	9	3	1	68	89.5	88.38	0.77					
14	10	3	2	85	94.5	92.00	0.92					
15	11	3	3	100	104.25	99.38	1.01					
16	12	3	4	125	114.25	109.25	1.14					
17	13	4	1	107	123.75	119.00	0.90					
18	14	4	2	125	132.25	128.00	0.98					
19	15	4	3	138	139.25	135.75	1.02					
20	16	4	4	159	146.75	143.00	1.11					
21	17	5	1	135	156	151.38	0.89					
22	18	5	2	155	164.25	160.13	0.97					
23	19	5	3	175								
24	20	5	4	192								
25	21	6	1			175.8806985		143.1206				
26	22	6	2			182.8194853		170.7204				
27	23	6	3			189.7582721		201.3509				
28	24	6	4			196.6970588		234.3355				

Рис. 66.1. Данные для прогнозирования по «отношению к скользящему среднему»

- **Вычисление скользящих средних.** Сначала для каждого квартала вычислите скользящее среднее за четыре квартала (четыре квартала устраняют сезонность) путем усреднения значений предыдущего квартала, текущего квартала и двух следующих кварталов. Для этого скопируйте формулу =CP3НАЧ(E5:E8) из F6 в F7:F22. Например, для второго квартала скользящее среднее равно  $0,25 \times (24 + 44 + 61 + 79) = 52$ .
- **Вычисление центрированных скользящих средних.** Скользящее среднее для второго квартала центрируется кварталом 2,5, а скользящее среднее для третьего квартала — кварталом 3,5. Усреднение этих двух скользящих средних дает центрированное скользящее среднее, оценивающее уровень процесса в конце третьего квартала. Скопируйте формулу =CP3НАЧ(F6:F7) из G7 в G8:G22 для вычисления уровня ряда в каждом периоде — без сезонности!
- **Подбор линии тренда к центрированным скользящим средним.** С помощью центрированных скользящих средних подберите линию тренда для оценки будущего уровня ряда.

В ячейке F1 я использовал формулу =НАКЛОН(G7:G22;B7:B22) для нахождения наклона линии тренда, а в ячейке F2 по формуле =ОТРЕЗОК(G7:G22;B7:B22) я вычислил начальную ординату линии тренда. Теперь можно рассчитать уровень ряда в квартале  $t$  как  $6,94t + 30,17$ . Скопируйте формулу =intercept+slope\*B23 из G25 в G26:G28 для вычисления уровня ряда в 21-м квартале и далее.

- **Вычисление индексов сезонности.** Напомню, что индекс сезонности для квартала, например 2,0, означает, что объем продаж в этом квартале в два раза превышает объем продаж для «среднего» квартала, а индекс сезонности 0,5 для квартала означает, что объем продаж в этом квартале составляет половину от объема продаж «среднего» квартала. Для определения индексов сезонности сначала вычислим отношение фактических продаж к центрированному скользящему среднему для каждого квартала. Для этого скопируйте формулу  $=E7/G7$  из ячейки H7 в H8:H22. Как видите, для каждого первого квартала продажи составили 71, 77, 90 и 89% от среднего, поэтому индекс сезонности для первого квартала можно вычислить как среднее для этих четырех величин (82%). Для вычисления начальной оценки индексов сезонности скопируйте формулу  $=СЗНАЧЕСЛИ(\$D\$7:\$D\$22;J3;\$H\$7:\$H\$22)$  из ячейки K3 в K4:K6. Эта формула усредняет четыре оценки, имеющиеся для индекса сезонности первого квартала.

К сожалению, полученные индексы сезонности не равны в среднем 1. Для обеспечения этого условия скопируйте формулу  $=K3/СРЗНАЧ(\$K\$3:\$K\$6)$  из L3 в L4:L6.

- **Прогнозирование продаж для кварталов 21–24.** Для создания прогноза продаж в каждом будущем квартале просто умножьте вычисленный уровень ряда для каждого квартала (из столбца G) на соответствующий индекс сезонности. Вычислите окончательный прогноз для кварталов 21–24, скопировав формулу  $=ВПР(D25;season;3)*G25$  из ячейки I25 в I26:I28.

Если тренд ряда недавно изменился, можно оценить его на основе самых последних данных. Например, можно получить новую оценку тренда с помощью центрированных скользящих средних для кварталов 13–18 по формуле  $=НАКЛОН(G17:G22;B17:B22)$ . В результате предполагаемый тренд составит 8,09 единицы за квартал. Если требуется предсказать продажи, например, в квартале 22, то для оценки уровня ряда в квартале 22 необходимо добавить  $4 \times (8,09)$  к последнему имеющемуся центрированному скользящему среднему (160,13) для квартала 18. Затем результат следует умножить на индекс сезонности второго квартала (0,933) и получить окончательный прогноз для объема продаж в этом квартале:  $(160,13 + 4 \times 8,09) \times 0,933 = 179,6$  единицы.

## Задание

В файле **Walmartdata.xlsx** содержатся данные о квартальных доходах Walmart за 1994–2009 гг. Используйте метод «по отношению к скользящему среднему» для прогнозирования доходов за третий и четвертый кварталы 2009 г. и за первый и второй кварталы 2010 г. На основе данных за 53–60-й кварталы создайте оценку тренда, которую будете использовать в прогнозах.

## ГЛАВА 67

# Прогноз для особых случаев

### Обсуждаемые вопросы

- Как определить, влияют ли конкретные факторы на поток клиентов?
- Как оценить точность прогноза?
- Как проверить, являются ли ошибки прогноза случайными?

Для студенческого проекта в начале 1990-х гг. (то есть до эры прямых депозитов) я вместе со студентами пытался спрогнозировать количество клиентов, ежедневно посещающих филиал Eastland Plaza Branch кредитного союза Университета Индианы. Благодаря встречам с директором филиала были выявлены следующие факторы, влияющие на количество клиентов:

- месяц года;
- день недели;
- день выдачи зарплаты профессорско-преподавательскому составу или служебному персоналу;
- день до или после праздничного дня.

## Ответы на вопросы

### ❓ Как определить, влияют ли конкретные факторы на поток клиентов?

Собранные данные находятся на листе **Original** в файле **Creditunion.xlsx** (рис. 67.1). Если на основе этих данных выполнить регрессионный анализ с фиктивными переменными (см. главу 60 «Включение качественных факторов во множественную регрессию»), то зависимой переменной будет количество обслуживаемых каждый день клиентов (данные в столбце **E**). Кроме того, потребуется 19 независимых переменных:

- одиннадцать для вычисления месяца (12 месяцев минус 1);
- четыре для вычисления дня недели (5 рабочих дней минус 1);
- две для вычисления типа дня выплаты ежемесячной зарплаты;
- две для вычисления дня, следующего за праздничным днем, и дня, предшествующего праздничному дню.



Но в Microsoft Excel допускается только 15 независимых переменных, поэтому у нас проблема.

	B	C	D	E	F	G	H	I	J
1								RSQ	0.77118606
2									
3	MONTH	DAYMON	DAYWEEK	CUST	SPECIAL	SP	FAC	BH	AH
4	1	2	2	1825	SP,FAC,AH	1	1	0	1
5	1	3	3	1257	0	0	0	0	0
6	1	4	4	969	0	0	0	0	0
7	1	5	5	1672	SP	1	0	0	0
8	1	8	1	1098	0	0	0	0	0
9	1	9	2	691	0	0	0	0	0
10	1	10	3	672	0	0	0	0	0
11	1	11	4	754	0	0	0	0	0
12	1	12	5	972	0	0	0	0	0
13	1	15	1	816	0	0	0	0	0
14	1	16	2	717	0	0	0	0	0
15	1	17	3	728	0	0	0	0	0
16	1	18	4	711	0	0	0	0	0
17	1	19	5	1545	SP	1	0	0	0
18	1	22	1	873	0	0	0	0	0
19	1	23	2	713	0	0	0	0	0
20	1	24	3	626	0	0	0	0	0
21	1	25	4	653	0	0	0	0	0
22	1	26	5	1080	0	0	0	0	0

**Рис. 67.1.** Данные для прогноза потока клиентов в кредитном союзе

Если регрессионная модель прогнозирования требует более 15 независимых переменных, то вычислить коэффициенты независимых переменных можно с помощью инструмента Excel Поиск решения (Solver). Кроме того, в Excel можно вычислить значения  $R^2$  между прогнозируемым и фактическим трафиками клиентов и значение стандартного отклонения для ошибок прогнозирования. Для анализа этих данных (рис. 67.2) я создал уравнение прогноза, используя таблицу поиска дня недели, месяца и других факторов. Затем я применил инструмент Поиск решения (Solver) для выбора коэффициентов каждого уровня каждого фактора, дающих минимальную сумму квадратичных ошибок. (Ошибка для каждого дня равна разности между фактическим и прогнозируемым количеством клиентов.) Рассмотрим пример подробно.

Начнем с создания индикаторных переменных (в столбцах G–J) для дня выплаты зарплаты персоналу (SP), дня выплаты зарплаты преподавателям (F AC), дня перед праздником (BH) и дня после праздника (AH) — см. рис. 67.1. Например, значение 1, введенное в ячейки G4, H4 и J4, указывает, что 2 января было днем выплаты зарплаты персоналу, днем выплаты зарплаты преподавателям и днем после праздника. В ячейке I4 стоит 0, показывающий, что 2 января — не день перед праздником.

Прогноз определяется постоянной (которая позволяет центрировать прогнозы для повышения точности) и влиянием каждого дня недели, каждого месяца, дня



выплаты зарплаты персоналу, дня выплаты зарплаты преподавателям, дня перед праздником и дня после праздника. Я ввел пробные значения для всех этих параметров (изменяемые ячейки инструмента Поиск решения) в диапазон ячеек O4:O26, показанный на рис. 67.2. Затем Поиск решения выберет значения, при которых модель наилучшим образом соответствует данным. Для каждого дня прогноз по количеству клиентов будет дан на основании следующего уравнения:

прогнозируемое количество клиентов = постоянная + (влияние месяца) + (влияние дня недели) + (влияние дня выплаты з/п персоналу, если есть) + (влияние дня выплаты з/п преподавателям, если есть) + (влияние дня перед праздником, если есть) + (влияние дня после праздника, если есть).

С помощью этой модели вы можете вычислить прогноз ежедневного количества клиентов, скопировав из K4 в K5:K257 следующую формулу:

= \$O\$26 + ВПР(B4;\$N\$14:\$O\$25;2) + ВПР(D4;\$N\$4:\$O\$8;2) + G4\*\$O\$9 + H4\*\$O\$10 + I4\*\$O\$11 + J4\*\$O\$12.

В ячейке O26 вычисляется постоянный член. Функция ВПР(B4;\$N\$14:\$O\$25;2) возвращает коэффициент влияния для текущего месяца, а функция ВПР(D4;\$N\$4:\$O\$8;2) — коэффициент влияния дня недели для текущей недели. Часть формулы G4\*\$O\$9 + H4\*\$O\$10 + I4\*\$O\$11 + J4\*\$O\$12 вычисляет влияние, если текущий день отмечен как SP, FAC, BH или AH.

Скопировав формулу =(E4-K4)^2 из L4 в L5:L257, я вычислил квадратичную ошибку для каждого дня. Затем в ячейке L2 я вычислил сумму квадратичных ошибок по формуле =СУММ(L4:L257).

	G	H	I	J	K	L	M	N	O	P	Q	R
1			RSQ	0.77118606		stdeverr	163.17722					
2					SSE	6736582.1						
3	SP	FAC	BH	AH	Forecast	Sq Err	Error	Day of Week			average	
4	1	1	0	1	1766.78	3389.5598	58.219926	1	103.3575		dayweek	0
5	0	0	0	0	709.6035	299642.97	547.39654	2	-139.192		month	-3.9E-09
6	0	0	0	0	745.698	49863.782	223.302	3	-150.342			
7	1	0	0	0	1557.221	13174.184	114.77885	4	-114.247			
8	0	0	0	0	963.303	18143.294	134.69705	5	300.4243			
9	0	0	0	0	720.7533	885.25683	-29.75327	SP	396.8513			
10	0	0	0	0	709.6035	1414.0205	-37.60346	FAC	394.8944			
11	0	0	0	0	745.698	68.923168	8.3019979	BH	205.2928			
12	0	0	0	0	1160.37	35483.189	-188.3698	AH	254.2811			
13	0	0	0	0	963.303	21698.16	-147.303	Month				
14	0	0	0	0	720.7533	14.087005	-3.753266	1	-110.69			
15	0	0	0	0	709.6035	338.43256	18.396537	2	-75.7154			
16	0	0	0	0	745.698	1203.9514	-34.698	3	-40.3409			
17	1	0	0	0	1557.221	149.35654	-12.22115	4	0.02839			
18	0	0	0	0	963.303	8154.6235	-90.30295	5	87.8157			
19	0	0	0	0	720.7533	60.113132	-7.753266	6	133.341			
20	0	0	0	0	709.6035	6989.5391	-83.60346	7	115.8034			
21	0	0	0	0	745.698	8592.9196	-92.698	8	28.77429			
22	0	0	0	0	1160.37	6459.3078	-80.36982	9	-87.5632			
23	0	0	0	0	963.303	98158.741	-313.303	10	-53.0016			
24	0	0	0	0	720.7533	5891.0638	-76.75327	11	-42.7611			
25	0	0	0	0	709.6035	8722.913	93.396537	12	44.30913	Snip		
26	0	1	0	0	1175.567	11328.018	106.43316	constant	970.6353			

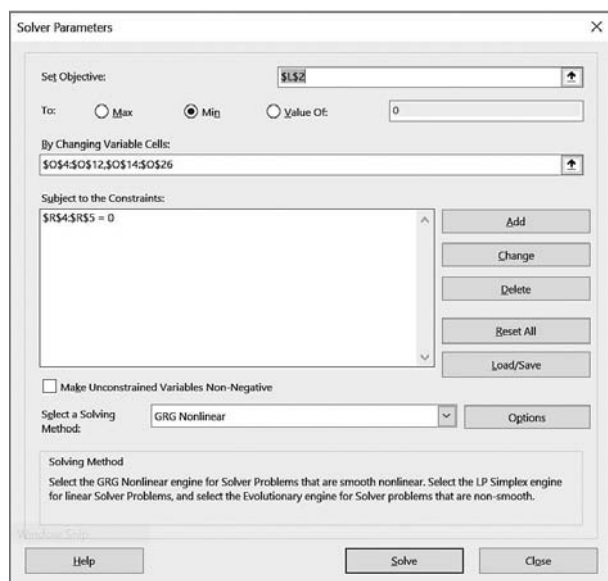
Рис. 67.2. Изменяемые ячейки и прогнозы по клиентам

В ячейке R4 я рассчитал среднее значение для изменяемых ячеек дня недели по формуле `=СРЗНАЧ(О4:О8)`, а в ячейке R5 — среднее значение для изменяемых ячеек месяца по формуле `=СРЗНАЧ(О14:О25)`. Позже я приму среднее влияние дня месяца и дня недели равным нулю; это гарантирует, что для месяца или дня недели с положительным влиянием количество клиентов будет больше, чем среднее количество клиентов, а для месяца или дня недели с отрицательным влиянием количество клиентов будет меньше среднего.

Вы можете использовать настройки инструмента Поиск решения (Solver), указанные на рис. 67.3, для выбора параметров прогнозирования, минимизирующих сумму квадратов ошибок. Для этого на вкладке Данные (Data) в группе Анализ (Analyze) щелкните Поиск решения (Solver).

## ПРИМЕЧАНИЕ

Для установки надстройки Поиск решения (Solver) на вкладке Файл (File) выберите команду Параметры (Options), затем среди параметров Excel выберите раздел Надстройки (Add-Ins). В нижней части диалогового окна в раскрывающемся списке Управление (Manage) выберите Надстройки Excel (Excel Add-Ins) и нажмите Перейти (Go). В диалоговом окне Надстройки (Add-Ins) установите флажок Поиск решения (Solver Add-In) и нажмите ОК. На вкладке Данные (Data) в группе Анализ (Analysis) появится кнопка Поиск решения (Solver). Подробную информацию см. в главе 29 «Введение в оптимизацию с надстройкой Поиск решения».



**Рис. 67.3.** Диалоговое окно Параметры поиска решения для определения параметров прогнозирования

В модели поиска решения коэффициенты для месяца, дня недели, ВН, АН, SP, FAS и постоянная изменяются в целях минимизации суммы квадратов ошибок. Можно также наложить ограничение (равенство нулю) на средние значения влияния дня недели и дня месяца. После этого Поиск решения (Solver) выдаст результаты, показанные на рис. 67.2. Например, самым загруженным днем недели является пятница, а самым загруженным месяцем — июнь. День выплаты зарплаты персоналу увеличивает прогнозируемое значение (при прочих равных условиях, на латыни — *ceteris paribus*) на 397 клиентов.

### ❓ Как оценить точность прогноза?

Чтобы оценить точность прогноза, вычислите в ячейке J1 значение  $R^2$  между прогнозируемым и фактическим числом клиентов по формуле =КВПИРСОН(E4:E257;K4:K257). Функция КВПИРСОН (RSQ) вычисляет процент фактического изменения количества клиентов, который объясняется моделью прогнозирования. Здесь независимые переменные объясняют 77% ежедневной вариации количества клиентов.

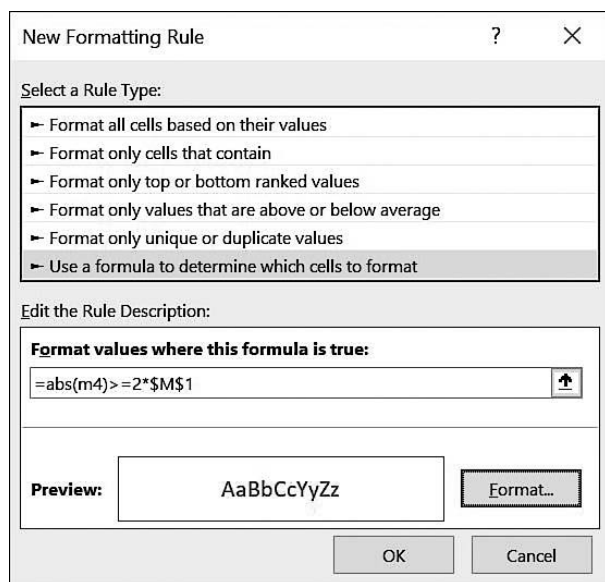
Вычислите ошибку для каждого дня в столбце М, скопировав формулу =E4-K4 из M4 в M5:M257. Достаточно хорошее приближение к стандартной погрешности прогноза дает стандартное отклонение ошибок. Это значение вычисляется в ячейке M1 по формуле =СТАНДОТКЛОН(M4:M257). Таким образом, приблизительно 68% прогнозов верны с точностью до 163 клиентов, 95% прогнозов — с точностью до 326 клиентов и т. д.

Попробуем найти выбросы. Напомню, что наблюдение является выбросом, если абсолютное значение ошибки прогноза превышает в два раза стандартную ошибку регрессии. Выделите диапазон M4:M257 (выделите ячейку M4 и нажмите Ctrl+Shift+J) и затем на вкладке Главная (Home) в списке Условное форматирование (Conditional Formatting) выберите Создать правило (New Rule). В диалоговом окне Создание правила форматирования (New Formatting Rule) щелкните на типе Использовать формулу для определения форматируемых ячеек (Use a formula to determine which cells to format). Заполните описание правила, как показано на рис. 67.4. (Подробнее об условном форматировании см. в главе 24 «Условное форматирование».)

После выбора формата с красным шрифтом любая ошибка, абсолютное значение которой в два раза превышает стандартное отклонение ошибок, будет отображена красным цветом. Как видите, модель часто занижает число клиентов для первых трех дней месяца. Также на второй неделе марта (весенние каникулы) модель завышает прогнозы, а за день до весенних каникул прогноз значительно занижен.

Чтобы это устранить, я добавил на листе 1st Three Days (файл Creditunion.xlsx) изменяемые ячейки для каждого из первых трех дней месяца, для весенних каникул и для дня перед весенними каникулами. В ячейки O26:O30 для этих новых эффектов я добавил пробные значения. Я включил эффект от первых трех дней месяца, скопировав из ячейки K4 в K5:K257 следующую формулу:

=O\$25+ВПР(В4;\$N\$13:O\$24;2)+ВПР(D4;\$N\$4:O\$8;2)+G4\*O\$9+H4\*O\$10+I4\*O\$11+J4\*O\$12+ ЕСЛИ(C4=1;O\$26;ЕСЛИ(C4=2;O\$27;ЕСЛИ(C4=3;O\$28;0))).



**Рис. 67.4.** Условное форматирование для обнаружения выбросов в прогнозе

Эффект от первых трех дней месяца вычисляет часть ЕСЛИ(C4=1;O\$26; ЕСЛИ(C4=2; O\$27;ЕСЛИ(C4=3;O\$28;0))). В ячейки K52:K57 я вручную ввел коэффициенты для весенних каникул. Например, в ячейке K52 я добавил к формуле +O29, а в ячейки K53:K57 я добавил +O30.

Включив в диалоговом окне **Параметры поиска решения (Solver Parameters)** новые изменяемые ячейки, я получил результаты, показанные на рис. 67.5. Обратите внимание, что для первых трех дней месяца количество клиентов возросло (вероятно, благодаря поддержке правительства и чекам от социальной защиты), а в весенние каникулы количество клиентов снизилось. Кроме того, как видно из рис. 67.5, точность прогнозирования также улучшилась. Значение  $R^2$  (рассчитанное при помощи функции **КВПИРСОН (RSQ)** в ячейке J1) повысилось до 87%, а стандартная ошибка уменьшилась до 122 клиентов.

Ошибки прогноза для недели с 24 по 31 декабря показывают (рис. 67.6), что модель значительно завышает прогноз количества клиентов для дней на этой неделе. Она также занижает прогнозируемое число клиентов для недели перед Рождеством. Дальнейшее изучение ошибок прогноза (часто называемых *остатками*) привело к следующим выводам.

- День благодарения отличается от обычных праздников тем, что в день после него кредитный союз загружен работой намного меньше, чем ожидалось.

	I	J	K	L	M	N	O	P	Q	R
1	RSQ	0.8715		stdeverr	122.28505					
2			SSE	3783269.1						
3	BH	AH	Forecast	Sq Err	Error	Day of Week			average	
4	0	1	1879.631	2984.5423	-54.63096		1	107.7063	dayweek	0
5	0	0	995.4	68434.536	261.59995		2	-138.9312	month	2.49E-14
6	0	0	722.9341	60548.436	246.06592		3	-153.3158		
7	0	0	1554.449	13818.155	117.55065		4	-115.0831		
8	0	0	945.7235	23188.13	152.27649		5	299.6238		
9	0	0	699.0859	65.38256	-8.085948	SP		416.8083		
10	0	0	684.7014	161.32604	-12.70142	FAC		96.64418		
11	0	0	722.9341	965.09124	31.065918	BH		196.4566		
12	0	0	1137.641	27436.946	-165.641	AH		299.1159		
13	0	0	945.7235	16828.189	-129.7235		1	-105.511		
14	0	0	699.0859	320.91325	17.914052		2	-81.76339		
15	0	0	684.7014	1874.7671	43.298581		3	-27.85624		
16	0	0	722.9341	142.42232	-11.93408		4	-7.28922		
17	0	0	1554.449	89.290215	-9.44935		5	83.84527		
18	0	0	945.7235	5288.7086	-72.72351		6	130.6715		
19	0	0	699.0859	193.60084	13.914052		7	106.6161		
20	0	0	684.7014	3445.8566	-58.70142		8	13.26009		
21	0	0	722.9341	4890.7759	-69.93408		9	-64.68717		
22	0	0	1137.641	3322.4867	57.64102		10	-68.30456		
23	0	0	945.7235	87452.393	-295.7235		11	-33.75333		
24	0	0	699.0859	3034.4617	-55.08595		12	54.7719		

**Рис. 67.5.** Параметры прогнозирования и прогнозы, включающие весенние каникулы и первые три дня месяца

- Накануне Страстной пятницы сотрудники действительно сильно загружены, так как многие уезжают из города на Пасху.
- День уплаты налогов (16 апреля) также является более загруженным днем, чем ожидалось.
- Неделя перед началом осеннего семестра в Университете Индианы (последняя неделя августа) не загружена, вероятно, потому, что многие сотрудники и преподаватели берут летом отпуск.

На листе **Christmas week** я добавил изменяемые ячейки для учета этих факторов. После добавления новых параметров как изменяемых ячеек я снова запустил

	A	B	C	D	K	L	M
1						stdeverr	122.28505
2					SSE	3783269.1	
3		MONTH	DAYMON	DAYWEEK	Forecast	Sq Err	Error
253		12	24	1	1302.463	130655.51	-361.463
254		12	26	3	1144.1	21054.062	-145.1002
255		12	27	4	883.2169	69810.592	-264.2169
256		12	28	5	1297.924	130266.04	-360.9239
257		12	31	1	1302.463	24480.674	-156.463

**Рис. 67.6.** Ошибки для рождественской недели

Поиск решения. Результаты представлены на рис. 67.7. Значение  $R^2$  повысилось до 92%, а стандартная ошибка снизилась до 98,61 клиента! Обратите внимание, что из-за влияния недели после Рождества ежедневное число клиентов снизилось на 359, в день перед Днем благодарения добавилось 607 клиентов, в день после Дня благодарения количество клиентов уменьшилось на 161 человека и т. д.

	I	J	K	L	M	N	O	P	Q	R
1	RSQ	0.9164307		stdeverr	98.614712		cutoff	110.5626225		
2			SSE	2460389.928			actual	125		
3	BH	AH	Forecast	Sq Err	Error	Day of Week		sign changes	average	
4	0	1	1981.09	24363.98781	-156.08968		1	108.1166	dayweek	-9.09E-14
5	0	0	976.0819	78914.96548	280.91808		2	-154.5804	1 month	1E-06
6	0	0	717.6763	63163.59728	251.32369		3	-164.7756	0	
7	0	0	1539.418	17577.98801	132.582		4	-121.0806	0	
8	0	0	946.8735	22839.22108	151.12651		5	332.32	0	
9	0	0	684.1765	46.55952932	6.8234544	SP		368.3411	0	
10	0	0	673.9813	3.925657514	-1.9813272	FAC		97.12028	1	
11	0	0	717.6763	1319.410473	36.32369	BH		272.6396	1	
12	0	0	1171.077	39631.62519	-199.07693	AH		477.8522	1	
13	0	0	946.8735	17127.8712	-130.87349		1	-111.147	0	
14	0	0	684.1765	1077.379156	32.823454		2	-82.12595	1	
15	0	0	673.9813	2918.01701	54.018673		3	-26.37362	0	
16	0	0	717.6763	44.57311188	-6.6763098		4	-34.81387	1	
17	0	0	1539.418	31.1587783	5.5820049		5	71.04197	1	
18	0	0	946.8735	5457.292987	-73.873493		6	127.4026	1	
19	0	0	684.1765	830.7915209	28.823454		7	93.98675	1	
20	0	0	673.9813	2302.207761	-47.981327		8	60.87196	1	
21	0	0	717.6763	4183.025043	-64.67631		9	-75.34943	0	
22	0	0	1171.077	8295.007695	-91.076933		10	-67.91572	0	
23	0	0	946.8735	88133.87092	-296.87349		11	-35.90971	0	
24	0	0	684.1765	1614.15482	-40.176546		12	80.33205	0	
25	0	0	673.9813	16645.81793	129.01867	constant		949.9039	1	
26	0	0	1387.864	11207.14342	-105.8638	d1		544.0462	1	
27	0	0	1922.039	14631.64768	120.96135	d2		353.5996	1	

**Рис. 67.7.** Окончательные параметры прогноза

Обратите также внимание, насколько усовершенствовалась модель прогнозирования при использовании выбросов. Если выбросы имеют нечто общее (как, например, первые три дня месяца), включите в качестве независимой переменной обобщающий фактор, и ошибка прогноза снизится.

### ❓ Как проверить, являются ли ошибки прогноза случайными?

Хороший метод прогнозирования должен создавать случайные ошибки прогноза, или остатки. Под *случайными ошибками* я понимаю ошибки, которые не носят закономерного характера. Если ошибки прогноза случайны, знак ошибок должен меняться (с плюса на минус или с минуса на плюс) примерно в половине случаев. Таким образом, широко применяемым тестом для оценки случайности ошибок прогноза является проверка перемены знака ошибок. Для  $n$  наблюдений неслучайность ошибок имеет место, если перемен знака меньше, чем

$$\frac{n-1}{2} | -\sqrt{n} |,$$

или больше, чем

$$\left| \frac{n-1}{2} + \sqrt{n} \right|.$$

На листе **Christmas week** (см. рис. 67.7) я определил число перемен знака в остатках, скопировав формулу `=ЕСЛИ(М5*М4<0;1;0)` из ячейки **P5** в **P6:P257**. Перемена знака в остатках происходит тогда и только тогда, когда произведение двух следующих друг за другом остатков отрицательно. Следовательно, эта формула дает в результате 1 при каждой перемене знака остатков. Всего было 125 перемен знака. В ячейке **P1** я вычислил значение перемен знака как отсечку для неслучайных остатков:

$$\frac{254-1}{2} - \sqrt{254} = 110,6.$$

Таким образом, здесь остатки носят случайный характер.

Такой же анализ я провел для прогнозирования ежедневного числа клиентов, ужинающих в крупной сети ресторанов. Конкретными факторами были праздники. Исследование показало, что «Супервоскресенье» (день Суперкубка НФЛ) было наименее загруженным днем, а самыми загруженными днями был День святого Валентина и День матери. Также самым загруженным днем недели в вечернее время стала суббота, а в дневное время — пятница.

## Задания

1. Как с помощью метода, описанного в данной главе, предсказать ежедневные объемы продаж авторучек на сайте **Staples.com**?
2. Как при наличии данных за несколько лет включить в анализ тренд?



## ГЛАВА 68

# Введение в теорию вероятности

### Обсуждаемые вопросы

- Что такое эксперимент, выборочное пространство и событие?
- Каковы аксиомы, которым должны удовлетворять вероятности наступления событий?
- Что такое правило дополнений?
- Что такое взаимоисключающие события?
- Что такое правило сложения вероятностей?
- Что такое независимые события?
- Что такое условная вероятность?
- Что такое формула полной вероятности?
- Что такое теорема Байеса?

Единственная вещь, которая нам известна со всей определенностью, — что мы живем в очень неопределенном мире. Чтобы разумно противостоять шквалу статистических и вероятностных данных, обрушивающихся на нас каждый день, необходимо понимание основ теории вероятностей.

## Ответы на вопросы

### ❓ Что такое эксперимент, выборочное пространство и событие?

*Эксперимент* — это любая процедура, которая может быть повторена много раз и имеет четко определенное множество исходов. Множество всех возможных исходов эксперимента называется *выборочным пространством*. Вот несколько примеров экспериментов и выборочных пространств.

Вы бросаете два правильных игральных кубика. На каждом кубике будет равномерно выброшено 1, 2, 3, 4, 5 или 6 очков. Выборочное пространство состоит из следующих 36 точек. Первое число в каждой точке — это число очков, выпавшее на первом кубике, а второе число — число очков на втором кубике:



(1,1) (1,2) (1,3) (1,4) (1,5) (1,6) (2,1) (2,2) (2,3) (2,4) (2,5) (2,6) (3,1) (3,2) (3,3) (3,4) (3,5) (3,6) (4,1) (4,2) (4,3) (4,4) (4,5) (4,6) (5,1) (5,2) (5,3) (5,4) (5,5) (5,6) (6,1) (6,2) (6,3) (6,4) (6,5) (6,6)

Вы подбрасываете две монеты. Выборочное пространство состоит из четырех равновероятных точек:

(O, O) (O, P) (P, O) (P, P)

*Событием* называется любое подмножество точек выборочного пространства. Вот два примера событий:

- Суммарное количество очков на двух кубиках равно восьми.
- При подбрасывании двух монет выпал один орел.

### ❓ Каковы аксиомы, которым должны удовлетворять вероятности наступления событий?

Вероятности наступления событий (вероятность обозначается  $P$ ) должны удовлетворять следующим аксиомам:

- **Аксиома 1:** для любого события  $E$   $0 \leq P(E) \leq 1$ .
- **Аксиома 2:** если событие  $E$  включает в себя все точки выборочного пространства, то  $P(E) = 1$ .

В случае подбрасывания двух кубиков каждая из точек выборочного пространства имеет одинаковую вероятность  $x$ . Поэтому по аксиоме 2 получаем:  $36 \cdot x = 1$  или  $x = 1/36$ .

Опираясь на этот факт, вы можете определить вероятность выбрасывания 8 очков в сумме на двух кубиках. Вот следующие пять точек в выборочном пространстве, которые дают в сумме 8: (2, 6), (3,5), (4,4), (5,3) и (6,2). Следовательно, вероятность выбрасывания 8 очков в сумме на двух кубиках равна  $5/36$ . Это означает, что если вы бросаете два кубика много раз, в среднем можно ожидать, что в  $5/36$  бросков суммарный результат будет равен 8.

### ❓ Что такое правило дополнений?

Для любого события  $A$  событие *не A* (противоположное, или дополнительное событие, обозначается  $\neg A$ ) состоит из всех точек пространства событий, которые не входят в  $A$ . Правило дополнений утверждает следующее:

$$P(\text{не } A) = 1 - P(A)$$

В качестве примера рассмотрим броски двух кубиков. Какова вероятность, что в сумме не выпадет 2? Пусть  $A$  = вероятность выпадения 2 в сумме. Так как единственная точка пространства событий, дающая в сумме 2, это (1,1), то  $P(A) = 1/36$ . Событие  $\text{не } A$  = Все\_суммы\_не\_равны\_2. Из правила дополнений получаем:

$$P(\text{Все\_суммы\_не\_равны\_2}) = 1 - P(\text{Все\_суммы\_равны\_2}) = 1 - 1/36 = 35/36.$$

Важность этого правила состоит в том, что если вы вычислите более «простой» элемент из  $P(A)$  и  $P(\text{He } A)$ , то вам известна и вторая вероятность.

### ❓ Что такое взаимоисключающие события?

Если два события не могут произойти одновременно, они называются *взаимоисключающими* (или *несовместными*). Рассмотрим следующие два примера:

- Если событие  $A$  = Сумма\_на\_двух\_кубиках\_равна\_4 и событие  $B$  = Сумма\_на\_двух\_кубиках\_равна\_8, то  $A$  и  $B$  — взаимоисключающие события.
- Если событие  $A$  = Индекс\_Доу — Джонса\_вырастет\_не\_менее\_чем\_на\_10%\_в\_2020\_г. и событие  $B$  = Индекс\_Доу — Джонса\_вырастет\_не\_менее\_чем\_на\_15%\_в\_2020\_г., то события  $A$  и  $B$  не являются взаимоисключающими.

### ❓ Что такое правило сложения вероятностей?

В общем случае для двух событий  $A$  и  $B$  правило сложения вероятностей состоит в следующем:

$$P(A \text{ ИЛИ } B) = P(A) + P(B) - P(A \text{ И } B).$$

То есть требуется вычесть  $P(A \text{ И } B)$  или в противном случае вероятность для точек выборочного пространства, общих для событий  $A$  и  $B$ , будет учтена дважды. В частном случае, когда  $A$  и  $B$  являются взаимоисключающими событиями,  $P(A \text{ И } B) = 0$  и правило сложения вероятностей приобретает вид:

$$P(A \text{ ИЛИ } B) = P(A) + P(B).$$

Два примера правила сложения вероятностей в действии:

- Если существует 50%-ная вероятность дождя в субботу и 50%-ная вероятность дождя в воскресенье, следует ли из этого, что есть 100%-ная вероятность дождя в выходные?

Пусть событие  $A$  = Дождь\_в\_субботу, а  $B$  = Дождь\_в\_воскресенье, тогда мы видим следующее:

$$P(\text{Дождь\_в\_выходные}) = P(\text{Дождь\_в\_субботу}) + P(\text{Дождь\_в\_воскресенье}) - P(\text{Дождь\_и\_в\_субботу\_и\_в\_воскресенье}) = 0,5 + 0,5 - P(\text{Дождь\_и\_в\_субботу\_и\_в\_воскресенье}) = 1 - P(\text{Дождь\_и\_в\_субботу\_и\_в\_воскресенье}).$$

Так как  $P(\text{Дождь\_и\_в\_субботу\_и\_в\_воскресенье}) > 0$ , мы получаем, что  $P(\text{Дождь\_в\_выходные}) < 1$ , то есть необязательно в выходные будет дождь (кроме как в Сиэтле).

В качестве второго примера на правило сложения рассмотрим броски двух кубиков. Какова вероятность, что хотя бы на одном из них выпадет 4? Если событие  $A$  = на\_первом\_кубике\_4, а  $B$  = на\_втором\_кубике\_4, мы получаем  $P(A) = 1/6$ ,  $P(B) = 1/6$  и  $P(A \text{ И } B) = 1/36$ . Тогда:

$$P(\text{на\_первом\_или\_на\_втором\_кубике\_4}) = (1/6) + (1/6) - (1/36) = 11/36.$$

### ❓ Что такое независимые события?

Два события,  $A$  и  $B$ , являются *независимыми*, если наступление одного из них не изменяет вероятность наступления другого. Следовательно, два события  $A$  и  $B$  являются независимыми тогда и только тогда, когда  $P(A \text{ И } B) = P(A) * P(B)$ . Для случая более двух событий строгое определение независимости событий выходит за рамки нашего рассмотрения, но если  $n$  событий  $A_1, A_2, \dots, A_n$  являются независимыми, верно следующее:

$$P(A_1 \text{ И } A_2 \text{ И } \dots \text{ И } A_n) = P(A_1) * P(A_2) * \dots * P(A_{n-1}) * P(A_n).$$

Следующие два «интуитивных» примера должны прояснить определение независимых событий:

- Пусть событие  $A$  = Индекс\_Доу – Джонса\_вырастет\_в\_2021 г., а событие  $B$  = Houston\_Texans\_выиграют\_Суперкубок\_в\_2021 г. Наступление одного из этих событий не изменяет вероятность наступления другого, поэтому эти события являются независимыми.
- Пусть событие  $A$  = Chicago\_Cubs\_выиграют\_Мировую\_серию\_2021, а  $B$  = Houston\_Astros\_выиграют\_Мировую\_серию\_2021. Ясно, что если наступит событие  $A$ , то вероятность события  $B$  падает до 0, поэтому события  $A$  и  $B$  не являются независимыми. Также ясно, что события  $A$  и  $B$  не могут быть и взаимоисключающими, и независимыми (см. задание 6).

А вот примеры для закрепления понимания концепции независимых событий.

- Предположим, вы подбрасываете монету и бросаете кубик. Какова вероятность того, что монета упадет вверх орлом и что на кубике выпадет 6 очков? Если определить событие  $A$  = монета упадет вверх орлом, а  $B$  = на кубике выпадет 6, то очевидно, что события  $A$  и  $B$  независимы. Так как  $P(A) = 1/2$  и  $P(B) = 1/6$ , то  $P(\text{монета упадет вверх орлом И на кубике выпадет 6}) = (1/2) * (1/6) = 1/12$ .
- Предположим, из колоды случайным образом выбрана карта. Вы выбрали одну карту. Являются ли независимыми события, что вытащена карта масти пик  $I$  что вытащен туз? Пусть событие  $A$  = масть карты пика, а  $B$  = вытащен туз. Так как в колоде из 52 карт 13 пик и 4 туза,  $P(A) = 13/52$  и  $P(B) = 4/52$ . Также одна карта в колоде и туз, и пика (туз пик), поэтому  $P(\text{пика И туз}) = 1/52$ . Поскольку  $P(\text{пика}) * P(\text{туз}) = (13/52) * (4/52) = 1/52$ , мы получаем, что  $A$  и  $B$  независимые события.
- Теперь предположим, что перед вытаскиванием карты мы убрали из колоды двойку пик. Остаются ли после этого независимыми события  $A$  и  $B$ ? Теперь  $P(A) = 12/51$ ,  $P(B) = 4/51$ , а  $P(A \text{ И } B) = 1/51$ . Так как  $(12/51) * (4/51)$  не равно  $(1/51)$ , события  $A$  и  $B$  не являются независимыми.
- Наконец, предположим, вы бросаете три кубика. Какова вероятность, что выпадет хотя одна шестерка? Пусть событие  $A_i$  = на кубике  $i$  не выпало 6. Тогда  $P(A_i) = 1 - (1/6) = 5/6$ . Тогда по правилу дополнений мы получаем:  
 $P(\text{по крайней мере 1 шестерка}) = 1 - P(0 \text{ шестерок}).$

Так как последовательные броски кубика являются независимыми событиями,  $P(0 \text{ шестерок}) = P(A1) * P(A2) * P(A3) = (5/6)^3 = 125/216$ . Следовательно, мы получим:

$$P(\text{по крайней мере 1 шестерка}) = 1 - (125/216) = 91/216 = 0,42.$$

### ❓ Что такое условная вероятность?

Часто требуется узнать, как наступление одного события изменяет вероятность наступления другого события. Формально мы записываем *условную вероятность* наступления события  $B$ , если известно, что событие  $A$  уже произошло, как  $P(B|A)$ . Эта запись читается как «вероятность наступления события  $B$  при условии, что событие  $A$  уже произошло».  $P(B|A)$  может быть вычислено из следующего уравнения:

$$P(B|A) = \frac{P(A \text{ И } B)}{P(A)}.$$

Это уравнение удобно переписать в виде  $P(A \text{ И } B) = P(A) * P(B|A)$  или  $P(A \text{ И } B) = P(B) * P(A|B)$ .

События  $A$  и  $B$  являются независимыми тогда и только тогда, когда  $P(B|A) = P(B)$ .

Следующие два примера раскрывают смысл условной вероятности:

- Предположим, вы бросаете два кубика. Определим событие  $A$  = выпала хотя бы 1 шестерка, а  $B$  = сумма очков на двух кубиках равна 10. Какова  $P(B|A)$ ?  $P(A \text{ и } B) = 2/36$  и  $P(A) = 1 - P(\text{нет шестерок}) = 1 - (5/6)^2 = 11/36$ . Следовательно, получим:

$$P(B|A) = (2/36)/(11/36) = 2/11.$$

- Предположим, вы вытянули туза из колоды карт. Если теперь вы вытягиваете вторую карту, какова вероятность, что это тоже будет туз? Определим событие  $A$  = первая карта туз, а  $B$  = вторая карта туз. После изъятия из колоды первой карты (туза) колода содержит 51 карту и 3 туза, поэтому  $P(B|A) = 3/51 = 1/17$ .

### ❓ Что такое формула полной вероятности?

*Формула полной вероятности* связана с вычислением вероятности события путем сложения вероятностей нескольких взаимоисключающих событий. Иногда требуется вычислить вероятность события, находящегося в зависимости от других событий. Например, предположим, что 5% подержанных машин пострадали во время наводнения и 80% этих машин впоследствии имеют проблемы с двигателем. Также предположим, что 10% машин, которые не пострадали во время наводнения, впоследствии имеют проблемы с двигателем. Какова вероятность того, что произвольно выбранный автомобиль будет иметь проблемы с двигателем?

Решая эту задачу, определим событие ПД = машина впоследствии имеет проблемы с двигателем, НВ = машина пострадала от наводнения. Теперь событие ПД

можно разложить на два взаимоисключающих события: «машина, пострадавшая от наводнения, имеет проблемы с двигателем» и «машина, не пострадавшая от наводнения, имеет проблемы с двигателем» и отразить это в следующей формуле:

$$P(\text{ПД}) = P(\text{ПД И НВ}) + P(\text{ПД И не НВ}) = P(\text{ПД} | \text{НВ}) * P(\text{НВ}) + P(\text{ПД} | \text{не НВ}) * P(\text{не НВ}) = (0,80) * (0,05) + (0,10) * (0,95) = 0,135.$$

Также вы можете создать *таблицу сопряженности признаков*, в которой отражаются все возможности. Ниже представлен вариант таблицы сопряженности:

	Проблемы с двигателем	Нет проблем с двигателем
НВ	0,04	0,01
не НВ	0,095	0,855

Числа в столбец **Нет проблем с двигателем** добавлены для демонстрации, какие слагаемые в первой строке дадут 0,05 и какие во второй дадут 0,95.

### ? Что такое теорема Байеса?

После того как вы разобрались с условной вероятностью и формулой полной вероятности, несложно понять суть *теоремы Байеса*. Во многих ситуациях мы пытаемся оценить вероятность различных состояний окружающего мира. Со временем мы получаем информацию, которая меняет наши вероятностные оценки. Например, представим себе 40-летнюю женщину без факторов риска заболевания раком молочной железы. Состояние окружающего мира можно определить событиями  $C$  = женщина имеет заболевание раком и  $NC$  = женщина не имеет заболевания раком. Без какой-либо дополнительной информации вероятности этих событий (известные как *априорные* вероятности) равны  $P(C) = 0,004$  и  $P(NC) = 0,996$ .

Теперь мы получаем дополнительную информацию (результаты маммографии), которая меняет наши оценки априорных вероятностей. Предположим, маммография выдает положительный (+) результат. Для изменения наших оценок вероятностей нам необходимо знать вероятность положительного результата теста для каждого из состояний мира. Вероятности положительного результата теста составляют  $P(+|C) = 0,80$  и  $P(+|NC) = 0,10$ . Теперь нам требуется изменить априорную вероятность рака (0,004) после получения положительного результата теста. Эта новая вероятность ( $P(C|+)$ ) называется *апостериорной* вероятностью. Применяя определение условной вероятности и правило полной вероятности, получаем:

$$P(C|+) = \frac{P(Cn+)}{P(+)} = \frac{P(+|C)P(C)}{P(+|C) + P(C) + P(+|NC) + P(NC)} = \frac{0,80 + 0,004}{0,80 + 0,004 + 0,996(0,10)} = 0,031.$$

Возможно, это удивительно, но даже после положительного результата теста существует только маленькая вероятность (слава богу!), что эта женщина имеет заболевание раком. Из-за того что большинство женщин не больны раком молочной

железы, многие из их маммографий дадут ложные положительные результаты. Другой способ убедиться в этом — взглянуть на типичную выборку из 10 000 женщин. Таблица сопряженности признаков показывает, что для 10 000 женщин заболевание и его отсутствие разделяется следующим образом:

	D	E	F
9	Cancer	$10,000 * (.004) * (.8) = 32$	$10,000 * (.004) * (1-.8) = 8$
10	No Cancer	$10,000 * (.996) * (.1) = 996$	$10,000 * (.996) * (1-.1) = 8964$

**Рис. 68.1.** Интуитивное объяснение примера с маммографией

В случае положительного результата теста мы получаем в столбце F 1028 женщин. Следовательно, при положительном результате теста для женщины вероятность заболевания раком равна  $32/1028 = 0,031$ .

В качестве последнего примера теоремы Байеса рассмотрим решение парадокса Монти Холла, представленного Мэрилин Вос Савант (Marilyn Vos Savant) в ее колонке «Спросите Мэрилин» в журнале *Parade*.

За одной из трех дверей находится машина, а за двумя другими — коза. Я выбираю дверь (например, дверь 1). После этого Монти Холл (ведущий передачи «Давайте заключим сделку») открывает дверь (дверь 2 или дверь 3), за которой находится коза. После этого вам предлагают поменять выбор двери. Сделаете ли вы это?

Предположим, Монти открыл дверь 2. Вот относящиеся к этой проблеме события:

- Состояния мира — это события  $D1, D2, D3$ , отражающие, что машина находится за дверью  $i$ .
- Определим, что  $S1, S2, S3$  — события; Монти говорит, что машина находится за дверью  $i$ .
- Нам известны априорные вероятности  $P(D1) = P(D2) = P(D3) = 1/3$ .
- Имеются следующие вероятности:  $P(S3|D1) = 1/2, P(S2|D1) = 1/2, P(S3|D2) = 1$  и  $P(S2|D3) = 1$ .

Мы знаем, что машина находится или за дверью 3, или за дверью 1. Применяя теорему Байеса, мы получаем следующее:

$$P(D3|S2) = \frac{P(D3 \text{ И } S2)}{P(D3 \text{ И } S2) + P(D1 \text{ И } S2)} = \frac{\left(\frac{1}{3}\right)^+1}{\left(\frac{1}{2}\right)^+1 + \left(\frac{1}{2}\right)^+\frac{1}{2}} = \frac{2}{3}.$$

Следовательно,  $P(D1|S2) = 1 - (2/3) = 1/3$ . Значит, мы должны изменить наш выбор на дверь 3!

**ПРИМЕЧАНИЕ**

Ответы на задания этой главы находятся в файле Word, а не Excel. См. файл Chapter68\_answers.docx в папке с ответами к главе 68.

**Задания**

1. Найдите вероятность каждой возможной суммы очков при броске двух кубиков.
2. При бросании трех кубиков определите вероятность выпадения в сумме 9 очков.
3. В Carver High School 20% студентов играют в бейсбол и 15% — в баскетбол. Пять процентов играет и в то и в другое. Какая часть студентов играет в бейсбол или в баскетбол?
4. Пусть событие  $A$  = на кубике выпало четное число очков, а  $B$  = на кубике выпало одно или два очка. Являются ли события  $A$  и  $B$  независимыми?
5. Являются ли события  $A$  и  $\text{не } A$  взаимоисключающими?
6. Могут ли взаимоисключающие события быть независимыми?
7. При игре в крэпс вы бросаете два кубика. Сумма очков 7 или 11 при первом броске приносит победу. Какова вероятность победы при первом броске?
8. Вы вытаскиваете из колоды две карты (без возвращения). Какова вероятность, что это будет две трефы?
9. На колесо рулетки нанесены числа 0, 00, 1, 2, ..., 36. Предположим, что при каждом из 25 пусков рулетки вы ставили на 00. Какова вероятность, что вы выиграли хотя бы раз из этих 25 пусков?
10. Предположим, что 10% взрослых людей смотрят шоу «Холостяк» (*The Bachelor*). Предположим, что 80% аудитории этого сериала составляют женщины и что мужчины составляют половину всех взрослых. Вычислите вероятность, что данная женщина или мужчина является зрителем этого сериала.
11. Какая доля мужчин и женщин смотрит шоу *The Bachelor*?
12. В примере с машинами, пострадавшими от наводнения, вам говорится, что машина имеет проблемы с двигателем. Какова вероятность, что машина до этого пострадала от наводнения?
13. В комодке два выдвижных ящика. В каждом ящике лежат две монеты; известно, что в одном ящике находятся две золотые монеты, а в другом — одна золотая и одна серебряная. Вы случайным образом выбираете ящик и вытаскиваете монету. Выбранная монета золотая. Какова вероятность, что оставшаяся в ящике монета тоже золотая?

14. 85% всех такси синего цвета, а остальные зеленого. Было определено, что такси, которое попало в аварию и водитель после этого скрылся, было зеленым. Люди определяют верно цвет такси в 80% случаев. При опросе большая часть студентов Стэнфордского университета утверждала, что вероятность того, что такси действительно было зеленым, равна 80%. Вы согласны с этим?
15. В урне находится девять шаров. На каждом из шаров написан номер от 1 до 9. Вытащите два шара (с возвращением). Какова вероятность, что вы вытащите шар с одним и тем же номером?
16. Бросьте кубик и пусть событие  $A$  = выпало нечетное число очков,  $aB$  = выпало число  $\geq 4$ . Найдите  $P(A|B)$ .
17. Каждый из двух механизмов работает с вероятностью 0,9. Работа или сбой механизмов происходят независимо. Если механизмы подключены последовательно, для работы системы требуется, чтобы работали оба механизма. Если механизмы подключены параллельно, для работы системы требуется, чтобы работал хотя бы один механизм. Какова вероятность, что система работает, при последовательном и при параллельном подключении?
18. Следующая таблица представляет статистику по приему мужчин и женщин в большой американский университет. Общее число мест для приема приведено в скобках, число принятых — вне скобок. Можно видеть, что процент принятых мужчин выше, чем процент принятых женщин. На основании применения статистики обоснуйте, есть ли в университете дискриминация женщин.

	Мужчины	Женщины
Простые специальности	864 (1386)	106 (133)
Сложные специальности	334 (1306)	451 (1702)

19. Компания имеет производство в Хьюстоне и в Далласе. Семьдесят процентов сотрудников работает в Хьюстоне, а 30% — в Далласе. Каждый год 3% хьюстонских и 5% далласских сотрудников получают травмы. Если случайным образом выбрать сотрудника, получившего в прошлом году травму, какова вероятность, что он работает в Хьюстоне?
20. Студентка МБА изучает финансы и маркетинг. Предположим, эта студентка имеет 90%-ную вероятность получения оценки  $A$  по финансам и 80%-ную вероятность получения оценки  $A$  по маркетингу. Если ее успешность в каждом из курсов не зависит от успешности в другом, какова вероятность, что она получит хотя бы одну оценку  $A$ ?
21. В вазе находится четыре красных и три синих шара. Из вазы вынимаются два шара (без возвращения). Известно, что второй шар синий. Какова вероятность, что первый шар синий?
22. Предположим, вы бросаете два кубика. Пусть событие  $A$  = на первом кубике выпало 3, а  $B$  = сумма очков на двух кубиках равна 8. Являются ли события  $A$  и  $B$  независимыми?



23. Среди страхователей страховой компании 80% являются высокорисковыми и 20% — низкорисковыми. Предположим, никто не получает более одной травмы в год. Также предположим, что 10% высокорисковых и 3% низкорисковых людей получают травму в течение года. Если вы случайным образом выбрали страхователя, получившего травму в прошлом году, какова вероятность того, что он относится к высокорисковым страхователям?
24. В ежегодном турнире университетских команд первого дивизиона по баскетболу есть четыре команды, которые могут с вероятностью 3% победить команду с первым номером посева. До начала чемпионата определите вероятность того, что хотя бы одна команда с 16-м номером посева обыграет команду с 1-м номером посева.
25. Предположим, один из 1000 людей является лжецом (неспособен говорить правду). Также предположим, что точность детектора лжи составляет 98%. Это означает, что если человек лжет, есть 98%-ная вероятность, что детектор укажет, что этот человек лжет. Кроме того, если человек не лжет, есть 98%-ная вероятность, что детектор укажет, что этот человек не лжет. Если детектор лжи показывает, что человек лжет, какова вероятность того, что человек действительно лжет?
26. После броска двух кубиков определим событие  $A$  = сумма очков является четным числом, а  $B$  = на первом кубике выпало пять очков. Являются ли события  $A$  и  $B$  независимыми?
27. В 40% недель супермаркет снижает цену на макароны и сыр. В 20% недель супермаркет выставляет макароны и сыр на рекламных стойках, и в 15% недель супермаркет снижает цену на макароны и сыр и выставляет макароны и сыр на рекламных стойках. В какой доле недель макароны и сыр находятся на распродаже или на рекламных стойках?
28. Известно, что из колоды вытащены две карты, и обе черви. Какова вероятность того, что первая карта была двойка червей?
29. В ящике лежат два нормальных четвертака и один четвертак, у которого с двух сторон орел. С закрытыми глазами вы вытаскиваете одну монету, ваш друг подбрасывает ее и говорит, что она упала орлом вверх. Какова вероятность того, что вы выбрали монету с двумя орлами?
30. Вы бросаете два кубика. Пусть событие  $A$  = сумма очков равна 10, а  $B$  = на первом кубике выпало нечетное число очков. Являются ли события  $A$  и  $B$  независимыми?
31. Среди страхователей страховой компании 20% являются высокорисковыми, 40% — низкорисковыми и 40% — среднерисковыми. Среди низкорисковых страхователей 2% получают травмы в течение года, среди среднерисковых — 4%, а среди высокорисковых — 20%. Если страхователь получил травму в течение года, какова вероятность того, что он относится к высокорисковым страхователям?

## ГЛАВА 69

# Введение в случайные величины

### Обсуждаемые вопросы

- Что такое случайная величина?
- Что такое дискретная случайная величина?
- Что такое среднее, дисперсия и стандартное отклонение случайной величины?
- Что такое непрерывная случайная величина?
- Что такое функция плотности вероятности?
- Что такое независимые случайные величины?

В современном мире мы уверены только в том, что ни в чем не можем быть уверены. В следующих девяти главах я покажу вам несколько эффективных методов включения неопределенности в бизнес-модели. Ключом к моделированию неопределенности является умение использовать *случайные величины*.

## Ответы на вопросы

### ? Что такое случайная величина?

Любая ситуация, исход которой не определен, называется *экспериментом*. Значение случайной величины основывается на (неопределенном) результате эксперимента. Например, бросок двух кубиков является экспериментом, и случайная величина может быть определена как сумма очков, выпавших на каждом кубике. В этом примере случайная величина могла быть равна 2, 3 и т. д. до 12. В качестве другого примера можно привести эксперимент с продажей новой игровой приставки, для которой случайная величина может быть определена как доля рынка этого нового продукта.

### ? Что такое дискретная случайная величина?

Случайная величина дискретна, если она может принимать конечное число возможных значений. Вот несколько примеров дискретных случайных величин:

- число потенциальных конкурентов;
- число тузов, выпавших в пятикарточном покере;
- количество дорожно-транспортных происшествий (надеюсь, 0!) в год;
- число точек на игральной кости;

- число штрафных бросков из 12, выполненных баскетболистом Кевином Дюрантом за игру.

### ❓ Что такое среднее, дисперсия и стандартное отклонение случайной величины?

В главе 44 «Обобщение данных с помощью описательной статистики» обсуждались среднее, дисперсия и стандартное отклонение для множества данных. В сущности, среднее случайной величины (обозначаемое как  $\mu$ ) — это среднее значение, которое можно было бы ожидать при многократном выполнении эксперимента. Среднее случайной величины часто называют *математическим ожиданием* случайной величины. Дисперсия случайной величины (обозначаемая как  $s^2$ ) — это среднее значение квадрата отклонения случайной величины от среднего, которое можно было бы ожидать при многократном выполнении эксперимента. Стандартное отклонение случайной величины (обозначаемое как  $\sigma$ ) — это просто корень квадратный из ее дисперсии. Как и для множеств данных, среднее случайной величины является суммарным показателем для типичного значения случайной величины, а дисперсия и стандартное отклонение оценивают разброс случайной величины вокруг ее среднего.

В качестве примера вычисления среднего, дисперсии и стандартного отклонения случайной величины рассмотрим предполагаемые доходы на фондовой бирже в следующем году, регулируемые следующими вероятностями:

Вероятность	Доход на фондовой бирже
0,40	+20%
0,30	0%
0,30	-20%

Выполненные вручную расчеты показывают следующее:

$$\mu = 0,40 \times (0,20) + 0,30 \times (0,00) + 0,30 \times (-0,20) = 0,02, \text{ или } 2\%;$$

$$s^2 = 0,4 \times (0,20 - 0,02) \times 2 + 0,30 \times (0,0 - 0,02) \times 2 + 0,30 \times (-0,20 - 0,02) \times 2 = 0,0276.$$

Тогда  $\sigma = 0,166$ , или 16,6%.

Эти расчеты я проверил в файле *Meanvariance.xlsx* (рис. 69.1).

Среднее для дохода на фондовой бирже я вычислил в ячейке C9 по формуле `=СУММПРОИЗВ(B4:B6;C4:C6)`. В этой формуле каждое значение случайной величины умножается на его вероятность, и затем произведения суммируются.

Для вычисления дисперсии дохода я определил квадрат отклонения каждого значения случайной величины от среднего, скопировав формулу `=(B4-$C$9)^2` из D4 в D5:D6. Затем в ячейке C10 я вычислил дисперсию дохода как среднее

	B	C	D	E
3	Value	Probability	Squared deviation	
4	0.2	0.4	0.0324	=(B4-\$C\$9)^2
5	0	0.3	0.0004	=(B5-\$C\$9)^2
6	-0.2	0.3	0.0484	=(B6-\$C\$9)^2
7				
8				
9	Mean	0.02	=SUMPRODUCT(B4:B6,C4:C6)	
10	Variance	0.0276	=SUMPRODUCT(C4:C6,D4:D6)	
11	Standard deviation	0.166132477	=SQRT(C10)	
12				

**Рис. 69.1.** Вычисление среднего, стандартного отклонения и дисперсии случайной величины

квадратов отклонений по формуле =СУММПРОИЗВ(C4:C6;D4:D6). Наконец, в ячейке C11 я вычислил стандартное отклонение дохода на фондовой бирже по формуле =КОРЕНЬ(C10).

### ? Что такое непрерывная случайная величина?

Непрерывная случайная величина — это случайная величина, которая может принимать очень большое или фактически бесконечное число значений, включающее все значения на некотором интервале. Вот несколько примеров непрерывных случайных величин:

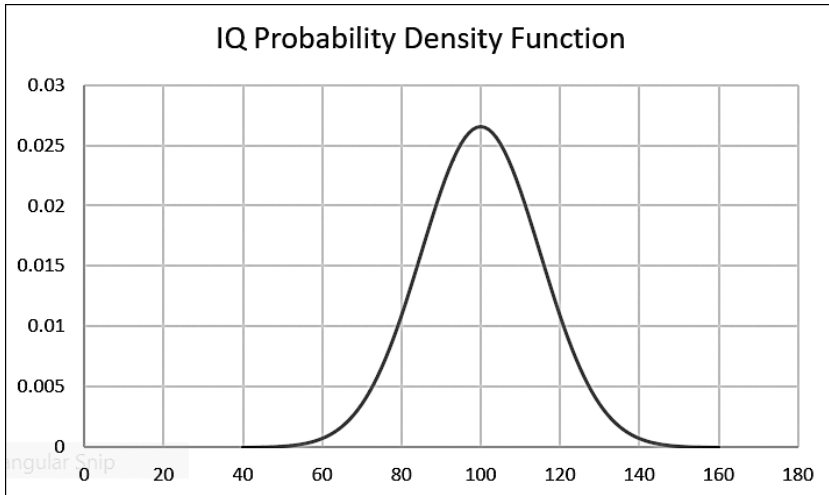
- цена акции Microsoft через год, начиная с этого момента;
- доля рынка для нового продукта;
- объем рынка для нового продукта;
- стоимость разработки нового продукта;
- вес новорожденного;
- чей-либо IQ;
- процент трехочковых бросков Дирка Новицки в следующем сезоне.

### ? Что такое функция плотности вероятности?

*Дискретная случайная величина* может быть задана набором значений и вероятностью появления каждого значения случайной величины. Поскольку непрерывная случайная величина может принимать бесконечное число значений, вероятность появления каждого значения непрерывной случайной величины перечислить невозможно. Непрерывная случайная величина полностью описывается своей *функцией плотности вероятности*. Например, функция плотности вероятности для IQ случайно выбранного человека показана на рис. 69.2.

Функция плотности вероятности (ФПВ) обладает несколькими свойствами.

- Значение ФПВ всегда больше или равно 0.
- Площадь под ФПВ равна 1.



**Рис. 69.2.** Функция плотности вероятности для IQ

- Высота ФПВ для значения  $x$  случайной величины пропорциональна вероятности того, что случайная величина примет значение в окрестностях  $x$ . Например, высота ФПВ для IQ = 83 составляет около половины высоты ФПВ для IQ = 100. Это говорит о том, что вероятность для IQ ниже 83 приблизительно в два раза меньше, чем вероятность для IQ около 100. Кроме того, поскольку пик плотности приходится на значение 100, значение IQ, приблизительно равное 100, наиболее вероятно.
- Вероятность того, что непрерывная случайная величина принимает диапазон значений, равна соответствующей площади под графиком функции плотности. Например, доля людей с IQ от 80 до 100 просто равна площади под графиком функции плотности от 80 до 100.
- Следует отметить, что дискретная случайная величина, принимающая много значений, часто моделируется непрерывной случайной величиной (см. главу 72 «Нормальная случайная величина и Z-оценка»). Например, несмотря на то что количество пакетов молока, продаваемых за день в небольшом продуктовом магазине, дискретно, гораздо удобнее моделировать эту дискретную случайную величину как непрерывную.

### ❓ Что такое независимые случайные величины?

Множество случайных величин независимо, если знание значений любого его подмножества ничего не говорит о значениях других случайных величин. Например, число игр, выигранных за год футбольной командой Университета Индианы, не зависит от процентного дохода по акциям Microsoft за тот же год. Знание о том, что дела у Индианы идут хорошо, не изменит взгляд на то, как провела год компания Microsoft.

Однако доходы по акциям Microsoft и Intel не являются независимыми. Если говорят о высоком доходе по акциям Microsoft в каком-либо году, то, по всей вероятности, продажи компьютеров в этом году также были значительными, что, очевидно, свидетельствует и о неплохом годе для Intel.

## Задания

1. Определите, являются ли следующие случайные величины дискретными или непрерывными:
  - число игр, которые «Сизл Сихокс» выиграет в следующем сезоне;
  - число, которое выпадет после остановки колеса рулетки;
  - объем продаж планшетных ПК в следующем году;
  - срок службы лампочки до того, как она перегорит.
2. Вычислите среднее, дисперсию и стандартное отклонение для числа очков при броске кубика.
3. Определите, являются ли следующие переменные независимыми:
  - ежедневная температура и продажи мороженого в магазине;
  - масть и достоинство карты, вытащенной из колоды;
  - инфляция и доход на фондовой бирже;
  - изменившаяся цена модели автомобиля и количество проданных автомобилей.
4. Текущая цена акции компании составляет \$20. Компания является объектом для поглощения. Если поглощение пройдет успешно, цена акции повысится до \$30. Если поглощения не будет, цена акции упадет до \$12. Определите диапазон значений для вероятности успешного поглощения, благодаря которому целесообразно приобрести акции сегодня. Предположим, что цель заключается в получении максимальной ожидаемой прибыли. Подсказка: используйте инструмент Подбор параметра (Goal Seek), рассмотренный в главе 18 «Инструмент Подбор параметра».
5. При вращении колеса рулетки возможными исходами являются номера секторов 0, 00, 1, 2, ..., 36. Если ставить на какой-либо сектор, то можно выиграть \$35, если выпадет этот сектор, или проиграть \$1 в противном случае. Каково среднее значение и стандартное отклонение выигрыша при однократном запуске рулетки?
6. В настоящий момент акции торгуются по \$40. Через месяц с 60%-ной вероятностью цена акций удвоится, и с 40%-ной вероятностью цена акции упадет на 50%. Через месяц вы собираетесь продать акции. Найдите среднее значение и стандартное отклонение вашего дохода (в долларах).
7. Вы поставили в рулетке на нечет. Если выпадет нечетное число, вы выигрываете \$1, а если нет, вы проигрываете \$1. Найдите среднее значение и стандартное отклонение вашего дохода.

## ГЛАВА 70

# Биномиальные, гипергеометрические и отрицательные биномиальные случайные величины

### Обсуждаемые вопросы

- Что такое биномиальная случайная величина?
- Как с помощью функций БИНОМ.РАСП и БИНОМ.РАСП.ДИАП вычислить вероятности биномиального распределения?
- Если одинаковое количество людей предпочитает кока-колу пепси-коле и наоборот, и у 100 человек спросили, предпочитают ли они кока-колу пепси-коле, какова вероятность того, что ровно 60 человек предпочитают кока-колу пепси-коле, и вероятность того, что от 40 до 60 человек предпочитают кока-колу пепси-коле?
- Из всех лифтовых направляющих, производимых компанией, 3% считаются дефектными. Компания собирается отгрузить клиенту партию из 10 000 направляющих. Перед приемкой партии клиент случайным образом отбирает 100 направляющих и проверяет на наличие дефектов. Если дефекты обнаружатся не более чем у двух направляющих, клиент примет партию. Как определить вероятность приемки партии?
- Авиакомпаниям не нравятся рейсы с неза занятыми местами. Предположим, что в среднем 95% всех купивших билет регистрируются на рейс. Если авиакомпания продает 105 билетов на 100-местный рейс, какова вероятность того, что произойдет избыточное бронирование рейса?
- В кафе Village Deli ежедневно приходит на ланч 1000 клиентов. В среднем 20% клиентов заказывают особый вегетарианский сэндвич. Эти сэндвичи готовят заранее. Сколько сэндвичей необходимо приготовить, если вероятность их нехватки должна быть равна 5%?
- Что такое гипергеометрическая случайная величина?
- Что такое отрицательная биномиальная случайная величина?

## Ответы на вопросы

### ❓ Что такое биномиальная случайная величина?

*Биномиальная случайная величина* — это дискретная случайная величина, используемая для вычисления вероятностей при выполнении следующих трех условий:

- проводится  $n$  независимых испытаний;
- каждое испытание приводит к одному из двух исходов: благоприятному либо неблагоприятному;
- в каждом испытании вероятность благоприятного исхода ( $p$ ) остается постоянной.

В такой ситуации для вычисления вероятностей, связанных с числом благоприятных исходов в заданном числе испытаний, может быть использована биномиальная случайная величина. Пусть  $x$  — случайная величина, обозначающая число благоприятных исходов в  $n$  независимых испытаниях, когда вероятность благоприятного исхода в каждом испытании равна  $p$ . Вот несколько примеров, в которых биномиальная случайная величина релевантна.

- **Кока-кола или пепси-кола.** Пусть одинаковое количество людей предпочитает кока-колу пепси-коле и наоборот. У 100 человек спросили, предпочитают ли они кока-колу пепси-коле. Требуется узнать вероятность того, что 60 человек предпочитают кока-колу пепси-коле, и вероятность того, что от 40 до 60 человек предпочитают кока-колу пепси-коле. В этой ситуации биномиальная случайная величина определяется следующим образом:

- испытание: опрос отдельных лиц;
- благоприятный исход: предпочтение отдано кока-коле;
- $p = 0,50$ ;
- $n = 100$ .

Пусть  $x$  — количество человек из выборки, предпочитающих кока-колу. Необходимо определить вероятность того, что  $x = 60$  и  $40 \leq x \leq 60$ .

- **Направляющие для лифтов.** Из всех лифтовых направляющих, производимых компанией, 3% считаются дефектными. Компания собирается отгрузить клиенту партию из 10 000 направляющих. Перед приемкой партии клиент случайным образом отбирает 100 направляющих и проверяет на наличие дефектов. Если дефекты обнаружатся не более чем у двух направляющих, клиент примет партию. Необходимо определить вероятность того, что партия будет принята.

Биномиальная случайная величина определяется следующим образом:

- испытание: осмотр выбранной направляющей;
- благоприятный исход: направляющая имеет дефекты;



- $p = 0,03$ ;
- $n = 100$ .

Пусть  $x$  — число дефектных направляющих в выборке. Требуется найти вероятность того, что  $x \leq 2$ .

○ **Избыточное бронирование в авиакомпании.** Авиакомпаниям не нравятся рейсы с незанятыми местами. Предположим, что в среднем 95% всех купивших билетов регистрируются на рейс. Если авиакомпания продает 105 билетов на 100-местный рейс, какова вероятность того, что произойдет избыточное бронирование рейса?

Биномиальная случайная величина определяется следующим образом:

- испытание: отдельные владельцы билетов;
- благоприятный исход: регистрация владельца билета;
- $p = 0,95$ ;
- $n = 105$ .

Пусть  $x$  — число зарегистрировавшихся на рейс владельцев билетов. Необходимо найти вероятность того, что  $x \geq 101$ .

❓ **Как с помощью функции БИНОМ.РАСП и БИНОМ.РАСП.ДИАП вычислить вероятности биномиального распределения?**

Microsoft Excel включает функции БИНОМ.РАСП (BINOM.DIST) и БИНОМ.РАСП.ДИАП (BINOM.DIST.RANGE), с помощью которых можно вычислить вероятности биномиального распределения. Если требуется вычислить вероятность не более  $x$  благоприятных исходов для биномиальной случайной величины при  $n$  испытаниях с вероятностью благоприятного исхода  $p$ , просто введите формулу =БИНОМ.РАСП( $x$ ;  $n$ ;  $p$ ; 1). Если требуется вычислить вероятность ровно  $x$  благоприятных исходов для биномиальной случайной величины при  $n$  испытаниях с вероятностью благоприятного исхода  $p$ , введите формулу =БИНОМ.РАСП( $x$ ;  $n$ ;  $p$ ; 0). Ввод значения 1 в качестве последнего аргумента функции БИНОМ.РАСП приводит к выводу накопленной вероятности; ввод значения 0 приводит к расчету функции распределения масс для любого частного значения. (Обратите внимание, что в качестве последнего аргумента вместо 1 можно указать значение ИСТИНА, а вместо значения 0 — ЛОЖЬ.)

Функция БИНОМ.РАСП.ДИАП( $n$ ;  $p$ ;  $s1$ ;  $s2$ ) вычисляет вероятность получения от  $s1$  до  $s2$  (включительно) благоприятных исходов при  $n$  независимых испытаниях с вероятностью благоприятного исхода  $p$  в каждом испытании.

Далее приведены примеры вычисления вероятностей, о которых шла речь выше в главе, с помощью функции БИНОМ.РАСП. Данные и анализ см. в файле Binomialexamples.xlsx, показанном на рис. 70.1.

❓ Если одинаковое количество людей предпочитает кока-колу пепси-коле и наоборот, и у 100 человек спросили, предпочитают ли они кока-колу пепси-коле, какова вероятность того, что ровно 60 человек предпочитают кока-колу пепси-коле, и вероятность того, что от 40 до 60 человек предпочитают кока-колу пепси-коле?

Здесь  $n = 100$  и  $p = 0,5$ . Требуется найти вероятность того, что  $x = 60$ , и вероятность того, что  $40 \leq x \leq 60$ , где  $x$  — это количество человек, предпочитающих кока-колу пепси-коле. Сначала найдем вероятность того, что  $x = 60$ , введя в ячейку C4 формулу =БИНОМ.РАСП(60;100;0,5;0). Результат — 0,011.

	B	C	D	E	F
3	Coke vs. Pepsi Probability exactly 60 people prefer Coke to Pepsi	Binom.dist			
4	Probability from 40 through 60 people prefer Coke to Pepsi	0.01084387 =BINOM.DIST(60,100,0.5,0)		BINOM.DIST.RANGE	
5		0.9647998 =BINOM.DIST(60,100,0.5,1)-BINOM.DIST(39,100,0.5,1)		0.9647998	=BINOM.DIST.RANGE(100,0.5,40,60)
6	Elevator Rails Probability 2 or fewer of 100 rails are defective				
7	Airline Overbooking Probability flight is overbooked if 105 tickets are sold				
8		0.41977508 =BINOM.DIST(2,100,0.03,1)		0.419775083	=BINOM.DIST.RANGE(100,0.03,0,2)
9					
10		0.39243372 =1-BINOM.DIST(100,105,0.95,1)		0.392433721	=BINOM.DIST.RANGE(105,0.95,101,105)

**Рис. 70.1.** Использование биномиальной случайной величины

Чтобы применить функцию БИНОМ.РАСП для вычисления вероятности для  $40 \leq x \leq 60$ , следует заметить, что эта вероятность равна вероятности для  $x \leq 60$  минус вероятность для  $x \leq 39$ . Таким образом, получить вероятность того, что от 40 до 60 человек предпочитают кока-колу, можно по формуле в ячейке C5: =БИНОМ.РАСП(60;100;0,5;1)-БИНОМ.РАСП(39;100;0,5;1). Результат — 0,9648. Итак, если кока-колу и пепси-колу предпочитают в равной степени, маловероятно, что в выборке из 100 человек кока-кола или пепси-кола будет иметь преимущество более 10%. Если выборка из 100 человек показывает 10%-ное преимущество кока-колы или пепси-колы, то возникают сомнения в том, что кока-колу и пепси-колу предпочитают в равной степени. Та же вероятность вычислена в ячейке E5 по формуле = БИНОМ.РАСП.ДИАП(100;0,5;40;60).

❓ Из всех лифтовых направляющих, производимых компанией, 3% считаются дефектными. Компания собирается отгрузить клиенту партию из 10 000 направляющих. Перед приемкой партии клиент случайным образом отбирает 100 направляющих и проверяет на наличие дефектов. Если дефекты обнаружатся не более чем у двух направляющих, клиент примет партию. Как определить вероятность приемки партии?

Если  $x$  — число дефектных направляющих в партии, это биномиальная случайная величина с  $n = 100$  и  $p = 0,03$ . Необходимо найти вероятность того, что  $x \leq 2$ . Вве-

дите формулу `=БИНОМ.РАСП(2;100;0,03;1)` в ячейку C8. Результат — 0,42. Таким образом, партия будет принята в 42% случаев. Та же вероятность вычислена в ячейке E7 по формуле `= БИНОМ.РАСП.ДИАП(100;0,03;0;2)`.

В реальности вероятность благоприятного исхода не равна ровно 3% в каждом испытании. Например, если первые 10 направляющих имеют дефекты, вероятность того, что следующая направляющая тоже имеет дефекты, снижается до 290/9990; если первые 10 направляющих не имеют дефектов, вероятность того, что следующая направляющая дефектная — 300/9990. Следовательно, вероятность благоприятного исхода в одиннадцатом испытании зависит от вероятности благоприятного исхода в одном из первых 10 испытаний. Несмотря на это, биномиальная случайная величина может быть использована в качестве приближения, если выборка сделана и объем выборки составляет менее 10% от генеральной совокупности. Здесь генеральная совокупность — 10 000 и объем выборки — 100. Точные вероятности, связанные с выборкой из конечной совокупности, могут быть вычислены с помощью *гипергеометрической случайной величины*, рассматриваемой далее в этой главе.

**?** **Авиакомпаниям не нравятся рейсы с незанятыми местами. Предположим, что в среднем 95% всех купивших билет регистрируются на рейс. Если авиакомпания продает 105 билетов на 100-местный рейс, какова вероятность того, что произойдет избыточное бронирование рейса?**

Пусть  $x$  — число обладателей билетов, зарегистрировавшихся на рейс. Здесь  $n = 105$  и  $p = 0,95$ . Требуется найти вероятность того, что  $x \geq 101$ . Обратите внимание, вероятность того, что  $x \geq 101$ , равна 1 минус вероятность того, что  $x \leq 100$ . Таким образом, для вычисления вероятности избыточного бронирования рейса введите в ячейку C10 формулу `=1-БИНОМ.РАСП(100;105;0,95;1)`. Результат 0,392 означает, что вероятность избыточного бронирования равна 39,2%. Та же вероятность вычислена в ячейке E10 по формуле `=БИНОМ.РАСП.ДИАП(105;0,95;101;105)`.

**?** **В кафе Village Deli ежедневно приходит на ланч 1000 клиентов. В среднем 20% клиентов заказывают особый вегетарианский сэндвич. Эти сэндвичи готовят заранее. Сколько сэндвичей необходимо приготовить, если вероятность их нехватки должна быть равна 5%?**

Начиная с Microsoft Excel 2013 функция `БИНОМ.ОБР (BINOM.INV)` с синтаксисом `БИНОМ.ОБР(число_испытаний; вероятность_успеха; альфа)` определяет наименьшее число  $x$ , для которого вероятность не более  $x$  благоприятных исходов равна, по меньшей мере,  $\alpha$ . В более ранних версиях Excel результаты, аналогичные результатам функции `БИНОМ.ОБР`, давала функция `КРИТБИНОМ (CRITBINOM)` с синтаксисом `КРИТБИНОМ(число_испытаний; вероятность_успеха; альфа)`. В этом примере число испытаний равно 1000, вероятность успеха — 0,2 и  $\alpha = 0,95$ . Как показано на рис. 70.2, если в кафе приготовят 221 сэндвич, вероятность того, что закажут не более 221 сэндвича, равна, по меньшей мере, 0,95. Вероятность того, что потребуются не больше 220 сэндвичей, составляет менее 0,95.

	G	H	I
4	Customers	1000	
5	fraction wanting veggie	0.2	
6	prob of filling demand	0.95	
7			
8	221	=BINOM.INV(H4,H5,H6)	
9	<=220	0.946142835	=BINOM.DIST.RANGE(H4,H5,0,220)
10	<=221	0.954070186	=BINOM.DIST.RANGE(H4,H5,0,221)

**Рис. 70.2.** Пример с функцией БИНОМ.ОБР

### ? Что такое гипергеометрическая случайная величина?

Гипергеометрическая случайная величина определяет, например, такую ситуацию:

- в урне находится  $N$  шаров;
- каждый шар относится к одному из двух типов (благоприятный или неблагоприятный исход);
- в урне  $s$  благоприятных исходов;
- из урны взята выборка объемом  $n$  шаров.

Давайте посмотрим на пример в файле `Hypergeom.dist.xlsx` (рис. 70.3). В формуле `=ГИПЕРГЕОМ.РАСП(х;n;s;N;0)` функция `ГИПЕРГЕОМ.РАСП` (`HYPERGEOM.DIST`) вычисляет вероятность  $x$  благоприятных исходов, если выбрано  $n$  шаров из урны, содержащей  $N$  шаров, из которых  $s$  помечены как благоприятный исход. В формуле `=ГИПЕРГЕОМ.РАСП(х;n;s;N;1)` вычисляется вероятность не более  $x$  благоприятных исходов, если взято  $n$  шаров из урны, содержащей  $N$  шаров, из которых  $s$  помечены как благоприятный исход. (Как и в случае с функцией `БИНОМ.РАСП`, вместо значения 1 можно указать `ИСТИНА`, а вместо 0 — `ЛОЖЬ`.)

Например, предположим, что 40 компаниями из списка `Fortune 500` руководят женщины. Пятьсот главных директоров являются аналогом шаров в урне ( $N = 500$ ), а 40 женщин представляют  $s$  благоприятных исходов в урне. Формула `=ГИПЕРГЕОМ.РАСП(С8;объем_выборки; женщины_в_ген.совокупности; объем_ген.совокупности; 0)`, скопированная из D8 в D9:D18, дает вероятность того, что в выборке из 10 компаний из списка `Fortune 500` женщины возглавляют 0, 1, 2, ..., 10 компаний. Здесь `объем_выборки = 10`, `женщины_в_ген.совокупности = 40` и `объем_ген.совокупности = 500`. Вместо значения в формуле 0 можно ввести `FALSE`.

Благоприятным исходом является факт, что директором является женщина. Для выборки из 10 компаний, например, вероятность отсутствия в выборке женщин-директоров равна 0,431. Кстати, эту вероятность можно приблизительно вычислить по формуле  $\text{=БИНОМ.РАСП}(0;10;0,08;0)$ , получив в результате 0,434, что очень близко к истинному значению вероятности 0,431. В ячейке F10 по формуле  $\text{=ГИПЕРГЕОМ.РАСП}(2; \text{объем\_выборки}; \text{женщины\_в\_ген.совокупности}; \text{объем\_ген.совокупности}; \text{ИСТИНА})$  вычислена вероятность того, что не более 2 человек в выборке из 10 человек являются женщинами. Таким образом, вероятность того, что директорами окажутся не более двух женщин, равна 96,2%. Разумеется, такой же результат можно было получить, сложив значения в выделенных ячейках (D8:D10).

	C	D	E	F	G	H	I	J	K
3	Population size	500							
4	Sample Size	10							
5	Population women	40		Binomial approximation					
6				0.434388454	=BINOMDIST(0,10,0.08,0)				
7	Number of women	Probability							
8	0	0.430956908	=HYPGEOM.DIST(C8,Sample_Size,Population_women,Population_size,FALSE)						
9	1	0.382223421		<=2 women in sample					
10	2	0.148407545		0.961587874					
11	3	0.033197861							
12	4	0.004734717		=HYPGEOM.DIST(2,Sample_Size,Population_women,Population_size,TRUE)					
13	5	0.000449538							
14	6	2.87533E-05							
15	7	1.2224E-06							
16	8	3.30288E-08							
17	9	5.11702E-10							
18	10	3.44843E-12							

Рис. 70.3. Гипергеометрическая случайная величина

### ❓ Что такое отрицательная биномиальная случайная величина?

Отрицательная биномиальная случайная величина применима к тем же ситуациям, что и биномиальная случайная величина, но отрицательная биномиальная случайная величина дает вероятность  $f$  неблагоприятных исходов до  $s$ -го благоприятного исхода. Таким образом, формула  $\text{=ОТРБИНОМ.РАСП}(f;s;p;0)$  вычисляет вероятность того, что будет достигнуто ровно  $f$  неблагоприятных исходов до  $s$ -го благоприятного исхода, если вероятность благоприятного исхода равна  $p$  для всех испытаний. А формула  $\text{=ОТРБИНОМ.РАСП}(f;s;p;1)$  вычисляет вероятность того, что не более  $f$  неблагоприятных исходов будет достигнуто до  $s$ -го благоприятного исхода, если вероятность благоприятного исхода равна  $p$  для всех испытаний. Возьмем для примера бейсбольную команду, выигрывающую 40% игр (см. файл Negbinom.dist.xlsx и рис. 70.4). Формула  $\text{=ОТРБИНОМ.РАСП}(D9;2;0,4;ЛОЖЬ)$ , скопированная из E9 в E34, дает вероятность для 0, 1, 2, ..., 25 неблагоприятных исходов перед вторым благоприятным. Здесь благоприятный исход эквивалентен выигрышу. Например, с вероятностью 19,2% команда проиграет одну игру до того, как выиграет две игры. В ячейке G10 я использовал формулу  $\text{=ОТРБИНОМ.РАСП}(3;2;0,4;ИСТИНА)$  для вычисления вероятности того, что команда потерпит не более трех поражений до того, как выиграет две игры (0,663). Конечно, тот же самый результат можно было получить, просто сложив значения в выделенных ячейках (E9:E12).

	D	E	F	G
4	40% chance of success			
5				
6				
7				
8	Losses before 2nd win	prob		Probability <=3 failures
9	0	0.16	=NEGBINOM.DIST(D9,2,0.4,FALSE)	before 2nd win
10	1	0.192	=NEGBINOM.DIST(D10,2,0.4,FALSE)	0.66304
11	2	0.173	=NEGBINOM.DIST(D11,2,0.4,FALSE)	
12	3	0.138	=NEGBINOM.DIST(D12,2,0.4,FALSE)	=NEGBINOM.DIST(3,2,0.4,TRUE)
13	4	0.104		
14	5	0.075		
15	6	0.052		
16	7	0.036		
17	8	0.024		
18	9	0.016		
19	10	0.011		
20	11	0.007		

**Рис. 70.4.** Отрицательная биномиальная случайная величина

## Задания

- Допустим, в среднем 4% всех дисководов для компакт-дисков, получаемых компьютерной компанией, имеют дефекты. Компания придерживается следующей политики: выбирает 50 дисководов из каждой партии и принимает партию, если ни один из выбранных дисководов не имеет дефектов. На основе этой информации определите следующее:
  - Какая часть партий будет принята?
  - Какая часть партий будет принята, если политика компании изменится следующим образом: партия принимается при наличии в выборке только одного дисковода с дефектами?
  - Какова вероятность того, что выборка объемом 50 дисководов содержит по меньшей мере, 10 дисководов с дефектами?
- На основе данных об избыточном бронировании в авиакомпании:
  - определите, как изменяется вероятность избыточного бронирования при изменении количества проданных билетов со 100 до 115. Подсказка: используйте однонаправленную таблицу данных;
  - покажите, как изменяется вероятность избыточного бронирования при изменении количества проданных билетов со 100 до 115 и изменении вероятности регистрации владельцев билетов от 80 до 95%. Подсказка: используйте двунаправленную таблицу данных.
- Пусть каждый год заданный открытый фонд имеет 50%-ный шанс превзойти фондовый индекс Standard&Poor's 500 Stock Index (индекс S&P 500). Како-

ва вероятность того, что в группе из 100 открытых фондов, по меньшей мере, 10 фондов превзойдут индекс S&P 500 в течение как минимум 8 лет из 10?

4. Профессиональный баскетболист Стив Нэш забивает штрафные броски с вероятностью 90%. Ответьте на следующие вопросы:
  - Какова вероятность непопадания мяча в корзину более 15 раз при выполнении 100 штрафных бросков?
  - Насколько хорошо Стив Нэш пробивал бы штрафные броски, если бы он попадал в корзину меньше 90 раз из 100 с вероятностью только 5%? Подсказка: используйте инструмент Подбор параметра (Goal Seek) в опциях Анализ "что, если" (What-If Analysis).
5. При тестировании на экстрасенсорное восприятие участников просят определить фигуру на карте из 25-карточной колоды. В колоде 5 фигур и 5 карт с каждой фигурой. Какой можно сделать вывод, если участник правильно определил фигуру на 12 картах?
6. Допустим, в группе из 100 человек 20 заболели гриппом, а 80 — нет. Если выбрать 30 человек случайным образом, какова вероятность того, что среди них окажется, по меньшей мере, 10 человек, заболевших гриппом?
7. Учащийся продает журналы в поддержку школьного благотворительного фонда. Существует 20%-ная вероятность покупки журнала в данном доме. Необходимо продать пять журналов. Определите вероятность необходимости посещения 5, 6, 7, ..., 100 домов для продажи пяти журналов.
8. Проверьте с помощью функции **БИНОМ.РАСП** правильность ответа с использованием функции **БИНОМ.ОБР** для примера с кафе Village Deli.
9. В XVII веке французские математики Пьер Ферма́ и Блез Паскаль, пытаясь решить «задачу об очках», смогли сформулировать положения современной теории вероятностей. Поясним на простом примере: пусть Паскаль и Ферма по очереди бросают монету. Ферма выигрывает очко, если выпадает орел, а Паскаль выигрывает, если выпадает решка. Если Паскаль впереди со счетом 8:7 и игрок, первым набравший 10 очков, выигрывает, каковы шансы Паскаля на выигрыш?
10. Допустим, если клуб «Хьюстон Астрос» играет с клубом «Нью-Йорк Янкиз», «Астрос» имеет 35%-ную вероятность победы в игре. Если «Астрос» играет с «Янкиз» 10 раз, какова вероятность, что «Астрос» выиграет не менее половины игр?
11. Результативность моих штрафных бросков составляет 55%. Если я сделаю 400 бросков, какова вероятность, что я попаду от 200 до 300 раз (включительно)?
12. Предположим, 10% степлеров неисправны. Если было продано 60 степлеров, какова вероятность, что не менее 10 из них неисправны?
13. На производстве имеется пять сборочных линий. Каждая линия остановлена (не работает)  $p$  процентов времени. Каким должно быть  $p$ , чтобы с вероятностью 95% по меньшей мере одна линия работала?

## ГЛАВА 71

# Пуассоновская и экспоненциальная случайные величины

### Обсуждаемые вопросы

- Что такое пуассоновская случайная величина?
- Как вычислить распределение пуассоновской случайной величины?
- Если количество клиентов, обращающихся в банк, описывается пуассоновской случайной величиной, какая случайная величина определяет время между обращениями?

## Ответы на вопросы

### ❓ Что такое пуассоновская случайная величина?

*Пуассоновская случайная величина* — это дискретная случайная величина, применяемая для описания вероятностей в ситуациях, в которых события (такие как обращение клиентов в банк или размещение заказов на продукт) имеют малую вероятность возникновения в короткий промежуток времени. Точнее говоря, в короткий промежуток времени, обозначаемый как  $t$ , произойдет одно событие или ни одного события, и вероятность одного события, происходящего в короткий промежуток времени длиной  $t$  (для некоторого  $\lambda$ ), определяется  $\lambda t$ . Здесь  $\lambda$  — среднее число событий в единицу времени.

Ситуации, к которым применима пуассоновская случайная величина:

- количество единиц продукта, требуемое в течение месяца;
- количество смертей в год от удара копытом в прусской армии;
- количество ДТП для одного водителя в год;
- количество экземпляров «Где находится душа», заказанное сегодня на Amazon.com;
- количество исков о компенсации ущерба, поданных сотрудниками компании в этом месяце;



- количество дефектов в 100 ярдах струны. (Здесь 1 ярд струны играет роль времени.)

### ? Как вычислить распределение пуассоновской случайной величины?

Для вычисления распределения пуассоновской случайной величины в Excel можно воспользоваться функцией ПУАССОН.РАСП (POISSON.DIST). В версиях Excel, предшествующих Excel 2013, пуассоновское распределение можно получить с помощью функции ПУАССОН (POISSON). Просто помните, что в промежуток времени  $t$  среднее значение пуассоновской случайной величины равно  $t$ . Синтаксис функции ПУАССОН.РАСП:

- функция ПУАССОН.РАСП( $x$ ;лямбда;ИСТИНА\_или\_1) вычисляет вероятность того, что пуассоновская случайная величина со средним значением, равным лямбда, не больше  $x$ ;
- функция ПУАССОН.РАСП( $x$ ;лямбда;ЛОЖЬ\_или\_0) вычисляет вероятность того, что пуассоновская случайная величина со средним значением, равным лямбда, равна  $x$ .

Вот несколько примеров вычисления пуассоновской случайной величины. Эти примеры вы найдете в файле Poisson.xlsx (рис. 71.1).

	B	C	D	E	F
1	Calls per hour	30			
2	Mean	60			
3					
4	Prob 60 calls in two hours	0.0514	=POISSON.DIST(60,C2,FALSE)		
5	Prob <= 60 calls in two hours	0.5343	=POISSON.DIST(60,C2,1)		
6	Prob between 50 and 100 calls(inclusive) in two hours	0.9156	=POISSON.DIST(100,C2,TRUE) - POISSON.DIST(49,C2,TRUE)		

**Рис. 71.1.** Использование пуассоновской случайной величины

Допустим, в мою консалтинговую фирму звонят 30 раз в час. Для двухчасового периода необходимо определить следующее:

- вероятность того, что за следующие два часа поступит ровно 60 звонков;
- вероятность того, что количество звонков за следующие два часа не превысит 60;
- вероятность того, что за следующие два часа будет получено от 50 до 100 звонков.

Среднее количество звонков за двухчасовой период равно 60. В ячейке C4 я вычислил по формуле =ПУАССОН.РАСП(60;C2;ЛОЖЬ) вероятность (0,05) 60 звонков в следующие два часа. В ячейке C5 по формуле =ПУАССОН.РАСП(60;C2;1) я нашел вероятность (0,534) получения не более 60 звонков за два часа. В ячейке C6 по формуле =ПУАССОН.РАСП(100;C2;ИСТИНА)-ПУАССОН.РАСП(49;C2;ИСТИНА) я рассчитал вероятность (0,916) получения от 50 до 100 звонков за два часа.

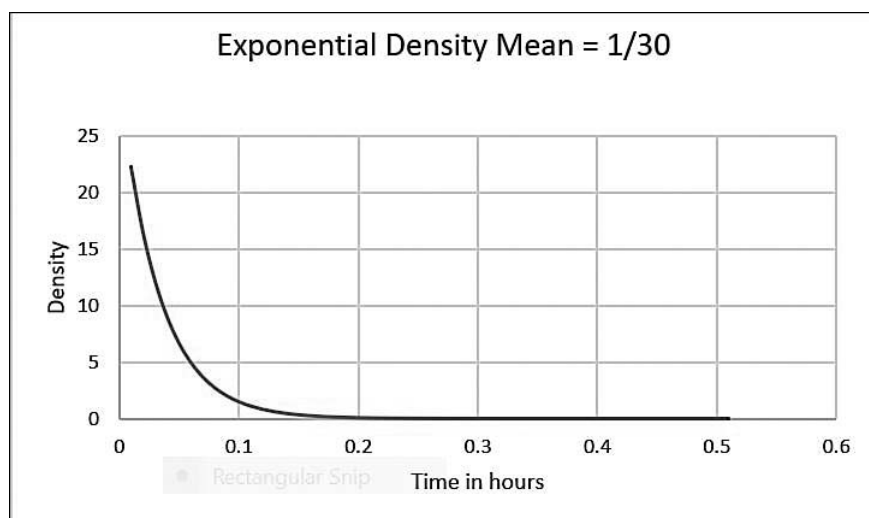
**ПРИМЕЧАНИЕ**

В качестве аргумента любой функции Excel вместо значения ИСТИНА всегда можно указать значение 1.

**?** Если количество клиентов, обращающихся в банк, описывается пуассоновской случайной величиной, какая случайная величина определяет время между обращениями?

Время между обращениями может быть любым, поэтому время между обращениями представляет собой непрерывную случайную величину. Если среднее число обращений  $\lambda$  происходит за единицу времени, время между обращениями соответствует экспоненциальной случайной величине с плотностью распределения вероятностей  $f(t) = \lambda e^{-\lambda t}$  (при  $t \geq 0$ ).

Эта случайная переменная имеет среднее значение, равное  $1/\lambda$ . График экспоненциальной функции плотности вероятности (ФПВ) для  $\lambda = 30$  показан на рис. 71.2. Вы найдете диаграмму и данные для этого примера на листе **Density** в файле **Exponentialdist.xlsx**.



**Рис. 71.2.** Экспоненциальная функция распределения плотности вероятностей

В главе 69 «Введение в случайные величины» говорилось о том, что для непрерывной случайной величины значение ФПВ в точках отражает вероятность, с которой случайная величина примет значение в окрестностях точки  $x$ . Как видно из рис. 71.2, чрезвычайно короткое время между обращениями в банк (например, меньше 0,05 ч) весьма вероятно, а для более длительных промежутков времени значение функции резко падает.

Даже если среднее время между обращениями составляет  $1/30 = 0,033$  ч, существует вероятность, что время между обращениями составит не более 0,20 ч. Функция ЭКСП.РАСП (EXPON.DIST) в формуле `=ЭКСП.РАСП(х;1/среднее;ИСТИНА_или_1)` дает вероятность того, что экспоненциальная случайная величина с заданным средним примет значение не больше  $x$ . Таким образом, второй аргумент функции ЭКСП.РАСП — это интенсивность возникновения событий. Например, для вычисления вероятности, с которой время между обращениями составит 5, 10 или 15 минут, скопируйте формулу `=1-ЭКСП.РАСП(C5;$D$2;ИСТИНА)` из ячейки D5 в D7. В более ранних версиях Excel такой же результат, как функция ЭКСП.РАСП, дает функция ЭКСПРАСП (EXPONDIST).

Обратите внимание, что сначала я преобразовал минуты в часы (5 минут =  $1/12$  часа и т. д.). Кроме того, среднее время между обращениями составляет 0,033 ч, поэтому я ввел в формулу значение  $1/\text{среднее} = 1/0,033 = 30$ , то есть количество обращений в единицу времени (рис. 71.3 и лист Computation).

	B	C	D	E
1		mean	0.033333333	
2		1/mean	30	
3				
4		x = Time between arrivals	Prob time >=x	
5	5 minutes	0.08333333	0.082084999	=1-EXPON.DIST(C5,\$D\$2,TRUE)
6	10 minutes	0.16666667	0.006737947	=1-EXPON.DIST(C6,\$D\$2,TRUE)
7	15 minutes	0.25	0.000553084	=1-EXPON.DIST(C7,\$D\$2,TRUE)

Рис. 71.3. Вычисление экспоненциального распределения

## Задания

- В пабе «У Ника» в Блумингтоне, штат Индиана, посетители заказывают в среднем 40 кружек пива в час.
  - Какова вероятность заказа не менее 100 кружек за два часа?
  - Какова вероятность того, что время между заказами составит не более 30 секунд?
- Предположим, что подростки-водители попадают в среднем в 0,3 ДТП в год.
  - Какова вероятность, что подросток попадет не более чем в одно ДТП в течение года?
  - Какова вероятность того, что время между ДТП составит не более шести месяцев?

3. Я следующий в очереди в ресторане быстрого питания, в котором покупатели образуют одну очередь, и время обслуживания покупателя соответствует экспоненциальному распределению со средним значением 3 минуты. Какова вероятность ожидания обслуживания не менее 5 минут?
4. С 1900 г. доля игр без хитов среди всех игр Высшей бейсбольной лиги составила 0,00124. Команда играет 162 игры за сезон. Какова вероятность того, что команда сыграет не менее двух игр без хитов за сезон?
5. В кофейню в среднем заходит 80 посетителей в час. Какова вероятность того, что за двухчасовой период зайдет не менее 150 посетителей?
6. Какая почасовая посещаемость кофейни обеспечит вероятность, равную 0,5, того, что не менее 150 посетителей посетит кофейню в течение трех часов?

## ГЛАВА 72

# Нормальная случайная величина и Z-оценка

### Обсуждаемые вопросы

- Каковы свойства нормальной случайной величины?
- Как в Excel найти распределение нормальной случайной величины?
- Как в Excel найти процентиля для нормальных случайных величин?
- Почему нормальная случайная величина соответствует многим реальным ситуациям?
- Что такое Z-оценка?

### Ответы на вопросы

#### ❓ Каковы свойства нормальной случайной величины?

Из главы 69 «Введение в случайные величины» вы узнали, как использовать непрерывные случайные величины для моделирования следующих величин:

- цена акций Microsoft через год, начиная с этого момента;
- доля рынка для нового продукта;
- объем рынка для нового продукта;
- стоимость разработки нового продукта;
- вес новорожденного ребенка;
- чей-либо IQ.

Напомню, что если дискретная случайная величина (такая как объем продаж биг-маков за 2025 г.) может принимать много возможных значений, ее можно также аппроксимировать с помощью непрерывной случайной величины. Как описано в главе 69, любая непрерывная случайная величина  $X$  характеризуется функцией плотности вероятностей (ФПВ). ФПВ непрерывной случайной величины является неотрицательной функцией со следующими свойствами ( $a$  и  $b$  — произвольные числа):

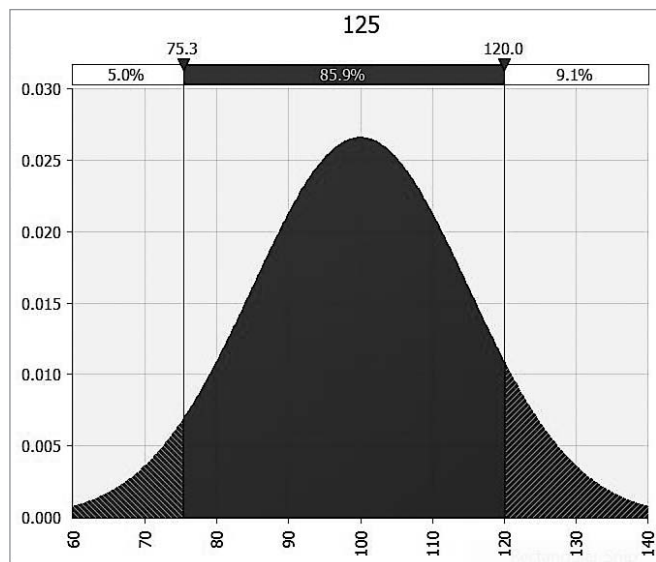
- площадь под кривой распределения вероятностей равна 1;
- вероятность  $X < a$  равна вероятности  $X \leq a$ . Эта вероятность представлена площадью под кривой распределения вероятностей слева от  $a$ ;

- вероятность  $X > b$  равна вероятности  $X \geq b$ . Эта вероятность представлена площадью под кривой распределения вероятностей справа от  $b$ ;
- вероятность  $a < X < b$  равна вероятности  $a \leq X \leq b$ . Эта вероятность равна площади под кривой распределения вероятностей между  $a$  и  $b$ .

Таким образом, вероятность численно равна площади под кривой распределения вероятностей непрерывной случайной величины. Кроме того, чем больше значение плотности в точке  $X$ , тем больше вероятность того, что случайная величина примет значение в окрестностях  $X$ . Например, если значение плотности распределения вероятностей случайной величины в точке 20 в два раза больше значения плотности распределения вероятностей случайной величины в точке 5, то случайная величина в два раза вероятнее примет значение около 20, чем около 5.

Для непрерывной случайной величины вероятность того, что  $X$  равно какому-либо значению, всегда равна 0. Например, некоторые люди имеют рост от 5,99999 до 6,00001 фута, но ни у одного человека рост не может быть ровно 6 футов. Это объясняет, почему в вероятностных высказываниях можно заменить знак « $<$ » знаком « $\leq$ ».

На рис. 72.1 показана ФПВ для  $X = IQ$  случайно выбранного человека. Площадь под этой кривой равна 1. Если требуется найти вероятность для IQ не больше 90 (0,252), просто найдите площадь слева от значения 90. Если требуется найти вероятность для IQ от 95 до 120 (0,656), найдите площадь под кривой от 95 до 120. Если требуется найти вероятность IQ больше 120 (0,091), найдите площадь под кривой плотности распределения вероятностей справа от 120. (См. лист IQ

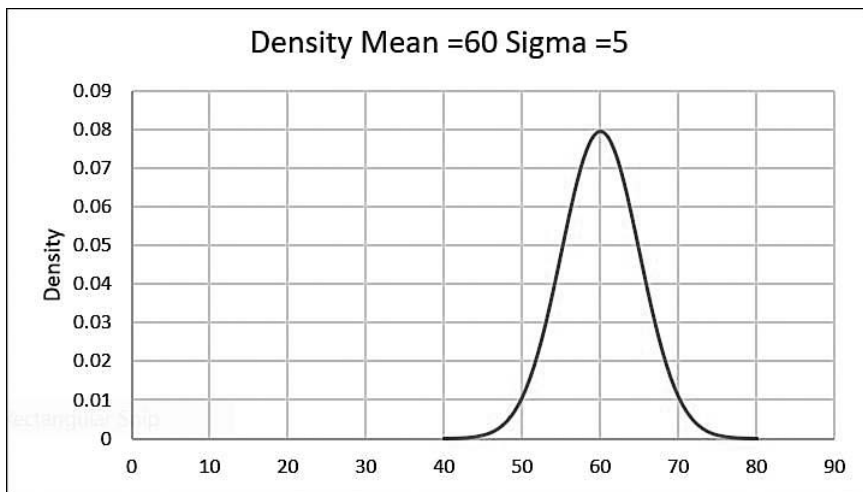


**Рис. 72.1.** Плотность распределения вероятностей для IQ

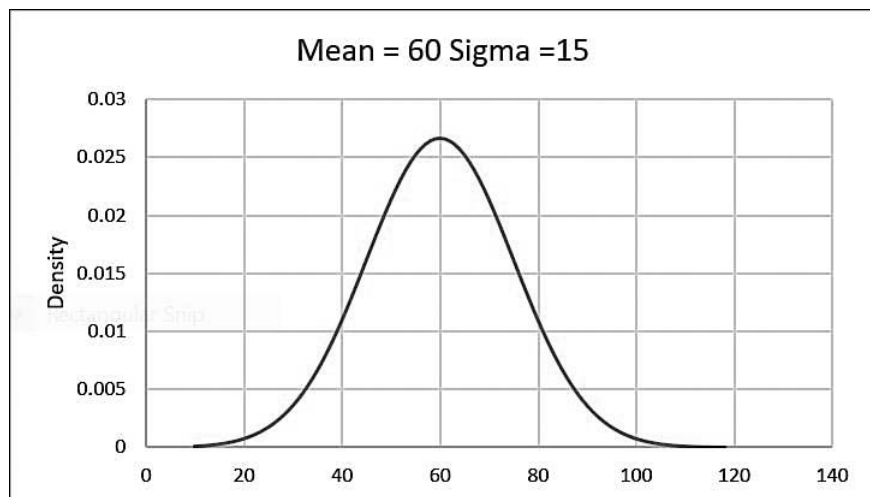
Density в *Normalexamples.xlsx*). В действительности плотность распределения вероятностей на рис. 72.1 — пример распределения *нормальной случайной величины*. Нормальная случайная величина определяется своим средним значением и стандартным отклонением. IQ соответствует нормальной случайной величине с  $\mu = 100$  и  $\sigma = 15$ . Плотность распределения для этой величины отображена на рис. 72.1. Нормальная случайная величина обладает следующими свойствами.

- Наиболее вероятным значением нормальной случайной величины является  $\mu$  (как показывает пик кривой плотности распределения в точке 100 на рис. 72.1).
- По мере удаления значения  $x$  случайной величины от  $\mu$  вероятность того, что случайная величина находится вблизи  $x$ , резко уменьшается.
- Нормальная случайная величина симметрична относительно своего среднего значения. Например, IQ вблизи 80 так же вероятен, как вблизи 120.
- Нормальная случайная величина с вероятностью 68% находится на расстоянии  $\sigma$  от своего среднего значения, с вероятностью 95% — на расстоянии  $2\sigma$  от среднего значения и с вероятностью 99,7% — на расстоянии  $3\sigma$  от среднего значения. Эти меры напоминают эмпирическое правило, описанное в главе 42. На самом деле данное эмпирическое правило основано на предположении, что данные выбраны из нормального распределения. Это объясняет, почему эмпирическое правило не работает так же успешно, если данные не соответствуют симметричной гистограмме.

Для больших значений  $\sigma$  нормальная случайная величина меньше сгруппирована вокруг своего среднего значения. Это положение проиллюстрировано на рис. 72.2 и 72.3. (См. лист *Sigma of 5 and 15* в файле *Normalexamples.xlsx*.)



**Рис. 72.2.** Кривая плотности распределения нормальной случайной величины для  $\mu = 60$  и  $\sigma = 5$



**Рис. 72.3.** Кривая плотности распределения нормальной случайной величины для  $\mu = 60$  и  $\sigma = 15$

### ? Как в Excel найти распределение нормальной случайной величины?

Рассмотрим нормальную случайную величину  $X$  со средним значением  $\mu$  и стандартным отклонением  $\sigma$ . Допустим, для любого числа  $x$  необходимо найти вероятность  $X \leq x$ , которая называется *нормальной интегральной функцией*. Для поиска в Excel вероятности того, что  $X \leq x$ , введите формулу `=НОРМ.РАСП( $x$ ;  $\mu$ ;  $\sigma$ ; 1)` (NORM.DIST). Для четвертого аргумента значение 1 можно заменить значением ИСТИНА.

Если для четвертого аргумента указано значение 1, вычисляется нормальная интегральная функция. Если последний аргумент функции равен 0, вычисляется фактическое значение функции распределения вероятностей нормальной случайной величины. Начиная с Microsoft Excel 2010 статистические функции (такие как НОРМ.РАСП) были или изменены, или полностью переработаны для обеспечения более высокой точности по сравнению с аналогами (например, с функцией НОРМРАСП) в предыдущих версиях Excel.

Для ответов на вопросы, касающиеся нормального распределения, рассмотрим в действии функцию НОРМ.РАСП. Примеры находятся в файле `Normalexamples.xlsx` на листе `Normal`, представленном на рис. 72.4, и в следующих трех сценариях.

- **У какой части населения IQ меньше 90?** Пусть  $X$  равен IQ случайно выбранного человека. Тогда необходимо найти вероятность того, что  $X < 90$ , которая эквивалентна вероятности того, что  $X \leq 90$ . Следовательно, в ячейку C3 на листе `Normal` можно ввести формулу `=НОРМ.РАСП(90;100;15;1)` и получить в результате 0,252. Таким образом, 25,2% всего населения обладают IQ меньше 90.



	B	C	D
2		Prob	
3	IQ<90	0.25249254	=NORM.DIST(90,100,15,1)
4	90<IQ<120	0.65629624	=NORM.DIST(120,100,15,1)-NORM.DIST(90,100,15,1)
5	IQ>120	0.09121122	=1-NORM.DIST(120,100,15,1)
6			
7	99 %ile Prozac	71.6317394	=NORM.INV(0.99,60,5)
8	10% Bloomington income	19747.5875	=NORM.INV(0.1,30000,8000)

**Рис. 72.4.** Вычисление вероятности нормального распределения

- **Какая часть населения имеет IQ от 90 до 120?** При поиске вероятности того, что  $a \leq X \leq b$ , применяется формула (площадь\_под\_кривой\_плотности слева от  $b$ ) – (площадь\_под\_кривой\_плотности слева от  $a$ ). Таким образом, вероятность того, что  $a \leq X \leq b$ , можно найти по формуле  $\text{=НОРМ.РАСП}(b;\mu;s;1)-\text{НОРМ.РАСП}(a;\mu;s;1)$ . Для ответа на вопрос с IQ от 90 до 120 введите формулу  $\text{=НОРМ.РАСП}(120;100;15;1)-\text{НОРМ.РАСП}(90;100;15;1)$  в ячейку C4 на листе Normal. Результат вычисления вероятности – 0,656, то есть 65,6% всего населения имеют IQ от 90 до 120.
- **У какой части населения IQ не менее 120?** При поиске вероятности того, что  $X \geq b$ , обратите внимание, что (вероятность для  $X \geq b$ ) эквивалентна ( $1 -$  вероятность для  $X < b$ ). Для вычисления вероятности того, что  $X \geq b$ , введите формулу  $\text{=}1-\text{НОРМ.РАСП}(b;\mu;\sigma;1)$ . Требуется найти вероятность для  $X \geq 120$ . Это эквивалентно ( $1 -$  вероятность для  $X < 120$ ). Введите формулу  $\text{=}1-\text{НОРМ.РАСП}(120; 100;15;1)$  в ячейку C5 на листе Normal. Вычисленная вероятность равна 0,091, то есть 9,1% всего населения обладает IQ не меньше 120.

### ❓ Как в Excel найти процентиля для нормальных случайных величин?

Рассмотрим заданную нормальную случайную величину  $X$  со средним и стандартным отклонением. Во многих ситуациях требуется получить ответы на вопросы, подобные указанным ниже.

- Фармацевтическая компания считает, что в следующем году спрос на популярный антидепрессант будет соответствовать нормальному закону распределения со средним, равным 60 млн дней лечения, и стандартным отклонением, равным 5 млн дней лечения. Сколько единиц препарата необходимо произвести в этом году, если риск остаться без препарата должен составить только 1%?
- Семейный доход в Блумингтоне, штат Индиана, имеет нормальное распределение со средним значением \$30 000 и стандартным отклонением \$8000. Самые бедные 10% семей от всех семей в Блумингтоне имеют право на федеральную помощь. Какой должна быть максимально допустимая величина дохода для получения помощи?

В первом примере требуется определить 99-й процентиль спроса на антидепрессант. Иначе говоря, необходимо найти такое число  $x$ , при котором спрос превысит

$x$  с вероятностью 1%, и с вероятностью 99% спрос будет меньше  $x$ . Во втором примере требуется найти 10-й перцентиль семейного дохода в Блумингтоне. Значит, необходимо найти такое число  $x$ , при котором семейный доход будет меньше  $x$  с вероятностью 10% и с вероятностью 90% семейный доход превысит  $x$ .

Допустим, требуется найти  $p$ -й перцентиль (выраженный как десятичная дробь) нормальной случайной величины  $X$  со средним значением ( $\mu$ ) и стандартным отклонением ( $\sigma$ ). Просто введите формулу `=НОРМ.ОБР(p;μ;σ)`. Формула возвратит такое число  $x$ , что вероятность для  $X \leq x$  будет равна указанному перцентилю. Теперь можно решить примеры. Вы найдете данные на листе **Normal** в файле **Normalexamples.xlsx**.

В примере с фармацевтической компанией пусть  $X$  равно годовому спросу на препарат. Требуется найти такое значение  $x$ , что вероятность для  $X \geq x$  будет равна 0,01 или вероятность для  $X < x$  равна 0,99. Таким образом, необходимо найти 99-й перцентиль спроса (в млн). Для этого введите формулу `=НОРМ.ОБР(0,99;60;5)` в ячейку C7. Excel возвратит 71,63, то есть компания должна произвести препарата для 71 630 000 дней лечения. Конечно, здесь подразумевается, что компания начинает год без запасов препарата. Если, например, начальные запасы препарата обеспечивают 10 млн дней лечения, то в текущем году необходимо произвести препарата для 61 630 000 дней лечения.

При определении дохода, максимально допустимого для получения федеральной помощи, пусть  $X$  равен семейному доходу в Блумингтоне. Тогда требуется найти такое значение  $x$ , что вероятность для  $X \leq x$  будет равна 0,10, или 10-й перцентиль семейного дохода в Блумингтоне. Найдите это значение по формуле `=НОРМ.ОБР(0,10;30000;8000)`. Вычисления показывают, что на федеральную помощь могут рассчитывать все семьи с доходом ниже \$19 749,59.

### ❓ Почему нормальная случайная величина соответствует многим реальным ситуациям?

Хорошо известная в математике *центральная предельная теорема* (ЦПТ) показывает, что сумма достаточно большого количества (как правило, не меньше 30) независимых случайных величин будет распределена по нормальному закону. Этот результат сохраняется даже в том случае, если отдельные случайные величины не являются нормально распределенными. Во многих случаях оцениваемые величины (например, ошибки измерений) формируются путем сложения большого количества независимых случайных величин. Эта схема объясняет, почему нормальная случайная величина так часто встречается в реальном мире. Вот еще несколько ситуаций, к которым применима ЦПТ.

- Общий объем спроса на пиццу в супермаркете в течение месяца подчиняется закону нормального распределения, даже если ежедневный спрос на пиццу не описывается этим законом.
- Сумма выигрыша, если вы 1000-ный раз играете в кости, распределена по нормальному закону, даже если сумма выигрыша в каждой отдельной игре не подчиняется этому закону.

Другим важным математическим результатом является возможность найти среднее, дисперсию и стандартные отклонения сумм независимых случайных величин. Если сложить вместе независимые случайные величины  $X_1, X_2, \dots, X_n$ , где среднее  $X_i = \mu_i$  и стандартное отклонение  $X_i = \sigma_i$ , то справедливо следующее:

1. Среднее  $(X_1 + X_2 + \dots + X_n) = \mu_1 + \mu_2 + \dots + \mu_n$ .
2. Дисперсия  $(X_1 + X_2 + \dots + X_n) = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2$ .
3. Стандартное отклонение  $(X_1 + X_2 + \dots + X_n) = (\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2)^{0.5}$ .

Следует отметить, что первое уравнение справедливо даже тогда, когда случайные величины не являются независимыми. Опираясь на центральную предельную теорему, можно с помощью уравнений 1–3 решать разнообразные сложные вероятностные задачи, например задачу прогнозирования спроса на пиццу в течение месяца.

В качестве примера применения ЦПТ предположим, что ежедневный спрос на пиццу не распределен по нормальному закону, но имеет среднее значение 45 и стандартное отклонение 12. Предположим, что спрос в различные дни является независимой величиной. Какова вероятность того, что за месяц будет продано не менее 1400 пицц? Если вы заказываете основу для пиццы на месяц в начале срока, сколько основ вы должны заказать, чтобы вероятность нехватки основы составляла 1%? Решение задачи про пиццу находится в файле **Normalexamples.xlsx** на листе **Central Limit**, который показан на рис. 72.5.

	C	D	E
2	Daily frozen pizza demand		
3	mean	45	
4	sigma	12	
5			
6	30 day		
7	mean	1350	=D3*30
8	variance	4320	=30*D4^2
9	sigma	65.7267069	=SQRT(D8)
10			
11	Probability more than 1400 sold	0.22568935	=1-NORM.DIST(1399.5,D7,D9,TRUE)
12	1% chance of running out	1502.90318	=NORM.INV(0.99,D7,D9)

**Рис. 72.5.** Использование центральной предельной теоремы

Даже если ежедневный спрос на замороженную пиццу не распределен по нормальному закону, согласно ЦПТ спрос на пиццу в течение 30 дней описывается нормальным законом распределения. Учитывая это, из формул 1–3 получаем следующее:

- из формулы 1: среднее для спроса в течение 30 дней равно  $30 \cdot 45 = 1350$ ;
- из формулы 2: дисперсия для спроса в течение 30 дней равна  $30 \cdot 12^2 = 4320$ ;
- из формулы 3: стандартное отклонение для спроса в течение 30 дней равно  $\sqrt{4320} = 65,73$ .

Таким образом, спрос на пиццу в течение 30 дней можно моделировать нормальной случайной величиной со средним значением 1350 и стандартным отклонением 65,73. В ячейке D11 по формуле  $=1-\text{НОРМ.РАСП}(1399,5;D7;D9;\text{ИСТИНА})$  вычислена вероятность продажи не менее 1400 пицц как вероятность того, что аппроксимация нормальным распределением составляет не менее 1399,5 (обратите внимание, что спрос 1399,6 округляется до 1400). Вероятность того, что спрос в течение 30 дней составит не менее 1400 пицц, равна 22,6%.

Количество пицц, которое необходимо заготовить для обеспечения 1%-ного риска остаться без пиццы, является 99-м перцентилем распределения спроса. 99-й перцентиль распределения спроса (1503) определен в ячейке D12 по формуле  $=\text{НОРМ.ОБР}(0,99;D7;D9)$ . Таким образом, в начале месяца необходимо запасти 1503 пиццы, чтобы риск остаться без пиццы составлял только 1%.

### ? Что такое Z-оценка?

Для заданного множества данных вам может потребоваться быстро оценить, насколько необычной является некоторая точка данных. Наиболее распространенный способ стандартизировать все точки данных — вычислить *Z-оценку* по следующей формуле:

$(\text{Значение} - \text{Среднее}) / (\text{Стандартное отклонение})$

По сути, Z-оценка — это мера отклонения точки данных от среднего, выраженная в единицах стандартного отклонения. Так как 95% наблюдений нормальной случайной величины находятся в пределах двух стандартных отклонений от среднего значения, любая точка данных с Z-оценкой выше 2 в абсолютном выражении считается выбросом.

Пример вычисления Z-оценок приведен в файле *Superbowlspreads.xlsx* (рис. 72.6), содержащем данные букмекерской конторы в Лас-Вегасе за 1967–2016 гг. с прогнозом числа очков, с которым фаворит выиграет игру, и с теми очками, которые фаворит в реальности получил. Например, в 2016 г. выигрыш «Каролина Пантерс» был оценен в 5 очков, а они проиграли 14 очков, поэтому они отклонились от разницы в счете на  $-14 - 5 = -19$  очков. Разница между разницей в счете и действительным исходом игры называется *остатком*, или ошибкой, для этой игры. Можно вычислить Z-оценки для остатков, которые покажут, насколько для каждой игры результат оказался неожиданным или необычным. Выполните следующие действия:

- Рассчитайте остаток для каждой игры, скопировав формулу  $=E6-D6$  из ячейки C6 в C7:C55.
- В ячейке C4 вычислите среднее остатков  $(-0,39)$  по формуле  $=\text{СРЗНАЧ}(C6:C55)$ .
- В ячейке I4 вычислите стандартное отклонение  $(16,10)$  ошибок по формуле  $=\text{СТАНДОТКЛОН}(C6:C55)$ .

Для вычисления Z-оценки для каждого наблюдения скопируйте из ячейки G6 в G7:G55 формулу  $=(C6-\$C\$4)/\$I\$4$ .

	C	D	E	F	G	H	I
2					Z Score		
3	Mean				Mean	2.89E-17	Stdev
4	-0.39				Sigma	1	16.10434
5	residual or error	Favorite	Result	Year	Z Score		
6	-19	5	-14	2016	-1.155588822		
7	4	0	4	2015	0.272597256		
8	-37.5	2.5	-35	2014	-2.30434719		
9	-7.5	4.5	-3	2013	-0.441495783		
10	-6.5	2.5	-4	2012	-0.379400736		
11	3	3	6	2011	0.210502209		
12	-19	5	-14	2010	-1.155588822		
13	-3	7	4	2009	-0.162068072		
14	-15	12	-3	2008	-0.907208635		
15	5	7	12	2007	0.334692303		
16	7	4	11	2006	0.458882396		
17	-4	7	3	2005	-0.224163119		
18	-4	7	3	2004	-0.224163119		
19	-31	4	-27	2003	-1.900729385		
20	-15	12	-3	2002	-0.907208635		
21	21	3	24	2001	1.328213053		
22	0	7	7	2000	0.024217068		
23	7.5	7.5	15	1999	0.48992992		
24	-18	11	-7	1998	-1.093493775		
25	0	14	14	1997	0.024217068		

**Рис. 72.6.** Z-оценки для исходов игр Суперкубка

Вы можете видеть, что в 2014 г. фаворит «Денвер Бронкос» сыграл на 2,3 стандартного отклонения ниже среднего (самый большой выброс за историю Суперкубка!), а в 1990 г. фаворит «Сан-Франциско, 49» сыграл на 2,07 стандартного отклонения выше среднего. Оба выброса выделены цветом с помощью условного форматирования.

Для любого множества данных среднее Z-оценок будет равно 0, а стандартное отклонение Z-оценок будет равно 1. Как показано в ячейках H3 и H4, среднее Z-оценок равно 0 и стандартное отклонение Z-оценок равно 1.

## Задания

- Допустим, вы можете установить среднее количество унций газированной воды, наливаемое в банку. Фактическое количество унций имеет стандартное отклонение 0,05 унции.
  - Если установлено среднее значение 12,03 унции и банка с газировкой считается годной, если содержит не менее 12 унций напитка, какая часть банок окажется годной?
  - Какая часть банок будет содержать менее 12,1 унции?

- Каким необходимо установить среднее значение, если не более 1% от всех банок должно содержать не более 12 унций? Подсказка: воспользуйтесь инструментом Подбор параметра (Goal Seek).
2. Годовой спрос на лекарственный препарат распределен по нормальному закону со средним 40 000 единиц и стандартным отклонением 10 000 единиц.
    - Какова вероятность того, что годовой спрос составит от 35 000 до 49 000 единиц?
    - Если риск остаться без препарата должен составлять только 5%, каким должен быть объем производства препарата?
  3. Вероятность выиграть в крэпс составляет 0,493. Если сыграть 10 000 игр и ставить одинаковую сумму на каждую игру, какова вероятность оказаться в выигрыше? Начните с определения среднего и стандартного отклонения выигрыша в одной игре. Затем примените центральную предельную теорему.
  4. Еженедельные продажи универсалов Volvo Cross Country распределены по нормальному закону со средним 1000 и стандартным отклонением 250.
    - Какова вероятность того, что в течение недели будет продано от 400 до 1100 универсалов?
    - Какое минимальное число универсалов, проданных в течение недели, соответствует 1%-ной вероятности продажи?
  5. Мое время при заплыве на 100 ярдов соответствует нормальному распределению со средним, равным 51 секунде, и стандартным отклонением в 1,5 секунды. Если я проплыву дистанцию со временем менее 49 секунд, я выиграю заплыв. Какова для меня вероятность выиграть заплыв?
  6. Федеральное правительство собирается ввести доплату к цене медицинской страховки для американцев 65 лет и старше, чей доход попадает в верхние 2% доходов всех американских семей. Если доходы американских семей соответствуют нормальному распределению со средним, равным \$60 000, и стандартным отклонением, равным \$20 000, каково должно быть пороговое значение дохода для назначения доплаты?
  7. Напряжение, поддерживаемое регулятором напряжения, следует нормальному распределению со средним, равным 200 вольт, и стандартным отклонением, равным 5 вольт. Регулятор соответствует спецификации, если он поддерживает напряжение между 185 и 210 вольт. Какая доля регуляторов соответствует спецификации?
  8. Техас собирается выдать чеки на оплату социальных нужд 5% беднейших техасских семей. Если доходы семей в Техасе соответствуют нормальному распределению со средним, равным \$50 000, и стандартным отклонением, равным \$15 000, какой уровень дохода должен быть гарантией назначения социальной помощи?

9. Стержень отвечает спецификации, если его диаметр находится в интервале от 0,98 до 1,02 дюйма. Средний диаметр стержня равен 1 дюйму. Чему должно равняться стандартное отклонение для диаметров стержней, чтобы 99% стержней соответствовали спецификации?
10. В файле **Problem10data.xlsx** приведено количество тачдаунов, выполненных каждым квотербеком в сезоне 2014 г. Определите Z-оценку для каждого квотербека и выделите выбросы с помощью условного форматирования.
11. Количество бутылок чая со льдом, продаваемых в супермаркете по понедельникам, следует нормальному распределению со средним, равным 100, и стандартным отклонением, равным 12. Какова вероятность для некоторого понедельника, что в супермаркете будет продано не менее 105 бутылок чая со льдом? Существует 10%-ная вероятность, что не более \_\_\_\_\_ (заполните пропуск) бутылок чая со льдом будет продано в понедельник.
12. Предположим, ежедневное среднее процентное изменение индекса Доу — Джонса составляет 1%, а стандартное отклонение — 1,5%. Будем считать, что в году 252 операционных дня. Какова вероятность того, что индекс Доу — Джонса вырастет на 20% или более в течение года?
13. В мой кондитерский магазин приходит ежедневно 500 потенциальных покупателей. Семьдесят процентов покупателей пирожных не покупают, 15% покупают одно пирожное, 10% покупают два пирожных и 5% покупают три пирожных. Какова вероятность продать как минимум 280 пирожных в течение дня?

## ГЛАВА 73

# Распределение Вейбулла и бета-распределение: моделирование надежности механизмов и продолжительности работы

### Обсуждаемые вопросы

- Как оценить вероятность безотказной работы станка в течение не менее 20 часов?
- Как оценить вероятность того, что продолжительность работ по монтажу перегородки из гипсокартона в здании составит более 200 часов?

Случайная величина Вейбулла является непрерывной случайной величиной, часто применяемой в моделировании сроков службы механизмов. Если имеются данные о продолжительности работы аналогичных механизмов в прошлом, можно вычислить два параметра (альфа и бета), определяющие случайную величину Вейбулла. Затем с помощью функции Microsoft Excel **ВЕЙБУЛЛ.РАСП (WEIBULL.DIST)** можно определить представляющие интерес вероятности, например оценить продолжительность безотказной работы станка.

Случайная величина с бета-распределением является непрерывной случайной величиной, часто применяемой в моделировании продолжительности деятельности. При заданных оценках минимальной, максимальной и средней продолжительности и стандартного отклонения продолжительности можно с помощью функции **БЕТА.РАСП (BETA.DIST)** определить представляющие интерес вероятности.

## Ответы на вопросы

- ❓ **Как оценить вероятность безотказной работы станка в течение не менее 20 часов?**

Предположим, что были проведены наблюдения за продолжительностью работы семи одинаковых станков. Собранные данные о станках находятся в файле **Weibullest.xlsx** (рис. 73.1).

Инженеры по надежности обнаружили, что для моделирования продолжительности безотказной работы механизмов, как правило, подходит случайная величина



	A	B	C
13	mean	18.6842857	
14	sigma	7.40088476	
15			
16		Machine 1	8.5
17		Machine 2	12.54
18		Machine 3	13.75
19		Machine 4	19.75
20		Machine 5	21.46
21		Machine 6	26.34
22		Machine 7	28.45

**Рис. 73.1.** Данные о продолжительности работы станков

Вейбулла. Случайная величина Вейбулла определяется двумя параметрами: альфа и бета. На основе полученных данных с помощью функций **СРЗНАЧ (AVERAGE)** и **СТАНДОТКЛОН.В (STDEV.S)** в ячейках **B13** и **B14** было определено, что в среднем станки работают без отказа 18,68 часа со стандартным отклонением 7,40 часа. Скопировав эти значения в ячейки **G6** и **G11** и запустив **Поиск решения (Solver)** с параметрами, приведенными на рис. 73.2, можно вычислить альфа и бета так, что среднее значение и стандартное отклонение случайной величины Вейбулла будут соответствовать данным.

## ПРИМЕЧАНИЕ

Для запуска надстройки **Поиск решения (Solver)** ее необходимо активировать. На вкладке **Файл (File)** выберите команду **Параметры (Options)**. В разделе **Надстройки (Add-Ins)** в раскрывающемся списке **Управление (Manage)** выберите **Надстройки Excel (Excel Add-Ins)**, нажмите кнопку **Перейти (Go)**. В диалоговом окне **Надстройки (Add-Ins)** установите флажок **Поиск решения (Solver Add-In)** и нажмите кнопку **ОК**. На вкладке **Данные (Data)** в группе **Анализ (Analysis)** появится кнопка **Поиск решения (Solver)**. Подробнее см. в главе 29 «Введение в оптимизацию с надстройкой **Поиск решения**».

Как видно из рис. 73.3, в этом примере альфа равна 2,725, а бета — 21,003. Любые значения, введенные в ячейки **E2** и **E3** для альфа и бета случайной величины Вейбулла, дают в результате среднее значение (вычисляемое в ячейке **E6**) и стандартное отклонение (вычисляемое в ячейке **E11**). В модели поиска решения альфа и бета изменяются до тех пор, пока среднее значение и стандартное отклонение распределения Вейбулла не станут равны среднему значению и стандартному отклонению продолжительности работы станка, вычисленным на основе собранных данных. Для решения этой задачи методом **ОПГ** я рекомендую использовать несколько начальных точек. Кроме того, метод **ОПГ** с несколькими начальными точками лучше всего работает с верхними и нижними границами для изменяемых ячеек. Для альфа и бета нижняя граница 0,010 работает всегда. Для альфа и бета я опробовал верхнюю границу 1000. Если при поиске решения альфа или бета превышают указанную верхнюю границу, ее необходимо расширить.

**Рис. 73.2.** Параметры поиска решения для определения параметров случайной величины Вейбулла

	A	B	C	D	E	F	G	H
1								
2				alpha	2.725265			
3				beta	21.00372			
4								
5							assumed	Sq Err
6				mean	18.68428	=	18.68429	1.48E-11
7				variance parts	441.1562			
8					0.915493			
9					0.88957			
10				variance	54.77313			
11				sigma	7.400887	=	7.400885	6.07E-12
12								
13	mean	18.6842857			SSE		2.09E-11	
14	sigma	7.40088476						
15								
16		Machine 1	8.5			Prob>=20 hours	0.416832	
17		Machine 2	12.54					
18		Machine 3	13.75			Prob between 15 and 30 hours	0.599417	
19		Machine 4	19.75					
20		Machine 5	21.46					
21		Machine 6	26.34					
22		Machine 7	28.45					
23								

**Рис. 73.3.** Вычисление альфы и беты для случайной величины Вейбулла

Синтаксис функции: ВЕЙБУЛЛ.РАСП: ВЕЙБУЛЛ.РАСП(х;альфа;бета;интегральная).

Если значение аргумента Интегральная равно ИСТИНА, по формуле вычисляется вероятность того, что случайная величина Вейбулла с параметрами альфа и бета

меньше или равна  $x$ . При замене **ИСТИНА** на **ЛОЖЬ** вычисляется высота функции плотности распределения Вейбулла. Вспомните, в главе 69 «Введение в случайные величины» мы узнали, что высота функции плотности вероятности (ФПВ) для любого значения  $x$  непрерывной случайной величины указывает вероятность того, что случайная величина примет значение в окрестности  $x$ . Таким образом, если плотность распределения Вейбулла для 20 часов в два раза больше плотности распределения для 10 часов, то вероятность того, что станок проработает 20 часов до отказа в два раза выше вероятности того, что он проработает 10 часов. Перед тем как ответить на вопросы о представляющих интерес вероятностях, отметим, что в предыдущих версиях Excel те же результаты дает функция **ВЕЙБУЛЛРАСП** (**WEIBULLDIST**).

- **Какова вероятность того, что станок проработает не менее 20 часов?** Эта вероятность (41,7%) вычисляется в ячейке G16 по формуле **=1-ВЕЙБУЛЛ.РАСП(20;альфа;бета;1)**. По существу, по этой формуле вычисляется площадь под ФПВ Вейбулла справа от отметки 20 часов путем вычитания из 1 площади слева от отметки 20 часов.
- **Какова вероятность того, что станок безотказно проработает от 15 до 30 часов?** Эта вероятность (59,9%) вычисляется в ячейке G18 по формуле **=ВЕЙБУЛЛ.РАСП(30;альфа;бета;ИСТИНА)-ВЕЙБУЛЛ.РАСП(15;альфа;бета;ИСТИНА)**. По этой формуле вычисляется площадь под ФПВ Вейбулла между отметками 15 и 30 часов как площадь слева от отметки 30 часов минус площадь слева от отметки 15 часов. После вычитания вероятности безотказной работы станка в течение менее 15 часов из вероятности безотказной работы станка не более 30 часов останется вероятность безотказной работы станка от 15 до 30 часов.

② **Как оценить вероятность того, что продолжительность работ по монтажу перегородки из гипсокартона в здании составит более 200 часов?**

Со времен разработки ракеты *Polaris* в 1950 г. руководители проектов моделируют продолжительность работ с помощью *случайной величины с бета-распределением*. Для вычисления случайной величины с бета-распределением необходимо определить минимальное и максимальное значение, а также два параметра (альфа и бета). Для вычисления параметров бета-распределения можно воспользоваться данными в файле **Beta.xlsx** (рис. 73.4).

Допустим, для монтажа перегородки в здании необходимо от 0 до 600 часов. Это минимальное и максимальное значения, введенные в ячейки C7 и C8. Диапазон ячеек F8:F22 содержит периоды времени, необходимые для монтажа перегородок в 15 одинаковых зданиях. В ячейке F5 я воспользовался функцией **СРЗНАЧ** (**AVERAGE**), чтобы вычислить среднее время (78,49 часа), необходимое для монтажа перегородки в этих 15 зданиях. В ячейке F6 с помощью функции **СТАНДОТКЛОН.Б** (**STDEV**) я определил стандартное отклонение (47,97 часа) для периода времени, необходимого для монтажа перегородки в этих зданиях. Любой выбор значений альфа и бета определяет форму бета-распределения ФПВ, а также среднее значение и стандартное отклонение для соответствующей случайной величины.

	B	C	D	E	F
3	Fit a beta				
4					
5	alpha	2.196262		mean	78.486
6	beta	14.59335		sigma	47.967
7	lower	0			Data
8	upper	600			26.257
9	mean	78.48647			91.840
10	sigma	47.96749			66.543
11	transformed mean	0.130811			53.893
12	transformed var	0.006391			222.437
13					72.054
14					75.357
15	Probability >=200 hours	0.020502			78.274
16	Probability <=80 hours	0.583238			99.044
17	Probability between 30 and 150 hours	0.771058			90.476
18					47.000
19					16.215
20					117.277
21					69.935
22					50.694

**Рис. 73.4.** Определение вероятностей с помощью случайной величины с бета-распределением

Если выбрать значения альфа и бета в соответствии со средним значением и стандартным отклонением времени монтажа перегородки, вычисленными на основе предоставленных данных, то кажется разумным, что эти значения альфа и бета дадут в результате вероятности, согласующиеся с данными наблюдения. Когда вы введете в ячейки C9 и C10 среднее значение и стандартное отклонение для времени монтажа перегородки, на листе будут вычислены значения для альфа (2,20) в ячейке C5 и для бета (14,59) в ячейке C6, которые обеспечат соответствие среднего значения и стандартного отклонения случайной величины с бета-распределением среднему значению и стандартному отклонению данных.

Функция БЕТА.РАСП (БЕТА.DIST) с синтаксисом БЕТА.РАСП(х; альфа; бета; ИСТИНА\_или\_1; lower; upper) определяет вероятность того, что случайная величина с бета-распределением, изменяющаяся от нижней границы до верхней границы, с параметрами альфа и бета принимает значение, меньшее или равное  $x$ . Последние два аргумента являются необязательными, и если они опущены, по умолчанию нижняя граница равна 0, а верхняя — 1. Функция БЕТА.РАСП(х; альфа; бета; ЛОЖЬ\_или\_0; lower; upper) возвращает ФПВ случайной величины с бета-распределением. Теперь с помощью функции БЕТА.РАСП можно вычислить интересующие нас вероятности.

Для вычисления вероятности того, что монтаж перегородки займет не менее 200 часов, введите в ячейку C15 формулу =1-БЕТА.РАСП(200; альфа; бета; ИСТИНА; lower; upper). Вероятность составит 2,1%. По этой формуле вероятность того, что монтаж займет не менее 200 часов, вычисляется как единица минус вероятность того, что монтаж займет не более 200 часов.

Вероятность того, что монтаж перегородки займет не более 80 часов (58,3%), вычисляется в ячейке C16 по формуле `=БЕТА.РАСП(80; альфа; бета; ИСТИНА; lower; upper)`. И для вычисления вероятности того, что на выполнение задания уйдет от 30 до 150 часов (77,1%), введите в ячейку C17 формулу `=БЕТА.РАСП(150;альфа;бета;ИСТИНА; lower; upper)-БЕТА.РАСП(30;альфа;бета;ИСТИНА; lower; upper)`. По этой формуле вероятность того, что монтаж займет от 30 до 150 часов, вычисляется как вероятность того, что монтаж займет не более 150 часов, минус вероятность того, что монтаж займет не более 30 часов. Разность между этими вероятностями учитывает только случаи, когда монтаж занимает от 30 до 150 часов.

## Задания

Данные к заданиям находятся в файле Ch73data.xlsx.

1. На листе **Problem 1** приведены данные о продолжительности безотказной работы станка.
  - Какова вероятность того, что станок проработает не менее 10 часов?
  - Какова вероятность того, что станок проработает от 1 до 5 часов?
  - Какова вероятность того, что станок сломается в течение 6 часов?
2. Вам нужно сегодня прибраться в доме. На листе **Problem 2** приведены данные о продолжительности предыдущих уборок. Если приступить к работе в 12 дня, каковы шансы закончить ее и уйти в 19 вечера в кино?

## ГЛАВА 74

# Создание вероятностных высказываний на основе прогнозов

### Обсуждаемые вопросы

- На основании чего вычисляются прогнозы?
- Фармацевтическая компания планирует продать в следующем году 60 млн единиц препарата. Какова вероятность того, что в следующем году будет продано более 65 млн единиц этого препарата?

Каждый день мы имеем дело с кучей прогнозов. Вот несколько примеров таких прогнозов.

- Правительство прогнозирует в следующем году рост ВВП на 4%.
- Рекламный отдел компании *Eli Lilly* прогнозирует, что в следующем году спрос на лекарственный препарат составит 400 000 000 дней лечения.
- Аналитики Уолл-стрит прогнозируют, что индекс Доу — Джонса поднимется на 20% в течение следующих 12 месяцев.
- Букмекеры предсказывают победу «Индиана Пэйсерс» (Indiana Pacers) над «Хьюстон Рокетс» (Houston Rockets) с преимуществом в 6 очков в первом матче сезона НБА 2015 г.

Даже если вам предоставили наилучшие прогнозы из возможных, почти наверняка они будут неправильными. Например, прогноз букмекеров о победе «Индиана Пэйсерс» с разрывом в 6 очков будет ошибочным, если только «Пэйсерс» не выиграет точно с таким разрывом. То есть любое единичное значение (или точечный прогноз) подразумевает распределение вероятностей для прогнозируемой величины. Как найти случайную величину, правильно моделирующую неопределенность, присущую точечному прогнозу? Ключом к распределению вероятностей для точечного прогноза должны быть данные о точности последних прогнозов по интересующей вас величине. Например, для индекса Доу — Джонса имеются прогнозы процентного изменения индекса, сделанные в январе каждого года за последние 10 лет, и фактические изменения индекса для каждого из этих 10 лет.

## Ответы на вопросы

### ? На основании чего вычисляются прогнозы?

Сначала следует проверить, не выявляют ли последние прогнозы какую-либо систематическую ошибку. Для каждого из последних прогнозов определите отношение фактического значения к прогнозируемому значению. Затем усредните эти соотношения. Если прогнозы объективны, среднее значение будет равно приблизительно 1. Любое значительное отклонение от 1 указывает на систематическую ошибку. Например, если среднее значение отношения факта к прогнозу равно 2, фактические результаты, как правило, в два раза превосходят прогноз. Для исправления систематической ошибки необходимо автоматически удвоить прогнозируемые значения. Если среднее значение отношения факта к прогнозу равно 0,5, фактические результаты составляют примерно половину от прогнозируемых. Таким образом, для устранения систематической ошибки необходимо прогнозируемые значения уменьшить наполовину. После устранения систематической ошибки изучите стандартное отклонение процентных ошибок для прогноза без систематической ошибки и смоделируйте прогнозируемую величину с помощью следующей нормальной случайной величины:

- среднее = прогноз без систематической ошибки;
- стандартное отклонение = (прогноз без систематической ошибки) × (стандартное отклонение процентных ошибок, связанных с прогнозом без систематической ошибки).

Критерием точности прошлых прогнозов является стандартное отклонение процентных ошибок. Очевидно, что меньшее стандартное отклонение процентных ошибок прогноза предпочтительнее большего стандартного отклонения.

В файле *Drugfore.xlsx* (рис. 74.1) проиллюстрированы понятия систематической ошибки прогноза и точности прогноза. Файл содержит данные о фактических и прогнозируемых объемах продаж лекарственного препарата (в миллионах дней лечения) с 2005 по 2012 г. Сначала я показываю прогнозы с систематической ошибкой, а затем удаляю систематическую ошибку.

Прежде всего, в ячейках F5:F12 я проверил, есть ли систематическая ошибка, вычислив для каждого года соотношение факта и прогноза. Для этого я скопировал формулу  $=D5/E5$  из F5 в F6:F12.

Затем в ячейке F2 я вычислил систематическую ошибку первоначальных прогнозов, усреднив отношение фактических значений к прогнозируемым для каждого года по формуле  $=CРЗНАЧ(F5:F12)$ . Наблюдается тенденция отставания фактических продаж от прогнозируемых на 8%. Для корректировки прошлых прогнозов с систематической ошибкой я умножил их на 0,92, скопировав формулу  $=F\$2*E5$  из G5 в G6:G12. В ячейках H5:H12 я вычислил для каждого года процентную ошибку прогноза без систематической ошибки, скопировав формулу  $=(D5-G5)/G5$  из H5 в H6:H12. Как видите, в 2005 г. фактические продажи были на 16% меньше

	D	E	F	G	H	I
1					mean	std dev
2		mean	0.91803		5E-17	0.11375
3						
4	Actual Sales	Forecast	A/F	Unbiased forecast	%age error	
5	17	22	0.77273	20.19668	-16%	
6	59	61	0.96721	55.9999	5%	
7	46	51	0.90196	46.81959	-2%	
8	85	86	0.98837	78.95067	8%	
9	98	103	0.95146	94.5572	4%	
10	94	118	0.79661	108.3277	-13%	
11	24	22	1.09091	20.19668	19%	
12	14	16	0.875	14.6885	-5%	
13						
14						
15		Mean 2013	55.0819			
16		Sigma 2013	6.2657			
17						
18	Probability	0.05671928				
19	we sell >65 million units	1-NORM.DIST(65,F15,F16,TRUE)				
20						

**Рис. 74.1.** Корректировка прогнозов с систематической ошибкой и измерение точности прогнозов

прогнозируемых без систематической ошибки, а в 2011 г. фактические продажи превзошли прогнозируемые без систематической ошибки на 19%.

Теперь оценим точность прогноза по ошибкам прогнозов в ячейке I2 путем вычисления стандартного отклонения процентных ошибок по формуле =СТАНДОТКЛОН.В(Н5:Н12). Стандартное отклонение для прошлых прогнозов без систематической ошибки составило примерно 11,4%.

**?** Фармацевтическая компания планирует продать в следующем году 60 млн единиц препарата. Какова вероятность того, что в следующем году будет продано более 65 млн единиц этого препарата?

Ошибки прогнозов часто следуют закону распределения нормальной случайной величины. Значит, после корректировки систематической ошибки можно предположить, что ошибки прогнозов соответствуют нормальной случайной величине с параметрами:

- среднее = прогноз без систематической ошибки;
- стандартное отклонение = (прогноз без систематической ошибки) × (стандартное отклонение процентных ошибок, связанных с прогнозом без систематической ошибки).

В моем примере это означает, что нужно смоделировать объемы продаж лекарственного препарата в следующем году как нормальную случайную величину со



средним значением, равным  $60 \times (0,918) = 55,08$  млн, и стандартным отклонением, равным  $(0,11375) \times 55,08 = 6,27$  млн. Теперь с помощью функции НОРМ.РАСП (NORM.DIST) вычислите в E18 (рис. 74.1) вероятность продажи не менее 65 млн единиц в следующем году по формуле  $=1-\text{НОРМ.РАСП}(65;F15;F16;\text{ИСТИНА})$ . Шанс продать более 65 млн единиц лекарственного препарата составляет 5,7%.

## Задания

1. В седьмой игре финала НБА в 2010 г. между «Бостон Селтикс» и «Лос-Анджелес Лейкерс» команда «Лейкерс» была фаворитом с преимуществом в 7 очков. Последние игры НБА показывают, что прогнозы букмекеров не содержат систематической ошибки. Кроме того, стандартное отклонение букмекерских прогнозов в последних играх НБА составляет 12 очков. Как бы вы оценили вероятность выигрыша «Лейкерс»?
2. Ведущая компания — производитель смартфонов предполагает продать в следующем году 3 млн единиц своей лидирующей модели. В прошлом фактические объемы продаж не содержали систематической ошибки. Ошибки прогнозов имели стандартное отклонение 10% от прогнозируемых объемов продаж. Какова вероятность продажи менее 2,5 млн смартфонов в следующем году?
3. В файле USC.xlsx для футбольной команды USC содержатся фактические показатели побед для 2005–2009 гг. и прогноз букмекеров (до игры). Указывают ли эти данные на отсутствие систематической ошибки в прогнозах букмекеров? Если букмекеры прогнозируют, что команда USC победит с разрывом в 10 очков, какова вероятность победы команды USC?

## ГЛАВА 75

# Логарифмически нормальная случайная величина в моделировании курса акций

### Обсуждаемые вопросы

- Что такое логарифмически нормальная случайная величина?
- По какой причине курс акций ведет себя подобно случайной величине с логарифмически нормальным распределением?
- Как смоделировать будущий курс любых акций как логарифмически нормальную случайную величину?
- Как вычислить вероятность того, что курс акций Microsoft превысит \$130 через шесть месяцев?
- Как вычислить вероятность того, что курс акций Microsoft не превысит \$90 через шесть месяцев?
- Как вычислить средний курс акций Microsoft через шесть месяцев?

Многие люди заинтересованы в моделировании будущего курса акций, цен на товары, например на нефть или на пшеницу, или будущего валютного курса. За последние 40 лет для моделирования курсов акций наиболее часто использовалась логарифмически нормальная случайная величина. Из этой главы вы узнаете, почему логарифмически нормальная случайная величина является приемлемой моделью для курсов акций и как определить соответствующие параметры логарифмически нормального распределения для любых акций. В конце главы показано, как использовать функции ЛОГНОРМ.РАСП (LOGNORM.DIST) и ЛОГНОРМ.ОБР (LOGNORM.INV) для вычисления вероятности будущего курса акций.

## Ответы на вопросы

### ❓ Что такое логарифмически нормальная случайная величина?

Случайная величина  $Y$  имеет логарифмически нормальное распределение, если  $\ln Y$  имеет нормальное распределение. При использовании функций Excel

ЛОГНОРМ.РАСП и ЛОГНОРМ.ОБР логарифмически нормальная случайная величина  $Y$  характеризуется двумя параметрами: средним *мю*, равным математическому ожиданию  $\ln Y$ , и стандартным отклонением *сигма*, равным стандартному отклонению  $\ln Y$ .

❓ **По какой причине курс акций ведет себя подобно случайной величине с логарифмически нормальным распределением?**

Пусть  $Y$  — курс акции через  $n$  дней, начиная с сегодняшнего дня,  $P$  — курс акции сегодня и  $X_t$  — относительное изменение курса акций за день  $t$ . Тогда  $Y = P \times X_1 \times X_2 \times \dots \times X_n$ . Поскольку логарифм произведения равен сумме логарифмов, справедливо, что  $\ln Y = \ln P + \ln X_1 + \ln X_2 + \dots + \ln X_n$ . Допустим, изменения курса в разные дни являются независимыми случайными величинами. В главе 72 «Нормальная случайная величина и Z-оценка» была представлена центральная предельная теорема, утверждающая, что сумма большого количества независимых величин является нормальной случайной величиной, даже если отдельные случайные величины не распределены по нормальному закону. Это объясняет, почему  $\ln Y$ , скорее всего, является нормальной случайной величиной, из чего следует, что  $Y$  является логарифмически нормальной величиной.

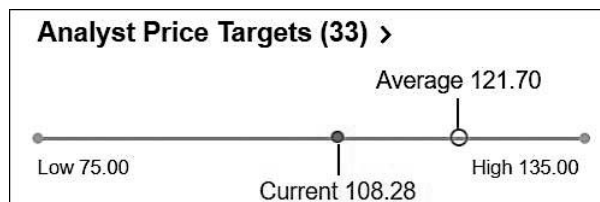
❓ **Как смоделировать будущий курс любых акций как логарифмически нормальную случайную величину?**

При закрытии торгов 14 августа 2018 г цена акций Microsoft составляла \$108,53. Мы хотим смоделировать курс акций Microsoft через шесть месяцев как логарифмически нормальную случайную величину. Как вычислить *мю* и *сигму*? К счастью, в интернете есть все, что нужно для этого. На сайте [IVolatility.com](http://IVolatility.com) (рис. 75.1) предоставляется оценка годовой волатильности акции (назовем ее *сигма*). Годовая волатильность является мерой рискованности или изменчивости акции. Как указано на сайте [IVolatility.com](http://IVolatility.com), при закрытии торгов 14 августа 2018 г. подразумеваемая волатильность акций Microsoft составляла 15,98%. Представим *сигма* в виде десятичной дроби: *сигма* = 0,1598. (Метод вычисления подразумеваемой волатильности рассматривается в главе 82 «Ценообразование опционов».)

Current		1 WK AGO	1 MO AGO	52 wk Hi/Date	52 wk Low/Date
HISTORICAL VOLATILITY ▢					
10 days	10.46%	23.67%	14.46%	57.23% - 04-Apr	4.73% - 19-Oct
20 days	19.68%	20.65%	17.72%	42.80% - 18-Apr	8.02% - 24-Oct
30 days	18.20%	18.85%	16.66%	37.82% - 02-May	9.94% - 26-Oct
IMPLIED VOLATILITY ▢					
IV Index call ▢	16.27%	16.15%	21.93%	37.81% - 02-Apr	13.54% - 27-Oct
IV Index put ▢	15.69%	15.80%	21.68%	35.08% - 02-Apr	13.48% - 27-Nov
IV Index mean ▢	15.98%	15.98%	21.81%	36.44% - 02-Apr	13.55% - 27-Oct

**Рис. 75.1.** Волатильность акций Microsoft на 14 августа 2018 г.

В разделе Investment сайта [finance.yahoo.com](http://finance.yahoo.com) можно найти единодушное мнение аналитиков о цене акций Microsoft через год (рис. 75.2).



**Рис. 75.2.** Аналитический прогноз для курса акций Microsoft на 14 августа 2018 г.

Пусть  $S_0$  — сегодняшний курс акций, а  $S_1$  — рассчитанный средний курс акций через год. Тогда  $\mu_0$ , мера годового среднего ожидаемого дохода по акциям, может быть вычислена как  $\mu_0 = \ln(S_1/S_0)$ .

По данным сайта [yahoo.com](http://yahoo.com) известно, что на закрытие торгов 14 августа 2018 г.  $S_0 = \$108,53$ , а  $S_1 = \$121,70$ . Следовательно (как вычислено в ячейке F7 файла Lognormal.xlsx; рис. 75.3),  $\mu_0 = \ln(121,70/108,53) = 0,114532$ .

	E	F	G
1			
2	<b>Microsoft August 14, 2018</b>		
3			
4	Price today	\$ 108.53	
5	Forecast in a year	\$ 121.70	
6	sigma_	0.1598	
7	Mu_	0.114532 =LN(Forecast_in_a_year/Price_today)	
8	time(years)	0.5	
9	Mu for t	4.737909 =LN(Price_today)+(Mu_-0.5*sigma_^2)*time_years	
10	Sigma for t	0.112996 =sigma_*SQRT(time_years)	
11			
12	In 6 months		
13	Chance price >=\$130	0.125655	=1-LOGNORM.DIST(130,Mu_for_t,Sigma_for_t,TRUE)
14	Chance price <=\$90	0.017552	=LOGNORM.DIST(90,Mu_for_t,Sigma_for_t,TRUE)
15	Median Price	114.1951	=LOGNORM.INV(0.5,Mu_for_t,Sigma_for_t)

**Рис. 75.3.** Моделирование курса акций Microsoft через шесть месяцев с помощью логарифмически нормального распределения

Если требуется смоделировать курс акций Microsoft через  $t$  лет, необходимо знать  $\mu_0$ , равное среднему значению для  $\ln$  от курса акций Microsoft через  $t$  лет, и  $\sigma_0$ , равное стандартному отклонению для  $\ln$  от курса акций Microsoft через  $t$  лет. Оказывается, что:

$$\mu_t = \ln S_0 + t(\mu_0 - 0,5\sigma_0^2) \text{ и } \sigma_t = \sigma_0 \times \sqrt{t}.$$

Для шести месяцев  $t = 0,5$  года. В ячейке E9 я вычислил  $\mu_{0,5}$  по формуле  $\ln(108,53) + 0,5*(0,114532) - 0,5*(0,15988^2) = 4,737909$ . В ячейке E10 вычисляется  $\sigma_{0,5} = 0,1598 \times \sqrt{0,5} = 0,112996$ .

Таким образом, можно смоделировать курс акций Microsoft через шесть месяцев как логарифмически нормальную случайную величину с  $\mu = 4,737909$  и  $\sigma = 0,112996$ .

**❓ Как вычислить вероятность того, что курс акций Microsoft превысит \$130 через шесть месяцев?**

Функция Excel ЛОГНОРМ.РАСП (LOGNORM.DIST) с синтаксисом ЛОГНОРМ.РАСП( $x$ ; среднее; стандартное\_отклонение; ИСТИНА\_или\_1) вычисляет вероятность того, что логарифмически нормальная случайная величина с параметрами  $\mu$  и  $\sigma$  меньше или равна  $x$ . При замене ИСТИНА (или 1) на ЛОЖЬ (или 0) функция возвращает высоту кривой распределения логарифмически нормальной случайной величины в точке  $x$ .

В версиях Excel до Excel 2013 те же результаты возвращала функция ЛОГНОРМ-РАСП (LOGNORMDIST). В ячейке F13 определяется вероятность того, что курс акций Microsoft превысит \$130. Она вычисляется как  $1 - (\text{вероятность для курса акции} \leq \$130)$ , поэтому я ввел формулу  $=1-\text{ЛОГНОРМ.РАСП}(130; \mu_{\text{for } t}; \sigma_{\text{for } t}; \text{ИСТИНА})$ . (В файле Lognormal.xlsx я использовал в формулах имена диапазонов вместо указаний ячеек, и имя  $\mu_{\text{for } t}$  соответствует ячейке F9, а  $\sigma_{\text{for } t}$  — ячейке F10.) Вероятность того, что через шесть месяцев акции Microsoft будут продаваться дороже \$130 за акцию, составляет 12,6%.

**❓ Как вычислить вероятность того, что курс акций Microsoft не превысит \$90 через шесть месяцев?**

В ячейке F14 по формуле  $=\text{ЛОГНОРМ.РАСП}(90; \mu_{\text{for } t}; \sigma_{\text{for } t}; \text{ИСТИНА})$  я определил вероятность того, что через шесть месяцев акции Microsoft будут продаваться не дороже \$90 за штуку. Вероятность того, что через шесть месяцев акции Microsoft будут продаваться не дороже \$90, составляет 1,8%.

**❓ Как вычислить средний курс акций Microsoft через шесть месяцев?**

Пусть требуется определить среднюю цену акций Microsoft через шесть месяцев. Это цена  $x$  в долларах, такая, что с вероятностью 50% через шесть месяцев цена акций Microsoft составит не более  $x$  долларов. В ячейке F15 я нашел среднюю цену акций через шесть месяцев по формуле  $=\text{ЛОГНОРМ.ОБР}(0,5; \mu_{\text{for } t}; \sigma_{\text{for } t})$ . Средняя цена через шесть месяцев составит \$114,20.

## Замечания

- В главе 82 рассматривается знаменитая формула ценообразования Блэка — Шоулза. Эта модель предполагает, что цена акций действительно имеет логарифмически нормальное распределение.
- В последние годы стало очевидно, что для многих акций курс принимает необычно высокие или низкие значения чаще, чем предсказывает логарифмиче-

ски нормальная случайная величина. Это побудило аналитиков искать другие случайные величины с «тяжелыми хвостами» распределения, с помощью которых можно было бы моделировать будущий курс акций. Для получения дополнительной информации по этой теме см. книгу Нассима Николаса Талеба «Черный лебедь» (Nicholas Taleb. *The Black Swan*. — New York: Random House, 2010). Кроме того, альтернативный метод моделирования курса акций, не подразумевающий для него логарифмически нормального распределения, рассматривается в главе 79 «Моделирование цен на акции и распределения средств между активами».

## Задания

Пусть сегодня акции продаются по \$30 за штуку. Годовая волатильность акции составляет 35%, и аналитики предсказывают, что ожидаемая цена акции через год составит \$35. Исходя из этой информации, ответьте на следующие вопросы:

1. Какова вероятность того, что через два года акции будут продаваться по цене выше \$45?
2. Какова вероятность того, что через два года акции будут продаваться по цене ниже \$25?
3. В каком диапазоне будет находиться цена акций через два года с вероятностью 95%? Подсказка: для нижних и верхних границ укажите в функции ЛОГНОРМ. ОБР в качестве первого аргумента 0,025 и 0,975 соответственно.

# Импорт в Excel истории торгов (загрузка биржевых данных)

### Обсуждаемые вопросы

- Как просто импортировать ретроспективные биржевые данные в Excel?
- Как рассчитать годовую доходность по моему портфелю акций?

## Ответы на вопросы

### ❓ Как просто импортировать ретроспективные биржевые данные в Excel?

В настоящее время более половины взрослых американцев вложили свои деньги в какие-нибудь акции. Чтобы помочь таким инвесторам отслеживать рост их портфеля, Самир Хан создал замечательный (и бесплатный) инструмент (вы можете найти его по адресу <http://investexcel.net/multiple-stock-quote-downloader-for-excel/>), с помощью которого можно с легкостью загружать в Excel ретроспективные сведения по самым различным акциям и/или биржевым индексам. Чтобы скачать этот замечательный инструмент, вам нужно просто перейти по ссылке С. Хана, найти внизу страницы блог и щелкнуть **Получить лист Excel для загрузки ретроспективных биржевых данных с Yahoo (Get Excel Spreadsheet To Download Bulk Historical Stock Data From Yahoo)**. Скачав zip-архив file, извлеките из него все файлы (Extract All). Файл **Multiple Stock Quote Downloader.xlsm** — это то, что вам нужно. Там уже находятся все необходимые ссылки и поля. Предположим, вы хотите загрузить ежедневную информацию за период с 20 августа 2017 г. по 20 августа 2018 г. по компаниям IBM, Southwest Airlines и индекс S&P. Их биржевыми условными обозначениями (тикерами) являются, соответственно, IBM, LUV и VFINX. Чтобы получить инвесторскую информацию по этим трем компаниям, заполняем первый лист в книге **IBM,SW,SandP.xlsm**, как показано на рис. 76.1. Щелкнув **Получить все котировки (Get Bulk Quotes)**, мы получаем ретроспективные показатели по каждому виду акций из выбранных: цена на момент открытия биржи, или начальная, Open Price, максимальная цена за день, High Price, минимальная цена за день, Low Price, цена закрытия, или последняя, Close Price, скорректированная цена закрытия, Adjusted Close Price, и объем сделок, Trading Volume.

	A	B	C	D	E
1	<b>Bulk Stock Quote Downloader</b>				
2	<a href="http://investexcel.net">http://investexcel.net</a>				
3					
4	<b>Parameters</b>				
5	Start Date	2017-08-20	<div>Get Bulk Quotes</div>		
6	End Date	2018-08-20			
7	Frequency	d			
8	<input type="checkbox"/> Write to CSV	c:\temp\			
9	<input checked="" type="checkbox"/> Collate Data				
10	Oldest First	▼			
11					
12	<b>Enter Tickers Below</b>		Enter a list of stock tickers and click the button to automatically retrieve historical stock quotes from Yahoo Finance		
13	IBM				
14	LUV				
15	VFINX				

**Рис. 76.1.** Настройки для загрузки данных по акциям IBM и Southwest Airlines и индексу S&P

	A	B	C	D
1	Date	IBM	LUV	VFINX
2	2017-08-21	134.6216	53.02791	220.8159
3	2017-08-22	135.274	53.0428	223.0258
4	2017-08-23	136.358	51.34519	222.2695
5	2017-08-24	137.1255	50.34251	221.8177
6	2017-08-25	137.8929	52.29824	222.2106
7	2017-08-28	136.7129	51.37497	222.3186
8	2017-08-29	137.3173	51.08707	222.5445
9	2017-08-30	136.7609	51.14664	223.6249
10	2017-08-31	137.2118	51.76215	224.9116
11	2017-09-01	138.2191	51.81178	225.383
12	2017-09-05	137.2214	50.76939	223.6839
13	2017-09-06	137.9697	51.29555	224.391
14	2017-09-07	137.0871	51.83164	224.391
15	2017-09-08	136.6554	52.48686	224.0767
16	2017-09-11	138.9673	53.21156	226.4929

**Рис. 76.2.** Данные о цене закрытия на акции

Мы выбрали временной интервал с 20 августа 2017 г. по 20 августа 2018 г. Выбрав частоту **d**, мы получим ежедневные сведения (для ежемесячных нужно выбрать **m**). Параметр **Сначала давние (Oldest First)** задает хронологический порядок расположения данных. Параметр **Сопоставить данные (Collate Data)** приведет к объединению показателей из выбранного вами временного интервала на одном листе. Итоговый вид скачанных нами данных представлен в файле **IBM,SW,SandP.xlsx**. В этой



книге на первом листе перечислены наши показатели. Далее следуют листы, объединяющие все типы показателей по избранным нами акциям, и листы, напротив, посвященные каждому отдельному показателю, содержащие все сведения по акциям, касающиеся этого показателя: цена открытия, цена закрытия, объем сделок и т. п. Например, как можно видеть на рис. 76.2, лист Adjusted Close Price содержит ежедневные цены закрытия по акциям каждой из компаний.

### ? Как рассчитать годовую доходность по моему портфелю акций?

Загрузив данные по скорректированной цене закрытия для акций, находящихся в вашем портфеле, вы без труда создадите лист, вычисляющий его годовую доходность. Чтобы показать как, мы скопировали данные по скорректированной цене закрытия в новую книгу (PortfolioIRR.xlsx). С проделанной работой можно ознакомиться на рис. 76.3.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Date	IBM	LUV	VFINX	IBM Return	LUV Return	Sand P Return	IBM Value	LUV Value	S and P Value	Final Value	Days	IRR
2	2017-08-21	135	53	220.8				\$100.00	\$200.00	\$500.00	\$800.00		
3	2017-08-22	135	53	223	0.005	0.0003	0.01	\$100.48	\$200.06	\$505.00	\$805.54	1	1144.18%
4	2017-08-23	136	51.3	222.3	0.008	-0.032	-0.003	\$101.29	\$193.65	\$503.29	\$798.23	2	-33.18%
5	2017-08-24	137	50.3	221.8	0.006	-0.02	-0.002	\$101.86	\$189.87	\$502.27	\$794.00	3	-59.99%
6	2017-08-25	138	52.3	222.2	0.006	0.0388	0.0018	\$102.43	\$197.25	\$503.16	\$802.84	4	38.11%
7	2017-08-28	137	51.4	222.3	-0.009	-0.018	0.0005	\$101.55	\$193.77	\$503.40	\$798.72	7	-8.00%
8	2017-08-29	137	51.1	222.5	0.004	-0.006	0.001	\$102.00	\$192.68	\$503.91	\$798.60	8	-7.70%
9	2017-08-30	137	51.1	223.6	-0.004	0.0012	0.0049	\$101.59	\$192.90	\$506.36	\$800.85	9	4.42%
10	2017-08-31	137	51.8	224.9	0.003	0.012	0.0058	\$101.92	\$195.23	\$509.27	\$806.42	10	33.90%
11	2017-09-01	138	51.8	225.4	0.007	0.001	0.0021	\$102.67	\$195.41	\$510.34	\$808.43	11	41.58%
12	2017-09-05	137	50.8	223.7	-0.007	-0.02	-0.008	\$101.93	\$191.48	\$506.49	\$799.91	15	-0.28%
13	2017-09-06	138	51.3	224.4	0.005	0.0104	0.0032	\$102.49	\$193.47	\$508.10	\$804.05	16	12.20%
14	2017-09-07	137	51.8	224.4	-0.006	0.0105	0	\$101.83	\$195.49	\$508.10	\$805.41	17	15.59%
15	2017-09-08	137	52.5	224.1	-0.003	0.0126	-0.001	\$101.51	\$197.96	\$507.38	\$806.85	18	18.88%
16	2017-09-11	139	53.2	226.5	0.017	0.0138	0.0108	\$103.23	\$200.69	\$512.85	\$816.78	21	43.43%
17	2017-09-12	140	53.6	227.3	0.006	0.0065	0.0034	\$103.87	\$202.00	\$514.61	\$820.48	22	52.12%
18	2017-09-13	140	54.3	227.4	0.002	0.0132	0.0008	\$104.03	\$204.66	\$515.01	\$823.71	23	58.95%
19	2017-09-14	140	54	227.3	-0.003	-0.006	-8E-04	\$103.71	\$203.50	\$514.61	\$821.83	24	50.58%
20	2017-09-15	139	54.3	227.7	-0.005	0.0068	0.0019	\$103.20	\$204.89	\$515.61	\$823.70	25	53.14%
21	2017-09-18	139	53.7	228.1	-0.002	-0.012	0.0015	\$103.01	\$202.53	\$516.39	\$821.93	28	42.26%
22	2017-09-19	139	53.2	228.3	-0.001	-0.01	0.0011	\$102.89	\$200.47	\$516.97	\$820.33	29	37.14%
23	2017-09-20	140	53.9	228.4	0.01	0.0133	0.0006	\$103.95	\$203.13	\$517.28	\$824.36	30	44.05%
24	2017-09-21	139	54.2	227.8	-0.004	0.007	-0.003	\$103.51	\$204.55	\$515.74	\$823.80	31	41.23%

**Рис. 76.3.** Расчет ежедневной внутренней ставки доходности от инвестиций (ВНД/IRR) для портфеля акций

Предположим, мы начинаем анализ в конце дня 21 августа 2017 г, и у нас вложено \$100 в IBM, \$200 в Southwest Airlines и \$500 в S&P. Покажем, как рассчитать годовую доходность от этого портфеля на конец каждого дня торгов. Пусть R = годовой коэффициент доходности для портфеля, а D = число дней, в течение которых мы осуществляем инвестирование.

$$(\text{Начальная стоимость}) \cdot (1+R)^{(D/365)} = \text{Конечная стоимость.}$$

Решив это уравнение для R, получаем:

$$R = [(Конечная\ стоимость / (Начальная\ стоимость))^{(365/D)} - 1].$$

Чтобы рассчитать годовой коэффициент доходности на каждый день для нашего портфеля, мы должны выполнить следующие шаги:

- Скопировав формулу  $= (B3 - B2) / B2$  из E3 в E3:G252, мы рассчитаем ежедневную доходность наших инвестиций.
- Скопировав из H3 в H3:J252 формулу  $= H2 * (1 + E3)$ , мы рассчитаем стоимость инвестиции в IBM на конец дня (аналогично рассчитывается стоимость для LUV и S&P).
- Скопировав из K3 в K4:K252 формулу  $= СУММ(H3:J3)$ , мы рассчитаем стоимость портфеля на конец дня.
- Скопировав из L3 в L4:L252 формулу  $= A3 - \$A\$2$ , мы рассчитаем количество дней, в течение которых мы осуществляем инвестирование по этому портфелю.
- Наконец, скопировав формулу  $= (K3 / \$K\$2)^{(365/L3)} - 1$  из M3 в M4:M253, мы получим годовую внутреннюю ставку доходности, IRR, нашего портфеля на текущую дату. Например, на конец дня 17 августа 2018 г. текущая ВНД равна 16,57%.

## Задания

1. Загрузите ежедневные биржевые данные для компаний WMT, Facebook и CVS за период со 2 января 2015 г. по 7 декабря 2017 г.
2. Рассчитайте годовой коэффициент доходности для портфеля, в который всегда входит по одной акции этих компаний.

## ГЛАВА 77

# Введение в моделирование по методу Монте-Карло

### Обсуждаемые вопросы

- Где применяется моделирование по методу Монте-Карло?
- Что происходит при вводе в ячейку формулы =СЛЧИС()?
- Как моделируют значения дискретной случайной величины?
- Как моделируют значения нормальной случайной величины?
- Как компания — производитель поздравительных открыток может определить, сколько открыток следует произвести?

Аналитикам данных часто требуется точно оценить вероятность неопределенных событий. Например, какова вероятность того, что денежные потоки для нового продукта будут демонстрировать положительную чистую приведенную стоимость (ЧПС)? Каков фактор риска инвестиционного портфеля? Метод Монте-Карло позволяет моделировать ситуации, демонстрирующие неопределенность, и затем воспроизводить их на компьютере тысячи раз.

### ПРИМЕЧАНИЕ

---

Термин «моделирование по методу Монте-Карло» пришел из моделирования, выполнявшегося на вычислительных машинах в 1930–1940-х гг. для оценки вероятности того, что цепная реакция, необходимая для взрыва атомной бомбы, произойдет. Физики, принимавшие участие в проекте, были большими поклонниками азартных игр, поэтому метод моделирования и получил кодовое название «Монте-Карло».

---

В следующих пяти главах я приведу примеры выполнения в Microsoft Excel моделирования по методу Монте-Карло.

## Ответы на вопросы

### ❓ Где применяется моделирование по методу Монте-Карло?

Во многих компаниях моделирование по методу Монте-Карло давно стало важной частью процесса принятия решений. Вот несколько примеров.

- Компании General Motors (GM), Procter&Gamble, Pfizer, Bristol-Myers Squibb и Eli Lilly с помощью моделирования оценивают и средний доход, и фактор риска для новой продукции. В этих компаниях информация используется для принятия решения о том, какой продукт выпустить на рынок.
- Компания GM использует моделирование для таких действий, как прогноз по чистой прибыли, по структурным затратам и затратам на приобретение, а также по чувствительности к различным видам риска (таким как изменения процентных ставок и колебания обменного курса).
- Компания Eli Lilly с помощью моделирования определяет оптимальную производственную мощность для каждого лекарственного препарата.
- Компания Procter&Gamble с помощью метода Монте-Карло моделирует и оптимально хеджирует валютные риски.
- Компания Sears с помощью моделирования определяет, сколько штук по каждой линии продуктов необходимо заказать у поставщиков, например количество брюк Dockers на этот год.
- Нефтяные и фармацевтические компании моделируют оценку реальных возможностей, например оценку возможности расширения, заключения сделки или отсрочки проекта.
- Специалисты по финансовому планированию используют метод Монте-Карло для определения оптимальных инвестиционных стратегий для выхода на пенсию своих клиентов.

### **❓ Что происходит при вводе в ячейку формулы =СЛЧИС()?**

При вводе в ячейку формулы =СЛЧИС() функция СЛЧИС (RAND) возвращает число, которое с равной вероятностью принимает любое значение от 0 до 1. Таким образом, примерно в 25% случаев должно быть возвращено число не более 0,25; примерно в 10% случаев число должно быть не менее 0,90 и т. д. Для демонстрации работы функции СЛЧИС см. файл Randdemo.xlsx (рис. 77.1).

На листе Sheet1 я скопировал формулу =СЛЧИС() из ячейки C3 в C4:C402 и присвоил диапазону C3:C402 имя Data. Затем в столбце F я отследил среднее значение для 400 случайных чисел (ячейка F2) и с помощью функции СЧЁТЕСЛИ (COUNTIF) подсчитал долю значений в каждом интервале: от 0 до 0,25, от 0,25 до 0,50, от 0,50 до 0,75 и от 0,75 до 1. Чтобы случайные числа пересчитались, нажмите F9.

Обратите внимание, что среднее для 400 чисел всегда приблизительно равно 0,5 и на каждый из интервалов приходится примерно 25% результатов. Эти результаты согласуются с определением случайного числа. Обратите внимание, что значения, генерируемые функцией СЛЧИС в разных ячейках, независимы. Например, если в ячейке C3 сгенерировано большое случайное число (скажем, 0,99), о значениях других сгенерированных случайных чисел ничего сказать невозможно.

	B	C	D	E	F	G
1						
2	<b>Trial</b>			<b>mean</b>	<b>0.510511</b>	
3	<b>1</b>	<b>0.929558</b>				
4	<b>2</b>	<b>0.410748</b>		<b>Fraction</b>		
5	<b>3</b>	<b>0.252281</b>		<b>0-.25</b>	<b>0.24</b>	<b>=COUNTIF(data,"&lt;=.25")/400</b>
6	<b>4</b>	<b>0.740772</b>		<b>.25-.50</b>	<b>0.2575</b>	<b>=(COUNTIF(data,"&lt;=.5")-400*F5)/400</b>
7	<b>5</b>	<b>0.664546</b>		<b>.50-.75</b>	<b>0.24</b>	<b>=(COUNTIF(data,"&lt;=.75")-400*(F5+F6))/400</b>
8	<b>6</b>	<b>0.880954</b>		<b>.75-1</b>	<b>0.2625</b>	<b>=(400-400*SUM(F5:F7))/400</b>
9	<b>7</b>	<b>0.59274</b>				
10	<b>8</b>	<b>0.944597</b>				
11	<b>9</b>	<b>0.274288</b>				
12	<b>10</b>	<b>0.737283</b>				
13	<b>11</b>	<b>0.976839</b>				
14	<b>12</b>	<b>0.662258</b>				

**Рис. 77.1.** Демонстрация функции СЛЧИС

### ПРИМЕЧАНИЕ

В файле Randdemo.xlsx на листе Sheet1 отсутствуют случайные числа, показанные на рис. 77.1. Функция СЛЧИС при открытии листа или при вводе новой информации на лист всегда выполняет перерасчет генерируемых чисел.

### ❓ Как моделируют значения дискретной случайной величины?

Предположим, что спрос на календари определяется дискретной случайной величиной, представленной ниже:

Спрос	Вероятность
10 000	0,10
20 000	0,35
40 000	0,3
69 000	0,25

Как смоделировать в Excel спрос на календари множество раз? Нужно связать каждое возможное значение функции СЛЧИС с возможным спросом на календари. Соответствия, представленные ниже, гарантируют, что спрос на 10 000 календарей будет встречаться в 10% случаев и т. д.

Спрос	Соответствующее случайное число
10 000	Меньше 0,10
20 000	Больше или равно 0,10 и меньше 0,45
40 000	Больше или равно 0,45 и меньше 0,75
69 000	Больше или равно 0,75

См. моделирование спроса на листе Sim в файле Discretesim.xlsx (показано на рис. 77.2).

	A	B	C	D	E	F	G
1						<b>Cutoffs</b>	<b>Demand</b>
2	<b>Trial</b>		<b>rand</b>			<b>0</b>	<b>10000</b>
3	<b>1</b>	<b>40000</b>	<b>0.564906</b>			<b>0.1</b>	<b>20000</b>
4	<b>2</b>	<b>40000</b>	<b>0.540969</b>			<b>0.45</b>	<b>40000</b>
5	<b>3</b>	<b>40000</b>	<b>0.744298</b>			<b>0.75</b>	<b>60000</b>
6	<b>4</b>	<b>40000</b>	<b>0.677778</b>				
7	<b>5</b>	<b>20000</b>	<b>0.187121</b>			<b>Fraction of time</b>	
8	<b>6</b>	<b>40000</b>	<b>0.678377</b>		<b>10000</b>	<b>0.1175</b>	
9	<b>7</b>	<b>40000</b>	<b>0.557394</b>		<b>20000</b>	<b>0.335</b>	
10	<b>8</b>	<b>10000</b>	<b>0.035359</b>		<b>40000</b>	<b>0.31</b>	
11	<b>9</b>	<b>20000</b>	<b>0.306731</b>		<b>60000</b>	<b>0.2375</b>	
12	<b>10</b>	<b>40000</b>	<b>0.590446</b>				
13	<b>11</b>	<b>60000</b>	<b>0.847604</b>				
14	<b>12</b>	<b>20000</b>	<b>0.385141</b>				
15	<b>13</b>	<b>60000</b>	<b>0.960034</b>				

**Рис. 77.2.** Моделирование дискретной случайной величины

Ключом к моделированию здесь является использование случайного числа для инициации поиска в табличном диапазоне F2:G5 (с именем Lookup). Случайные числа от 0 до 0,10 соответствуют спросу на 10 000 календарей; случайные числа от 0,10 (включительно) до 0,45 — спросу на 20 000 календарей; случайные числа от 0,45 (включительно) до 0,75 — спросу на 40 000 календарей; и случайные числа от 0,75 (включительно) соответствуют спросу на 60 000 календарей. С помощью формулы =СЛЧИС(), скопированной из C3 в C4:C402, получено четыреста случайных чисел. Я сгенерировал 400 испытаний, или итераций, для спроса на календари, скопировав формулу =ВПР(C3;lookup;2) из B3 в B4:B402. Эта формула обеспечивает для любого случайного числа до 0,10 генерацию спроса на 10 000 календарей; для любого случайного числа от 0,10 до 0,45 генерацию спроса на 20 000 календарей и т. д. В ячейках F8:F11 я использовал функцию СЧЁТЕСЛИ (COUNTIF) для подсчета долей от 400 итераций, соответствующих каждому объему спроса. Когда я нажал F9 для перерасчета случайных чисел, моделируемые вероятности стали близки к предполагаемым вероятностям спроса. На листе Frozen, чтобы скопировать 400 значений спроса с листа Sheet1 на Frozen, нам пришлось воспользоваться командами Редактировать (Edit), Копировать (Copy) и Специальная вставка значений (Paste Special Values).

### ❓ Как моделируют значения нормальной случайной величины?

Если вы введете в любую ячейку формулу =НОРМ.ОБР(СЛЧИСЛ(); среднее; стандартное\_отклонение), то смоделируете значение нормальной случайной величины с параметрами среднее и стандартное\_отклонение. Эта процедура показана на листе Sim в файле Normalsim.xlsx. (См. рис. 77.3 с версией с листа Frozen.)

	A	B	C	D	E	F	G
3	<b>Trial</b>	<b>Normal Rv</b>	<b>Rand</b>				
4	1	35299.87077	0.319172892			sim mean	39928.16988
5	2	34057.11223	0.276159492			sim sigma	10349.91542
6	3	32891.87632	0.238600258				
7	4	26206.24775	0.083889547				
8	5	57666.30611	0.961354944				
9	6	57372.92915	0.958832259				
10	7	52334.20906	0.891290615				
11	8	32472.69553	0.225805955				
12	9	51180.01558	0.868216836				
13	10	35543.73668	0.327933593				
14	11	29878.96132	0.155744197				
15	12	49808.67811	0.836671034				

**Рис. 77.3.** Моделирование нормальной случайной величины

Предположим, что требуется смоделировать 400 испытаний для нормальной случайной величины со средним значением 40 000 и стандартным отклонением 10 000. (Я ввел эти значения в ячейки E1 и E2 с именами **mean** и **sigma** соответственно.) Скопировав формулу =СЛЧИС() из C4 в C5:C403, я сгенерировал 400 случайных чисел. Скопировав формулу =НОРМ.ОБР(C4,mean,sigma) из B4 в B5:B403, я сгенерировал 400 пробных значений нормальной случайной величины со средним 40 000 и стандартным отклонением 10 000. При нажатии F9 для перерасчета случайных чисел среднее значение остается близким к 40 000, а стандартное отклонение — к 10 000.

По существу, для случайного числа  $p$  по формуле =НОРМ.ОБР( $p$ ; среднее; стандартное\_отклонение) генерируется  $p$ -й процентиль нормальной случайной величины с параметрами **среднее** и **стандартное\_отклонение**. Например, для случайного числа 0,32 из ячейки C4 (см. рис. 77.3) в ячейке B4 генерируется приблизительно 32-й процентиль нормальной случайной величины со средним значением 40 000 и стандартным отклонением 10 000.

**?** Как компания — производитель поздравительных открыток может определить, сколько открыток следует произвести?

Здесь я покажу, как использовать метод Монте-Карло в качестве инструмента для принятия решений. Допустим, спрос на открытки ко Дню святого Валентина определяется дискретной случайной величиной, представленной в таблице ниже.

Спрос	Вероятность
10 000	0,10
20 000	0,35
40 000	0,3
60 000	0,25



Поздравительная открытка стоит \$4,00, а переменные затраты на производство каждой открытки составляют \$1,50. Нераспроданные открытки необходимо сбыть по цене \$0,20 за штуку. Сколько открыток нужно напечатать?

Сначала много раз (например, тысячу) моделируется каждый возможный объем производства (10 000, 20 000, 40 000 или 69 000). Затем по тысяче испытаний определяется, какой размер заказа соответствует максимальной средней прибыли. Данные вы найдете на листе **Sim** в файле **Valentine.xlsx**. (На рис. 77.4 показана версия с листа **Frozen**.) Имена в ячейках **B1:B11** присвоены ячейкам **C1:C11**. Диапазон ячеек **G3:H6** я присвоил имя **Lookup**. В ячейки **C4:C6** введена продажная цена и затраты.

	A	B	C
1		produced	40000
2		rand#	0.075175785
3		demand	10000
4		unit prod cost	\$ 1.50
5		unit price	\$ 4.00
6		unit disp cost	\$ 0.20
7			
8		revenue	\$ 40,000.00
9		total var cost	\$ 60,000.00
10		total disposing cost	\$ 6,000.00
11		profit	\$ (26,000.00)

**Рис. 77.4.** Моделирование для производства открыток ко Дню святого Валентина

Я ввел в ячейку **C1** пробный объем производства (40 000 в этом примере). Затем в ячейке **C2** по формуле **=СЛЧИС()** я создал случайное число. Как было описано ранее, спрос на открытки можно смоделировать в ячейке **C3** по формуле **=ВПР(случайное\_число;Lookup;2)**. (В этой формуле **случайное\_число** — имя, присвоенное ячейке **C3**.)

Количество проданных единиц продукции меньше, чем объем производства и спрос. В ячейке **C8** я вычислил выручку по формуле **=МИН(произведено;спрос)\*цена\_открытки**. В ячейке **C9** я вычислил общие производственные затраты по формуле **=произведено\*себестоимость\_открытки**.

Если открыток произведено больше, чем требуется, количество непроданных открыток равно **произведено – спрос**; в противном случае непроданных открыток не останется. В ячейке **C10** по формуле **=затраты\_на\_ликвидацию\_открытки\*ЕСЛИ(произведено>спрос;произведено-спрос;0)** вычисляются затраты на ликвидацию. Наконец, в ячейке **C11** прибыль вычислена как **выручка – общ\_переменные\_затраты – общ\_затраты\_на\_ликвидацию**.

Было бы неплохо, если бы можно было не напрягаясь нажать **F9** множество раз (например, тысячу) для каждого объема производства, а также подсчета ожидаемой прибыли для каждого объема. В этой ситуации на помощь приходит двуполосная таблица данных. (О таблицах данных см. в главе 17 «Анализ чув-



ствительности с помощью таблиц данных». ) Таблица данных для этого примера представлена на рис. 77.5.

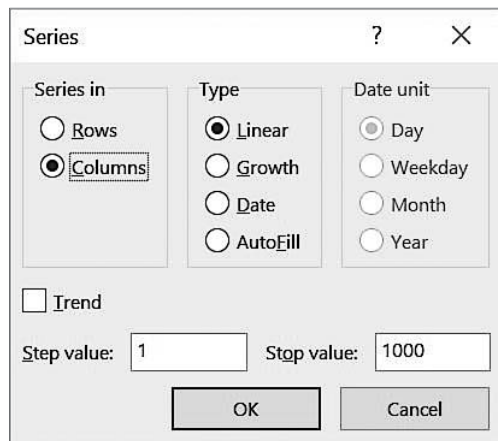
	A	B	C	D	E
13	mean	25000	45506	58042	43614
14	st dev	0	12989.25266	47856	72526.3
15	-26000	10000	20000	40000	60000
16	1	25000	50000	-26000	-18000
17	2	25000	50000	-26000	150000
18	3	25000	8000	16000	-18000
19	4	25000	8000	100000	-18000
20	5	25000	50000	100000	66000
21	6	25000	50000	100000	66000
22	7	25000	50000	100000	150000
23	8	25000	50000	16000	-18000
24	9	25000	50000	16000	66000
25	10	25000	50000	100000	-60000
26	11	25000	50000	16000	-18000
27	12	25000	50000	100000	-18000
28	13	25000	50000	16000	66000
29	14	25000	50000	-26000	-18000
30	15	25000	50000	-26000	66000
31	16	25000	50000	16000	-18000
32	17	25000	50000	16000	66000
33	18	25000	8000	100000	-60000
34	19	25000	50000	100000	-18000
35	20	25000	50000	16000	66000
36	21	25000	8000	16000	-18000
37	22	25000	50000	16000	66000
38	23	25000	8000	16000	150000
39	24	25000	50000	100000	-60000

**Рис. 77.5.** Двухнаправленная таблица данных для моделирования примера с поздравительными открытками

В ячейках A16:A1015 я ввел номера от 1 до 1000 (соответствующие тысяче испытаний). Простой способ это сделать — ввести 1 в ячейку A16, затем выделить ячейку, на вкладке Главная (Home) в группе Редактирование (Editing) открыть список Заполнить (Fill) и выбрать Прогрессия (Series). В открывшемся диалоговом окне Прогрессия (Series), показанном на рис. 77.6, укажите для шага значение 1, а для предельного значения — 1000. Установите переключатель Расположение (Series In Area) в положение по столбцам (Columns) и нажмите ОК. В столбец A, начиная с ячейки A16, будут введены числа от 1 до 1000.

Введите в ячейки B15:E15 возможные объемы производства (10 000, 20 000, 40 000, 69 000). Для каждого номера испытаний (от 1 до 1000) и каждого объема производства необходимо вычислить прибыль. Укажите ссылку на формулу для прибыли (вычисляемую в ячейке C11) в левой верхней ячейке таблицы данных (A15) путем ввода =profit (ячейке C11 было присвоено имя profit).

Теперь можно запустить в Excel моделирование тысячи испытаний для спроса при каждом объеме производства. Выделите таблицу (A15:E1015) и затем на вкладке Данные (Data) в группе Прогноз (Forecast) в списке Анализ "что, если" (What



**Рис. 77.6.** Диалоговое окно Прогрессия для ввода номеров испытаний от 1 до 1000

If Analysis) выберите Таблица данных (Data Table). Установите двунаправленную таблицу, выбрав объем производства (ячейку C1) в поле Подставлять значения по столбцам в (Row Input Cell) и любую пустую ячейку (здесь выбрана ячейка I14) в поле Подставлять значения по строкам в (Column Input Cell). Теперь нажмите кнопку ОК для моделирования в Excel одной тысячи значений спроса для каждого объема производства.

Чтобы понять, почему это работает, рассмотрим значения, помещенные с помощью таблицы данных в ячейки C16:C1015. Для каждой из этих ячеек при подстановке по столбцам используется значение 20 000 из ячейки C1. Для ячейки C16 при подстановке по строкам в пустую ячейку помещается значение 1, и случайное число в ячейке C2 пересчитывается. Затем в ячейку C16 записывается соответствующая прибыль. Далее для подстановки по строкам в пустую ячейку записывается значение 2, и случайное число в ячейке C2 пересчитывается еще раз. Соответствующая прибыль записывается в ячейку C17.

Скопировав формулу =СРЗНАЧ(B16:B1015) из ячейки B13 в C13:E13, вы можете вычислить среднюю смоделированную прибыль для каждого объема производства. Скопировав формулу =СТАНДОТКЛОН.В(B16:B1015) из ячейки B14 в C14:E14, вы можете вычислить стандартное отклонение смоделированной прибыли для каждого объема производства. Каждый раз при нажатии F9 для каждого объема производства моделируется 1000 итераций спроса. Для объема производства 40 000 открыток всегда в результате получается самая высокая ожидаемая прибыль, поэтому такой объем производства, по-видимому, является правильным решением.

## Влияние риска на решение

Если вы произведете 20 000 открыток вместо 40 000, ожидаемая прибыль снизится примерно на 22%, но и риск (судя по стандартному отклонению прибыли)

снизится почти на 73%. Следовательно, при общем нерасположении к риску правильным решением будет отпечатать 20 000 открыток. Между прочим, производство 10 000 открыток всегда имеет стандартное отклонение 0, поскольку 10 000 открыток будут проданы в любом случае, без остатков.

## ПРИМЕЧАНИЕ

В файле Valentine.xlsx установлен параметр Автоматически, кроме таблиц данных (Automatic Except For Tables). Он выбирается на вкладке Формулы (Formulas) в группе Вычисление (Calculation) в списке Параметры вычислений (Calculation). Этот параметр допускает перерасчет таблиц данных только при нажатии F9, что целесообразно, поскольку при вводе какой-либо информации на лист перерасчет больших таблиц данных замедляет работу. Обратите внимание, что в этом примере при каждом нажатии F9 средняя прибыль изменяется. Это происходит потому, что спрос для каждого объема производства генерируется с помощью другой последовательности из 1000 случайных чисел.

## Доверительный интервал для средней прибыли

В такой ситуации естественно задать вопрос: в какой интервал с вероятностью 95% попадает истинная средняя прибыль? Этот интервал называется *95%-ным доверительным интервалом для средней прибыли*. 95%-ный доверительный интервал для среднего значения любого конечного результата моделирования вычисляется по следующей формуле:

$$\text{средняя\_прибыль} \pm \frac{1,96 \times \text{стандартное\_отклонение\_прибыли}}{\sqrt{\text{число\_итераций}}}$$

В ячейке J11 по формуле =D13-1,96\*D14/КОРЕНЬ(1000) я вычислил нижний предел 95%-ного доверительного интервала для средней прибыли при производстве 40 000 календарей. В ячейке J12 я вычислил верхний предел 95%-ного доверительного интервала по формуле =D13+1,96\*D14/КОРЕНЬ(1000). Результаты вычислений показаны на рис. 77.7.

	I	J	K
9	95% CI for mean profit		
10	ordering 40,000 calendars		
11	Lower	55075.85	=D13-1.96*D14/SQRT(1000)
12	Upper	61008.15	=D13+1.96*D14/SQRT(1000)

**Рис. 77.7.** 95%-ный доверительный интервал для средней прибыли при производстве 40 000 календарей

С вероятностью 95% при производстве 40 000 календарей средняя прибыль составит от \$55 076 до \$61 008.

## Задания

1. Автодилер считает, что спрос на модели 2015 г. распределяется по нормальному закону со средним значением 200 и стандартным отклонением 30. Он приобретает Envoy за \$25 000 и продает за \$40 000. Половину автомобилей, не проданных по полной цене, можно продать по \$30 000. Дилер рассматривает варианты заказа на 200, 220, 240, 269, 280 или 300 штук. Сколько автомобилей он должен заказать?
2. Магазин пытается определить, сколько экземпляров журнала *People* необходимо заказывать каждую неделю. Владелец считает, что спрос на журнал соответствует следующей дискретной случайной величине:

Спрос	Вероятность
15	0,10
20	0,20
25	0,30
30	0,25
35	0,15

Магазин платит за каждый экземпляр \$1,00 и продает его по \$1,95. Каждый непроданный экземпляр может быть возвращен за \$0,50. Сколько экземпляров журнала еженедельно должен заказывать магазин?

3. Ваш супермаркет пытается определить, сколько мясных рулетов требуется изготовить в кулинарии в понедельник. Спрос на рулеты по понедельникам имеет нормальное распределение со средним, равным 100, и стандартным отклонением, равным 20. Себестоимость производства мясного рулета составляет \$2,00, а продаете вы его за \$7,00. Утилизация не проданного до конца дня рулета стоит \$0,60. Если возможно производить только 100, 110, 120 или 130 рулетов, какое их количество стоит произвести?
4. Спрос по понедельникам на круассаны с миндальным кремом соответствует нормальному распределению со средним, равным 100, и стандартным отклонением, равным 25. Себестоимость круассана — \$1,00, а продажная цена — \$4,00. Непроданные круассаны должны быть утилизированы по цене \$0,40 за штуку. Кондитерская может производить 80, 100, 120 или 140 круассанов. С помощью метода Монте-Карло определите объем производства, обеспечивающий для кондитерской максимальную прибыль.
5. Спрос на капустный салат (в фунтах) по воскресеньям соответствует нормальному распределению со средним, равным 500, и стандартным отклонением, равным 100. Себестоимость фунта капустного салата — \$3,00, а продажная цена — \$8,00. В конце дня остатки салата выбрасываются. Исходя из уровней

производства в 500, 520, 540, 560, 580 или 600 фунтов, определите объем производства, обеспечивающий максимальную прибыль.

6. Компания — производитель поздравительных открыток пытается определить, сколько выпустить открыток ко Дню святого Валентина. Стоимость печати  $x$  открыток вычисляется по формуле:  $\$2 \text{ млн} + 0,70x$ . То есть печать 1 млн открыток стоит  $\$2,7 \text{ млн}$ . Спрос на открытки следует нормальному распределению со средним, равным 2 млн, и стандартным отклонением, равным 400 000. Открытки продаются по цене  $\$4,00$ , а остатки — по  $\$0,05$ . Какой из возможных объемов производства — 2,4; 2,6; 2,8; 3 или 3,2 млн — обеспечит максимальную прибыль компании?

## ГЛАВА 78

# Вычисление оптимальной цены предложения

### Обсуждаемые вопросы

- Как смоделировать случайную величину с биномиальным распределением?
- Как определить, должна ли непрерывная случайная величина моделироваться как нормальная случайная величина?
- Как с помощью моделирования определить оптимальную тендерную заявку для проекта строительства?

Когда вы вступаете с конкурентами в борьбу за проект, двумя основными источниками неопределенности являются количество конкурентов и цены, предлагаемые каждым конкурентом. При высокой цене предложения можно заработать на проекте много денег, но самих проектов будет мало. При низкой цене предложения проектов будет немало, но денег каждый проект принесет немного. Оптимальная цена предложения находится где-то посередине. Удобным инструментом для определения цены предложения, максимально увеличивающей предполагаемую прибыль, является моделирование по методу Монте-Карло.

## Ответы на вопросы

### ❓ Как смоделировать случайную величину с биномиальным распределением?

Формула `=БИНОМ.ОБР(п;р;СЛЧИС())` моделирует количество благоприятных исходов в  $n$  независимых испытаниях, каждое из которых имеет вероятность благоприятного исхода, равную  $p$ . Как объяснялось в главе 77 «Введение в моделирование по методу Монте-Карло», функция `СЛЧИС (RAND)` генерирует число, с равной вероятностью принимающее любое значение от 0 до 1. В файле `Binomialsim.xlsx` (рис. 78.1) показано, что при нажатии F9 формула `=БИНОМ.ОБР(100;0,9;D3)`, введенная в ячейку C3, моделирует число штрафных бросков, выполняемых Стивом Нэшем (с вероятностью 90% забивающим штрафные в НБА) в 100 попытках. Формула `=БИНОМ.ОБР(100;0,5;D4)` в ячейке C4 моделирует количество выпавших орлов при подбрасывании монеты 100 раз. В ячейке C5 формула `=БИНОМ.ОБР(3;0,4;D5)` моделирует количество конкурентов, появившихся на рынке в течение года, при трех возможных участниках рынка и 40%-ной вероятности для каждого конкурента выйти на рынок. Разумеется, в ячейки D3:D5 я ввел формулу `=СЛЧИС()`.

	B	C	D	E
2			Rand#	
3	Number of free throws Steve Nash makes out of 100	88	0.211714	=BINOM.INV(100,0.9,D3)
4	Number of heads in 100 tosses	51	0.583189	=BINOM.INV(100,0.5,D4)
5	Number of competitors entering market; p = 4 n = 3	2	0.893061	=BINOM.INV(3,0.4,D5)

**Рис. 78.1.** Моделирование случайной величины с биномиальным распределением

❓ **Как определить, должна ли непрерывная случайная величина моделироваться как нормальная случайная величина?**

Допустим, вы считаете, что наиболее вероятная тендерная заявка конкурента составляет \$50 000. Напомню, что нормальная ФПВ является симметричной относительно среднего значения. Следовательно, для определения возможности применения нормальной случайной величины в моделировании цены предложения конкурента необходимо проверить симметрию относительно среднего значения цены предложения. Если цена предложения конкурента проявляет симметрию относительно среднего значения \$50 000, цены предложения \$40 000 и \$60 000, \$45 000 и \$55 000 и т. д., должны быть приблизительно одинаково вероятны. Если предположение о симметрии выглядит разумным, то цену предложения каждого конкурента можно смоделировать как нормальную случайную величину со средним значением \$50 000.

Как оценить стандартное отклонение цены предложения каждого конкурента? Напомню, что согласно правилу из главы 44 «Обобщение данных с помощью описательной статистики», наборы данных с симметричными гистограммами имеют примерно 95% данных в пределах двух стандартных отклонений от среднего значения. Аналогично нормальная случайная величина с вероятностью 95% находится в пределах двух стандартных отклонений от своего среднего значения. Предположим, что с вероятностью 95% цена предложения конкурента составляет от \$30 000 до \$70 000. Это означает, что  $2 \times$  (стандартное отклонение цены предложения конкурента) равно \$20 000 или стандартное отклонение цены предложения конкурента равно \$10 000.

Если предположение о симметрии разумно, цену предложения конкурента можно смоделировать с помощью формулы =НОРМ.ОБР(СЛЧИС(); 50000; 10000). (О моделировании нормальных случайных величин с помощью функции НОРМ.ОБР см. главу 77.)

Кроме того, нормальное распределение имеет нулевые *асимметрию* и *эксцесс*. Если вы по достаточно большой выборке получаете, что асимметрия и/или эксцесс в абсолютном выражении больше 1, то маловероятно, что данные соответствуют нормальному распределению.

### ❓ Как с помощью моделирования определить оптимальную тендерную заявку для проекта строительства?

Предположим, что вы участвуете в тендере на проект строительства, который обойдется вам в \$25 000. Подготовка тендерной заявки стоит \$1000. У вас имеется шесть потенциальных конкурентов, и вы подсчитали, что с вероятностью 50% каждый конкурент примет участие в тендере. Если конкурент подает тендерную заявку, то цена его предложения следует закону нормального распределения со средним значением \$50 000 и стандартным отклонением \$10 000. Предположим также, что вы подготавливаете заявки только с ценой, кратной \$5000. Какова должна быть цена вашего предложения для получения максимальной ожидаемой прибыли? Напомню, что выигрывает самая низкая тендерная заявка! Данные к этой задаче находятся в файле *Bidsim.xlsx* (рис. 78.2 и 78.3).

	D	E	F	G
1	costproject	25000		
2	cost bid	1000	rand#	
3	Number bidders	3	0.599314	
4	mybid	40000		
5				
6				
7				
8	Bidder #	In	Bid	rand#
9	1	yes	50150.27	0.505995
10	2	yes	41587.4	0.200101
11	3	yes	42536.87	0.227739
12	4	no	100000	0.684822
13	5	no	100000	0.206349
14	6	no	100000	0.51067
15				
16	Do I win?			
17	yes			
18	Profit			
19	14000			

**Рис. 78.2.** Модель цены предложения

Ваша стратегия должна быть следующей.

- Сгенерируйте количество конкурентов.
- Для каждого потенциального конкурента, который действительно подает заявку, смоделируйте предложение с помощью нормальной случайной величины. Если потенциальный конкурент не подает заявку, назначьте для него большую цену предложения (например, \$100 000), которая гарантирует ему поражение.



	D	E	F	G	H	I	J	K
21	mean	3695	7200	8600	5900	3450	1610	190
22	14000	30000	35000	40000	45000	50000	55000	60000
23	1	4000	9000	14000	-1000	-1000	-1000	-1000
24	2	4000	9000	-1000	-1000	-1000	-1000	-1000
25	3	4000	9000	14000	-1000	24000	-1000	-1000
26	4	4000	9000	-1000	-1000	-1000	-1000	-1000
27	5	4000	9000	14000	19000	-1000	29000	-1000
28	6	4000	9000	14000	-1000	-1000	-1000	-1000
29	7	4000	9000	14000	-1000	-1000	-1000	-1000
30	8	4000	9000	14000	19000	-1000	-1000	-1000
31	9	4000	-1000	-1000	19000	-1000	-1000	-1000
32	10	4000	9000	14000	-1000	24000	-1000	-1000
33	11	-1000	9000	14000	19000	-1000	-1000	-1000
34	12	4000	9000	14000	19000	-1000	-1000	-1000
35	13	4000	9000	-1000	-1000	24000	29000	-1000
36	14	4000	-1000	14000	-1000	-1000	-1000	-1000
37	15	4000	9000	-1000	-1000	-1000	-1000	-1000
38	16	4000	-1000	-1000	-1000	-1000	-1000	-1000
39	17	4000	9000	14000	19000	-1000	-1000	-1000
40	18	4000	9000	-1000	-1000	-1000	-1000	-1000
41	19	4000	9000	14000	-1000	-1000	-1000	-1000
42	20	4000	9000	14000	19000	-1000	-1000	-1000
43	21	4000	9000	14000	-1000	-1000	-1000	-1000
44	22	4000	9000	14000	-1000	-1000	-1000	-1000
45	23	4000	9000	14000	-1000	24000	-1000	-1000
46	24	4000	-1000	14000	-1000	-1000	-1000	-1000

**Рис. 78.3.** Таблица данных для моделирования цены предложения

- Определите, является ли цена вашего предложения самой низкой.
- Если вы предлагаете самую низкую цену, то получите прибыль, равную предложению минус стоимость проекта минус \$1000 (стоимость заявки). Если цена вашего предложения не самая низкая, то вы теряете \$1000 — стоимость заявки.
- Для моделирования каждой возможной цены предложения (например, \$30 000, \$35 000, ..., \$60 000) 1000 раз используйте двунаправленную таблицу данных и затем выберите цену предложения с наибольшей ожидаемой прибылью.

Сначала я присвоил имена из ячеек D1:D4 ячейкам E1:E4. В ячейке E3 я определил количество конкурентов по формуле =БИНОМ.ОБР(6;0,5;F3). Ячейка F3 содержит формулу =СЛЧИС(). Затем я определил, какой из потенциальных конкурентов действительно подает заявку, скопировав формулу =ЕСЛИ(D9<=количество\_предложений;"да"; "нет") из E9 в E10:E14.

Далее я сгенерировал заявки для каждого конкурента (конкурентам, не подающим заявки, назначил цену \$100 000), скопировав формулу =ЕСЛИ(E9="да";НОРМ.ОБР(G9;50000;10000);100000) из F9 в F10:F14. Каждая ячейка в диапазоне G9:G14 содержит функцию СЛЧИС. В ячейке D17 я определил, даю ли я самую низкую цену и выиграю ли тендер, по формуле =ЕСЛИ(мое\_предложение<=МИН(F9:F14); "да"; "нет"). В ячейке D19 я вычислил прибыль по формуле =ЕСЛИ(D17="да"; мое\_предложение-стоимость\_проекта-стоимость\_предложения; -стоимость\_предложения), пока-

зывающую, что в случае выигрыша я получаю сумму предложения и оплачиваю расходы на проект.

Теперь я могу использовать двунаправленную таблицу (см. рис. 78.3) и смоделировать 1000 предложений от \$30 000 до \$60 000. Я скопировал прибыль в ячейку D22, введя формулу  $=D19$ . Затем выделил табличный диапазон D22:K1022. На вкладке **Данные (Data)** в группе **Прогноз (Forecast)** в списке **Анализ "что, если" (What If Analysis)** я выбрал **Таблица данных (Data Table)** и указал входные значения для таблицы данных. Можно указать любую пустую ячейку в поле **Подставлять значения по строкам в (column input cell)**. А в поле **Подставлять значения по столбцам в (row input cell)** я указал ячейку E4 (местоположение предложения). В диалоговом окне **Таблица данных (Data Table)** я нажал ОК, и прибыль для каждого предложения была смоделирована 1000 раз.

Средняя прибыль для каждого предложения вычисляется по формуле  $=CPЗНАЧ(E23:E1022)$ , скопированной из ячейки E21 в F21:K21. Каждый раз при нажатии F9 средняя прибыль для 1000 испытаний максимальна для цены предложения \$40 000.

## Задания

1. Как изменится оптимальная цена предложения при наличии 12 конкурентов?
2. Вы участвуете в тендере на нефтяную скважину, которая, по вашей оценке, должна принести \$40 млн (включая стоимость разработки и добычи нефти). В тендере участвуют еще три конкурента, и предложение каждого конкурента предположительно соответствует закону нормального распределения со средним значением \$30 млн и стандартным отклонением \$4 млн. Какое предложение вы должны сделать (с точностью до \$1 млн)?
3. Часто используется непрерывная случайная величина, имеющая *равномерное распределение*. Случайная величина с равномерным распределением, записываемая как  $U(a, b)$ , с равной вероятностью может принимать любое значение между двумя заданными числами:  $a$  и  $b$ . Объясните, почему формулу  $=a+(b-a)*\text{СЛЧИС}()$  можно применять для моделирования  $U(a, b)$ .
4. Инвестор Питер Фишер предлагает цену за биотехнологическую компанию. Компания с равной вероятностью может стоить от \$0 до \$200 за акцию. Только сама компания знает свою истинную стоимость. Питер — настолько хороший инвестор, что рынок немедленно оценит стоимость компании на 50% больше ее истинной стоимости. Какую цену должен предложить Питер за акцию компании?
5. Робинсон Кано (Robinson Cano), игрок бейсбольной команды «Сиэтл Маринерс», приглашен на арбитраж по повышению своей зарплаты. Процедура арбитража в Высшей бейсбольной лиге проходит следующим образом: игрок представляет на рассмотрение размер зарплаты, которую, по его мнению, он

должен получать, то же самое делает команда. Арбитр (не видя размера зарплат, представленных игроком или командой) оценивает справедливую зарплату. Игрок получает ту из предложенных зарплат, которая ближе к оценке арбитра. Например, предположим, что Кано представил на рассмотрение зарплату в \$12 млн, а команда — \$7 млн. Если арбитр скажет, что справедливая зарплата составляет \$10 млн, Кано получит свои \$12, а если арбитр скажет, что справедливая зарплата составляет \$9 млн, Кано получит только \$7. Предположим, что оценка арбитра с равной вероятностью находится где-то между \$8 и \$11 млн, а предложение команды с равной вероятностью составляет от \$6 до \$9 млн. Какую зарплату должен предложить Кано с точностью до \$1 млн?

6. Две компании по производству напитков борются за эксклюзивные права на продажу газировки в Хьюстонском университете в течение следующих 10 лет. Компания 1 получит 50 центов дохода с каждой проданной бутылки (за вычетом отчислений университету) и предполагает, что число проданных бутылок в течение следующих 10 лет будет иметь нормальное распределение со средним, равным 6 млн, и стандартным отклонением, равным 1,5 млн. Компания 1 предполагает, что предложение компании 2 составляет с равной вероятностью любое значение между \$1 млн и \$2,5 млн. Каким должно быть предложение компании 1, чтобы оно обеспечило максимальную предполагаемую прибыль? Для нахождения целых чисел, равновероятно выпадающих в интервале между целыми  $n_1$  и  $n_2$  (включительно), можно воспользоваться функцией Excel =СЛУЧМЕЖДУ( $n_1, n_2$ ) (RANDBETWEEN). Исходя из предложений в \$1; \$1,5; \$2; \$2,5 и \$3 млн, какое из этих предложений принесет максимальную ожидаемую прибыль?
7. Вы участвуете в тендере на ямочный ремонт улицы Вестхаймер-роуд в Хьюстоне. (Удачи!) Вы оцениваете стоимость заделки ям в \$4 млн. У вас есть четыре конкурента. Предложение каждого конкурента соответствует закону нормального распределения со средним в \$6 млн и стандартным отклонением, равным \$0,5 млн. Какое предложение — \$4,5; \$5; \$5,5 или \$6 млн — принесет наибольшую ожидаемую прибыль?

## ГЛАВА 79

# Моделирование цен на акции и распределения средств между активами

### Обсуждаемые вопросы

- Я недавно купил 100 акций компании GE. Какова вероятность того, что в течение следующего года эта инвестиция принесет доход более 10%?
- Я пытаюсь распределить свой инвестиционный портфель между акциями, казначейскими векселями и облигациями. Какое распределение средств между активами в рамках пятилетнего горизонта планирования принесет ожидаемый доход не менее 8% и минимизирует риск?

Последние несколько лет показали, что будущие доходы от инвестиций являются весьма неопределенными. В главе 75 «Логарифмически нормальная случайная величина в моделировании курса акции» я рассказал, как моделировать цены на акции с помощью логарифмически нормальной случайной величины. Многие финансовые аналитики критикуют использование логарифмически нормальной случайной величины для моделирования цен на акции, поскольку недооценивается вероятность экстремальных событий (которые часто называют *черными лебедями*). В этой главе я объясню относительно простой подход к оценке неопределенности будущих инвестиционных доходов. Этот подход основан на идее бутстрэппинга. По сути, идея бутстрэппинга при моделировании будущих инвестиционных доходов состоит в предположении, что будущее подобно прошлому. Например, если необходимо смоделировать цену акций GE через год, можно предположить, что ежемесячное относительное изменение цены — это с равной вероятностью относительное изменение цены в один из предыдущих 60 месяцев. Этот метод позволяет быстро сгенерировать тысячи сценариев для будущей стоимости инвестиций. Кроме того, в сценариях, предполагающих, что будущие изменчивость и среднее значение доходов подобны прошлым, можно легко отразить позицию, что будущие инвестиционные доходы будут меньше или больше, чем в недавнем прошлом.

После создания будущих сценариев для инвестиционных доходов можно решить задачу распределения средств между активами легко и просто, применив Поиск решения (Solver). То есть вы должны решить задачу о том, как распределить инве-

стиции так, чтобы это принесло вам ожидаемые доходы желаемых размеров, но с минимальным риском.

Простота и мощь бутстрэппинга демонстрируются на следующих двух примерах.

## Ответы на вопросы

**?** Я недавно купил 100 акций компании GE. Какова вероятность того, что в течение следующего года эта инвестиция принесет доход более 10%?

Допустим, что акции GE в настоящее время торгуются по \$28,50 за акцию. Данные о ежемесячных доходах по акциям GE (а также по акциям Microsoft и Intel) для периода с августа 1997 г. по июль 2002 г. находятся в файле `GESim.xlsx` (рис. 79.1). Например, в месяце, заканчивающемся 2 августа 2002 г (по сути, это июль 2002 г), доходы по акциям GE снизились на 12,1%. Эти доходы включают дивиденды (если таковые имеются), выплачиваемые каждой компанией.

	B	C	D	E	F
3		mean	0.014746	0.006237	0.008828
4	Code		MSFT	INTC	GE
5	1	8/2/2002	-0.08316	-0.15397	-0.12112
6	2	7/2/2002	-0.12285	0.028493	0.108434
7	3	6/2/2002	0.074445	-0.33853	-0.06139
8	4	5/2/2002	-0.02583	-0.03464	-0.01276
9	5	4/2/2002	-0.13348	-0.05894	-0.15658
10	6	3/2/2002	0.033768	0.064867	-0.02849
11	7	2/2/2002	-0.08429	-0.18514	0.041088
12	8	1/2/2002	-0.03834	0.114295	-0.07314
13	9	12/1/2001	0.031771	-0.03709	0.045898
14	10	11/1/2001	0.104213	0.337433	0.057167
15	11	10/1/2001	0.136408	0.194417	-0.02102
16	12	9/1/2001	-0.10307	-0.26889	-0.08653
17	13	8/1/2001	-0.13809	-0.06181	-0.05957
18	14	7/1/2001	-0.09329	0.019172	-0.10944
19	15	6/1/2001	0.055218	0.082654	0
20	16	5/1/2001	0.021107	-0.12601	0.009701
21	17	4/1/2001	0.238801	0.174658	0.159413
22	18	3/1/2001	-0.07305	-0.07886	-0.09653
23	19	2/1/2001	-0.03374	-0.22808	0.011168

**Рис. 79.1.** Данные по акциям GE, Microsoft и Intel

Цена на акции GE через год не определена. Каким образом можно получить представление о диапазоне разброса цены на акции GE в течение года начиная с текущего момента? Бутстрэппинг позволяет оценить доход по акциям GE для каждого из следующих 12 месяцев на основе предположения о том, что доход в каждом месяце с одинаковой вероятностью может быть равен доходу в любом из перечисленных 60 месяцев. Другими словами, доход по акциям GE в следующем месяце

с одинаковой вероятностью равен любому числу из диапазона ячеек F5:F64. Для реализации этой идеи выберите сценарий для каждого из 12 следующих месяцев с помощью формулы `=СЛУЧМЕЖДУ(1;60)`. Например, если функция `СЛУЧМЕЖДУ` (`RANDBETWEEN`) возвратит 7 для следующего месяца, возьмите в качестве дохода для следующего месяца доход по акциям GE в ячейке F11 (4,1%), которая является седьмой ячейкой в диапазоне. Результаты показаны на рис. 79.2. (В файле вы увидите другие значения, поскольку при открытии листа функция `СЛУЧМЕЖДУ` (`RANDBETWEEN`) автоматически пересчитывает случайные числа.)

Сначала введите текущую цену акций GE (\$28,50) в ячейку J6. Затем создайте сценарии для каждого из следующих 12 месяцев, скопировав формулу `=СЛУЧМЕЖДУ(1;60)` из K6 в K7:K17. Используйте таблицу поиска, чтобы получить доход по акциям GE на основе вашего сценария. Для этого просто скопируйте формулу `=ВПР(K6;lookup;5)` из L6 в L7:L17. Как видно из формулы, диапазон B5:F64 называется `lookup`, а доходы по акциям GE указаны в пятом столбце диапазона поиска. В сценариях, показанных на рис. 79.2, можно видеть, например, что доход по акциям GE через шесть месяцев будет равен доходу в точке данных 1/04/1999, то есть 4,7%. (Помните, что при каждом открытии файла вы будете видеть разный набор сценариев.)

	I	J	K	L	M
4	GE				
5	Month	Start price	Scenario	Return	End price
6	1	28.5	29	0.010352379	28.79504
7	2	28.7950428	19	0.011168193	29.11663
8	3	29.1166314	49	-0.105336105	26.0496
9	4	26.0495989	28	0.003940887	26.15226
10	5	26.1522574	46	0.032932011	27.0135
11	6	27.0135038	41	-0.047338936	25.73471
12	7	25.7347133	37	0.030415009	26.51743
13	8	26.5174348	45	0.128556736	29.92643
14	9	29.9264297	45	0.128556736	33.77367
15	10	33.7736738	40	-0.034989709	32.59194
16	11	32.5919428	8	-0.073139975	30.20817
17	12	30.2081689	38	-0.035371538	29.13966

**Рис. 79.2.** Моделирование цены акций GE через год

Для определения цены акций GE на конец каждого месяца скопируйте формулу `=(1+L6)*J6` из M6 в M7:M17. Эта формула имеет вид  $(1 + \text{доход за месяц}) \times (\text{начальная цена GE})$ . Наконец, по формуле `=M6`, скопированной из J7 в J8:J17, вычислите начальную цену акций для каждого месяца как равную цене акций на конец предыдущего месяца.

Теперь с помощью таблицы данных можно сгенерировать 1000 сценариев для цены акций GE через год и годовой доход от инвестиций в процентах. Таблица данных представлена на рис. 79.3. В ячейку J19 скопируйте конечную цену по



формуле =M17. В ячейку K19 введите формулу =(M17-\$J\$6)/\$J\$6 для расчета годового дохода в виде (конечная\_цена\_акций\_GE – начальная\_цена\_акций\_GE)/начальная\_цена\_акций\_GE.

	I	J	K	L	M
18		End Price	End Return		
19		29.1396595	0.022444193	Mean	0.106965
20	1	39.0103499	0.368784206	Prob lose money	0.388
21	2	37.7908528	0.325994837	Prob make more than 10%	0.458
22	3	54.4774682	0.911490113	Make between 0 and 10%	0.154
23	4	17.1784002	-0.397249115	Lose between 0 and 10%	0.115
24	5	21.4253592	-0.248233011	Lose More than 10%	0.273
25	6	31.2594539	0.096822943		
26	7	21.7861276	-0.235574471		
27	8	45.5626088	0.598688028		
28	9	38.8959902	0.364771587		
29	10	16.3155196	-0.427525627		
30	11	29.1597675	0.023149737		
31	12	31.4274492	0.102717517		
32	13	31.8482638	0.117482941		
33	14	38.736722	0.359183228		
34	15	22.2880993	-0.217961427		

**Рис. 79.3.** Таблица данных для моделирования цены акции GE

Далее выделите таблицу (J19:K1019), на вкладке Данные (Data) в группе Прогноз (Forecast) в списке Анализ "что, если" (What If Analysis) выберите Таблица данных (Data Table). Задайте однонаправленную таблицу данных, указав любую пустую ячейку в поле Подставлять значения по строкам в (Column Input Cell). После нажатия ОК в диалоговом окне Таблица данных (Data Table) Excel сгенерирует 1000 сценариев для цены акций GE через год. (Для этого листа на вкладке Формулы (Formulas) я выставил параметр Автоматически, кроме таблиц данных (Automatic Except For Tables). Нажмите F9, если хотите увидеть изменение смоделированных цен.)

В ячейках M20:M24 я применил функцию СЧЁТЕСЛИ (COUNTIF) (см. главу 20 «Функции СЧЁТЕСЛИ, СЧЁТЕСЛИМН, СЧЁТ, СЧЁТЗ и СЧИТАТЬПУСТОТЫ») и обобщил данные по возможным доходам через год. Например, в ячейке M20 я вычислил вероятность потерять деньги через год по формуле =СЧЁТЕСЛИ(returns;"<0")/1000. (Я назвал диапазон с 1000 смоделированных доходов returns.) Результаты моделирования на основе данных за 1997–2002 гг. показывают, что вероятность потери денег в следующем году от инвестиции в акции GE составляет 39%. Для 1000 смоделированных доходов верно следующее:

- доход превысит 10% с вероятностью 46%;
- доход составит от 0 до 10% с вероятностью 15%;
- потери от инвестиций составят от 0 до 10% с вероятностью 12%;

- потери от инвестиций составят более 10% с вероятностью 28%;
- средний доход в следующем году составит приблизительно 10,7%.

Многие знатоки считают, что в будущем доходность акций окажется не такой высокой, как в недавнем прошлом. Предположим, вы считаете, что в следующем году GE будет в среднем на 5% менее эффективна, чем в период 1997–2002 гг., для которого имеются данные. Это предположение можно быстро включить в модель, изменив формулу конечной цены для акций GE в ячейке M17 на формулу  $=(1+L17)*J17-0,05*J6$ . По этой формуле конечная цена акций GE просто уменьшается на 5% от цены в начале года, что снижает доход в следующем году на 5%. Эти результаты показаны в файле GEsimless5.xlsx и на рис. 79.4.

	L	M
19	Mean	0.074622
20	Prob lose money	0.481
21	Prob make more than 10%	0.406
22	Make between 0 and 10%	0.113
23	Lose between 0 and 10%	0.132
24	Lose More than 10%	0.349

**Рис. 79.4.** Более пессимистичный взгляд на будущее акций GE

Обратите внимание, что теперь вероятность падения цены акций GE в следующем году составляет 48%. Средняя цена акций ниже не ровно на 5% по сравнению с предыдущей моделью, поскольку каждый раз, когда вы запускаете 1000 итераций, смоделированные значения изменяются. Поэтому когда вы откроете файл, то увидите немного другие числа, не такие, как на рис. 79.4.

**?** Я пытаюсь распределить свой инвестиционный портфель между акциями, казначейскими векселями и облигациями. Какое распределение средств между активами в рамках пятилетнего горизонта планирования принесет ожидаемый доход не менее 8% и минимизирует риск?

Ключевым решением, принимаемым отдельными лицами, менеджерами взаимных фондов и другими инвесторами, является распределение средств между различными классами активов с учетом будущей неопределенности доходов для этих классов активов. Разумный подход к распределению средств между активами следующий: применить бутстрэппинг для генерации 1000 смоделированных значений для будущих стоимостей каждого класса активов, а затем применить Поиск решения (Solver) для определения распределения средств между активами, приносящего ожидаемый доход с минимальным риском. Пусть для примера даны годовые доходы по акциям, казначейским векселям и облигациям в период с 1973 по 2012 г. Вы инвестируете в рамках пятилетнего горизонта планирования и на основе исторических данных хотите узнать, какое распределение средств между активами дает ожидаемый годовой доход не менее 8% с минимальным риском (если



судить по стандартному отклонению). Эти данные находятся в файле Assetallsim.xlsx (рис. 79.5). (Показана только часть данных.)

	A	B	C	D
6		Annual Returns on Investments in		
7	Year	Stocks	T.Bills	T.Bonds
8	1973	-14.31%	5.07%	3.66%
9	1974	-25.90%	7.45%	1.99%
10	1975	37.00%	7.15%	3.61%
11	1976	23.83%	5.44%	15.98%
12	1977	-6.98%	4.35%	1.29%
13	1978	6.51%	6.07%	-0.78%
14	1979	18.52%	9.08%	0.67%
15	1980	31.74%	12.04%	-2.99%
16	1981	-4.70%	15.49%	8.20%
17	1982	20.42%	10.85%	32.81%
18	1983	22.34%	7.94%	3.20%
19	1984	6.15%	9.00%	13.73%
39	2004	10.88%	8.51%	1.02%
40	2005	4.91%	7.81%	1.20%
41	2006	15.79%	1.19%	2.98%
42	2007	5.49%	9.88%	4.66%
43	2008	-37.00%	25.87%	1.60%
44	2009	26.46%	-14.90%	0.10%
45	2010	14.82%	0.13%	8.46%
46	2011	2.07%	0.03%	16.04%
47	2012	15.83%	0.05%	2.97%

**Рис. 79.5.** Исторические данные по акциям, казначейским векселям и облигациям

Сначала с помощью бутстрэппинга сгенерируйте 1000 смоделированных значений для акций, казначейских векселей и облигаций через пять лет. Допустим, для каждого класса активов текущая цена составляет \$1 (рис. 79.6).

Для каждого класса активов введите начальную цену \$1 в ячейки H10:J10. Затем, скопировав формулу =СЛУЧМЕЖДУ(1973;2012) из K10 в K11:K14, создайте сценарий для каждого года из пяти последующих лет. Например, для данных на рис. 79.6 следующий год будет похож на 1989-й, следующий за ним — на 1995-й и т. д. По формуле =H10\*(1+ВПР(\$K10;lookup;L\$8)), скопированной из L10 в L10:N14, для каждого класса активов генерируется цена на конец года. Например, для акций по этой формуле вычисляется следующее:

цена акции на конец года  $t = (\text{цена акции на начало года } t) \times (1 + \text{доход по акции в год } t)$ .

Скопируйте формулу =L10 из H11 в H11:J14 для вычисления цены каждого класса активов на начало каждого последующего года.

Теперь вы можете использовать однонаправленную таблицу данных и сгенерировать 1000 сценариев для значений акций, казначейских векселей и облигаций за

	G	H	I	J	K	L	M	N
9	Year	Stock value	Bill value	Bond value	Scenario	End stock	End bill	End bond
10	1	1	1	1	1978	1.0651	1.0607	0.9922
11	2	1.0651	1.0607	0.9922	1981	1.01504	1.225002	1.07356
12	3	1.01504	1.22500243	1.07356	2004	1.125477	1.32925	1.084511
13	4	1.125477	1.329250137	1.084511	2006	1.303189	1.345068	1.116829
14	5	1.303189	1.345068213	1.116829	1980	1.716822	1.507014	1.083436
15			5 yr stock	5 year bill	5 yr bond			
16			1.716821786	1.507014	1.083436			
17		1	0.984913402	1.407977	1.750108			
18		2	2.113731575	1.229216	1.352769			
19		3	0.508550251	1.94635	1.187041			
39		23	1.070908577	1.316186	1.395109			
40		24	1.442456654	1.337122	1.380632			
41		25	1.883655596	1.325864	1.912206			
42		26	1.575362983	1.309547	1.762178			
43		27	1.882561059	1.098241	1.248533			
44		28	1.266377383	1.378358	1.096124			
45		29	1.458521484	1.6187	1.502827			
46		30	0.65321012	1.533285	1.461328			
47		31	0.870558594	1.110473	1.229864			
48		32	1.338427978	1.394857	1.600366			

**Рис. 79.6.** Моделирование доходов по акциям, казначейским векселям и облигациям на пять лет

пять лет. Начните с копирования цены на конец года 5 для каждого класса активов в ячейки I16:K16. Выделите таблицу (H16:K1015), на вкладке **Данные (Data)** в группе **Прогноз (Forecast)** в списке **Анализ "что, если" (What If Analysis)** выберите **Таблица данных (Data Table)**. Задайте однонаправленную таблицу данных, указав любую пустую ячейку в поле **Подставлять значения по строкам в (Column Input Cell)**. После нажатия **ОК** в диалоговом окне **Таблица данных (Data Table)** вы получите 1000 смоделированных цен для акций, казначейских векселей и облигаций через пять лет. Важно отметить, что цены акций, казначейских векселей и облигаций при таком подходе к моделированию не изменяются независимо. Для каждого года из пяти лет доходы по акциям, векселям и облигациям всегда выбираются из одного и того же ряда данных. Это позволяет бутстрэппингу отразить взаимозависимость доходов по этим классам активов, проявленную в недавнем прошлом. (Конкретные доказательства того, что бутстрэппинг должным образом моделирует взаимозависимость между доходами по этим трем классам активов, см. в задании 7 в конце этой главы.)

Теперь вы готовы найти оптимальное распределение средств между активами, которое я вычислил в файле *Assetallocationopt.xlsx* (рис. 79.7). Сначала я скопировал 1000 смоделированных стоимостей активов за пять лет на пустой лист (я выбрал диапазон ячеек C4:E1003). В ячейки C2:E2 я ввел пробные доли распределения средств между акциями, векселями и облигациями соответственно. В ячейке F2 я сложил эти доли распределения средств по формуле **=СУММ(C2:E2)**. Далее я добавил ограничение **F2=1** в модель поиска решения, которое обеспечит 100%-ное инвестирование денежных средств в один из трех классов активов.

	B	C	D	E	F	G	H	I
1		stock	bill	bond	total			
2		0.253439	0.72496	0.021598	1			
3	Iteration#	5 yr stock	5 yr bill	5 yr bond	final value	annual return	mean	0.075872
4	1	2.013289	1.27984	1.361632	1.4674917	0.079729904	stdev	0.023885
5	2	1.697591	1.34639	1.300374	1.4344032	0.074816315		
6	3	0.64513	1.99583	1.359452	1.6397654	0.103967646		
7	4	2.498326	1.15132	1.510334	1.5004558	0.084537671		
8	5	1.733605	1.25654	1.189519	1.3760024	0.065918104		
9	6	1.554756	1.5456	1.590621	1.5488894	0.091450594		
10	7	1.597563	1.25608	1.514132	1.3482021	0.061575772		
11	8	1.7993	1.26951	1.237904	1.4030957	0.070082989		
12	9	1.802803	1.31441	1.677732	1.4460318	0.076553382		
13	10	2.877351	1.42034	1.616723	1.7938469	0.123976099		
14	11	2.806566	1.10423	1.300825	1.5399178	0.090183264		
15	12	2.784766	1.43937	1.61732	1.7841901	0.122763344		
16	13	1.651119	1.4176	1.35738	1.475484	0.080903445		

**Рис. 79.7.** Модель оптимального распределения средств между активами

Далее нам нужно определить окончательную стоимость портфеля для каждого сценария. Для этих вычислений подойдет следующая формула:

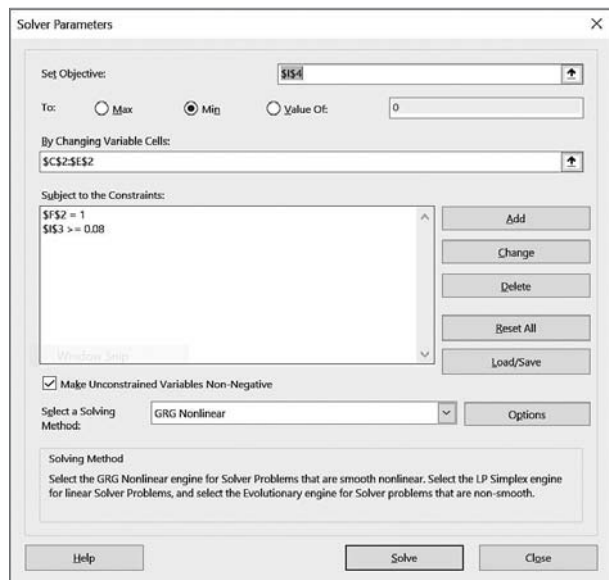
(окончательная стоимость портфеля) = (окончательная стоимость акций) + (окончательная стоимость векселей) + (окончательная стоимость облигаций).

Для окончательного распределения активов для каждого сценария я скопировал формулу =СУММПРОИЗВ(C4:E4;\$C\$2:\$E\$2) из ячейки F4 в F5:F1003.

Следующим шагом является определение годового дохода за пятилетний моделируемый период для каждого созданного сценария. Обратите внимание, что  $(1 + \text{годовой доход})^5 = (\text{окончательная стоимость портфеля}) / (\text{начальная стоимость портфеля})$ . Так как начальная стоимость портфеля равна \$1, то  $\text{годовой доход} = (\text{окончательная стоимость портфеля})^{1/5} - 1$ .

Скопировав формулу  $=(F4/1)^(1/5)-1$  из ячейки G4 в G5:G1003, я вычислил годовой доход для каждого сценария за пятилетний моделируемый период. Я присвоил диапазону G4:G1003 (содержащему смоделированные годовые доходы) имя **returns** и вычислил средний годовой доход в ячейке I3 по формуле =СРЗНАЧ(returns) и стандартное отклонение годовых доходов в ячейке I4 по формуле =СТАНДОТКЛОН.В(returns).

Теперь все готово для использования инструмента Поиск решения (Solver). Он посчитает доли распределения, которое принесет ожидаемый годовой доход не менее 8% с минимальным стандартным отклонением годовых доходов. Диалоговое окно Параметры поиска решения (Solver Parameters) для выполнения этих вычислений представлено на рис. 79.8.



**Рис. 79.8.** Диалоговое окно Параметры поиска решения для моделирования распределения средств между активами

В группе Анализ (Analyze) на вкладке Данные (Data) я нажимаю Поиск решения (Solver) и в диалоговом окне Параметры поиска решения выставляю вот какие настройки:

- Стандартное отклонение годовой доходности инвестиционного портфеля минимизируется в ячейке I4.
- Изменяемые ячейки — это доли распределения средств между активами (ячейки C2:E2).
- Необходимо распределить 100% денежных средств между тремя классами активов (F2=1).
- Ожидаемый годовой доход должен составить не менее 8% (I3>=0,08).
- Я предполагаю, что продажи без покрытия недопустимы, и это моделируется с помощью неотрицательных долей денежных средств в каждом классе активов. Для этого я включил флажок Сделать переменные без ограничений неотрицательными (Make Unconstrained Variables Non-Negative).
- Получено следующее распределение средств между активами с минимальным риском: 33,7% — акции, 60,7% — векселя и 5,7% — облигации. Такой портфель принесет ожидаемый годовой доход 8% и годовое стандартное отклонение 2,9%.

Предположим, вы считаете, что в ближайшие пять лет доход по акциям в среднем будет ниже на 5%, чем был в предыдущий 30-летний период. Такие ожидания можно быстро отразить в модели. (См. задание 4 в конце главы.)

## ПРИМЕЧАНИЕ

Если на вкладке Данные (Data) отсутствует надстройка Поиск решения (Solver), ее необходимо активировать. Для этого на вкладке Файл (File) выберите команду Параметры (Options), затем среди параметров Excel выберите раздел Надстройки (Add-Ins). На странице Надстройки (Add-Ins) нажмите Перейти (Go), установите флажок Поиск решения (Solver Add-In) и нажмите ОК. Подробную информацию см. в главе 29 «Введение в оптимизацию с надстройкой Поиск решения».

## Задания

Данные к заданиям 1–3 находятся в файле *GESim.xlsx*.

1. Пусть текущая цена акций Microsoft составляет \$28. Какова вероятность того, что через два года акции Microsoft будут стоить не менее \$35?
2. Выполните задание 1 снова, но теперь с предположением, что в течение следующих двух лет доход по акциям Microsoft в среднем будет на 6% в год выше, чем в период с 1997 по 2002 г., для которого имеются данные.
3. Пусть текущая цена акций Intel составляет \$20. Какова вероятность того, что в течение следующих трех лет ваш доход от приобретения акций Intel составит не менее 30% (для трехлетнего периода)?

Данные к заданиям 4–7 находятся в файле *Assetallsim.xlsx*.

4. Вы считаете, что в течение следующих пяти лет акции принесут доход в среднем на 5% в год меньше, чем по данным за 1973–2012 гг. Вычислите распределение средств между акциями, векселями и облигациями, которое принесет ожидаемый годовой доход не менее 6% с минимальным риском.
5. Вы считаете, что в два раза более вероятно, что доходы от инвестиций для каждого года из следующих пяти лет больше похожи на доходы в период 1992–2001 гг., чем на доходы в период 1972–1991 гг. Например, вероятность того, что следующий год будет похож на 1993 г., в два раза выше, чем вероятность того, что следующий год будет похож на 1980 г. Из-за этой убежденности следует в бутстрэппинге придать больший вес недавнему прошлому. Как включить это предположение в модель оптимизации портфеля?
6. Многие взаимные фонды и инвесторы хеджируют риск падения акций покупкой пут-опционов. (О пут-опционах см. в главе 82 «Ценообразование опционов».) Как с помощью модели распределения средств между активами определить оптимальную стратегию хеджирования с использованием пут-опционов?
7. Определите корреляции (опираясь на данные за 1973–2012 гг.) между годовыми доходами по акциям, векселям и облигациям. Затем определите корреляции (на основе 1000 сценариев, созданных с помощью бутстрэппинга) между окончательной стоимостью акций, векселей и облигаций. Не кажется ли вам, что бутстрэппинг позволяет быстро определить взаимозависимость между доходами по акциям, векселям и облигациям?

## ГЛАВА 80

# Развлечения и игры: моделирование вероятностей для азартных игр и спортивных соревнований

### Обсуждаемые вопросы

- Какова вероятность выиграть в крэпс?
- Какова вероятность получить «тройку»<sup>1</sup> в пятикарточном дро-покере?
- Какова вероятность выигрыша в Финале Четырех 2018 г. для каждой баскетбольной команды Национальной студенческой спортивной ассоциации?

Азартные игры и просмотр спортивных соревнований — популярный способ провести свободное время. Азартные игры и спорт вызывают такой интерес, поскольку никогда не известно, что будет дальше. Для аппроксимирования вероятностей для азартных игр и спортивных соревнований подходит такой эффективный инструмент, как моделирование по методу Монте-Карло. По сути, он оценивает вероятность путем многократного повторения игровой или соревновательной ситуации. Если, например, в Microsoft Excel сыграть в крэпс 10 000 раз и выиграть 4900 раз, то вероятность выигрыша можно будет оценить как  $4900/10\,000$ , или 49%. Если симитировать баскетбольный мужской Финал Четырех Национальной студенческой спортивной ассоциации 2018 г. 1000 раз и при этом команда У университета Вилла-нова выиграет 300 раз, то вероятность победы в чемпионате составит  $300/1000$ , или 30%.

## Ответы на вопросы

### ❓ Какова вероятность выиграть в крэпс?

По правилам этой игры игрок бросает два кубика. Если выпадет 2, 3 или 12, то он проиграл. Если выпадет 7 или 11 — выиграл. Если выпала другая комбинация, игрок продолжает бросать кости до тех пор, пока не выпадет число очков, выпавшее при первом броске (называемое «пойнт»), или 7. Если «пойнт» выпало рань-

---

<sup>1</sup> Три карты одного ранга. — *Примеч. ред.*

ше 7, игрок выиграл. Если 7 выпало раньше — проиграл. Путем сложных вычислений можно показать, что вероятность выигрыша в крэпс составляет 0,493. Для аппроксимирования этой вероятности можно использовать Excel и симитировать игру в крэпс множество раз (я выбрал 2000 раз).

Важно помнить, что в этой игре заранее неизвестно, сколько бросков нужно будет сделать. Я покажу, что вероятность того, что игра потребует выбросить кости более 50 раз, крайне маловероятна, поэтому предположим, что игра состоит из 50 бросков. После каждого броска состояние игры отслеживается следующим образом:

- 0 — игра проиграна;
- 1 — игра выиграна;
- 2 — игра продолжается.

Выходная ячейка содержит состояние игры после 50 бросков, где 1 обозначает выигрыш, а 0 — проигрыш. Результаты проделанной мной работы вы найдете в файле Craps.xlsx (рис. 80.1).

	A	B	C	D	E	F	G	H	AX	AY
1	TOSS#	1	2	3	4	5	6	7	49	50
2	Die Toss 1	4	4	3	4	2	2	3	6	3
3	Die Toss 2	6	5	2	2	1	1	5	4	1
4	Total	10	9	5	6	3	3	8	10	4
5	GAME STATUS	2	2	2	2	2	2	2	0	0
6		WIN??	0							
7										
8				prob win	0.4895					
9					0					
10				1	1					
11				2	0					
12				3	1					
13				4	0					
14				5	1					
15				6	0					
16				7	1					
17				8	0					
18				9	1					
19				10	0					
20				11	1					

**Рис. 80.1.** Имитация игры крэпс

В ячейке B2 я использовал функцию СЛУЧМЕЖДУ (RANDBETWEEN) и сгенерировал число, выпавшее на первой кости после первого броска по формуле =СЛУЧМЕЖДУ(1;6). Функция СЛУЧМЕЖДУ обеспечивает равную вероятность для каждого возвращаемого значения, так что на каждой кости с равной вероятностью (1/6) может выпасть 1, 2, 3, 4, 5 или 6. Скопированная в диапазон B2:AY3, эта формула генерирует 50 бросков. (На рис. 80.1 броски с 8 по 48 скрыты.)

В ячейках B4:AY4 я вычислил число возможных комбинаций для каждого из 50 бросков, скопировав формулу =СУММ(B2:B3) из B4 в C4:AY4. В ячейке B5 я определил



состояние игры после первого броска по формуле `=ЕСЛИ(ИЛИ(B4=2;B4=3;B4=12);0; ЕСЛИ(ИЛИ(B4=7;B4=11);1;2))`. Напомню, что 2, 3 или 12 означает проигрыш (в ячейку записывается 0), 7 или 11 — выигрыш (1), а любой другой результат ведет к продолжению игры (2).

В ячейке C5 я вычислил состояние игры после второго броска по формуле `=ЕСЛИ(ИЛИ(B5=0;B5=1);B5;ЕСЛИ(C4=$B4;1;ЕСЛИ(C4=7;0;2)))`. Если игра заканчивается после первого броска, записывается соответствующее состояние игры. Далее, если выпадает «пойнт», записывается выигрыш (1). Если выпадает 7, записывается проигрыш. В остальных случаях игра продолжается. В этой формуле знак доллара в ссылке на ячейку в столбце B (`$B4`) обеспечивает сравнение выпавших после каждого броска очков с количеством очков на первом броске. Эта формула, скопированная из C5 в D5:AY5, записывает состояние игры после бросков 2–50.

Результат игры, находящийся в ячейке AY5, я скопировал в ячейку C6 для удобства. Затем я воспользовался однонаправленной таблицей данных для имитации игры в крэпс 2000 раз. В ячейку E9 я ввел формулу `=C6`, отслеживающую исход игры (0 в случае проигрыша или 1 в случае выигрыша). Затем выделил таблицу D9:E2009) и на вкладке Данные (Data) в группе Прогноз (Forecast) в списке Анализ "что, если" (What If Analysis) выбрал Таблица данных (Data Table). Здесь для однонаправленной таблицы данных в поле Подставлять значения по строкам в (Column Input Cell) можно указать любую пустую ячейку. После нажатия F9 Excel симулирует крэпс 2000 раз.

В ячейке E8 по формуле `=СРЗНАЧ(E10:E2009)` можно вычислить долю имитаций с выигрышем. Из 2000 итераций выигрышными были 49%. При большем числе испытаний (например, в случае 10 000 итераций) процент выигрыша приблизится к истинной вероятности выигрыша в крэпс (49,3%). (Об однонаправленной таблице данных см. в главе 17 «Анализ чувствительности с помощью таблиц данных».)

### **❓ Какова вероятность получить «тройку» в пятикарточном дро-покере?**

Стандартная колода карт содержит по четыре карты каждого ранга, от четырех двоек до четырех тузов. Для вычисления вероятности конкретной покерной комбинации можно присвоить тузу значение 1, двойке значение 2 и т. д. до валета, которому присваивается значение 11, даме присваивается значение 12 и королю — значение 13.

В пятикарточном дро-покере вы получаете на руки пять карт. Здесь представляют интерес многие вероятности, но ограничимся оценкой вероятности «тройки», то есть у вас на руках должно быть три карты одного ранга, а оставшиеся две не должны образовывать пару — в противном случае у вас на руках будет «фул-хаус». Для имитации игры в пятикарточный дро-покер необходимо выполнить следующие шаги. (См. файл `Poker.xlsx` и рис. 80.2.)

1. Свяжите с каждой картой в колоде случайное число.
2. Пять выбранных карт будут пятью картами, связанными с самыми маленькими случайными числами, благодаря чему каждая карта может быть выбрана с равной вероятностью.



3. Подсчитайте, сколько карт каждого ранга (от туза до короля) выбрано.

	A	B	C	D	E	F	G	H	I	J
2		Drawn hand	Rank		Rand#				How many	
3	1	6	25	1	0.3476				1	0
4	2	8	13	1	0.1784				2	0
5	3	6	30	1	0.4519				3	0
6	4	13	9	1	0.1242				4	1
7	5	4	14	2	0.1823				5	0
8			42	2	0.6796				6	2
9			46	2	0.826				7	0
10			27	2	0.3703				8	1
11			39	3	0.6273				9	0
12			16	3	0.2044				10	0
13			6	3	0.0864				11	0
14			29	3	0.45				12	0
15			47	4	0.829				13	1
16			44	4	0.7738					
17			5	4	0.0673				three of kind?	0
18			19	4	0.2566					
19			22	5	0.2794					0
20			48	5	0.8309		prob 3 of kind		1	0
21			20	5	0.2642		0.0225		2	0
22			51	5	0.9694				3	0
23			1	6	0.0019				4	0
24			3	6	0.0259				5	0
25			40	6	0.6376				6	0
26			33	6	0.4731				7	0
27			35	7	0.5192				8	0
28			10	7	0.1426				9	0
29			37	7	0.5658				10	0

**Рис. 80.2.** Оценка вероятности «тройки» в покере

Сначала я перечислил все карты колоды в ячейках D3:D54: четыре единицы (тузы), четыре двойки (2) и т. д. до четырех чисел 13 (представляющих королей). Для связывания каждой карты со случайным числом я скопировал формулу =СЛЧИС() из E3 в E4:E54. Для получения ранга (от меньшего к большему) каждого случайного числа я скопировал формулу =РАНГ.РВ(E3;\$E\$3:\$E\$54;1) из C3 в C4:C54. Например, как показано на рис. 80.2, первая тройка в колоде (ряд 11) связана с 38-м наименьшим случайным числом. (Вы увидите на листе другие результаты, поскольку при открытии листа случайные числа автоматически пересчитываются.)

Функция РАНГ.РВ (RANK.EQ) имеет синтаксис РАНГ.РВ(число;массив;1\_или\_0). Если последний аргумент функции РАНГ.РВ равен 1, функция возвращает ранг числа в массиве, при этом наименьшее число получает ранг 1, второе наименьшее число — ранг 2 и т. д. Если последний аргумент функции РАНГ.РВ равен 0, функция возвращает ранг числа в массиве, при этом наибольшее число в массиве получает ранг 1, второе наибольшее число — ранг 2 и т. д.

В рейтинге случайных чисел не может возникнуть никаких связей, потому что тогда случайные числа должны были бы совпасть по 16 цифрам.

Предположим, например, что ранг присваивается числам 1, 2, 3 и 4 и что последний аргумент функции РАНГ.РВ равен 1. Функция Excel возвратит следующие ранги:

Число	Ранг числа (наименьшее число имеет ранг 1)
1	1
2	2
3	2
4	4

Поскольку число 3 является вторым наименьшим числом, ему будет присвоен ранг 2. Второму числу 3 тоже присвоен ранг 2. Поскольку 4 является четвертым наименьшим числом, ему присвоен ранг 4. Без понимания того, как функция РАНГ.РВ обрабатывает связи, вы не выполните задание 1 в конце главы.

Скопируйте формулу =ВПР(А3;lookup;2;ЛОЖЬ) из ячейки В3 в В4:В7 для получения пяти карт из колоды. По этой формуле из колоды выбираются пять карт, соответствующие пяти наименьшим случайным числам. (Имя lookup присвоено таблице в ячейках С3:Д54.) В функции ВПР указан аргумент ЛОЖЬ, поскольку ранги не должны быть в порядке возрастания.

После присвоения имени drawn выбранным картам (В3:В7) определите по формуле =СЧЁТЕСЛИ(drawn;I3), скопированной из J3 в J4:J15, какие карты находятся на руках. В ячейке J17 по формуле =ЕСЛИ(И(МАКС(J3:J15)=3;СЧЁТЕСЛИ(J3:J15;2)=0);1;0) определите, есть ли на руках «тройка». Формула возвращает 1 тогда и только тогда, когда на руках имеется «тройка» и нет пары.

Теперь с помощью однонаправленной таблицы данных смоделируйте 4000 покерных раздач. Скопируйте результат из ячейки J17 по формуле =J17 в ячейку J19. Выделите таблицу (I19:J4019). Затем на вкладке Данные (Data) в группе Прогноз (Forecast) в списке Анализ "что, если" (What If Analysis) выберите Таблица данных (Data Table). Определите однонаправленную таблицу, указав в поле Подставлять значения по строкам в (Column Input Cell) любую пустую ячейку. После нажатия ОК будут смоделированы 4000 покерных раздач. В ячейке G21 вычислите вероятность «тройки» по формуле =СРЗНАЧ(J20:J4019). Вероятность составляет 2,0%. (С помощью теории вероятностей можно показать, что истинная вероятность собрать «тройку» равна 2,1%.)

**?** Какова вероятность выигрыша в Финале Четырех 2018 г. для каждой баскетбольной команды Национальной студенческой спортивной ассоциации?

Рейтинг баскетбольных студенческих команд можно оценить по методике, описанной в главе 35 «Поиск решения при оценке спортивных команд», но сложно получить результаты всех прошлых игр. Пользующиеся большим уважением рейтинги Сагарина (см. сайт *sagarin.com*) всегда предоставляют актуальные рейтинги для всех студенческих баскетбольных команд. Накануне Финала Четырех 2018 года рейтинги Сагарина были следующими: 96 для Университета Вилланова; 92 для Канзасского университета; 89 для Университета Мичигана и 82 для

Университета Лойолы. На основе этой информации для оценки шансов каждой команды на выигрыш можно сыграть Финал Четырех несколько тысяч раз.

Средний прогноз для количества очков, которое выиграет команда хозяев, равен рейтинг фаворита – рейтинг аутсайдера. В Финале Четырех нет команды хозяев, но если бы она была, к рейтингу команды хозяев следовало бы добавить пять очков. (В профессиональном баскетболе хозяевам добавляют три очка, и в студенческом и профессиональном футболе — три очка.) Затем с помощью функции НОРМ.ОБР (NORM.INV) можно смоделировать результат каждой игры. (О моделировании нормальной случайной величины с помощью функции НОРМ.ОБР см. в главе 72 «Нормальная случайная величина и Z-оценка».)

Вероятный результат Финала Четырех 2018 года я вычислил в файле Final4sim.xlsx (рис. 80.3). В полуфиналах встретились «Канзас» — «Вилла-нова» и «Мичиган» — «Лойола».

	C	D	E	F	G	H	I	J	K	L	M
1											
2			Semifinals								
3		Rating	Outcome	rand#							
4	Kansas	92	-13.061248	0.182435	Final	rand#	Forecast	Outcome	Winner		
5	Villanova	96			Villanova	0.1336	96	-4.094951	Michigan		
6					Michigan		89				
7											
8	Michigan	89	16.411573	0.826688							
9	Loyola	82									
10											
11							odds		PROB WIN		Michigan
12							3.329004	Kansas	23.10%	1	Michigan
13							0.890359	Villanova	52.90%	2	Kansas
14							3.494382	Michigan	22.25%	3	Michigan
15							56.14286	Loyola	1.75%	4	Villanova
16										5	Michigan
17										6	Michigan
18										7	Villanova
19										8	Kansas
20										9	Villanova
21										10	Villanova
22										11	Kansas

**Рис. 80.3.** Моделирование результата Финала Четырех 2018 года Национальной студенческой спортивной ассоциации

Сначала я ввел в ячейки C4:D5 и C8:D9 названия и рейтинги всех команд. В ячейку F4 с помощью функции СЛЧИС (RAND) я ввел случайное число для игры «Вилла-нова — Канзас» и в ячейку F8 — случайное число для игры «Мичиган — Лойола». Смоделированный результат всегда указывается относительно первой команды.

В ячейке E4 я определил результат игры «Канзас — Вилла-нова» (с позиции «Канзаса») по формуле =НОРМ.ОБР(F4;D4-D5;10). Обратите внимание, что «Канзас» является фаворитом по разнице очков D4-D5. В ячейке E8 я определил результат игры «Мичиган — Лойола» (с позиции «Мичигана») по формуле =НОРМ.ОБР(F8;D8-D9;10). (Стандартное отклонение для преимущества в студенческих баскетбольных играх по количеству очков составляет 10 очков.)

В ячейках G5 и G6 обеспечивается проход в финал победителя каждой полуфинальной игры. Если результат полуфинала больше 0, выиграла команда, указанная первой; в противном случае выиграла вторая команда. Таким образом, в ячейку G5 вводится победитель первой игры по формуле =ЕСЛИ(E4>0;C4;C5). В ячейку G6 вводится победитель второй игры по формуле =ЕСЛИ(E8>0;C8;C9).

Для моделирования результата игры за чемпионский титул я ввел в ячейку H5 случайное число. Скопировав формулу =ВПР(G5;\$C\$4:\$D\$9;2;ЛОЖЬ) из I5 в I6, я получил рейтинг каждой команды в чемпионской игре. Затем в ячейке J5 я вычислил результат финала (с позиции первой указанной команды в ячейке G5) по формуле =НОРМ.ОБР(H5;I5-I6;10). Наконец, в ячейке K5 определен фактический чемпион по формуле =ЕСЛИ(J5>0;G5;G6).

Теперь я могу использовать однонаправленную таблицу данных для моделирования Финала Четырех 2000 раз. Смоделированные победители указаны в диапазоне M12:M2011. Формула =СЧЁТЕСЛИ(\$M\$12:\$M\$2011;J12)/2000, скопированная из K12 в K13:K15, вычисляет прогнозируемую вероятность победы каждой команды: 23% для «Канзаса», 53% для «Вилла-новы», 22% для «Мичигана» и 15% для «Лойолы». Эти вероятности могут быть переведены в букмекерские ставки по следующей формуле:

$$\text{перевес} = \frac{\text{вероятность\_поражения\_команды}}{\text{вероятность\_победы\_команды}}.$$

Например, ставки на проигрыш «Канзаса» составляют 3,33 к 1:

$$(1 - 0,23) / 0,23 = 3,333.$$

То есть если вы поставили на победу «Канзаса» \$1, и букмекер заплатил \$3,33, то ставка справедливая. Разумеется, букмекер слегка понизит ставки, чтобы не остаться без заработка. (См. задание 6 в конце главы.)

Эта методика может быть легко расширена до моделирования всего турнира Национальной студенческой спортивной ассоциации. Просто используйте утверждения ЕСЛИ для обеспечения продвижения каждого победителя и функции ВПР для нахождения рейтинга каждой команды. Моделирование сезона 2003 г см. в файле Ncaa2003.xlsx. Анализ отдавал команде Сиракузского университета (действительному чемпиону) победу с вероятностью 4% на старте соревнований.

На этом листе я использовал примечания к работе, показанные на рис. 80.4. Вот несколько советов по использованию примечаний.

- Для вставки примечания в ячейку сначала выделите ячейку . На вкладке Рецензирование (Review) в группе Примечания (Comments) выберите Создать примечание (New Comment) и введите текст примечания. Вы увидите небольшую красную отметку в правом верхнем углу ячейки с примечанием.
- Для редактирования примечания щелкните правой кнопкой мыши на ячейке и выберите команду Изменить примечание (Edit Comment).

- Чтобы примечание всегда было видимым, щелкните на ячейке правой кнопкой мыши и выберите **Показать или скрыть примечание (Show/Hide Comments)**. Если отображаемое примечание необходимо скрыть, выберите **Скрыть примечание (Hide Comment)**. Оно будет отображаться только при наведении курсора на ячейку с примечанием.
- Если вы хотите распечатать примечания, на вкладке **Разметка страницы (Page Layout)** щелкните на кнопке вызова диалогового окна **Параметры страницы (Page Setup)** — на маленьком значке в правом нижнем углу группы. В открывшемся диалоговом окне на вкладке **Лист (Sheet, в разделе Печать)** в поле со списком **Примечания (Comments)** можно указать, должны ли примечания быть напечатаны так, как на листе, или внизу листа. Нажмите **ОК** для сохранения настроек.

	A	B	C	D	E	F	G
1		1		4	5	6	7
2	Arizona	V	Tech. Services: enter teams in each region in order of draw sheet in paper; 1 seed plays 16, then 8 plays 9 etc.	DNZ	ND	WIS MIL	ILL
3		1		4	5	6	7
4	93.1			84.18	88.52	82.71	88.8
5	West						
6		1			3		4
7	31.20983				-12.4284	4	
8							
9		1	Tech. Services: enter Sagarin ratings at start of tourney in row 4.		5		7
10	10.06963				-1.98011	7	
11							
12		1		7		9	
13		12.79732	1			-8.87394	15
14							
15			1		15		
16			0.817106	1			

**Рис. 80.4.** Показ примечаний на листе

## Задания

1. В комнате находится 30 человек. Какова вероятность того, что по крайней мере двое из них родились в один день?
2. Какова вероятность собрать одну пару в пятикарточном дро-покере?
3. Какова вероятность собрать две пары в пятикарточном дро-покере?
4. В игре кено перемешивают 80 шаров (пронумерованных от 1 до 80) и затем случайным образом вытаскивают 20 шаров. Перед вытаскиванием 20 шаров игрок выбирает 10 различных номеров. Если вытасчено хотя бы пять шаров с этими номерами, игрок выигрывает. Какова вероятность выигрыша?
5. Для финала НБА 2003 года рейтинговая система (см. главу 35) поставила клуб «Сан-Антонио Спёрс» на три очка выше клуба «Нью-Джерси Нетс». Команды играют до четырех побед. Первые две игры проводятся в Сан-Антонио, следу-

- ющие три игры — в Нью-Джерси и две последние игры запланированы в Сан-Антонио. Какова вероятность того, что «Сан-Антонио» выиграет серию?
6. Какие ставки на победу «Канзаса» в Финале Четырех должен предложить букмекер, если хочет заработать в среднем 10 центов с каждого доллара ставки?
  7. Какова вероятность собрать флэш (пять карт одной масти) в пятикарточном дро-покере?
  8. Если палочку длиной 1 дюйм сломать в двух случайно выбранных местах, то получится три маленькие палочки. Какова вероятность того, что из них можно составить треугольник?
  9. Предположим, что каждый хитер в бейсбольной команде может с вероятностью 50% выполнить хоумран и с вероятностью 50% выбить в аут . Сколько в среднем перебежек заработает эта команда в одном иннинге?
  10. Кэти и Тэйлор подбрасывают монету 100 раз. Каждый раз, когда выпадает орел, Кэти платит Тэйлору \$1; каждый раз, когда выпадает решка, Тэйлор платит Кэти \$2. Какова вероятность того, что Тэйлор ни разу не будет в выигрыше?
  11. Вы подбрасываете правильную монету 20 раз. Оцените вероятность того, что вы выбросите не менее четырех орлов подряд.
  12. В начале 2015 года вы вложили \$100, купив акции. Каждый год с вероятностью 60% стоимость акций растет на 50%, с вероятностью 20% стоимость акций растет на 10% и с вероятностью 20% стоимость акций падает на 50%. Какова вероятность, что до начала 2065 г. стоимость вашего актива достигнет (хотя бы раз) \$10 000?
  13. При игре в блэкджек дилер берет карты, пока сумма очков на них не будет больше или равна 16. После этого он останавливается. Предположим, туз считается за 1 очко, а валет, дама и король — за 10. Если дилер вытаскивает карты с возвратом, какова вероятность, что общая сумма очков будет 19, 20 или 21?
  14. Вы собираетесь бросать кубик 50 раз. Если выпадает четное число очков, вам выплачивается столько долларов, сколько очков выпало на кубике. Если выпало нечетное число очков, вы теряете \$1. Какова вероятность того, что ваш выигрыш в какой-то момент будет не менее \$75?
  15. Возьмем три колоды карт. Первая колода содержит 10 карт: 3 двойки, 4 пятерки и 3 девятки. Вторая колода содержит четыре карты с числами от 1 до 4. Колода 3 содержит шесть карт с числами от 1 до 6. Грегори вытаскивает одну карту из колоды 1, Тэйлор вытаскивает одну карту из колоды 2 и одну карту из колоды 3. Какова вероятность того, что на карте Грегори число больше, чем сумма чисел на двух картах Тэйлора?

## ГЛАВА 81

# Анализ данных с помощью повторной выборки

### Обсуждаемый вопрос

- Девять партий продукции производятся при высокой температуре, а семь партий — при низкой. Какова вероятность того, что выход продукта выше при высокой температуре?

Аналитикам данных часто приходится решать следующие задачи:

- Какова вероятность того, что новая методика обучения повысит успеваемость учащихся?
- Какова вероятность того, что аспирин снижает частоту инфарктов?
- Какова вероятность того, что машина 1 является самой производительной из трех машин?

Анализировать подобные данные позволяет простой, но эффективный метод, известный как *повторная выборка*. Для получения статистических выводов с использованием повторной выборки данные генерируются множество раз путем выборки с возвращением. Выборка с возвращением означает, что одна и та же точка данных может быть выбрана несколько раз. Затем на основе результатов повторной выборки делаются выводы. Ключевым инструментом для реализации повторной выборки является функция `СЛУЧМЕЖДУ (RANDBETWEEN)`. Формула `=СЛУЧМЕЖДУ(a;b)` дает с равной вероятностью любое целое число между *a* и *b* (включительно). Например, формула `=СЛУЧМЕЖДУ(1;9)` с равной вероятностью даст в результате одно из чисел от 1 до 9 (включительно).

## Ответ на вопрос

- ② Девять партий продукции производятся при высокой температуре, а семь партий — при низкой. Какова вероятность того, что выход продукта выше при высокой температуре?

В файле `Resampleyield.xlsx` (рис. 81.1) содержатся данные о выходе продукта для девяти партий, произведенных при высокой температуре, и для семи партий, произведенных при низкой температуре.



Средний выход продукта при высокой температуре составляет 39,74; средний выход продукта при низкой температуре — 32,27. Однако эта разница не доказывает, что средний выход продукта при высокой температуре превосходит средний выход продукта при низкой температуре. На основе данных выборки необходимо узнать вероятность того, что выход продукта при высокой температуре выше, чем при низкой. Для ответа на этот вопрос можно сгенерировать случайным образом девять целых чисел от 1 до 9, что дает повторную выборку выхода продукта при высокой температуре. Например, если сгенерировано случайное число 4, в повторно выбранные данные для выхода годной продукции при высокой температуре будет включено значение 41,40 и т. д. Затем случайным образом генерируются семь целых чисел от 1 до 7, с помощью которых создается повторная выборка для выхода продукта при низкой температуре. Далее для повторно выбранных данных проверяется, превосходит ли средний выход продукта, произведенного при высокой температуре, средний выход продукта, произведенного при низкой температуре. И, наконец, с помощью таблицы данных этот процесс повторяется несколько сотен раз. (Я повторил 400 раз.) В повторно выбранных данных доля случаев, когда средний выход продукта при высокой температуре превосходит средний выход продукта при низкой температуре, оценивает вероятность того, что процесс производства при высокой температуре эффективнее процесса производства при низкой температуре.

	C	D	E
1			
2			
3	mean	39.74	32.27
4	Number	High Temp	Low Temp
5	1	17.55	37.37
6	2	39.93	43.20
7	3	48.98	34.85
8	4	41.40	31.22
9	5	35.70	26.59
10	6	42.24	42.67
11	7	54.75	10.00
12	8	46.96	
13	9	30.19	

**Рис. 81.1.** Выход продукта при высокой и низкой температуре

Сначала сгенерируйте повторную выборку данных для высокой температуры, скопировав формулу =СЛУЧМЕЖДУ(1;9) из ячейки C16 в C17:C24, как показано на рис. 81.2. Исходные наблюдения могут быть выбраны несколько раз или не выбраны ни разу. Скопируйте формулу =ВПР(C16;lookup;2) из ячейки D16 в D17:D24 (lookup — это имя диапазона ячеек C4:E13) для генерации выходов продукта, соответствующих случайной повторной выборке данных. Далее сгенерируйте повторную выборку из данных о выходе продукта при низкой температуре. Скопируйте формулу =СЛУЧМЕЖДУ(1;7) из E16 в E17:E22 для генерации повторной выборки из семи случайных чисел для получения наблюдений на основе исходных данных



о выходе продукта при низкой температуре. Для повторного выбора семи выходов продукта при низкой температуре скопируйте формулу =ВПР(E16;lookup;3) из F16 в F17:F22.

	C	D	E	F
15	Resampled	High Temp	Resampled	Low Temp
16	4	41.40	5	26.59
17	8	46.96	1	37.37
18	4	41.40	5	26.59
19	2	39.93	7	10.00
20	7	54.75	6	42.67
21	7	54.75	4	31.22
22	6	42.24	2	43.20
23	3	48.98		
24	8	46.96		
25				
26	Average	46.38		31.09
27				
28				
29	High better	1		
30				Prob high better
31		1		0.9225
32	1	1		
33	2	1		
34	3	1		
35	4	1		
36	5	1		
37	6	1		
38	7	1		
39	8	1		
40	9	1		
41	10	1		
42	11	1		
43	12	1		

**Рис. 81.2.** Реализация повторной выборки

В ячейке D26 я вычислил для повторно выбранных данных средний выход при высокой температуре по формуле =СРЗНАЧ(D16:D24). Аналогично в ячейке F26 я вычислил средний выход для низкой температуры по формуле =СРЗНАЧ(F16:F22). В ячейке D29 по формуле =ЕСЛИ(D26>F26;1;0) я определил для повторно выбранных данных, превосходит ли средний выход продукта при высокой температуре средний выход продукта при низкой температуре.

Провести повторную выборку 400 раз можно с помощью однонаправленной таблицы данных. Я поместил номера итераций от 1 до 400 в ячейки C32:431. О том, как быстро создать список номеров итераций с помощью команды Заполнить (Fill) и диалогового окна Прогрессия (Series), см. в главе 77 «Введение в моделирование по методу Монте-Карло». Я ввел в ячейку D31 =D29 и создал формулу, записывающую в выходную ячейку таблицы данных то, превосходит ли средний выход продукта при высокой температуре средний выход продукта при низкой температуре. Выделите таблицу (C31:D431), затем на вкладке Данные (Data) в группе Прогноз (Forecast) в списке Анализ "что, если" (What If Analysis) выберите Таблица данных (Data Table). Укажите в поле Подставлять значения по строкам в (Column Input Cell) любую пустую ячейку. Теперь Excel выполнит повторную выборку 400 раз.

Каждая итерация со значением 1 указывает, что в повторной выборке средний выход продукта выше при высокой температуре. Каждая итерация со значением 0 указывает, что в повторной выборке средний выход продукта выше при низкой температуре. В ячейке F31 по формуле `=СРЗНАЧ(D32:D431)` я определил долю случаев, когда средний выход продукта выше при высокой температуре. На рис. 78.2 применение повторной выборки показывает, что с вероятностью 94% средний выход продукта при высокой температуре выше, чем при низкой. Разумеется, при нажатии F9 будет сгенерирован другой набор из 400 повторных выборок, который даст немного другую оценку вероятности того, что средний выход продукта при высокой температуре выше, чем при низкой.

## Задания

1. Тестируется новое лекарство от гриппа. Из 24 больных гриппом пациентов, получивших новое лекарство, 20 почувствовали себя лучше, а 4 — хуже. Из 9 больных гриппом пациентов, получивших плацебо, 6 почувствовали себя лучше, а 3 — хуже. Какова вероятность того, что лекарство эффективнее плацебо?
2. С восьмью рабочими была проведена беседа о вреде повышенного холестерина. У каждого рабочего до беседы и через какое-то время после нее был проверен уровень холестерина. Результаты проверки приведены в таблице. Какова вероятность того, что рабочие стали питаться правильнее после беседы и их уровень холестерина снизился?

Холестерин до беседы	Холестерин после беседы
220	210
195	198
250	210
200	199
220	224
260	212
175	179
198	184

3. Бета для акции — это значение наклона наилучшей подобранной линии, используемой для прогнозирования ежемесячного дохода от акции по ежемесячному доходу, приведенному в индексе Standard&Poor (S&P) 500. Бета больше 1 показывает, что цикличность цены акции выше, чем цикличность цен на рынке, а бета меньше 1 показывает, что цикличность цены акции ниже,

чем цикличность цен на рынке. Файл `Betaresampling.xlsx` содержит данные более чем за 12 лет о ежемесячном доходе по акциям Microsoft (MSFT), Pfizer (PFE) и другим акциям, а также индекс S&P 500. Исходя из этих данных, определите вероятность того, что бета для Microsoft ниже беты для Pfizer. При оценке беты для каждой итерации повторной выборки воспользуйтесь функцией `НАКЛОН (SLOPE)`.

4. В файле `Lawdata.xlsx` даны баллы юридического теста LSAT и средний бал GPA для 15 студентов юридической школы. На основе этих данных определите, в каких пределах находится корреляция между баллами LSAT и GPA с вероятностью 95%?

## ГЛАВА 82

# Ценообразование опционов

### Обсуждаемые вопросы

- Что такое колл-опцион и пут-опцион?
- В чем разница между американским и европейским опционами?
- Как выглядят выигрыши по европейским колл- и пут-опционам, если рассматривать их как функцию курса акций на дату исполнения?
- Какими параметрами определяется цена опциона?
- Как оценить волатильность акций на основе исторических данных?
- Как реализовать в Excel формулу Блэка — Шоулза?
- Как изменение ключевых параметров изменяет цену колл- и пут-опционов?
- Как оценить волатильность акций по формуле Блэка — Шоулза?
- В мои точные формулы ценообразования опционов никто не должен вносить изменения. Как защитить формулы на листе от изменений?
- Как с помощью ценообразования опционов компания может принимать лучшие инвестиционные решения?

В начале 1970-х гг. экономисты Фишер Блэк (Fischer Black), Майрон Шоулз (Myron Scholes) и Роберт Мертон (Robert Merton) вывели формулу ценообразования опционов Блэка — Шоулза, которая позволяет получить оценку европейских колл- и пут-опционов. За свою работу Шоулз и Мертон были удостоены Нобелевской премии по экономике в 1997 г. (Блэк умер до 1997 г., а Нобелевская премия не присуждается посмертно.) Их научный труд произвел революцию в области корпоративных финансов. В этой главе я расскажу об основных положениях их важной научной работы.

### ПРИМЕЧАНИЕ

Превосходное формальное исследование опционов см. в книге: David G. Luenberger, Investment Science, 2nd ed. — Oxford University Press, 2013.

## Ответы на вопросы

### ? Что такое колл-опцион и пут-опцион?

*Колл-опцион* дает владельцу опциона право купить акцию по цене исполнения.  
*Пут-опцион* дает владельцу опциона право продать акцию по цене исполнения.

❓ В чем разница между американским и европейским опционами?

*Американский опцион* может быть исполнен не позднее даты, известной как *дата исполнения* (часто называемая *датой истечения срока действия* опциона). *Европейский опцион* может быть исполнен только в дату исполнения.

❓ Как выглядят выигрыши по европейским колл- и пут-опционам, если рассматривать их как функцию курса акций на дату исполнения?

Давайте рассмотрим денежные потоки для шестимесячного европейского колл-опциона на акции IBM с ценой исполнения \$110. Пусть  $P$  — это курс акций IBM через шесть месяцев. Прибыль от колл-опциона на эти акции составит \$0, если  $P \leq 110$ , и  $(P - 110)$ , если  $P > 110$ . Если  $P$  меньше \$110, владелец опциона может не исполнять опцион. Если  $P$  больше \$110, владелец может исполнить опцион, купив акции по \$110 и немедленно продав их по  $P$ , и тем самым получить прибыль  $(P - 110)$ . На рис. 82.1 представлена прибыль от этого колл-опциона. Коротко говоря, колл-опцион платит \$1 за каждый доллар, на который курс акций превышает цену исполнения. Прибыль от колл-опциона можно записать как  $\max(0, P - 110)$ . Обратите внимание, что график колл-опциона на рис. 82.1 (см. лист *Call* в файле *Optionfigures.xlsx*) имеет наклон 0 для значений  $P$  меньше цены исполнения. Наклон графика равен 1 для значений  $P$  больше цены исполнения.

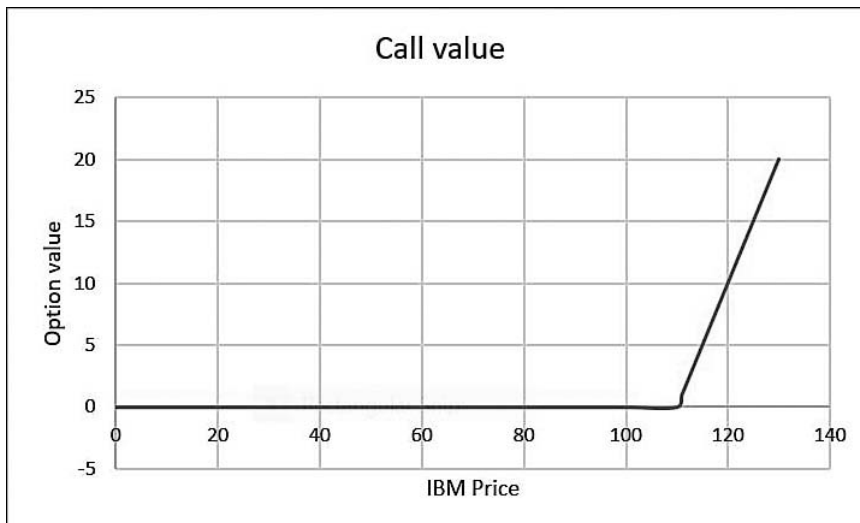
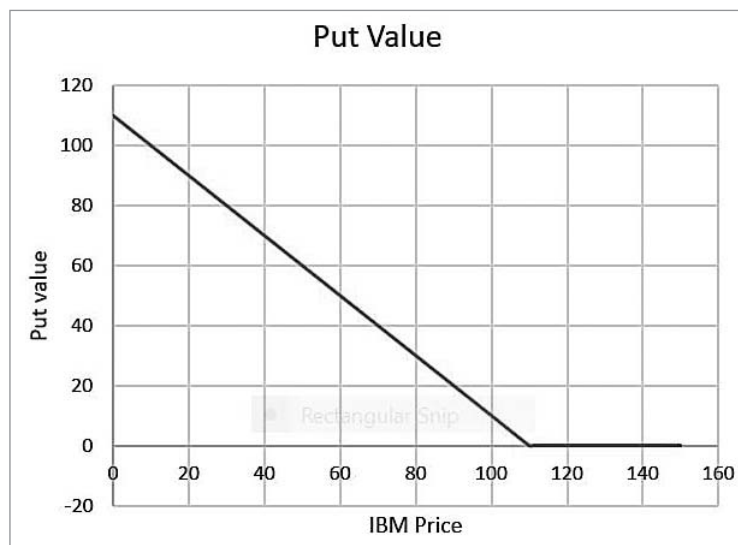


Рис. 82.1. Денежные потоки колл-опциона

Можно показать, что если акции не приносят дивидендов, нецелесообразно исполнять американский колл-опцион досрочно. Таким образом, цена американского и европейского колл-опционов для акций, по которым не выплачиваются дивиденды, одинакова.

Теперь рассмотрим денежные потоки для шестимесячного европейского пут-опциона на акции IBM с ценой исполнения \$110. Пусть  $P$  — курс акций IBM через шесть месяцев. Прибыль от пут-опциона составит \$0, если  $P \geq 110$ , и  $(P - 110)$ , если  $P < 110$ . Для  $P$  меньше \$110 владелец опциона может купить акции за  $P$  и сразу продать их за \$110 и тем самым получить прибыль  $(110 - P)$ . Если  $P$  больше \$110, покупать акции по  $P$  и продавать за \$110 невыгодно, так что владелец может не исполнять опцион для продажи акций по \$110.

На рис. 82.2 показана прибыль от этого пут-опциона (см. лист **Put** в файле **Optionfigures.xlsx**). Коротко говоря, пут-опцион платит \$1 за каждый доллар, на который курс акций ниже цены исполнения опциона. Прибыль от пут-опциона можно записать как  $\max(0, 110 - P)$ . Обратите внимание, что график пут-опциона имеет наклон  $-1$  для значений  $P$  меньше цены исполнения и 0 для значений  $P$  больше цены исполнения.



**Рис. 82.2.** Денежные потоки пут-опциона

Американский пут-опцион может быть исполнен досрочно, поэтому денежные потоки для американского пут-опциона нельзя определить без знания цены акций в период до даты истечения срока действия опциона.

### **?** Какими параметрами определяется цена опциона?

В модели ценообразования опционов Блэка — Шоулза Блэк, Шоулз и Мертон показали, что цена колл- и пут-опциона зависит от следующих параметров:

- ☐ Текущая цена акций.
- ☐ Цена исполнения опциона.

- Время (в годах) до истечения срока действия опциона (называемое *сроком* опциона).
- Процентная ставка (в год с учетом сложных процентов) на безрисковые инвестиции (как правило, в казначейские векселя) в течение всего срока инвестирования. Эта ставка называется *безрисковой ставкой*. Например, если трехмесячные казначейские векселя приносят 5% дохода, безрисковая ставка вычисляется как  $\ln(1 + 0,05)$ . (Взятие логарифма преобразует простую процентную ставку в ставку с учетом сложных процентов.) Сложные проценты означают, что в каждый момент времени проценты приносят проценты.
- Годовая ставка (в процентах от курса акций), по которой выплачиваются дивиденды. Если акции приносят каждый год 2% от своей стоимости в качестве дивидендов, ставка дивиденда составляет 0,02.
- *Волатильность* (нестабильность) акций (измеряемая на ежегодной основе). Годовая волатильность, например 30%, означает, что (приблизительно) стандартное отклонение относительного годичного изменения курса акций составит примерно 30%. Во время пузыря доткомов в конце 1990-х гг волатильность акций многих интернет-компаний превышала 100%. Мы рассмотрим два способа вычисления этого важного параметра.

В формуле ценообразования Блэка — Шоулза курс акций должен соответствовать логарифмически нормальной случайной величине (подробнее о ней см. в главе 75 «Логарифмически нормальная случайная величина в моделировании курса акций»).

### ❓ Как оценить волатильность акций на основе исторических данных?

Для оценки волатильности акций на основе данных о ежемесячной доходности акций нужно предпринять следующее.

- Определить ежемесячную доходность акций за несколько лет.
- Определить для каждого месяца  $\ln(1 + \text{ежемесячный доход})$ .
- Определить стандартное отклонение для  $\ln(1 + \text{ежемесячный доход})$ . Это вычисление дает ежемесячную волатильность.
- Умножить ежемесячную волатильность на квадратный корень из 12 для преобразования ежемесячной волатильности в годовую.

Эта процедура проиллюстрирована в файле *Dellvol.xlsx*, где я вычисляю годовую волатильность акций Dell на основе ежемесячных цен в период с августа 1988 г по май 2001 г. (См. рис. 82.3, на котором часть строк скрыта.)

Скопируйте формулу  $=(B2-B3)/B3$  из C2 в C3:C154 для вычисления ежемесячной доходности акций Dell. Затем скопируйте формулу  $=1+C2$  из D2 в D3:D154. Далее вычислите  $\ln(1 + \text{ежемесячный доход})$  для каждого месяца, скопировав формулу  $=LN(D2)$  из E2 в E3:E154, и вычислите ежемесячную волатильность в ячейке H3 по формуле  $=СТАНДОТКЛОН.В(E2:E154)$ . Наконец, вычислите годовую волатильность

акций Dell по формуле  $=\text{КОРЕНЬ}(12)*\text{H3}$ . Годовая волатильность акций Dell составляет 57,8%.

	A	B	C	D	E	F	G	H
1	Date	Dell Price	Return	1+Return	ln(1+Return)			
2	5/1/01	24.36	-0.07165	0.9283537	-0.074342521			
3	4/1/01	26.24	0.021509	1.0215085	0.021280472		monthly vol	0.16688
4	3/1/01	25.6875	0.174286	1.1742857	0.16066006		annual vol	0.57808
5	2/1/01	21.875	-0.16268	0.8373206	-0.177548278			
148	3/1/89	0.0742	-0.09512	0.904878	-0.099955097			
149	2/1/89	0.082	-0.17172	0.8282828	-0.188400603			
150	1/1/89	0.099	-0.0499	0.950096	-0.051192279			
151	12/1/88	0.1042	-0.08032	0.9196823	-0.083727039			
152	11/1/88	0.1133	-0.08407	0.9159256	-0.087820111			
153	10/1/88	0.1237	0.15824	1.1582397	0.146901353			
154	9/1/88	0.1068	0.282113	1.2821128	0.248509377			
155	8/1/88	0.0833						

**Рис. 82.3.** Вычисление исторической волатильности акций Dell

### ❓ Как реализовать в Excel формулу Блэка — Шоулза?

Для применения формулы Блэка — Шоулза в Microsoft Excel необходимо ввести значения для следующих параметров:

- $S$  — текущий курс акций;
- $t$  — срок опциона (в годах);
- $X$  — цена исполнения;
- $r$  — годовая безрисковая ставка (предполагается, что эта ставка постоянно вычисляется с учетом сложных процентов);
- $\sigma$  — годовая волатильность акций;
- $y$  — процент от стоимости акций, выплачиваемый в виде дивидендов.

С учетом этих входных значений цена европейского колл-опциона по формуле Блэка — Шоулза вычисляется следующим образом. Определим  $d_1$  и  $d_2$ :

$$d_1 = \frac{\ln\left(\frac{S}{X}\right) + \left(r - y + \frac{\sigma^2}{2}\right)t}{\sigma\sqrt{t}},$$

$$d_2 = d_1 - \sigma\sqrt{t}.$$

Тогда цена колл-опциона  $C$  определяется по формуле:

$$C = Se^{-yt} N(d_1) - Xe^{-rt} N(d_2).$$



Здесь  $N(x)$  — вероятность того, что нормальная случайная величина со средним значением 0 и стандартным отклонением  $\sigma$ , равным 1, меньше или равна  $x$ . Например,  $N(-1) = 0,16$ ,  $N(0) = 0,5$ ,  $N(1) = 0,84$  и  $N(1,96) = 0,975$ . Нормальная случайная величина со средним значением 0 и стандартным отклонением 1 называется *нормированной случайной величиной, распределенной по нормальному закону*. Нормальное интегральное распределение вычисляется с помощью функции НОРМ.СТ.РАСП (NORM.S.DIST). Формула =НОРМ.СТ.РАСП( $x$ ;ИСТИНА\_или\_1) возвращает вероятность того, что нормированная случайная величина, распределенная по нормальному закону, меньше или равна  $x$ . Например, формула =НОРМ.СТ.РАСП(-1;ИСТИНА), введенная в ячейку, дает в результате 0,16. Это значение показывает, что нормальная случайная величина со средним значением 0 и стандартным отклонением 1 с вероятностью 16% примет значение менее -1.

Цену европейского пут-опциона  $P$  можно записать как:

$$P = Se^{-rt}(N(d_1) - 1) - Xe^{-rt}(N(d_2) - 1).$$

В файле Bstemp.xlsx (рис. 82.4) я создал шаблон, вычисляющий цену для европейских колл- и пут-опционов. Введите значения параметров в B5:B10 и получите цену европейского колл-опциона в D13, а цену европейского пут-опциона в D14.

	A	B	C	D	E
4	Input data				
5	Stock price	20			
6	Exercise price	24			
7	Duration	7			
8	Interest rate	0.04879			
9	dividend rate	0			
10	volatility	0.5			
11					
12				Predicted	
13	Call price			\$10.64	
14	put			\$7.69	
15					
16					
17	Other quantities for option price				
18	d1	0.781789		N(d1)	0.7828
19	d2	-0.54109		N(d2)	0.2942

**Рис. 82.4.** Определение цены европейских колл- и пут-опционов

## ПРИМЕЧАНИЕ

Определение цены американских опционов в этой книге не рассматривается. Заинтересованные читатели должны обратиться к превосходному учебнику Луенбергера Investment Science.

В качестве примера предположим, что акции Cisco сегодня продаются по \$20 и вы выпустили семилетний европейский колл-опцион. Пусть годовая волатильность

акций Cisco равна 50%, и безрисковая ставка в течение семилетнего периода исчисляется исходя из 5% в год. С учетом сложных процентов она преобразуется в  $\ln(1 + 0,05) = 0,04879$ . Cisco не выплачивает дивиденды, так что годовая ставка дивиденда равна 0. Цена колл-опциона составляет \$10,64. Семилетний пут-опцион с ценой исполнения \$24 будет стоить \$7,69.

### ❓ Как изменение ключевых параметров изменяет цену колл- и пут-опционов?

- Как правило, влияние изменения входного параметра на цену колл- или пут-опциона соответствует влиянию, указанному в следующей таблице:

Параметр	Европейский колл-опцион	Европейский пут-опцион	Американский колл-опцион	Американский пут-опцион
Курс акций	+	–	+	–
Цена исполнения	–	+	–	+
Срок опциона	?	?	+	+
Волатильность	+	+	+	+
Безрисковая ставка	+	–	+	–
Дивиденды	–	+	–	+

- Повышение сегодняшнего курса акций всегда повышает цену колл-опциона и снижает цену пут-опциона.
- Повышение цены исполнения всегда повышает цену пут-опциона и снижает цену колл-опциона.
- Увеличение срока опциона всегда повышает цену американского опциона. В случае выплаты дивидендов увеличение срока опциона может либо повысить, либо понизить цену европейского опциона.
- Увеличение волатильности всегда повышает цену опциона.
- Повышение безрисковой ставки повышает цену колл-опциона, поскольку более высокие ставки, как правило, увеличивают темпы роста курса акций (что хорошо для колл-опциона). Эта ситуация более чем уравнивает тот факт, что из-за более высокой процентной ставки прибыль от опциона уменьшается. Повышение безрисковой ставки всегда понижает цену пут-опциона, поскольку более высокие темпы роста курса акций, как правило, наносят ущерб пут-опциону, и будущая прибыль от пут-опциона уменьшается.
- Дивиденды, как правило, снижают темпы роста курса акций, поэтому увеличение дивидендов снижает цену колл-опциона и повышает цену пут-опциона.

Используя однонаправленные и двунаправленные таблицы данных, вы можете при необходимости исследовать конкретные эффекты меняющихся параметров на цену колл- и пут-опционов (о работе с таблицами данных см. в главе 17 «Анализ чувствительности с помощью таблиц данных»).

### ? Как оценить волатильность акций по формуле Блэка — Шоулза?

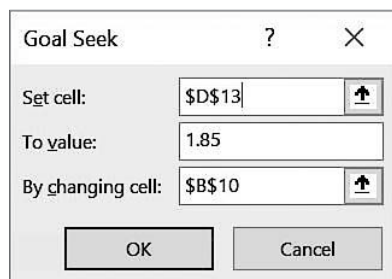
Ранее в этой главе я показал, как на основе исторических данных оценить годовую волатильность акций. Проблема с оценкой исторической волатильности состоит в том, что анализ представляет собой взгляд в прошлое. А нам нужен взгляд в будущее. Подход на основе *подразумеваемой волатильности* просто оценивает волатильность акций как значение волатильности, при котором цена по формуле Блэка — Шоулза соответствует рыночной цене опциона. Коротко говоря, подразумеваемая волатильность получает значение волатильности, подразумеваемое рыночной ценой опциона.

Инструмент Подбор параметра (Goal Seek) с описанными ранее входными параметрами позволяет легко и быстро вычислить подразумеваемую волатильность. 22 июля 2003 г. акции Cisco торговались по \$18,43 за штуку. В октябре 2003 г. колл-опцион с ценой исполнения \$17,50 продавался за \$1,85. Срок действия этого опциона истек 18 октября (89 дней спустя). Таким образом, срок действия опциона составил  $89/365 = 0,2438$  года. Cisco не выплачивает дивиденды. Пусть ставка для казначейских векселей составляет 5% и соответствующая безрисковая ставка равна  $\ln(1 + 0,05) = 0,04879$ . Для определения волатильности акций Cisco, которая подразумевается ценой опциона, введите в ячейки B5:B10 файла Ciscoimpvol.xlsx параметры, как показано на рис. 82.5.

	A	B	C	D	E
4	<b>Input data</b>				
5	<b>Stock price</b>	<b>18.43</b>			
6	<b>Exercise price</b>	<b>17.5</b>			
7	<b>Duration</b>	<b>0.24384</b>			
8	<b>Interest rate</b>	<b>0.04879</b>			
9	<b>dividend rate</b>	<b>0</b>			
10	<b>volatility</b>	<b>0.34036</b>			
11					
12				<b>Predicted</b>	
13	<b>Call price</b>			<b>\$1.85</b>	
14	<b>put</b>			<b>\$0.71</b>	
15					
16					
17	<b>Other quantities for option price</b>				
18	<b>d1</b>	<b>0.4629</b>		<b>N(d1)</b>	<b>0.678281</b>
19	<b>d2</b>	<b>0.29483</b>		<b>N(d2)</b>	<b>0.615937</b>

**Рис. 82.5.** Использование подразумеваемой волатильности для вычисления волатильности акций Cisco

Далее в диалоговом окне Подбор параметра (Goal Seek), показанном на рис. 82.6, определите волатильность (значение в ячейке B10), при которой цена опциона (формула в D13) достигнет \$1,85. На вкладке Данные (Data) в группе Прогноз (Forecast) выберите в опциях Анализ "что, если" (What-If Analysis) инструмент Подбор параметра (Goal Seek).



**Рис. 82.6.** Настройки в диалоговом окне Подбор параметра для поиска подразумеваемой волатильности

Как показано на рис. 82.5, этот опцион подразумевает годовую волатильность для Cisco в размере 34%.

## ПРИМЕЧАНИЕ

На сайте <https://finance.yahoo.com> вы найдете оценки волатильности любых акций.

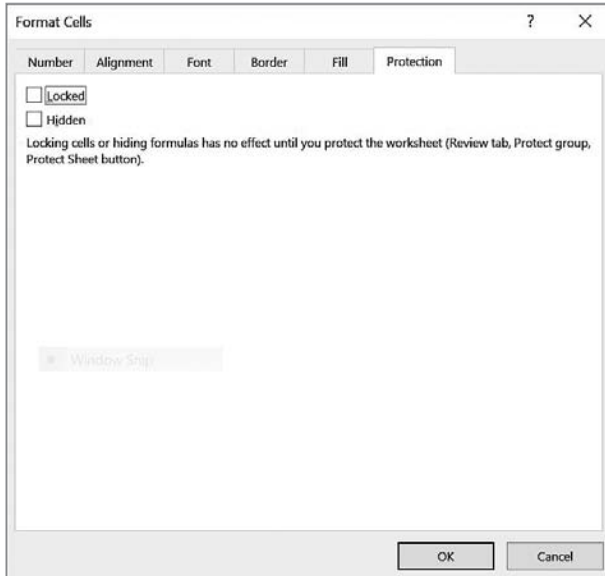
### ❓ В мои точные формулы ценообразования опционов никто не должен вносить изменения. Как защитить формулы на листе от изменений?

Я вполне уверен, что не один раз бывало так, что вы посылали сотруднику файл со своими расчетами, а тот менял выстраданные вами формулы. Нужно уметь защитить лист таким образом, чтобы другие пользователи могли только вводить данные, но не менять формулы на нем. Посмотрим, как защитить все формулы в шаблоне Блэка — Шоулза на следующем примере. (См. файлы Bstemp.xlsx и Bstempprotected.xlsx.)

Начнем с разблокировки всех ячеек на листе, а затем заблокируем ячейки, которые хотим защитить. Сначала щелкните на сером треугольнике в левом верхнем углу листа, где пересекаются заголовки строк и столбцов. После щелчка по этому треугольнику любые изменения формата будут применены на всем листе. Например, если выбрать полужирный шрифт, во всех ячейках листа будет использоваться полужирный шрифт.

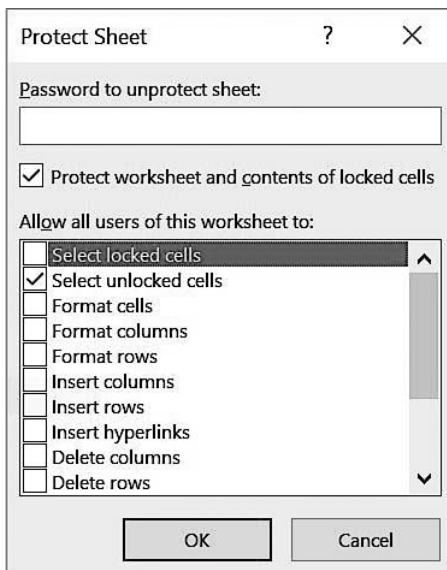
После выделения всего листа на вкладке Главная (Home) в группе Шрифт (Font) щелкните на маленьком значке вызова диалогового окна с параметрами шрифта. Откроется диалоговое окно Формат ячеек (Format Cells), показанное на рис. 82.7. На вкладке Защита (Protection) снимите флажок Защищаемая ячейка (Locked) и нажмите ОК. Теперь все ячейки на листе разблокированы. Это означает, что даже если лист защищен, к ячейкам можно получить доступ.

Далее выделите все формулы на листе. Для этого нажмите F5, открывающую диалоговое окно Переход (Go To). Нажмите Выделить (Special), выберите Формулы (Formulas) и нажмите ОК. Снова откройте диалоговое окно Формат ячеек (Format



**Рис. 82.7.** Диалоговое окно Формат ячеек

Cells) и на вкладке **Защита (Protection)** установите флажок **Защищаемая ячейка (Locked)**. Теперь этот флажок блокирует все формулы.



**Рис. 82.8.** Обеспечение доступа пользователей к незаблокированным ячейкам

Теперь вы можете защитить лист, что предотвратит изменение формул другими пользователями. На вкладке **Рецензирование (Review)** в группе **Изменения (Changes)** выберите **Защитить лист (Protect Sheet)**. В диалоговом окне **Защита листа (Protect Sheet)** установите флажок **выделение незаблокированных ячеек (Select Unlocked Cells)** — рис. 82.8 — и нажмите **ОК**. Это позволит пользователям шаблона выделять незаблокированные ячейки, но не формулы.

Теперь вы не сможете увидеть или изменить содержимое любой формулы. Попробуйте испортить какую-либо формулу! Получившийся защищенный лист сохранен в файле `Bstempprotected.xlsx`.

### **?** Как с помощью ценообразования опционов компания может принимать лучшие инвестиционные решения?

*Ценообразование опционов* можно использовать для улучшения бюджетирования капиталовложений или процесса принятия финансовых решений. Использование ценообразования опционов для оценки фактических инвестиционных проектов называется *реальными опционами*. Идею реальных опционов приписывают Джуди Левен (Judy Lewent), в прошлом финансовому директору Merck. По сути, реальные опционы позволяют назначить явную цену для управленческой гибкости, которую часто упускают из вида при традиционном бюджетировании капиталовложений. Концепция реальных опционов демонстрируется на следующих двух примерах.

### **ПРИМЕЧАНИЕ**

Для получения подробной информации о реальных опционах см. книгу Д. Луенбергера.

Пусть вы — владелец нефтяной скважины. Сегодня ваше наилучшее предположение заключается в том, что из скважины можно добыть нефти на \$50 млн. Через 5 лет (оставаясь владельцем скважины) вам предстоит принять решение о разработке этой нефтяной скважины, которая обойдется вам в \$70 млн. Один рискованный нефтяник готов купить скважину сегодня за \$10 млн. Следует ли продать скважину?

Конечно, цена нефти через 5 лет может вырасти. Даже если предположить, что цена нефти будет расти на 5% в год, через 5 лет нефть будет стоить всего \$63,81 млн. Традиционное бюджетирование капиталовложений говорит, что нефть обесценится, поскольку стоимость ее разработки превышает стоимость нефти в скважине. Но не торопитесь — через 5 лет цена нефти в скважине будет другой, так как многие факторы (такие как мировая цена на нефть) могут измениться. Существует вероятность, что через 5 лет нефть будет стоить не меньше \$70 млн. Если через 5 лет нефть будет стоить \$80 млн, разработка скважины через 5 лет принесет \$10 млн прибыли.

По сути, вы владеете пятилетним европейским колл-опционом на эту скважину, так как прибыль от скважины через 5 лет такая же, как прибыль от европейского колл-опциона с курсом акций \$50 млн, ценой исполнения \$70 млн и сроком

опциона 5 лет. Можно предположить, что годовая волатильность подобна волатильности акций типичной нефтяной компании (например, 30%). Если использовать ставку по казначейскому векселю 5%, соответствующую безрисковой ставке 4,879%, в файле *Oilwell.xlsx* (рис. 82.9) можно определить, что цена такого колл-опциона составляет \$11,47 млн. Из этого следует, что вы не должны продавать скважину за \$10 млн.

Конечно, фактическая волатильность для этой нефтяной скважины неизвестна. Поэтому с помощью однонаправленной таблицы данных определим, как цена опциона зависит от оценки волатильности (см. рис. 82.9). Как видите, пока волатильность скважины составляет по меньшей мере 27%, «опцион» на нефтяную скважину стоит более \$10 млн.

	A	B	C	D	E	F	G	H
4	Input data							
5	Stock price	50						
6	Exercise price	70						
7	Duration	5						
8	Interest rate	0.05						
9	dividend rate	0					Volatility	\$11.47
10	volatility	0.3					0.1	2.635248
11							0.11	3.064456
12				Predicted			0.12	3.498668
13	Call price			\$11.47			0.13	3.936575
14	put			\$16.32			0.14	4.377212
15							0.15	4.819845
16							0.16	5.263906
17	Other quantities for option price						0.17	5.708942
18	d1	0.2		N(d1)	0.5783		0.18	6.15459
19	d2	-0.47		N(d2)	0.318		0.19	6.600548
20							0.2	7.046562
21							0.21	7.492419
22							0.22	7.937931
23							0.23	8.382936
24							0.24	8.82729
25							0.25	9.270862
26							0.26	9.713537
27							0.27	10.15521
28							0.28	10.59577
29							0.29	11.03514
30							0.3	11.47322

**Рис. 82.9.** Реальный опцион на нефтяную скважину

В качестве второго примера рассмотрим биотехнологическую компанию, разрабатывающую лекарственный препарат для фармацевтической фирмы. В биотехнологической компании в настоящее время предполагают, что цена разрабатываемого препарата составляет \$50 млн. Безусловно, цена препарата со временем может снизиться. Для защиты от падения цены биотехнологической компании требуется, чтобы препарат через 5 лет имел гарантированную цену \$50 млн. Если страховая компания собирается подписать такое обязательство, какая справедливая цена должна быть назначена?

По сути, биотехнологическая компания просит заплатить ей \$1 млн за каждый \$1 млн, на который цена препарата упадет ниже \$50 млн, через 5 лет. Это эквивалентно пятилетнему пут-опциону на цену препарата. Предположим, что ставка по казначейскому векселю составляет 5% и годовая волатильность для акций сопоставимых компаний равна 40% (см. файл *Drugabandon.xlsx*, показанный на рис. 82.10), тогда цена этого опциона — \$10,51 млн. Такой тип опциона часто называют *опционом на отказ от проекта*, но он эквивалентен пут-опциону. (На рис. 82.10 односторонняя таблица данных показывает, как цена опциона отказа от проекта зависит от предполагаемой волатильности, составляющей 30–45% от цены препарата.)

	A	B	C	D	E	F	G	H
1	<b>Call with dividends</b>							
2								
3								
4	<b>Input data</b>							
5	Stock price	50						
6	Exercise price	50						
7	Duration	5						
8	Interest rate	0.049						
9	dividend rate	0					<b>Volatility</b>	<b>\$10.51</b>
10	volatility	0.4					0.3	7.03399
11							0.31	7.383411
12				<b>Predicted</b>			0.32	7.732875
13	Call price			\$21.34			0.33	8.082211
14	put			\$10.51			0.34	8.431265
15							0.35	8.779894
16							0.36	9.127966
17	<b>Other quantities for option price</b>						0.37	9.475362
18	d1	0.72		N(d1)	0.764		0.38	9.821968
19	d2	-0.174		N(d2)	0.431		0.39	10.16768
20							0.4	10.51241
21							0.41	10.85605
22							0.42	11.19853
23							0.43	11.53976
24							0.44	11.87967
25							0.45	12.21819

**Рис. 82.10.** Вычисление опциона отказа от проекта

## Задания

- На основе данных о ежемесячной доходности акций в файле *Volatility.xlsx* определите оценки годовой волатильности для акций Intel, Microsoft и GE.
- Сегодня акции продаются по \$42. Годовая волатильность акций составляет 40%, а годовая безрисковая ставка равна 10%.
  - Какова справедливая цена шестимесячного европейского колл-опциона с ценой исполнения \$40?



- Насколько текущая цена акций должна вырасти, чтобы покупатель колл-опциона достиг безубыточности через шесть месяцев?
  - Какова справедливая цена для шестимесячного европейского пут-опциона с ценой исполнения \$40?
  - Насколько текущая цена акций должна снизиться, чтобы покупатель пут-опциона достиг безубыточности через шесть месяцев?
  - При каком уровне волатильности 40-долларовый колл-опцион будет продаваться за \$6? Подсказка: воспользуйтесь инструментом Подбор параметра (Goal Seek).
3. 25 сентября 2000 г. акции JDS Uniphase продавались по \$106,81. В этот же день 100-долларовый европейский пут-опцион, срок действия которого истек 20 января 2001 г., продавался за \$11,875. На основе этой информации вычислите подразумеваемую волатильность для акций JDS Uniphase. Используйте 5%-ную ставку по казначейскому векселю.
  4. 9 августа 2002 г. акции Microsoft продавались по \$48,58. Тридцатипятидолларовый европейский колл-опцион, срок действия которого истек 17 января 2003 г., продавался за \$13,85. На основе этой информации вычислите подразумеваемую волатильность для акций Microsoft. Используйте 4%-ную ставку по казначейскому векселю.
  5. Есть опцион на покупку нового самолета через три года за \$25 млн. По вашим оценкам текущая цена самолета составляет \$21 млн. Годовая волатильность цены самолета — 25%, безрисковая ставка равна 5%. Сколько стоит опцион на покупку самолета?
  6. Текущая цена меди составляет 95 центов за фунт. Годовая волатильность для цен на медь равна 20%, безрисковая ставка — 5%. Через год имеется возможность потратить \$1,25 млн на добычу 8 млн фунтов меди. Медь может быть продана через год. Добыча фунта меди из недр обходится в 85 центов. Выгодна ли инвестиция в медь при таких условиях?
  7. У вас есть авторские права на биотехнологический лекарственный препарат. Наилучшая оценка текущей цены на эти права составляет \$50 млн. Пусть годовая волатильность биотехнологических компаний составляет 90% и безрисковая ставка равна 5%. Какова цена опциона на продажу прав на лекарственный препарат за \$40 млн через пять лет от текущего момента?
  8. В компании Merck обсуждается вопрос о том, стоит ли инвестировать в новый биотехнологический проект. По оценкам компании стоимость проекта составляет —\$56 млн. Инвестирование в новый проект даст компании Merck возможность поучаствовать в более крупном технологическом проекте, который будет доступен через 4 года. Если Merck не примет участие в новом проекте, она не сможет получить более крупный проект. Для более крупного проекта через 4 года, начиная с текущего момента, потребуется \$1,5 млрд. В настоящее

время Merck оценивает чистую приведенную стоимость (ЧПС) денежных потоков для более крупного проекта в \$597 млн. Если безрисковая ставка равна 10% и годовая волатильность для более крупного проекта составляет 35%, как должна поступить компания Merck? (С этой задачи началось применение реальных опционов!)

9. Разработайте лист со следующими входными данными для расчета годовой прибыли:

- ежегодные постоянные затраты;
- себестоимость единицы продукции;
- цена за единицу продукции;
- годовой спрос =  $10\,000 - 100 \times (\text{цена})$ .

Защитите ячейки, в которых вычисляется годовой спрос и годовая прибыль.

# Определение потребительской ценности

### Обсуждаемые вопросы

- Компания, работающая с кредитными картами, в настоящее время имеет коэффициент удержания клиентов, равный 80. Насколько возрастет прибыльность компании, если коэффициент удержания клиентов повысится до 90% и выше?
- Междугородная телефонная компания предоставляет клиентам конкурентов бонус при смене компании. Насколько значительным должен быть бонус?

Многие компании недооценивают своих клиентов. При оценке клиента компания должна учитывать чистую приведенную стоимость (ЧПС) для долгосрочной прибыли, которую принесет клиент. (О чистой приведенной стоимости см. в главе 8 «Оценка инвестиций по чистой приведенной стоимости».) Из-за отказа от учета долгосрочной значимости клиента компании часто принимают неправильные решения. Например, компания может сократить отдел обслуживания клиентов на 10%, чтобы сэкономить \$1 млн, но из-за снижения качества обслуживания потерять гораздо больше \$1 млн в *потребительской ценности*, когда клиент переведет свой бизнес в другое место. Это, в свою очередь, приведет к снижению прибыли компании. Способы определения потребительской ценности рассматриваются на следующих двух примерах.

## Ответы на вопросы

❓ Компания, работающая с кредитными картами, в настоящее время имеет коэффициент удержания клиентов, равный 80. Насколько возрастет прибыльность компании, если коэффициент удержания клиентов повысится до 90% и выше?

Этот пример основан на книге Фредерика Райхельда «Эффект лояльности»<sup>1</sup>. Данные к примеру вы найдете в файле *Loyalty.xlsx* (рис. 83.1). Райхельд оценивает прибыльность клиента — держателя кредитной карты на основе количества лет, в течение которых клиент является держателем карты. Например, в течение первого

---

<sup>1</sup> Frederick Reichheld. *The Loyalty Effect*. — Harvard Business School Press, 2001.

года держатель карты приносит прибыль –\$40, вследствие затрат на привлечение клиента и на создание клиентского счета. В течение каждого последующего года прибыль, приносимая клиентом, растет, и клиент, являющийся держателем карты 20 лет и больше, приносит прибыль \$161 в год.

Кредитная компания хочет определить, как потребительская ценность зависит от коэффициента удержания клиентов. В настоящее время коэффициент удержания клиентов составляет 80%. Это означает, что в конце каждого года 20% ( $1 - 0,80$ ) всех клиентов не возобновляют свои карты. (Эти 20% клиентов могут быть названы *годовым коэффициентом «текучести» клиентов*.) Кредитной компании необходимо определить долгосрочную ценность клиента для следующих значений коэффициента удержания клиентов: 80, 85, 90, 95 и 99%.

	A	B	C	D	E	F	G
4				npv per customer	141.7181		
5							
6		retention rate	0.8				
7		Interest rate	0.15				
8	Year	Mean Profit(if still here)	Number	Profit		retention rate	141.7181
9	1	(\$40.00)	100	(\$4,000.00)		0.8	141.7181
10	2	\$66.00	80	\$5,280.00		0.85	193.1495
11	3	\$72.00	64	\$4,608.00		0.9	269.3474
12	4	\$79.00	51.2	\$4,044.80		0.95	390.7125
13	5	\$87.00	40.96	\$3,563.52		0.99	548.5771
14	6	\$92.00	32.768	\$3,014.66			
15	7	\$96.00	26.2144	\$2,516.58			
16	8	\$99.00	20.97152	\$2,076.18			
17	9	\$103.00	16.77722	\$1,728.05			
18	10	\$106.00	13.42177	\$1,422.71			
19	11	\$111.00	10.73742	\$1,191.85			
20	12	\$116.00	8.589935	\$996.43			
21	13	\$120.00	6.871948	\$824.63			
22	14	\$124.00	5.497558	\$681.70			
23	15	\$130.00	4.398047	\$571.75			
24	16	\$137.00	3.518437	\$482.03			
25	17	\$142.00	2.81475	\$399.69			
26	18	\$148.00	2.2518	\$333.27			
27	19	\$155.00	1.80144	\$279.22			
28	20	\$161	1.441152	\$232.03			
29	21	\$161	1.152922	\$185.62			
30	22	\$161	0.922337	\$148.50			

**Рис. 83.1.** Ценность клиентов — держателей кредитных карт

Для определения долговременной ценности клиента начнем, например, с когорты из 100 клиентов. (*Когорта* — это группа лиц, имеющих общий статистический фактор. Размер группы, 100 человек, был выбран произвольно, но круглые числа позволяют легче следить за ходом анализа.) Затем для каждого года определим, сколько клиентов остается клиентами компании, по формуле (клиенты для года  $t + 1$ ) = (коэффициент удержания клиентов) × (клиенты для года  $t$ ). Будем считать, что клиенты уходят только в конце года. Затем с помощью функции ЧПС (NPV) определим общую ЧПС (с учетом 15%-ной учетной ставки), полученную от исходной когорты из 100 клиентов. 15%-ная учетная ставка означает, что \$1, заработанный через год от текущего момента, стоит так же, как \$1,00/\$1,15, заработанный сей-

час. Разделив это число на количество клиентов в исходной когорте (100), получим ценность отдельного клиента.

Прежде всего я присвоил имена, указанные в диапазоне B6:B7, ячейкам C6:C7. Затем я ввел исходное количество клиентов (100) в ячейку C9. Затем скопировал формулу  $\text{=коэфф\_сохран\_клиентов} \times C9$  из C10 в C11:C38 для вычисления количества клиентов, оставшихся на следующий год. Например, на год 2 осталось 80 клиентов.

Я вычислил прибыль, приносимую каждый год, умножив количество оставшихся клиентов на прибыль от каждого клиента. Для этого скопируйте формулу  $\text{=C9} \times B9$  из ячейки D9 в D10:D38. В ячейке E4 я вычислил среднюю ЧПС, полученную с отдельного клиента, по формуле  $\text{=(1+процентная\_ставка) \times ЧПС(процентная\_ставка; D9:D38)/100}$ . Допустим, денежные потоки рассчитываются на начало года и годовая учетная ставка равна 15%. Часть формулы  $\text{ЧПС(процентная\_ставка; D9:D38)/100}$  вычисляет среднюю ЧПС, полученную от отдельного клиента, для денежных потоков на конец года. ЧПС для денежных потоков на конец года преобразуются в ЧПС для денежных потоков на начало года путем умножения на  $(1 + \text{процентная\_ставка})$ .

При коэффициенте удержания клиентов 80% ценность среднего клиента составляет \$141,72. Чтобы определить, как ценность отдельного клиента варьирует вместе с изменением годового коэффициента удержания клиентов, я использовал однонаправленную таблицу данных. В ячейки F9:F13 я ввел соответствующие годовые коэффициенты удержания клиентов. В ячейку G8 я ввел формулу  $\text{=E4}$ , по которой с помощью таблицы будет рассчитываться ЧПС для отдельного клиента. Затем я выделил табличный диапазон (F8:G13) и на вкладке Данные (Data) в списке Анализ "что, если" (What If Analysis) выбрал Таблица данных (Data Table). Указав ячейку C6 в поле Подставлять значения по строкам в (Column Input Cell), я получил расчеты прибыли, показанные ранее на рис. 83.1. Обратите внимание, что повышение коэффициента удержания клиентов с 80 до 90% почти удваивает ценность каждого клиента, то есть крайне важно обращаться с такими клиентами вежливо и не экономить пенни на их обслуживании. Осознание ценности клиента дает большинству компаний преимущество и позволяет повысить прибыль.

**❓ Междугородная телефонная компания предоставляет клиентам конкурентов бонус при смене компании. Насколько значительным должен быть бонус?**

Междугородная телефонная компания, клиенты которой тратят в среднем \$400 в год, получает 10% прибыли с каждого потраченного доллара. В конце каждого года 50% клиентов компании переходят к конкурентам и 30% клиентов конкурентов без какого-либо привлечения переходят в эту компанию. Компания рассматривает вопрос о предоставлении клиентам конкурентов одноразового бонуса для смены компании. Насколько значительным может быть стимул без убытка для компании?

Ключом к решению этой проблемы (см. файл *Phoneloyalty.xlsx* и рис. 83.2) является внимательное изучение ЧПС при наличии двух условий:

- **Условие 1.** Сто клиентов вначале подключаются к конкурентам.
- **Условие 2.** Компания платит 100 клиентам конкурентов определенную сумму за переход.

Для отслеживания каждой ситуации в течение определенного времени (например, в течение 20 лет) можно с помощью инструмента Подбор параметра (Goal Seek) определить сумму  $x$  в долларах (выплачиваемую клиенту за переход в компанию), которая нивелирует влияние на прибыль следующих условий:

- **Условие 1.** Компания платит каждому из 100 нелояльных клиентов  $x$  долларов за переход.
- **Условие 2.** Рынок состоит из 100 нелояльных клиентов.

Пусть анализ начинается 30 июня 2014 г. и клиенты меняют компанию не чаще одного раза в год. Присвойте имена, указанные в ячейках A2:A6, ячейкам B2:B6. Ключевым шагом анализа является понимание того, что (клиенты компании в год  $t + 1$ ) =  $0,3 \times$  (клиенты конкурентов в год  $t$ ) +  $0,5 \times$  (клиенты компании в год  $t$ ). Аналогично (клиенты конкурентов в год  $t + 1$ ) =  $0,7 \times$  (клиенты конкурентов в год  $t$ ) +  $0,5 \times$  (клиенты компании в год  $t$ ).

	A	B	C	D	E	F	G	H	I
2	switch fee	34.22193			pay-no pay	0			
3	probleave	0.5							
4	probcome	0.3							
5	annrevenue	400	NPV	13674.36			13674.36		
6	profitmargin	0.1		Pay them			Do not pay them		
7		Date	Year	Number with us	Number with them	Profit	Number with us	Number with them	Profit
8		initial							
9		6/30/2014	1	100	0	577.81	30	70	1200
10		6/30/2015	2	50	50	2000	36	64	1440
11		6/30/2016	3	40	60	1600	37.2	62.8	1488
12		6/30/2017	4	38	62	1520	37.44	62.56	1498
13		6/30/2018	5	37.6	62.4	1504	37.488	62.512	1500
14		6/30/2019	6	37.52	62.48	1500.8	37.4976	62.5024	1500
15		6/30/2020	7	37.504	62.496	1500.2	37.49952	62.5005	1500
16		6/30/2021	8	37.5008	62.4992	1500	37.4999	62.5001	1500
25		6/30/2030	17	37.5	62.5	1500	37.5	62.5	1500
26		6/30/2031	18	37.5	62.5	1500	37.5	62.5	1500
27		6/30/2032	19	37.5	62.5	1500	37.5	62.5	1500
28		6/30/2033	20	37.5	62.5	1500	37.5	62.5	1500

**Рис. 83.2.** Анализ бонуса для телефонной компании

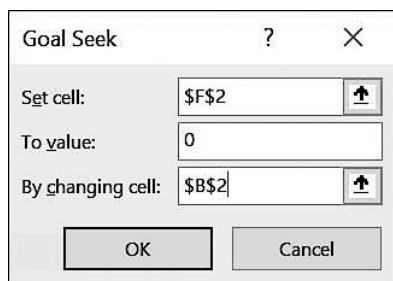
В ячейку D9 (клиенты компании) я ввел 100 и 0 в ячейку E9 (клиенты конкурентов). Такое распределение клиентов соответствует ситуации сразу после того, как

бонус был предложен 100 клиентам. Предположим, что клиенты, получившие бонус, должны оставаться клиентами компании, по крайней мере, один год. Скопировав формулу  $= (1 - \text{вероятность\_ухода}) * D9 + \text{вероятность\_прихода} * E9$  из ячейки D10 в D11:D28, я вычислил количество клиентов в каждом году. (На рис. 83.2 данные для 2022–2029 гг. скрыты.) Количество клиентов конкурентов в каждом году я вычислил по формуле  $= \text{вероятность\_ухода} * D9 + (1 - \text{вероятность\_прихода}) * E9$ , скопировав ее из E10 в E11:E28.

В ячейке F9 я вычислил прибыль, полученную за первый год, по формуле  $= D9 * \text{годовая\_выручка} * \text{коэфф\_прибыльности} - \text{бонус} * 100$ . Обратите внимание, что из прибыли я вычел сумму, выплаченную 100 клиентам за переход от конкурентов. Для вычисления прибыли в последующие годы я скопировал формулу  $= D10 * \text{годовая\_выручка} * \text{коэфф\_прибыльности}$  из F10 в F11:F28. В ячейке D5 вычислена ЧПС прибыли, связанная с бонусом, по формуле  $= \text{ЧИСТНЗ}(0, 1; F9:F28; B9:B28)$ . (Функция ЧИСТНЗ (XNPV) рассматривается в главе 8.)

Аналогичным образом в ячейках G8:I28 я вычислил прибыль, полученную для случая, когда 100 клиентов первоначально были подключены к конкурентам. 30 июня 2014 г. 30 из этих 100 клиентов перешли в компанию (даже без бонуса). В ячейке F2 я вычислил разницу между ЧПС с бонусом и ЧПС без бонуса.

Наконец, с помощью инструмента Подбор параметра (Goal Seek) я варьировал размер бонуса (ячейка B2) так, чтобы значение ячейки F2 стало равно 0. Диалоговое окно Подбор параметра (Goal Seek) показано на рис. 82.3. ЧПС для двух ситуаций одинакова, если размер бонуса составляет \$34,22. Таким образом, компания могла бы предоставлять клиентам бонус до \$34,22 за переход и по-прежнему иметь повышенную прибыль.



**Рис. 83.3.** Диалоговое окно Подбор параметра для определения максимального размера бонуса, повышающего прибыльность

## Задания

1. Компания Whirlswim Appliance рассматривает возможность предоставить всем клиентам бесплатное техническое обслуживание для каждого проданного видеомаягнитофона (такие существовали до того, как кино стали снимать на

цифру). По оценкам компании, за каждый проданный видеомагнитофон придется платить в среднем \$2,50 (стоимость в сегодняшних долларах). Рынок состоит из 72 000 потребителей, последняя покупка которых была в компании Whirlswim, и 86 000 потребителей, которые последний раз покупали у конкурентов. В указанном году 40% всех потребителей приобрели видеомагнитофоны. Если последняя покупка потребителя была в Whirlswim, с вероятностью 60% его следующая покупка также будет в Whirlswim. Если последняя покупка была не в Whirlswim, с вероятностью 30% следующая покупка будет в Whirlswim. Покупка в текущем году дает прибыль \$20. Вклад в прибыль от покупателя (и расходы на техническое обслуживание для каждого покупателя) растет на 5% в год. Коэффициент дисконтирования для прибыли (относительно 30-летнего горизонта) составляет 10% в год.

Предположим, что компания предоставляет бесплатное техническое обслуживание. Если последняя покупка клиента была в Whirlswim, вероятность того, что в следующий раз он опять купит в Whirlswim, увеличится на неизвестную величину в пределах от 0 до 10%. Аналогично, если компания предоставляет бесплатное техническое обслуживание и последняя покупка клиента была не в Whirlswim, вероятность того, что следующая покупка будет в Whirlswim, увеличится на неизвестную величину в пределах от 0 до 10%. Следует ли компании Whirlswim предлагать клиентам бесплатное техническое обслуживание?

2. Супермаркет Mr. D пытается определить, следует ли предоставить потребителям карту клиента. В настоящее время 30% всех покупателей лояльны к супермаркету. Лояльный клиент покупает в Mr. D в 80% случаев. Нелояльный к Mr. D клиент покупает в этом супермаркете в 10% случаев. Стандартный покупатель тратит \$150 в неделю, и супермаркет Mr. D работает из 4% прибыли.

Карта клиента обойдется супермаркету Mr. D в среднем в \$0,01 на каждый потраченный доллар. Вы считаете, что доля лояльных к Mr. D покупателей увеличится на неизвестную величину в пределах от 2 до 10%. Вы также считаете, что доля случаев, когда лояльный клиент покупает в Mr. D, увеличится на неизвестную величину в пределах от 2 до 12%. Следует ли супермаркету Mr. D предоставлять потребителям карту клиента? Следует ли супермаркету Mr. D предоставлять карту, если прибыль составляет 8%, а не 4%?



# Оптимальный размер заказа в модели управления запасами

### Обсуждаемые вопросы

- Магазин электроники продает 10 000 смартфонов в год. Каждый раз при закупке смартфонов магазин несет расходы в размере \$10 за закупку. Магазин платит \$100 за каждый смартфон, и стоимость хранения смартфона на складе в течение года составляет \$20. Когда магазин закупает смартфоны, какого размера должна быть закупка?
- Завод производит 10 000 серверов в год. Себестоимость производства каждого сервера составляет \$2000. Стоимость запуска производства партии серверов составляет \$200, а стоимость хранения сервера на складе в течение года — \$500. При необходимости завод может производить 25 000 серверов в год. Каким должен быть размер партии при производстве серверов?

Когда магазин заказывает один и тот же товар многократно, естественным образом возникает вопрос: какое количество товара магазин должен заказывать каждый раз? Если магазин закажет слишком много, его ждут непомерные затраты на хранение запасов. Если магазин закажет слишком мало товара, возникнут затраты на повторные заказы. Где-то должна пролежать золотая середина, которая минимизирует годовые затраты на хранение запасов и на заказы.

С такой же позиции рассмотрим ситуацию с заводом, производящим продукцию партиями. Какой размер партии минимизирует годовые затраты на хранение и на запуск производства партии продукции? В данной главе на двух примерах показано, как формула оптимального размера заказа (предложенная в 1913 г. Ф. Харрисом (F. Harris) из Westinghouse Corporation) позволяет ответить на эти вопросы.

## Ответы на вопросы

❓ Магазин электроники продает 10 000 смартфонов в год. Каждый раз при закупке смартфонов магазин несет расходы в размере \$10 за закупку. Магазин платит \$100 за каждый смартфон, и стоимость хранения смартфона на складе в течение года составляет \$20. Когда магазин закупает смартфоны, какого размера должна быть закупка?

Размер заказа на поставку, минимизирующий годовые затраты на хранение и заказы на поставку, можно определить, если известны следующие параметры:

- $K$  — затраты на закупки;
- $h$  — затраты на хранение единицы продукции в течение года;
- $D$  — годовой спрос на продукт.

Работа с этими параметрами представлена в файле `Eoq.xlsx` на листе `EOQ` (рис. 84.1).

	A	B	C	D
1				
2	cost/order	K	10	
3	annual holding cost per unit	h	20	
4	annual demand	D	10000	
5	order quantity	EOQ	100	=SQRT(2*K*D/h)
6	holding cost per year	annhc	\$ 1,000.00	
7	order cost per year	annoc	\$ 1,000.00	
8	total annual cost (excluding purchasing)	anncost	\$ 2,000.00	=annhc+annoc
9	orders per year	annorders	100	=D/EOQ

**Рис. 84.1.** Шаблон для вычисления оптимального размера заказа (EOQ)

Пусть  $q$  — размер заказа на поставку, тогда годовые затраты на хранение продукции равны  $0,5qh$ . (Назовем это уравнение уравнением 1.) В уравнении 1 мы приняли средний уровень запасов ( $0,5q$ ) равным половине от максимального уровня запасов. Чтобы убедиться, что средний уровень запасов равен  $0,5q$ , можно вычислить средний уровень запасов для цикла (периода времени между поставкой заказов). В начале цикла закупка прибывает, и уровень запасов равен  $q$ . В конце цикла все распродано, и уровень запасов равен 0. Поскольку спрос постоянен, средний уровень запасов в течение цикла просто равен среднему для 0 и  $q$ , или  $0,5q$ . Максимальный уровень запасов равен  $q$ , так как подразумевается, что следующий заказ поступает в тот момент, когда уровень запасов снижается до 0.

Поскольку в год делается  $D/q$  заказов на поставку, годовые затраты на заказы равны  $(D/q) \times K$ . (Я назвал это уравнение уравнением 2.) С помощью инструмента Поиск решения (Solver) можно показать, что годовая сумма затрат на хранение и заказы минимальна для значения  $q$ , равного оптимальному размеру заказа (EOQ), который вычисляется по следующей формуле (уравнение 3):

$$EOQ = \sqrt{\frac{2KD}{h}}.$$

Эта формула показывает:

- увеличение спроса или затрат на заказы увеличивает EOQ;
- увеличение затрат на хранение уменьшает EOQ.

В файле *Eoq.xlsx* я использовал уравнение 3, чтобы определить EOQ в ячейке C5 по формуле  $=\text{SQRT}(2*K*D/h)$ . Годовые затраты на хранение я определил в ячейке C6 по уравнению 1 и, соответственно, формуле  $=0,5*EOQ*h$ . Годовые затраты на заказы в ячейке C7 вычислены по уравнению 2. Обратите внимание, что для оптимального размера заказа EOQ годовые затраты на заказы всегда равны годовым затратам на хранение. В ячейке C8 я определил общие годовые затраты (без учета стоимости приобретения, которая не зависит от стратегии закупок) по формуле  $=C6+C7$ .

Конечно, можно воспользоваться однонаправленной или двунаправленной таблицей данных, чтобы определить чувствительность EOQ и различных затрат к изменению  $K$ ,  $h$  и  $D$ . В данном примере  $K = \$10$ ,  $D = 10\,000$  смартфонов в год и  $h = \$20$  за хранение одного смартфона. После ввода этих значений в ячейки C2:C4 получим следующее:

- размер каждого заказа должен составлять 100 смартфонов;
- и годовые затраты на хранение, и затраты на заказы составляют \$1000. Для оптимального размера заказа годовые затраты на хранение всегда равны годовым затратам на заказы;
- общие годовые затраты (без учета затрат на приобретение) составляют \$2000.

При работе с EOQ необходимо учитывать следующее.

- Наличие оптовых скидок лишает нахождение EOQ смысла, поскольку в таком случае годовые затраты на приобретение зависят от размера заказа.
- EOQ предполагает, что спрос равномерен в течение года. Оптимальный размер заказа не следует использовать для продуктов, на которые есть сезонный спрос.
- Предполагается, что годовые затраты на хранение, как правило, составляют 10–40% от затрат на приобретение единицы продукции.
- Лист *EOQ Protected* в файле *Eoq.xlsx* — это защищенная версия листа EOQ. Когда лист защищен, изменения в формулы внести невозможно. (О том, как защитить лист, см. в главе 82 «Ценообразование опционов».)

## ПРИМЕЧАНИЕ

О построении моделей управления запасами см. в книге Уэйна Винстона (Wayne L. Winston. *Operations Research: Applications and Algorithms*. — Duxbury Press, 2007).

**?** Завод производит 10 000 серверов в год. Себестоимость производства каждого сервера составляет \$2000. Стоимость запуска производства партии серверов составляет \$200, а стоимость хранения сервера на складе в течение года — \$500. При необходимости завод может производить 25 000 серверов в год. Каким должен быть размер партии при производстве серверов?

Модель EOQ предполагает, что по запросу закупка прибывает моментально. Когда компания производит продукт, а не закупает его, заказ на изготовление не мо-

жет быть исполнен мгновенно. В таких ситуациях нужно определить не размер заказа, минимизирующий затраты, а размер партии, минимизирующий затраты. Когда компания производит продукцию, а не закупает ее, размер партии, сводящей затраты к минимуму, зависит от следующих параметров:

- $K$  — затраты на запуск производства партии продукта;
- $h$  — затраты на хранение единицы продукции в запасе в течение года;
- $D$  — годовой спрос на продукт;
- $R$  — годовой показатель производства продукта. Например, производственная мощность IBM составляет 25 000 серверов в год.

Пусть  $q$  — размер каждой партии продукта, тогда годовые затраты на хранение равны  $0,5 \times (q/R) \times (R - D) \times h$ . (Я назвал это уравнение уравнением 4.) Уравнение 4 имеет такой вид, потому что производство каждой партии продолжается  $q/R$  лет и во время производственного цикла запас увеличивается как  $R - D$ . Максимальный уровень запасов, который имеет место после завершения производства партии, можно вычислить как  $(q/R) \times (R - D)$ . Таким образом, средний уровень запасов равен  $0,5 \times (q/R) \times (R - D)$ .

Поскольку в год производится  $D/q$  партий, годовые затраты на запуск производства равны  $KD/q$  (я назвал это уравнением 5). С помощью инструмента Excel Поиск решения (Solver) можно показать, что размер партии, минимизирующий годовые затраты на запуск производства и на хранение, можно вычислить по следующему уравнению (я назвал его уравнением 6). Эта модель называется оптимальным размером партии (EOB):

$$\text{EOB} = \sqrt{\frac{2KDR}{h(R-D)}}.$$

Исходя из этой формулы можно определить следующее:

- увеличение  $K$  или  $D$  увеличивает EOB;
- увеличение  $h$  или  $R$  уменьшает EOB.

В файле *Contrateeq.xlsx* на листе *Cont Rate EOQ* я создал шаблон для вычисления EOB, годовых затрат на запуск производства и затрат на хранение. Этот лист показан на рис. 84.2.

Для этого примера  $K = \$200$ ,  $h = \$500$ ,  $D = 10\,000$  единиц в год и  $R = 25\,000$  единиц в год. После ввода значений этих параметров в ячейки C2:C5 получим следующее.

- Размер партии, минимизирующий затраты, составляет 115,47 сервера. Таким образом, компания должна производить партию из 115 или 116 серверов.
- Годовые затраты на хранение и на запуск производства составляют по \$17 320,51. Как и в предыдущем случае, для оптимального размера партии годовые затраты на хранение всегда равны годовым затратам на запуск производства.

	A	B	C	D
2	cost per batch	K	\$ 200.00	
3	annual holding cost per unit	h	\$ 500.00	
4	annual demand	D	10000	
5	annual production rate	rate	25000	
6	batch size	EOB	115.470054	$=\text{SQRT}(2*K*D*rate/(h*(rate-D)))$
7	holding cost per year	annhc	\$ 17,320.51	$=(EOB/rate)*(rate-D)*0.5*h$
8	order cost per year	annoc	\$ 17,320.51	$=\text{annbatches}*K$
9	total annual cost (excluding purchasing)	anncost	\$ 34,641.02	$=\text{annhc}+\text{annoc}$
10	batches per year	annbatches	86.6025404	$=D/EOB$

Рис. 84.2. Шаблон для вычисления EOB

- Общие годовые затраты (без учета переменных затрат на производство) составляют 34 641,02 доллара.
- В год нужно производить 86,6 партии.

При работе с EOB необходимо учитывать следующее.

- Если переменные затраты на производство единицы продукции зависят от размера партии, модель EOB является недействительной.
- Модель EOB предполагает, что спрос относительно постоянен в течение года. Эту модель не следует использовать для продукции, на которую есть сезонный спрос.
- Предполагается, что годовые затраты на хранение, как правило, составляют 10–40% от затрат на производство единицы продукции.
- Лист Protected в файле Contrateeq.xlsx — это защищенная версия листа EOB. О том, как защитить лист, см. в главе 82.

## Задания

- Магазин бытовой техники продает плазменные телевизоры. Годовой спрос оценивается в 1000 единиц. Стоимость хранения одного телевизора в течение года составляет \$500, а стоимость закупки телевизоров — \$400.
  - Сколько телевизоров следует заказывать в одной закупке?
  - Сколько закупок в год нужно сделать?
  - Каковы годовые затраты на хранение телевизоров и закупку?
- Компания Waterford Crystal может произвести до 100 кувшинов для чая со льдом в день. Допустим, завод работает 250 дней в году и годовой спрос составляет 20 000 кувшинов. Стоимость хранения кувшина в течение года составляет \$10, а стоимость запуска производства кувшинов — \$40.
  - Каков рекомендуемый размер партии кувшинов?
  - Сколько партий в год нужно произвести?
  - Каковы годовые затраты на производство и хранение кувшинов?

## ГЛАВА 85

# Построение моделей управления запасами для неопределенного спроса

### Обсуждаемые вопросы

- При каком уровне запасов следует делать закупку, если необходимо минимизировать годовые затраты, связанные с дефицитом, хранением и закупками?
- Что означает термин «95%-ный уровень обслуживания»?

В главе 84 «Оптимальный размер заказа в модели управления запасами» мы узнали, как вычислять оптимальный размер заказа на поставку и оптимальный размер партии товара. В примерах подразумевалось, что спрос на продукт постоянен. Таким образом, если годовой спрос составляет, например, 1200 единиц в год, ежемесячный спрос равен 100 единицам. Пока спрос держится на постоянном уровне, оптимальный размер заказа является достаточно хорошим приближением для размера заказа, минимизирующего затраты.

В действительности спрос в любой период времени не определен, поэтому возникает естественный вопрос, до какой величины может понизиться уровень запасов, прежде чем потребуется следующая закупка. Уровень запасов, при котором возникает необходимость в закупке, называется *точкой заказа*. Очевидно, что высокая точка заказа снизит потери от дефицита и увеличит затраты на хранение. Аналогично низкая точка заказа увеличит потери от дефицита и снизит затраты на хранение. В некой промежуточной точке заказа сумма потерь от дефицита и затрат на хранение будет минимальной. В первом примере этой главы показано, как определить точку заказа, минимизирующую затраты, связанные с дефицитом, закупкой и хранением, на основе двух следующих предположений.

- Каждая недостающая единица продукции является невыполненным заказом клиента, и компания несет потери от дефицита  $C_B$ . Эти потери в первую очередь являются мерой неудовлетворенности клиента, вызванной задержкой в получении заказа.
- Каждая недостающая единица продукции приводит к упущенной продаже, и компания несет потери  $C_{LS} > C_B$ . Потери упущенной продажи включают прибыль, не полученную вследствие упущенной продажи, а также потери от дефицита  $C_B$ .

Второй пример показывает, как определить оптимальную точку заказа на основе уровня обслуживания. Например, 95%-ный уровень обслуживания означает, что точка заказа устанавливается на уровне, обеспечивающем в среднем удовлетворение 95% всех клиентов в срок. Как правило, стоимость потерь от дефицита и в случае невыполненного заказа, и в случае упущенной продажи определить трудно. По этой причине в большинстве компаний устанавливают точки заказа на основе уровня обслуживания.

## Ответы на вопросы

❓ При каком уровне запасов следует делать закупку, если необходимо минимизировать годовые затраты, связанные с дефицитом, хранением и закупками?

Как указано в главе 84, EOQ зависит от следующих параметров:

- $K$  — затраты на закупки;
- $h$  — затраты на хранение одной единицы продукции в запасе в течение года;
- $D$  — годовой спрос на продукт. Поскольку теперь спрос точно не известен, пусть  $D$  обозначает ожидаемый годовой спрос на продукт.

## Невыполненный заказ

Данные для этого примера находятся в файле `Reorderpoint_backorder.xlsx` (рис. 85.1). Сначала предположим, что каждый дефицит приводит к отложенному заказу. Другими словами, дефицит не приводит к потере спроса. Также подразумевается, что каждая недостающая единица продукции приводит к потерям от дефицита  $C_B$ .

	A	B	C
2	cost/order	K	\$ 50.00
3	annual holding cost per unit	h	\$ 10.00
4	mean annual demand	D	1000
5	order quantity	EOQ	100
6	orders per year	annorders	10
7	unit stockout cost	SOC	\$ 20.00
8	annual sigma	annsig	40.8
9	mean lead time	meanLT	0.0384615
10	sigma lead time	sigmaLT	0
11	mean lead time demand	meanLTD	38.461538
12	sigma lead time demand	sigmaLTD	8.0015383
13	probability of stockout	probout	0.05
14	reorder point	RP	51.622898
15	safety stock	SS	13.161359

**Рис. 85.1.** Определение точки закупки для случая отсрочки выполнения заказа из-за дефицита

В этом случае точка заказа зависит от следующих величин:

- $EOQ$  — оптимальный размер заказа (закупаемое каждый раз количество);
- $K$  — затраты на закупку;
- $h$  — годовые затраты на хранение единицы продукции;
- $D$  — среднегодовой спрос;
- $SOC$  — потери от недостающей единицы продукции;
- $annsig$  — стандартное отклонение годового спроса;
- $meanLT$  — средний срок выполнения заказа, то есть среднее время между размещением и получением заказа;
- $sigmaLT$  — стандартное отклонение срока выполнения заказа.

Предположим, магазин пытается определить оптимальную стратегию хранения для заказанных миксеров. Он располагает следующей информацией:

- заказ на поставку миксеров обходится в \$50;
- хранение миксера на складе в течение года обходится в \$10;
- в среднем магазин продает 1000 миксеров в год;
- все клиенты, которые пытаются приобрести миксеры, когда склад опустел, возвращаются позже и покупают миксеры после новой поставки. Магазин штрафует на \$20 за каждую единицу дефицита;
- годовой спрос на миксеры (на основе исторических данных) имеет стандартное отклонение 40,8;
- срок выполнения отложенного заказа всегда составляет две недели (0,038 года), со стандартным отклонением 0.

После ввода  $K$ ,  $h$  и  $D$  в ячейки C2:C4 в ячейке C5 вычисляется оптимальный размер заказа (100 миксеров). После ввода значений  $SOC$ ,  $annsig$ ,  $meanLT$  и  $sigmaLT$  в ячейки C7:C10 в ячейке C14 вычисляется точка заказа, минимизирующая ожидаемые годовые затраты, связанные с хранением и дефицитом (51,63 миксера). Таким образом, магазин должен закупать 100 миксеров, как только запасы понизятся до 51,62 (или 52) миксера.

*Безопасный уровень запаса*, связанный с заданной точкой заказа, равен (точка заказа) – (средний спрос за срок выполнения заказа).

Универмаг поддерживает безопасный уровень запаса в  $51,62 - 38,46 = 13,16$  миксера (вычислен в ячейке C15). По сути, безопасный запас всегда находится на складе, что приводит к дополнительным затратам на хранение. Но при этом высокий безопасный уровень запаса уменьшает затраты, связанные с дефицитом.

## Упущенные продажи

Теперь предположим, что каждый дефицит приводит к упущенной продаже. Потери, связанные с упущенными продажами, как правило, оцениваются как штраф



за невыполненный заказ плюс прибыль, связанная с проданной единицей продукции. Пусть магазин получает \$20 прибыли от продажи каждого миксера. Тогда потери от дефицита для случая упущенной продажи составляют \$40 (\$20 упущенной прибыли + \$20 штрафа за дефицит).

	A	B	C
2	cost/order	K	\$ 50.00
3	annual holding cost per unit	h	\$ 10.00
4	mean annual demand	D	1000
5	order quantity	EOQ	100
6	orders per year	annorders	10
7	lost sales cost	LSC	\$ 40.00
8	annual sigma	annsig	40.8
9	mean lead time	meanLT	0.038461538
10	sigma lead time	sigmaLT	0
11	mean lead time demand	meanLTD	38.46153846
12	sigma lead time demand	sigmaLTD	8.001538314
13	probability of stockout	probout	0.024390244
14	reorder point	RP	54.22861214
15	safety stock	SS	15.76707368

**Рис. 85.2.** Определение точки заказа для случая упущенной продажи

В файле Reorderpoint\_lostsales.xlsx (рис. 85.2) вы видите мои расчеты точки заказа для случая упущенной продажи. Введя потери от упущенной продажи (\$40) в ячейку C7, я нашел, что оптимальная политика хранения состоит в том, чтобы закупать по 100 миксеров и делать новую закупку, когда запасы снижаются до 54,23 миксера. Безопасный уровень запаса составляет 15,77 миксера, и спрос на миксеры не будет удовлетворен на 2,4%. Обратите внимание, что предположение об упущенной продаже увеличило точку заказа и уменьшило вероятность дефицита. Это произошло потому, что из-за увеличения потерь от дефицита (с \$20 до \$40) магазин прилагает больше усилий для избегания дефицита.

Увеличение неопределенности значительно повышает точку заказа. Например, для случая упущенной продажи, если стандартное отклонение для срока выполнения заказа составляет одну неделю (0,019 года), а не 0, точка закупки повышается до 79,50 миксера, а безопасный запас увеличивается более чем в два раза по сравнению с тем случаем, когда срок исполнения заказа был точно известен.

### ❓ Что означает термин «95%-ный уровень обслуживания»?

Как было сказано ранее в этой главе, 95%-ный уровень обслуживания означает удовлетворение 95% спроса в срок. Поскольку оценить штраф из-за дефицита и/или штраф от упущенной продажи достаточно трудно, многие компании устанавливают безопасный уровень запаса в зависимости от уровня обслуживания. В файле Servicelevelreorder.xlsx (рис. 85.3) точку заказа можно определить в соответствии с любым желаемым уровнем обслуживания.

	A	B	C	D	E	F
1	Service level	SL	0.95			
2	cost/order	K	\$ 50.00			
3	annual holding cost per unit	h	\$ 10.00			
4	mean annual demand	D	1000			
5	order quantity	EOQ	100			
6	orders per year	annorders	10			
7	annual sigma	annsig	69.28			
8	mean lead time	meanLT	0.08333			
9	sigma lead time	sigmaLT	0			
10	mean lead time demand	meanLTD	83.3333			
11	sigma lead time demand	sigmaLTD	19.9994			
12	reorder point	ROP	90.2301			
13	standardized reorder point	SROP	0.34485			DIFF
14	normal loss for stand. ROP	NLSTANDROP	0.25001	=	0.25	6E-08
15	safety stock	SS	6.89674			

**Рис. 85.3.** Определение точки заказа в соответствии с уровнем обслуживания

В качестве примера возьмем аптеку, которая пытается определить оптимальную стратегию хранения лекарства. Аптека должна удовлетворять 95% спроса на лекарственный препарат в срок. Для расчета важны следующие параметры:

- заказ на поставку лекарства стоит \$50;
- затраты на хранение единицы препарата на складе в течение года составляют \$10;
- среднегодовой спрос на препарат составляет 1000 единиц;
- стандартное отклонение годового спроса составляет 69,28 единицы;
- время, затрачиваемое на выполнение заказа, всегда равно одному месяцу (0,083 года).

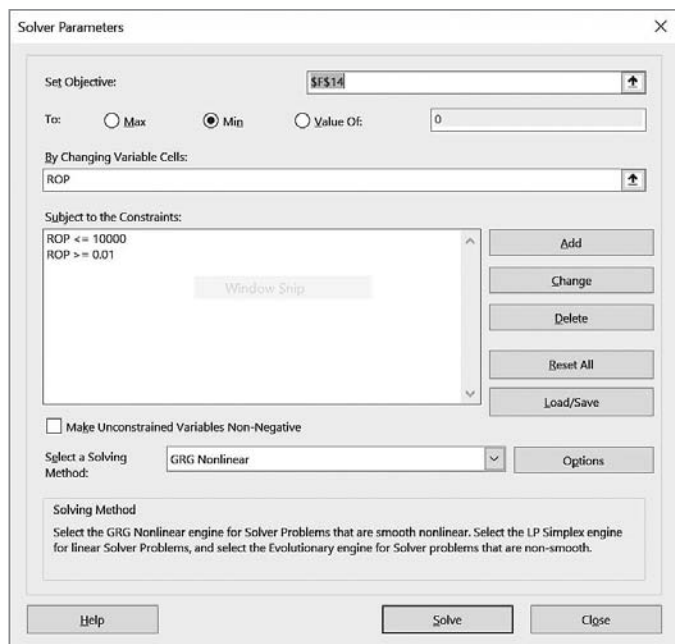
Введите в ячейку C1 требуемый уровень обслуживания (0,95), а в ячейки C2:C4 и C7:C9 остальные параметры. Для определения точки заказа, соответствующей желаемому уровню обслуживания, на вкладке Данные (Data) в группе Анализ (Data Analysis) выберите Поиск решения (Solver).

## ПРИМЕЧАНИЕ

Для установки надстройки Поиск решения (Solver) на вкладке Файл (File) выберите команду Параметры (Options), затем среди параметров Excel выберите раздел Надстройки (Add-Ins). В нижней части диалогового окна нажмите Перейти (Go). В диалоговом окне Надстройки (Add-Ins) установите флажок Поиск решения (Solver Add-In) и нажмите ОК.

В модели поиска решения (рис. 85.4) точка заказа подбирается до тех пор, пока процент удовлетворения спроса в срок не будет соответствовать желаемому уровню обслуживания. Нажмите Параметры (Options) и на вкладке Поиск решения нелинейных задач методом ОНР (GRG Nonlinear) включите флажок Использовать несколько

начальных точек (Multistart) — см. главу 36 «Расположение складов по методу ОПГ с несколькими начальными точками и согласно эволюционному поиску решения». Затем установите для точки заказа нижнюю границу 0,001 (чтобы точке заказа не присвоилось отрицательное значение) и верхнюю границу 10 000 (достоточно произвольно). Если инструмент Поиск решения (Solver) превысит верхнюю границу, мы ее передвинем. Как показано на рис. 85.3, аптека должна заказывать по 100 единиц лекарства каждый раз, когда уровень запаса падает до 90,23 единицы. Эта точка заказа соответствует безопасному уровню запаса 6,90 единицы.



**Рис. 85.4.** Диалоговое окно Параметры поиска решения для определения точки заказа при 95%-ном уровне обслуживания

В таблице ниже приведены точки заказа и безопасные уровни запаса, соответствующие уровням обслуживания от 80 до 99%.

Уровень обслуживания, %	Точка заказа, единицы продукции	Безопасный запас, единицы продукции
80	65,34	–17,99
85	71,85	–11,48
90	79,57	–3,76
95	90,23	6,90
99	108,44	25,11

Обратите внимание, что при повышении уровня обслуживания с 80 до 99% точка заказа увеличивается почти на 67%! Также следует отметить, что 90%-ный уровень обслуживания может быть достигнут при точке заказа, меньшей чем средний спрос за время выполнения заказа (см. ячейку C10 на рис. 85.3). 90%-ный уровень обслуживания приводит к отрицательному безопасному уровню запаса. Это возможно, поскольку дефицит возникает только во время выполнения заказа, а время выполнения заказа охватывает небольшую часть года.

## Задания

В заданиях 1 и 2 предполагается, что в ресторане подается в среднем 5000 бутылок вина в год. Стандартное отклонение годового спроса на вино составляет 1000 бутылок. Ежегодные затраты на хранение бутылки вина составляют \$1. Закупка вина обходится в \$10 и доставка вина в среднем занимает три недели (со стандартным отклонением в одну неделю):

1. Предположим, когда в ресторане заканчивается вино, он получает штраф \$5 как результат потери репутации. Также ресторан получает прибыль в размере \$2 за каждую бутылку вина. Определите оптимальную политику закупки вина.
2. Определите политику хранения вина, которая дает 99%-ный уровень обслуживания.
3. Политику определения точки заказа часто называют политикой двух бункеров<sup>1</sup>. Как может быть реализована политика определения точки заказа, если для хранения запасов используются два бункера?

---

<sup>1</sup> До использования компьютерных информационных систем запасы хранили в двух бункерах. В одном — количество запасов, равное точке заказа, в другом — остальные запасы. Запасы берутся из последнего до тех пор, пока не закончатся, что служит сигналом для следующего заказа на поставку. — *Примеч. ред.*

## ГЛАВА 86

# Теория массового обслуживания (теория очередей)

### Обсуждаемые вопросы

- Какие факторы влияют на длину очереди и время, проведенное клиентами в очереди?
- Какие условия должны выполняться, чтобы можно было анализировать среднее количество присутствующих клиентов или среднее время, проведенное клиентом в очереди?
- Почему изменчивость снижает производительность системы очередей?
- Можно ли быстро определить среднее время прохождения через систему контроля безопасности в аэропорту или среднее время ожидания в очереди в банке?

### Какие условия следует сформулировать, прежде чем приступить к анализу среднего числа посетителей или среднего времени, проводимого в очереди?

По необъяснимому недоразумению, все мы тратим много времени в очередях, и, как скоро будет показано, небольшое увеличение пропускной способности в обслуживании часто может значительно сократить очередь. При ведении бизнеса важно гарантировать, что ваши клиенты не тратят слишком много времени на ожидание. Поэтому деловые люди должны знать теорию очередей, обычно называемую *теорией массового обслуживания*. В этой главе показано, как определить пропускную способность, необходимую для высокого качества обслуживания.

## Ответы на вопросы

### ❓ Какие факторы влияют на длину очереди и время, проведенное клиентами в очереди?

В этой главе рассматриваются проблемы очередей, в которых все прибывающие клиенты ждут первого освободившегося специалиста в одной очереди. Эта модель дает довольно точное представление о распространенных ситуациях, когда вы ждете своей очереди в банке, на почте или при контроле безопасности в аэропорту. Кстати, идея формирования одной очереди возникла примерно в 1970 г.,

когда в банковских и почтовых отделениях осознали, что хотя ожидание в одной очереди не уменьшает среднее время ожидания для клиентов, оно снижает изменчивость времени, проводимого клиентами в очереди, что создает более справедливую систему.

На время, которое клиенты проводят в очереди, влияют три фактора.

- **Количество специалистов.** Очевидно, что чем больше специалистов, тем меньше времени в среднем проводят клиенты в очереди и меньше людей в среднем стоит в очереди.
- **Среднее и стандартное отклонение времени между прибытиями клиентов.** Время между прибытиями клиентов называется *временным интервалом между прибытиями*. Если средний временной интервал между прибытиями увеличивается, количество прибытий уменьшается, что приводит к более короткой очереди и меньшему времени, проведенному клиентом в очереди. Как вы вскоре увидите, увеличение стандартного отклонения временного интервала между прибытиями увеличивает среднее время, проведенное клиентом в очереди, и среднее количество присутствующих клиентов.
- **Среднее значение и стандартное отклонение времени, необходимого для обслуживания клиента.** Если среднее время обслуживания увеличивается, то повышается среднее время, проведенное клиентом в системе массового обслуживания, и количество присутствующих клиентов. Увеличение стандартного отклонения времени обслуживания увеличивает среднее время, проведенное клиентом в системе массового обслуживания, и среднее количество присутствующих клиентов.

❓ **Какие условия должны выполняться, чтобы можно было анализировать среднее количество присутствующих клиентов или среднее время, проведенное клиентом в очереди?**

При анализе времени ожидания клиентов в очередях математики говорят о характеристиках *устойчивого состояния* системы. Устойчивое состояние подразумевает, что система работает в течение длительного времени. Если говорить более конкретно, аналитикам необходимо выяснить значения следующих величин в устойчивом состоянии:

- $W$  — среднее время, проведенное клиентом в системе;
- $W_q$  — среднее время ожидания в очереди до обслуживания;
- $L$  — среднее количество клиентов, присутствующих в системе;
- $L_q$  — среднее количество клиентов, ожидающих в очереди.

При этом всегда справедливы равенства:  $L = (1/\text{среднее время между прибытиями}) \times W$  и  $L_q = (1/\text{среднее время между прибытиями}) \times W_q$ .

Для корректного обсуждения устойчивого состояния системы очередей должно иметь место следующее.

- Среднее значение и стандартное отклонение времени между прибытиями и времени обслуживания мало изменяются с течением времени. С технической точки зрения это означает, что распределение для времени между прибытиями и времени обслуживания является постоянным во времени.
- $(1/\text{среднее время обслуживания}) \times (\text{количество специалистов}) > (1/\text{среднее время между прибытиями})$ . Это уравнение 1.

По сути, если уравнение 1 выполняется, за час можно обслужить больше клиентов, чем число новоприбывших. Например, если среднее время обслуживания составляет 2 минуты (или  $1/30$  часа) и среднее время между прибытиями составляет 1 минуту (или  $1/60$  часа), из уравнения 1 следует, что  $30 \times (\text{количество специалистов}) > 60$  или что для устойчивого состояния системы количество специалистов должно быть больше или равно 3. Если клиенты не обслуживаются быстрее, чем прибывают новые, в конце концов накапливается отставание, которое никогда не будет наверстано, что приведет к бесконечной очереди.

### ❓ Почему изменчивость снижает производительность системы очередей?

Чтобы понять, почему изменчивость снижает производительность системы очередей, рассмотрим систему с одним специалистом, в которой клиенты прибывают каждые 2 минуты и время обслуживания также всегда составляет 2 минуты. В такой системе никогда не будет больше одного клиента. Теперь предположим, что клиенты прибывают каждые 2 минуты, но половина всего времени обслуживания составляет 0,5 минуты, а другая половина — 3,5 минуты. Даже если прибытия полностью предсказуемы, неопределенность во времени обслуживания означает, что, в конце концов, специалист перестанет справляться и сформируется очередь. Например, если на обслуживание каждого из первых четырех клиентов уйдет 3,5 минуты, через 12 минут в очереди будут ожидать четыре клиента (см. табл.).

Время, минуты	Событие	Находится в системе после события, кол-во человек
0	Прибытие	1
2	Прибытие	2
3,5	Обслуживание завершено	1
4	Прибытие	2
6	Прибытие	3
7	Обслуживание завершено	2
8	Прибытие	3
10	Прибытие	4
10,5	Обслуживание завершено	3
12	Прибытие	4

**?** Можно ли быстро определить среднее время прохождения через систему контроля безопасности в аэропорту или среднее время ожидания в очереди в банке?

На листе Model в файле Queuingtemplate.xlsx находится шаблон, который вы можете использовать для определения приблизительных значений  $L$ ,  $W$ ,  $L_q$  и  $W_q$  (как правило, в пределах 10% от их истинных значений). Этот лист представлен на рис. 86.1.

	A	B	C
3	Arrival rate	0.077734	per minute
4	Service rate	0.01297	per minute
5	s(servers)	6	
6	Mean interarrival time	12.864	
7	Mean service time	77.102	
8	Standard deviation of interarrival times	4.43908	
9	Standard deviation of service times	48.05051	
10	CV arrive	0.119079	
11	CV service	0.388387	
12	u	5.993412	
13	ro	0.998902	
14	R(s,mu)	0.73554	
15	E <sub>c</sub> (s,mu)	0.996956	
16	W <sub>q</sub>	2960.658	
17	L <sub>q</sub>	230.1434	
18	W	3037.76	
19	L	236.1368	

**Рис. 86.1.** Шаблон для системы очередей

Когда вы введете следующие данные в шаблон, он вычислит значения  $W_q$ ,  $L_q$ ,  $W$  и  $L$ . Параметры в ячейках B6:B9 могут быть легко рассчитаны на основе данных на листе Queuing Data.

- Количество специалистов (ячейка B5).
- Среднее время между прибытиями (ячейка B6).
- Среднее время обслуживания (ячейка B7).
- Стандартное отклонение времени между прибытиями (ячейка B8).
- Стандартное отклонение времени обслуживания (ячейка B9).

А теперь — пример шаблона в действии. Вам нужно определить, как рабочие характеристики системы безопасности в аэропорту в рабочее время с 9:00 до 17:00 зависят от количества сотрудников. На листе Queuing Data в файле Queuingtemplate.xlsx (рис. 86.2) в таблице приведено время между прибытиями и время обслуживания. (Часть строк данных скрыта.)

Скопировав формулу =CPЗНАЧ(B4:B62) из ячейки B1 в C1, вы найдете среднее время между прибытиями (12,864 с) и среднее время обслуживания (77,102 с).



	A	B	C	D
1	mean	12.86440678	77.10169492	seconds!
2	sigma	4.43908047	48.05051039	
3		Interarrival times	Service Time	
4		5	95	
5		17	240	
6		12	71	
7		18	68	
8		9	90	
9		16	117	
10		15	291	
11		15	116	
12		10	107	
13		11	100	
14		9	28	
15		15	119	
16		19	98	
17		9	72	

**Рис. 86.2.** Время между прибытиями и время обслуживания для примера с системой контроля безопасности аэропорта

Поскольку среднее время обслуживания почти в шесть раз больше среднего времени между прибытиями, для обеспечения устойчивого состояния требуется минимум шесть сотрудников. Скопировав формулу =СТАНДОТКЛОН.В(B4:B62) из ячейки B2 в C2, вы определите стандартное отклонение времени между прибытиями (4,439 с) и стандартное отклонение времени обслуживания (48,051 с).

Вернемся к шаблону системы очередей на листе Model. Когда вы введете полученные значения в ячейки B6:B9 и 6 специалистов в ячейку B5, то выясните, что катастрофа неизбежна. В устойчивом состоянии системы в очереди будут находиться около 236 человек (ячейка B19). Возможно, вам приходилось оказываться в такой ситуации в аэропорту.

Чтобы определить, как изменение числа сотрудников влияет на производительность системы, я воспользовался однонаправленной таблицей данных (рис. 86.3). В ячейки F9:F13 я ввел предположительное число сотрудников (от 6 до 10). Для вычисления  $L$  я ввел в ячейку G8 формулу =B19, а для вычисления  $W$  — в ячейку H8 формулу =B18. Выделил таблицу (F8:H13) и затем на вкладке Данные (Data) в группе Прогноз (Forecast) в списке Анализ "что, если" (What If Analysis) выбрал Таблица данных (Data Table). После выбора ячейки B5 (количество сотрудников) в поле Подставлять значения по строкам в (Column Input Cell) я получил таблицу, как на рис. 86.3. Обратите внимание, что добавление всего лишь одного сотрудника к исходным шести уменьшает ожидаемое количество клиентов в очереди с 236 до 7. Добавление седьмого сотрудника уменьшает среднее время нахождения клиента в системе с 3038 с (50,6 мин) до 89 с (1,5 минуты). Этот пример показывает, что небольшое увеличение пропускной способности может значительно повысить производительность системы очередей.

	E	F	G	H	I	J	K
7			L	W			
8		Servers	236.1368	3037.76			
9		6	236.1368	3037.76			
10		7	6.917396	88.9882			
11		8	6.262791	80.56709			
12		9	6.092038	78.37046			
13		10	6.031636	77.59342			
14							
15					servers		
16		3037.76	6	7	8	9	10
17		40	2342.086	86.1952	79.75282	78.07234	77.47788
18	service time	50	3225.354	89.74136	80.78666	78.45085	77.62457
19	sigma	60	4304.905	94.07556	82.05026	78.91348	77.80387
20		70	5580.737	99.1978	83.54359	79.46023	78.01577
21		80	7052.851	105.1081	85.26668	80.09109	78.26027
22		90	8721.247	111.8064	87.2195	80.80607	78.53736

**Рис. 86.3.** Анализ чувствительности для системы контроля безопасности аэропорта

В ячейках F16:K22 я использовал двунаправленную таблицу данных для исследования чувствительности среднего времени, проведенного клиентом в системе ( $W$ ), к изменению количества сотрудников и изменению стандартного отклонения времени обслуживания. Укажите в поле Подставлять значения по столбцам в (Row Input Cell) ячейку B5, а в поле Подставлять значения по строкам в (Column Input Cell) ячейку B9. Если работают семь сотрудников, увеличение стандартного отклонения времени обслуживания с 40 до 90 с приводит к увеличению на 29% среднего времени, проведенного клиентом в системе (с 86,2 до 111,8 с).

#### ПРИМЕЧАНИЕ

Читатели, заинтересованные в более широком обсуждении теории массового обслуживания, могут обратиться к книге Уэйна Винстона (Wayne L. Winston. Operations Research: Applications and Algorithms. — Duxbury Press, 2007).

## Задания

В банке работают шесть операционистов. Выполните задания 1–4 на основе следующей информации:

- ☐ среднее время обслуживания составляет 1 минуту;
- ☐ среднее время между прибытиями составляет 25 секунд;
- ☐ стандартное отклонение времени обслуживания составляет 1 минуту;
- ☐ стандартное отклонение времени между прибытиями составляет 10 секунд.

Выполните задания:

1. Определите среднее время ожидания клиента в очереди.
2. Сколько клиентов в среднем присутствуют в очереди?
3. Рекомендовали бы вы подключить к работе дополнительных операционистов?
4. Допустим, работа операциониста стоит \$20 в час, а время клиента банк оценивает в \$15 в час. Сколько операционистов должно работать в банке?
5. На пятом этаже бизнес-школы работают 30 женщин. Каждая женщина посещает туалет три раза в день (офис работает с 8:00 до 17:00) и проводит в туалете каждый раз в среднем 180 секунд со стандартным отклонением 90 секунд. Предположим, что стандартное отклонение времени между посещениями туалета составляет 5 минут. Сколько кабинок вы бы порекомендовали установить в туалете?

## ГЛАВА 87

# Оценка кривой спроса

### Обсуждаемые вопросы

- Что необходимо знать при назначении цены продукта?
- Что такое эластичность спроса?
- Существует ли простой способ оценить кривую спроса?
- Каким образом кривая спроса показывает готовность покупателя платить за продукт?

В любом бизнесе необходимо определить цену для каждого продукта. Оценить продукт адекватно достаточно сложно. В главах 88 «Ценообразование продуктов с сопутствующими товарами» и 89 «Ценообразование продуктов с помощью субъ-ективно определяемого спроса» я опишу несколько простых моделей, помогающих назначить цену на продукт, максимально увеличивающую его прибыльность. Для более полного представления о ценообразовании см. прекрасную книгу Роберта Долана и Германа Саймона<sup>1</sup>.

## Ответы на вопросы

### ❓ Что необходимо знать при назначении цены продукта?

Возьмем такой продукт, как шоколадный батончик. Для определения цены, приносящей максимальную прибыль, необходимо знать две вещи:

- переменные затраты на производство каждой единицы продукции (я назвал эту величину  $UC$ );
- кривую спроса на продукт. Кривая спроса показывает количество единиц продукта, на которое будет спрос, в зависимости от конкретной цены. Таким образом, если назначить цену единицы продукции  $\$ p$ , то по кривой спроса можно определить число  $D(p)$ , эквивалентное количеству единиц продукции, которое будет востребовано по цене  $\$ p$ . Безусловно, кривая спроса постоянно меняется и нередко зависит от факторов, находящихся вне контроля компании (таких как состояние экономики и цены конкурентов).

---

<sup>1</sup> Robert J. Dolan, Hermann Simon. *Power Pricing*. — Free Press, 1997.

Если значение  $UC$  и кривая спроса известны, прибыль, соответствующая цене в  $\$p$ , просто равна  $(p - UC) \times D(p)$ . Если известно уравнение для  $D(p)$ , вычисляющее количество востребованного продукта для каждой цены, можно с помощью инструмента Поиск решения (Solver) найти цену, обеспечивающую максимальную прибыль, как будет показано в главах 88 и 89.

### ? Что такое эластичность спроса?

В соответствии с кривой спроса *ценовая эластичность* спроса — это процентное сокращение спроса в результате повышения цены на 1%. Если эластичность выше 1%, спрос является эластичным по цене. Когда спрос эластичен по цене, снижение цены увеличивает выручку. Если эластичность меньше 1%, спрос является неэластичным по цене. Когда спрос неэластичен по цене, снижение цены снижает выручку. Далее приведены отслеженные оценки эластичности спроса на некоторые продукты.

- Соль — 0,1 (крайне неэластичный).
- Кофе — 0,25 (неэластичный).
- Гонорары адвокатов — 0,4 (неэластичный).
- Телевизоры — 1,2 (незначительно эластичный).
- Ресторанная еда — 2,3 (эластичный).
- Поездки за границу — 4,0 (очень эластичный).

Например, снижение стоимости поездки за границу на 1% приводит к увеличению спроса на такие поездки на 4%.

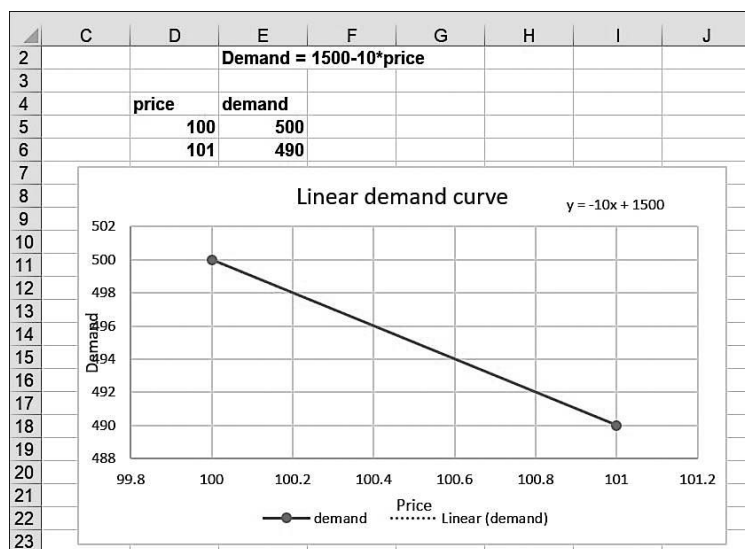
### ? Существует ли простой способ оценить кривую спроса?

Пусть  $q$  — величина спроса на продукт. Для оценки кривой спроса наиболее часто используются две формы.

- **Линейная кривая спроса.** В этом случае спрос выражается линейным соотношением вида  $q = a - bp$ . Например,  $q = 10 - p$  — линейная кривая спроса. (Здесь  $a$  и  $b$  могут быть определены по методу, который я описываю далее в этой главе.) Если кривая спроса линейна, эластичность постоянно меняется.
- **Степенная кривая спроса.** В этой ситуации кривая спроса описывается степенной кривой вида  $q = ap^b$ , где  $b < 0$  (см. главу 57 «Степенная кривая»). Как и в предыдущем случае,  $a$  и  $b$  можно определить по методу, который я описываю далее в этой главе. Уравнение  $q = 100p^{-2}$  является примером степенной кривой спроса. Если спрос соответствует степенной кривой, эластичность равна  $-b$  для любой цены. Таким образом, для кривой спроса  $q = 100p^{-2}$  ценовая эластичность спроса всегда равна 2.

Предположим, что кривая спроса на продукт соответствует линейной или степенной кривой спроса. Если известна текущая цена и спрос на продукт, а также ценовая эластичность спроса, определить кривую спроса для продукта проще простого. Вот два примера.

Пусть в настоящее время продукт продается за \$100, и спрос на него составляет 500 единиц. Ценовая эластичность спроса для продукта равна 2. Предположим, что кривая спроса является линейной. Необходимо определить уравнение кривой спроса. Решение находится в файле *Linearfit.xlsx* (рис. 87.1).



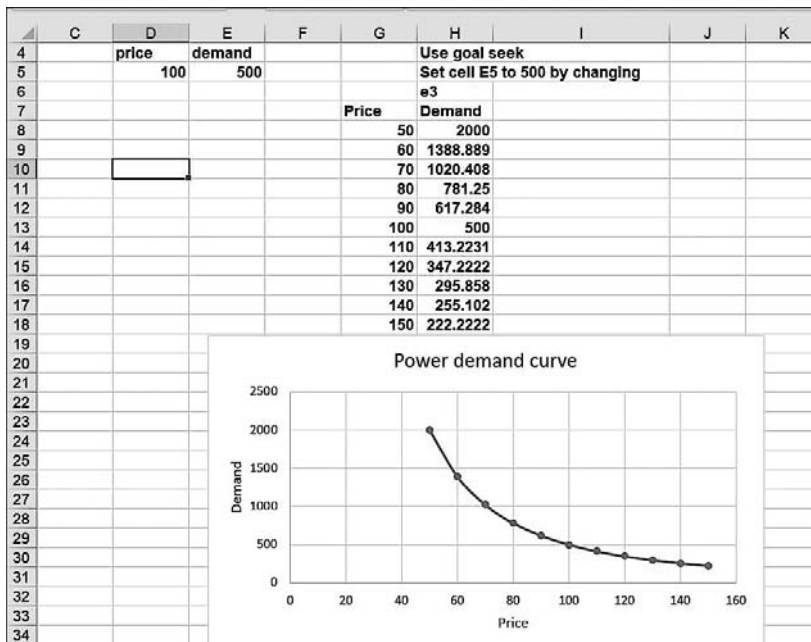
**Рис. 87.1.** Подбор линейной кривой спроса

Если заданы две точки, то через них можно провести только одну прямую. Для кривой спроса действительно известны две точки, и первая из них —  $p = 100$  и  $q = 500$ . Поскольку эластичность спроса равна 2, повышение цены на 1% приводит к уменьшению спроса на 2%. Следовательно, если  $p = 101$  (повышение на 1%), спрос падает на 2% от 500 (10 единиц) до 490. Таким образом,  $p = 101$  и  $q = 490$  — это вторая точка кривой спроса. Теперь с помощью линии тренда в Excel можно найти прямую линию, проходящую через точки (100,500) и (101,490).

Начнем с ввода этих точек в ячейки D5:E6, как показано на рис. 87.1. Затем выделите диапазон D4:E6 и на вкладке Вставка (Insert) в группе Диаграммы (Charts) выберите Точечная или пузырьковая диаграмма (Insert Scatter (X, Y) Or Bubble Chart) и затем последний тип точечной диаграммы — Точечная диаграмма с прямыми отрезками и маркерами (Scatter With Straight Lines). График получился с положительным наклоном. То есть повышение цены ведет к повышению спроса, а это неверно. Проблема в следующем: при наличии только двух точек данных Excel предполагает, что точки данных, по которым строится график, находятся в отдельных столбцах, а не в отдельных строках. В подтверждение того, что отдельные точки находятся в разных строках, щелкните в области диаграммы и затем в секции Работа с диаграммами (Chart Tools) откройте вкладку Конструктор (Design). На этой вкладке в группе Данные (Data) выберите Строка/столбец (Switch Row/Column). Обрати-

те внимание, что после нажатия **Выбрать данные (Select Data)** можно изменить исходные данные, на основе которых создается диаграмма. Теперь щелкните правой кнопкой мыши на одной из точек, выберите **Добавить линию тренда (Add Trendline)**, затем установите переключатель в положение **Линейная (Linear)** и отметьте флажок **Показывать уравнение на диаграмме (Display Equation On Chart)**. Появится прямолинейный график, а также уравнение прямой (вверху справа от графика), как показано на рис. 87.1. Так как  $x$  равен цене, а  $y$  — спросу, уравнение для кривой спроса имеет вид  $q = 1500 - 10p$ . Это уравнение означает, что каждое повышение цены на \$1 приводит к потере 10 единиц спроса. Разумеется, спрос не может быть линейным для всех значений цены, поскольку для больших значений  $p$  линейная кривая спроса приведет к негативному спросу. Однако для цен, близких к текущей цене, линейная кривая спроса является достаточно хорошим приближением к истинной кривой спроса на продукт.

Для второго примера давайте снова предположим, что в настоящее время продукт продается за \$100 и спрос на него составляет 500 единиц. Ценовая эластичность спроса для продукта равна 2. Теперь подберем к этим данным степенную кривую спроса. См. файл Powerfit.xlsx и рис. 87.2.



**Рис. 87.2.** Степенная кривая спроса

Введите в ячейку E3 пробное значение  $a$ . Затем в ячейку D5 введите текущую цену \$100. Так как эластичность спроса равна 2, кривая спроса имеет вид  $q = ap^{-2}$ , где  $a$  неизвестно. В ячейку E5 введите спрос для цены \$100, соответствующий зна-

чению  $a$  в ячейке E3 по формуле  $=a*D5^{-2}$ . Теперь с помощью инструмента Подбор параметра (Goal Seek) (см. главу 18) определите значение  $a$ , при котором спрос для цены \$100 составляет 500 единиц. Установите в ячейке E5 значение 500, изменив ячейку E3. Оказывается, спросу в 500 единиц при цене \$100 соответствует  $a$ , равное 5 000 000. Таким образом, кривая спроса (графически представленная на рис. 87.2) описывается уравнением  $q = 5000000p^{-2}$ . Для любой цены ценовая эластичность спроса на этой кривой равна 2.

### ❓ Каким образом кривая спроса показывает готовность покупателя платить за продукт?

Вы пытаетесь продать ПО крупной корпорации. Пусть  $q$  — это количество копий ПО, которое требуется компании, а  $p$  — цена ПО. Вы вычислили, что кривая спроса на ПО описывается уравнением  $q = 400 - p$ . Очевидно, что клиент хотел бы платить меньше за каждую дополнительную копию ПО. В этой кривой спроса заключена информация о том, сколько компания готова платить за каждую копию ПО. Эта информация является ключевой для максимального увеличения прибыльности продаж.

Перепишем уравнение кривой спроса как  $p = 400 - q$ . Таким образом, если  $q = 1$ ,  $p = \$399$  и т. д. Теперь попытаемся определить ценность, которую клиент придает каждой из первых двух копий ПО. Предположим, что клиент прагматичен, то есть будет покупать копию ПО тогда и только тогда, когда ценность этой копии превысит запрошенную цену. При цене \$400 спрос равен 0, так что первая копия не может стоить \$400. Однако при цене \$399 образуется спрос на одну копию ПО. Следовательно, первая копия ПО должна стоить где-то между \$399 и \$400. Аналогично, вторую копию ПО клиент за \$399 не купит. Однако при цене \$398 клиент покупает две копии, то есть клиент покупает вторую копию. Следовательно, потребительская ценность второй копии лежит где-то между \$399 и \$398.

Можно показать, что наилучшим приближением стоимости  $i$ -й единицы продукта, покупаемой клиентом, является цена, при которой спрос равен  $i - 0,5$ . Например, если установить для  $q$  значение 0,5, то стоимость первой копии равна  $400 - 0,5 = \$399,50$ . Аналогично, если установить  $q = 1,5$ , стоимость второй копии равна  $400 - 1,5 = \$398,50$ .

## Задания

1. Вы оценили изобретенную вами настольную игру в \$60 и продали 3000 экземпляров за последний год. Известно, что ценовая эластичность для настольных игр равна 3. На этой основе определите линейную и степенную кривые спроса.
2. Для каждого ответа из задания 1 определите потребительскую стоимость двухтысячного экземпляра игры.



## ГЛАВА 88

# Ценообразование продуктов с сопутствующими товарами

Обсуждаемый вопрос

- Как покупка клиентами вместе с бритвенным станком лезвий способствует установлению максимально увеличивающей прибыль цены на бритву?

Покупка некоторых потребительских товаров часто приводит к покупке сопутствующих товаров, или товаров принудительного ассортимента. В следующей таблице приведено несколько примеров таких покупок.

Первичная покупка	Сопутствующий товар
Бритвенный станок	Лезвия
Мужской костюм	Рубашка и/или галстук
Компьютер	Руководство по программному обеспечению
Игровая приставка	Видеоигра

Кривую спроса для первичного продукта можно легко определить с помощью методов, описанных в главе 87 «Оценка кривой спроса». Затем вы можете использовать Поиск решения (Solver) и рассчитать цену первичного продукта, максимизирующую сумму прибылей, полученных от исходного и сопутствующего товаров. Следующий пример демонстрирует, как проводится такой анализ.

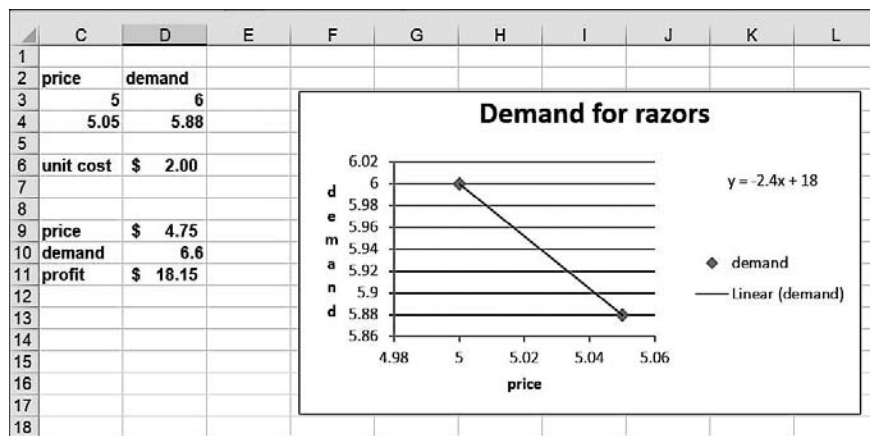
## Ответ на вопрос

- ❓ **Как покупка клиентами вместе с бритвенным станком лезвий способствует установлению максимально увеличивающей прибыль цены на бритву?**

Пусть в настоящее время вы берете за бритвенный станок \$5,00 и продаете 6 млн бритвенных станков. Предположим, что переменные затраты на производство одного станка составляют \$2,00. Наконец, допустим, что ценовая эластичность спроса на бритвенные станки равна 2. Какую цену следует назначить на станок?

Предположим (ошибочно), что покупатели бритвенных станков не покупают лезвия. Кривая спроса (в случае линейной кривой спроса) показана на рис. 88.1.

(См. данные и диаграмму на листе **No Blades** в файле **Razorsandblades.xlsx**.) Двумя точками на кривой спроса являются цена = \$5,00, спрос = 6 млн бритв и цена = 5,05 (увеличенная на 1%), спрос = 5,88 млн (на 2% меньше 6 млн). После создания диаграммы и вставки линейной линии тренда можно, как показано в главе 87 «Оценка кривой спроса», найти уравнение кривой спроса:  $y = 18 - 2,4x$ . Поскольку  $x$  — это цена, а  $y$  — спрос, уравнение кривой спроса на бритвенные станки можно записать в следующем виде: спрос (в млн) =  $18 - 2,4(\text{цена})$ .



**Рис. 88.1.** Определение цены на бритвенные станки, максимизирующей прибыль

Я связал имена в ячейках C6 и C9:C11 с ячейками D6 и D9:D11 соответственно. Затем я ввел пробную цену в ячейку D9 и определил спрос для этой цены в ячейке D10 по формуле  $=18-2,4*\text{цена}$ . Далее я определил в ячейке D11 прибыль от бритвенных станков по формуле  $=\text{спрос}*(\text{цена}-\text{себестоимость\_бритвы})$ .

Затем я применил Поиск решения (Solver) со вкладки Данные (Data) для определения цены, максимизирующей прибыль. Диалоговое окно Параметры поиска решения (Solver Parameters) показано на рис. 88.2.

## ПРИМЕЧАНИЕ

Если вы еще не установили надстройку Поиск решения (Solver), на вкладке Файл (File) выберите команду Параметры (Options), затем среди параметров Excel выберите раздел Надстройки (Add-Ins). В нижней части диалогового окна нажмите Перейти (Go). В диалоговом окне Надстройки (Add-Ins) установите флажок Поиск решения (Solver Add-In) и нажмите ОК.

Я максимизировал прибыль (ячейка D11), изменив цену (ячейка D9). Эта модель нелинейная, поскольку в целевой ячейке перемножаются две величины — спрос и (цена-себестоимость\_бритвы), каждая из которых зависит от значения в изменяемой ячейке. Инструмент Поиск решения (Solver) выполнит поиск цены за бритвен-

ный станок, максимизирующей прибыль, и получит \$4,75. (Максимальная прибыль составляет \$18,15 млн.)

Теперь предположим, что покупатель бритвенного станка приобретает в среднем 50 лезвий и каждое купленное лезвие приносит прибыль \$0,15. Каким образом

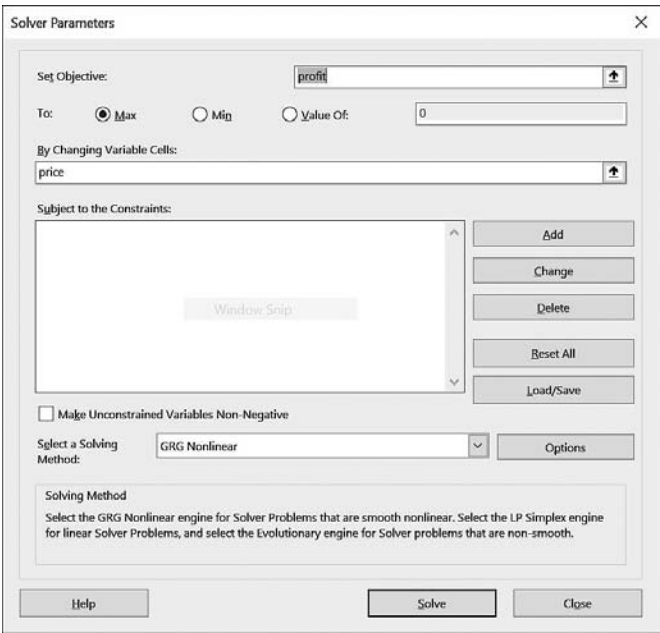


Рис. 88.2. Диалоговое окно Параметры поиска решения для поиска цены бритвы, максимизирующей прибыль

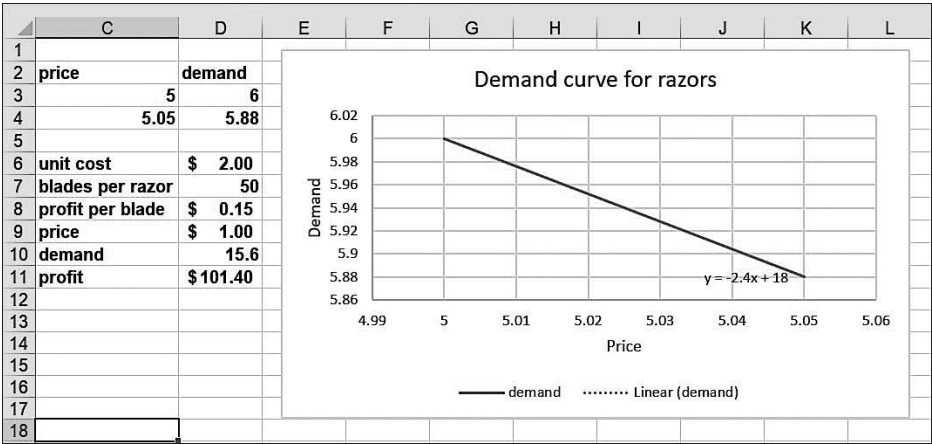


Рис. 88.3. Цена на станки с включенной прибылью от лезвий

должна измениться цена, назначаемая на бритвенный станок? Предположим, что цена на бритвенный станок фиксированная. (В задании 3 в конце этой главы цена станка меняется.) Анализ этой ситуации представлен на листе **Blades** и на рис. 88.3.

На вкладке **Формулы (Formulas)** в группе **Определенные имена (Defined Names)** я выбрал **Создать из выделенного (Create From Selection)** для присвоения имен из ячеек **C6:C11** ячейкам **D6:D11**. (Например, ячейка **D10** называется **Demand (спрос)**.)

## ПРИМЕЧАНИЕ

Внимательные читатели помнят, что ячейка **D10** на листе **No Blades** также носит имя «спрос». Что сделает Excel, если вы укажете имя **Спрос** в формуле? Он просто использует ячейку с именем **Спрос** на текущем листе. Другими словами, если указано имя **Спрос** на листе **Blades**, Excel обращается к ячейке **D10** на этом листе, а не к ячейке **D10** на листе **No Blades**.

В ячейки **D7** и **D8** я ввел соответствующие данные о лезвиях. В ячейку **D9** я ввел пробное значение цены бритвенного станка и в ячейке **D10** вычислил спрос по формуле  $=18-2,4*\text{цена}$ . Далее в ячейке **D11** по формуле  $=\text{спрос}*(\text{цена}-\text{себестоимость\_бритвы})+\text{спрос}*\text{лезвия\_для\_бритвы}*\text{прибыль\_от\_лезвия}$  я вычислил общую прибыль от бритвенных станков и лезвий. Обратите внимание, что часть формулы  $\text{спрос}*\text{лезвия\_для\_бритвы}*\text{прибыль\_от\_лезвия}$  вычисляет прибыль от лезвий.

Диалоговое окно **Параметры поиска решения (Solver Parameters)** выглядит так же, как было показано на рис. 88.2. Изменяйте цену для получения максимальной прибыли. Безусловно, теперь формула прибыли включает прибыль, полученную от лезвий. Решение показывает, что прибыль максимальна при цене всего лишь \$1,00 (половина переменных затрат!) за бритвенный станок. Такая цена является результатом получения большой прибыли от лезвий. Намного выгоднее продать бритвенные станки множеству клиентов, даже если вы теряете по \$1 на каждом проданном станке. Многие компании не понимают важность прибыли от сопутствующих товаров. Они завышают цену на первичный продукт и лишают себя возможности получить максимальную общую прибыль.

## Задания

Во всех заданиях предполагается линейная кривая спроса.

1. Определите цену игровой приставки, максимизирующую прибыль. В настоящее время на приставку назначена цена \$180, и объем продаж составляет 2 млн приставок в год. Стоимость производства приставки — \$150, и ценовая эластичность спроса на приставку равна 3. Какую цену следует назначить на приставку?
2. Теперь предположим, что покупатель игровой приставки покупает в среднем 10 видеоигр, и прибыль от каждой игры составляет \$10. Какова правильная цена на игровую приставку?

3. Предположим, что в примере с бритвами и лезвиями стоимость производства лезвия составляет \$0,20. Если за лезвие назначена цена \$0,35, клиент покупает в среднем 50 лезвий. Пусть ценовая эластичность спроса на лезвия равна 3. Какую цену следует назначить на станок и на лезвие?
4. Кинотеатр может обслужить 8000 зрителей в неделю. Спрос, текущая цена и эластичность для продаж билетов, попкорна, лимонада и конфет приведены на рис. 88.4. Кинотеатр оставляет себе 45% выручки от билетов. Даны также себестоимость билета, продаж попкорна, продаж конфет и продаж лимонада. Каким образом кинотеатр может получить максимальную прибыль при условии линейной кривой спроса? Спрос на продукты питания — это доля клиентов, покупающих указанные продукты. См. данные на рис. 88.4.

	B	C	D	E	F	G	H	I
2								
3			elasticity	current price	demand	cost	ticket percentage	
4	keep 45%	ticket	3	8	3000	0	0.45	
5		popcorn	1.3	3.5	0.5	0.4		
6		soda	1.5	3	0.6	0.6		
7		candy	2.5	2.5	0.2	1		

**Рис. 88.4.** Данные к заданию 4

5. Лекарства, отпускаемые по рецепту, производятся в США и продаются по всему миру. Производство каждой единицы лекарственного препарата стоит \$60. На немецком рынке препарат продается по 150 евро. Текущий обменный курс составляет 0,667 доллара США за евро. Текущий спрос на препарат составляет 100 единиц, а расчетная эластичность равна 2,5. Определите соответствующую продажную цену (в евро) на препарат при условии линейной кривой спроса.
6. Предположим, спрос на костюмы в магазине мужской одежды соответствует линейной кривой, эластичность цены равна 5. При текущей цене в \$500 продается 60 костюмов в неделю. Себестоимость костюма составляет \$200. Предположим, к костюму клиент покупает два галстука и одну рубашку. Прибыль от продажи галстука составляет \$15, а от продажи рубашки — \$25. Какую цену на костюм должен установить магазин?
7. В настоящее время региональный агент по продаже спортивных автомобилей продает за год 200 машин по цене \$150 000. Дилерская цена закупки автомобиля составляет \$60 000. Предположим, за время эксплуатации автомобиля дилер зарабатывает \$25 000 на его обслуживании. Годовой спрос на автомобили следует линейной кривой с эластичностью 4. Какую цену на автомобиль должен установить дилер, чтобы получить максимальную прибыль от продаж и обслуживания?

## ГЛАВА 89

# Ценообразование продуктов с помощью субъективно определяемого спроса

### Обсуждаемые вопросы

- Иногда ценовая эластичность продукта неизвестна. Или невозможно определить, какая из кривых спроса, линейная или степенная, является релевантной. Можно ли все равно оценить кривую спроса и воспользоваться инструментом Поиск решения, чтобы определить цену, максимизирующую прибыль?
- Как небольшой аптеке определить цену на помаду, максимизирующую прибыль?

## Ответы на вопросы

**?** Иногда ценовая эластичность продукта неизвестна. Или невозможно определить, какая из кривых спроса, линейная или степенная, является релевантной. Можно ли все равно оценить кривую спроса и воспользоваться инструментом Поиск решения, чтобы определить цену, максимизирующую прибыль?

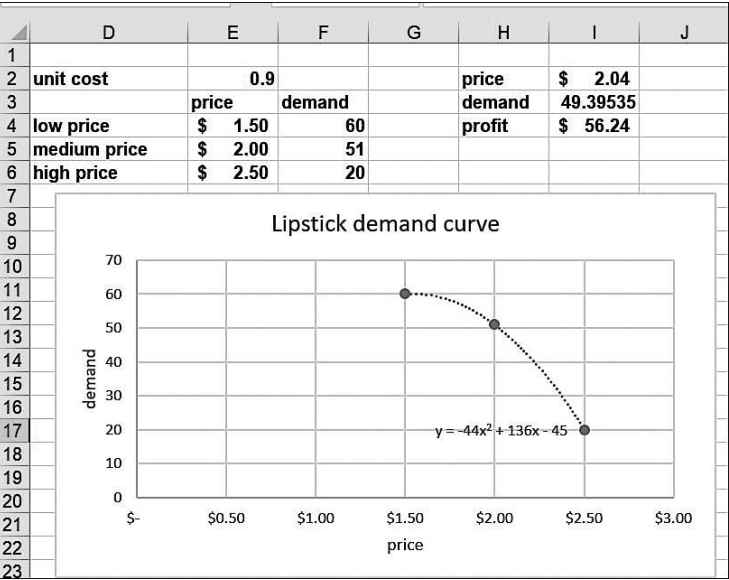
Если ценовая эластичность для продукта неизвестна или если невозможно выбрать, на линейную или на степенную кривую спроса полагаться, эффективным способом определения кривой спроса является выявление самой низкой и самой высокой разумной цены. Затем следует оценить спрос на продукт при высокой, низкой и промежуточной цене. При наличии этих трех точек на кривой спроса на продукт можно с помощью линии тренда в Microsoft Excel подобрать квадратичскую кривую спроса по следующей формуле (уравнение 1):

$$\text{спрос} = a(\text{цена})^2 + b(\text{цена}) + c.$$

Для любых трех указанных точек на кривой спроса существуют значения  $a$ ,  $b$  и  $c$ , при которых уравнение 1 точно соответствует трем указанным точкам. Поскольку уравнение 1 соответствует трем точкам на кривой спроса, по-видимому, разумно предположить, что уравнение даст точное представление о спросе и для других цен. Затем вы можете использовать уравнение 1 и инструмент Поиск решения (Solver), чтобы определить максимальную прибыль по формуле (цена-себестоимость)\*спрос. В следующем примере показано, как это сделать.

**?** Как небольшой аптеке определить цену на помаду , максимизирующую прибыль?

Предположим, что аптека приобретает помаду за \$0,90 за штуку. Для помады рассматривается цена от \$1,50 до \$2,50. В аптеке полагают , что при цене \$1,50 будет продаваться 60 штук в неделю (см. рис. 89.1 и файл Lipstickprice.xlsx), при цене \$2,00 — 51 штука в неделю и при цене \$2,50 — 20 штук в неделю. Какую цену на помаду должна назначить аптека?

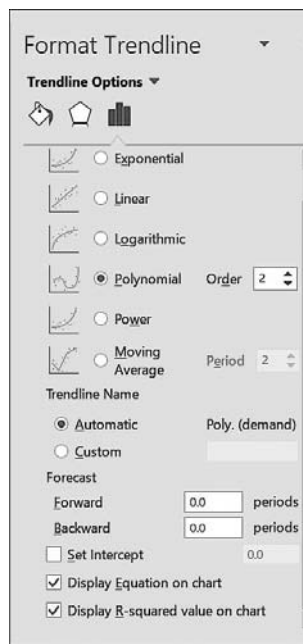


**Рис. 89.1.** Модель ценообразования на помаду

Начните с ввода в ячейки E3:F6 трех точек, по которым будет строиться кривая спроса. Выделите ячейки E3:F6, на вкладке Вставка (Insert) в группе Диаграммы (Charts) выберите Точечная или пузырьковая диаграмма (Insert Scatter (X, Y) Or Bubble Chart) и затем выберите первый тип точечной диаграммы (Scatter). Затем щелкните правой кнопкой мыши на точке данных и в контекстном меню выберите Добавить линию тренда (Add Trendline). В диалоговом окне Формат линии тренда (Format Trendline), показанном на рис. 89.2, установите переключатель в положение Полиномиальная (Polynomial) и в поле со списком Степень (Order) выберите значение 2 (для получения квадратичной кривой по уравнению 1). Установите флажок Показывать уравнение на диаграмме (Display Equation on chart).

Появится диаграмма, показанная на рис. 89.1. Уравнение вычисленной кривой спроса (уравнение 2):

$$\text{спрос} = -44 \times \text{цена}^2 + 136 \times \text{цена} - 45.$$



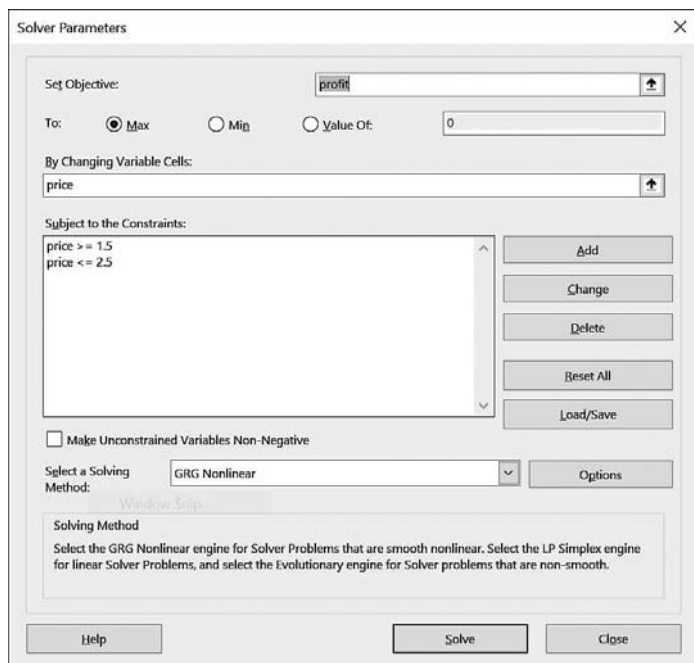
**Рис. 89.2.** Диалоговое окно Формат линии тренда для полиномиальной кривой спроса

Далее введите пробное значение цены в ячейку I2. В ячейке I3 в соответствии с уравнением 2 вычислите спрос на продукт по формуле  $-44 * \text{цена}^2 + 136 * \text{цена} - 45$ . (Ячейке I2 присвоено имя **цена**.) Потом в ячейке I4 по формуле  $=\text{спрос} * (\text{цена} - \text{себестоимость})$  вычислите еженедельную прибыль от продаж помады. (Ячейке E2 присвоено имя **себестоимость**, а ячейке I3 — имя **спрос**.) Затем с помощью инструмента **Поиск решения (Solver)** определите цену, приносящую максимальную прибыль. Диалоговое окно **Параметры поиска решения (Solver Parameters)** показано на рис. 89.3. Обратите внимание, что цена ограничена самой низкой и самой высокой ценой (от \$1,50 до \$2,50). Если так не сделать, квадратичная кривая спроса может пойти вверх, и тогда получится, что более высокая цена приводит к более высокому спросу. Такой результат не имеет смысла, поэтому цену необходимо ограничить.

Итак, аптека должна назначить на помаду цену \$2,04. Такая цена приведет к продаже 49,4 штуки в неделю и еженедельной прибыли \$56,24.

Подход к ценообразованию, изложенный в этой главе, не требует знания концепции эластичности спроса. Инструмент **Поиск решения** по умолчанию рассматривает эластичность для каждой цены, когда определяет цену, максимизирующую прибыль. Такой подход легко может быть применен в организациях, продающих тысячи разных продуктов. Единственные данные, которые требуется указать для каждого продукта, — это себестоимость и три точки на кривой спроса.





**Рис. 89.3.** Диалоговое окно Параметры поиска решения для вычисления цены на помару

## Задания

1. Предположим, что производство одной игровой приставки стоит \$250. Рассматривается цена от \$200 до \$400. Предполагаемый спрос на игровые приставки представлен в таблице:

Цена	Спрос (в млн)
\$200	2
\$300	0,9
\$400	0,2

Какую цену следует назначить на игровую приставку?

2. Для выполнения этого задания возьмите данные о спросе из задания 1. Каждый владелец игровой приставки покупает в среднем 10 видеоигр. Прибыль от каждой видеоигры составляет \$10. Какую цену следует назначить на игровую приставку?

3. Необходимо определить правильную цену на новый еженедельный журнал. Переменные затраты на печать и распространение экземпляра журнала составляют \$0,50. Рассматривается цена от \$0,50 до \$1,30 за экземпляр. Расчетные еженедельные продажи журнала представлены в таблице:

Цена	Спрос (в млн)
\$0,50	2
\$0,90	1,2
\$1,30	0,3

В дополнение к выручке от продаж журнала можно получить по \$30 за каждую проданную 1000 экземпляров каждой из 20 страниц рекламы, размещаемой в еженедельнике. Какую цену следует назначить на журнал?

4. Аптеке требуется определить, какую цену установить на шоколадный батончик. Аптека платит за каждый батончик 60 центов и планирует продавать их по цене от \$1 до \$2 за штуку. Предположительный спрос, соответствующий трем уровням цены, приведен в таблице ниже. Какую цену на батончик следует установить?

Цена	Спрос
\$1,00	88
\$1,50	72
\$2,00	40

5. Магазин женской одежды закупает костюмы по \$200 за штуку. Спрос за неделю оценивается следующим образом:

Цена	Спрос
250	32
300	25
350	23

Предположим, что каждая женщина, приобретая костюм, дополнительно покупает аксессуаров на \$80, 50% от которых составляют прибыль магазина. Какую цену на костюм должен назначить магазин?

6. Подписчики рекомендательной службы получают доступ к рекомендациям о местных поставщиках услуг, таких как врачи, сантехники и т. д. Предполагается, что число новых подписчиков, подключающихся к службе ежемесячно, зависит от стоимости подписки следующим образом:

Цена	Спрос
\$10	20 000
\$20	15 000
\$30	6000

Предположим, что каждый подписчик остается таковым в течение года. Кроме платы за подписку, прибыль приносит щелчок подписчика по рекламному объявлению. Средний подписчик просматривает 10 рекламных объявлений в год. Какая цена подписки будет приносить максимальную прибыль?

7. Число струйных принтеров, продаваемых магазином электроники в неделю, оценивается следующей зависимостью от цены:

Цена	Спрос
\$70	128
\$100	90
\$130	20

Магазин платит за каждый струйный принтер \$50. В среднем каждый, кто покупает принтер, покупает 10 картриджей, на каждом из которых магазин зарабатывает \$8. Какая цена на принтер обеспечит максимальную общую прибыль?

## ГЛАВА 90

# Нелинейное ценообразование

### Обсуждаемые вопросы

- Что такое линейное ценообразование?
- Что такое нелинейное ценообразование?
- Что такое объединение в набор и как оно может повысить прибыльность?
- Как составить план нелинейного ценообразования с максимальной прибылью?

## Ответы на вопросы

### ❓ Что такое линейное ценообразование?

В главах 88 «Ценообразование продуктов с сопутствующими товарами» и 89 «Ценообразование продуктов с помощью субъективно определяемого спроса» показано, как определить цену продукта, максимизирующую прибыль. Однако в примерах к этим главам подразумевается, что независимо от количества единиц приобретаемого товара покупатель платит одну и ту же сумму за каждую единицу. Такая модель известна как *линейное ценообразование*, поскольку стоимость покупки  $x$  единиц товара является прямолинейной функцией  $x$ , а именно:

стоимость  $x$  единиц товара = (цена единицы товара)  $\times x$ .

В этой главе показано, что *нелинейное ценообразование* зачастую может значительно увеличить прибыль компании.

### ❓ Что такое нелинейное ценообразование?

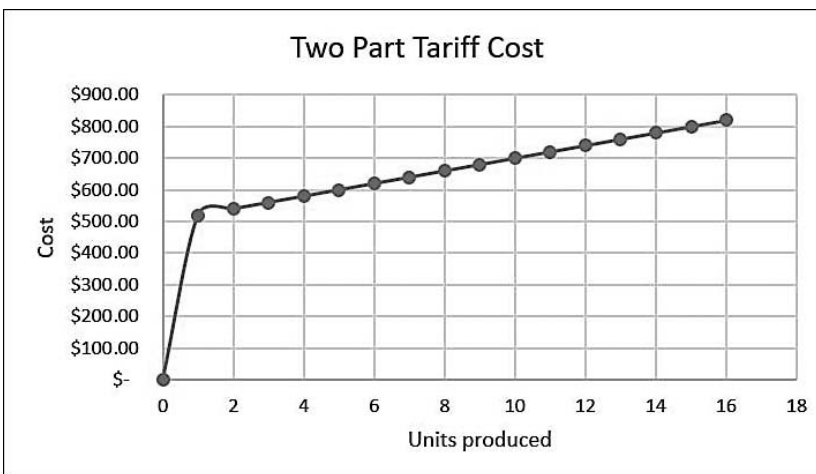
Нелинейная схема ценообразования просто означает, что стоимость покупки  $x$  единиц товара не является прямолинейной функцией  $x$ . С нелинейными стратегиями ценообразования сталкивался каждый. Вот два примера.

- **Скидка за количество.** Первые пять единиц товара могут стоить по \$20, а остальные единицы — по \$12. Скидки за количество, как правило, предоставляются компаниями, продающими компьютеры и компьютерный софт. Пример стоимости покупки  $x$  единиц товара показан на листе *nonlinear pricing examples* в файле *Nlp.xlsx* и на рис. 90.1. Обратите внимание, что график имеет наклон 20 для пяти и менее единиц товара и наклон 12 для большего количества приобретенных единиц товара.



**Рис. 90.1.** План стоимости со скидкой на количество

○ **Двухчастный тариф.** При вступлении в загородный клуб обычно платят фиксированный взнос, а затем взнос за каждую игру в гольф. Предположим, что загородный клуб взимает членский взнос в размере \$500 в год и \$20 долларов за каждую игру. Такой тип стратегии ценообразования называется *двухчастным тарифом*. Стоимость покупки заданного количества игр для такой политики ценообразования показана на рис. 90.2. Как и в предыдущем случае, см. лист *nonlinear pricing examples* в файле *Nlp.xlsx*. Обратите внимание, что график имеет наклон 520 от нуля до одной приобретенной единицы и наклон 20 для большего количества приобретенных единиц. Прямая линия должна всегда



**Рис. 90.2.** Стоимость при двухчастном тарифе

иметь одинаковый наклон, а двухчастный тариф, как видно из рисунка, является нелинейным.

### ❓ Что такое объединение в набор и как оно может повысить прибыльность?

При *объединении в набор* клиенту предлагается купить набор продуктов по цене меньшей, чем сумма цен входящих в него продуктов. Для анализа феномена объединения в набор необходимо понять, как рациональный потребитель принимает решения. Для каждой возможной комбинации продуктов рациональный потребитель определяет для себя ценность того, что продается, и вычитает из нее реальную стоимость покупки. Эта разница называется *потребительской выгодой*. Рациональный потребитель не купит ничего, если потребительская выгода для каждого доступного варианта является отрицательной. В противном случае потребитель покупает комбинацию продуктов, имеющую самую большую потребительскую выгоду.

Так как же объединение в набор может повысить прибыльность? Предположим, что вы продаете компьютеры и принтеры двум клиентам. Ценность, которую каждый клиент придает компьютеру и принтеру, представлена в таблице:

Клиент	Ценность компьютера	Ценность принтера
1	\$1000	\$500
2	\$500	\$1000

Если компьютер и принтер предлагаются к продаже только по отдельности, то, назначив цену \$1000 на принтер и на компьютер, вы продадите один принтер и один компьютер и получите \$2000 выручки. Теперь предположим, что вы предлагаете такой набор: компьютер и принтер за \$1500. Каждый клиент купит и компьютер, и принтер, и вы получите \$3000 выручки. При объединении в набор компьютера и принтера можно извлечь больше выгоды из общей потребительской ценности. Объединение в набор действует эффективнее, если корреляция для потребительской ценности продуктов, включенных в набор, является отрицательной. В приведенном примере корреляция между ценностью продуктов отрицательная, поскольку клиент, который высоко оценивает принтер, дает низкую оценку компьютеру, а клиент, для которого принтер малоценен, высоко оценивает компьютер.

Когда вы идете в парк развлечений, например в Диснейленд, вы не покупаете билет на каждый аттракцион. Вместо этого вы покупаете билет на вход в парк. Это пример чистого объединения в набор, поскольку потребитель не имеет возможности заплатить за подмножество предлагаемых продуктов. Такой подход уменьшает очереди (представьте очередь на каждый аттракцион) и приводит к дополнительной прибыли.

Чтобы понять причину увеличения прибыльности от объединения в набор, предположим, что имеется только один клиент, и количество аттракционов, которые клиент собирается посетить, определяется следующей кривой спроса: (количество аттракционов) =  $20 - 2 \times (\text{цена аттракциона})$ . Из главы 87 «Оценка кривой спроса»

вы знаете, что ценность, которую клиент придает  $i$ -му аттракциону, является ценой, при которой спрос равен  $i - 0,5$ . Таким образом,  $i - 0,5 = 20 - 2 \times$  (ценность аттракциона  $i$ ) или (ценность аттракциона  $i$ ) =  $10,25 - (i/2)$ . Первый аттракцион стоит \$9,75, второй аттракцион — \$9,25 и т. д. до двадцатого аттракциона, который стоит \$0,25.

Предположим, что за аттракцион назначена постоянная цена и переменные затраты составляют \$2 за аттракцион. Требуется найти схему линейного ценообразования с максимальной прибылью. Вычисление максимизирующей прибыль цены за аттракцион показано на листе OnePrice в файле Nlp.xlsx и на рис. 90.3.

	B	C	D
6	Linear		
7	Pricing		
8		price	\$ 6.00
9		Demand	\$ 8.00
10		unit cost	\$ 2.00
11			
12		profit	\$ 32.00

**Рис. 90.3.** Схема линейного ценообразования с максимальной прибылью

Я связал имена в ячейках C8:C10 с ячейками D8:D10. Затем я ввел пробное значение цены в ячейку D8 и вычислил количество приобретенных билетов на аттракционы в ячейке D9 по формуле  $=20-(2*D8)$ . Прибыль вычислена в ячейке D12 по формуле  $=\text{спрос}*(\text{цена}-\text{себестоимость})$ . Затем я применил Поиск решения (Solver) и нашел максимум значения в ячейке D12 (прибыль), изменив ячейку D8 (цена). Цена в \$6 приведет к покупке восьми билетов на аттракционы, что позволит получить максимальную прибыль в \$32.

Теперь представим, что наш парк — Диснейленд, который предлагает только набор из 20 билетов на аттракционы. Установим цену, равную сумме потребительской ценности для каждого аттракциона:  $(9,75 + 9,25 + \dots + 0,75 + 0,25 = \$100,00)$ . Потребительская ценность всех 20 аттракционов составляет \$100,00, поэтому клиент купит входной билет в парк за \$100,00. Полученная прибыль составит  $100,00 - 2,00 \times (20) = \$60,00$ , что почти в два раза превышает прибыль при линейном ценообразовании.



### Как составить план нелинейного ценообразования с максимальной прибылью?

В этом разделе я покажу, как определить план ценообразования для двухчастного тарифа с максимальной прибылью на примере парка развлечений. Выполните следующие шаги:

- Предложите пробные значения для фиксированного взноса и цены за билет на аттракцион.
- Определите ценность каждого аттракциона для клиента: ценность аттракциона  $i = 10 - 0,5(i - 0,5) = 10,25 - 0,5i$ .

- Определите общую ценность покупки  $i$  билетов на аттракционы.
- Определите цену для  $i$  билетов на аттракционы: (фиксированный взнос) +  $i \times$  (цена за билет на аттракцион).
- Определите потребительскую выгоду для покупки  $i$  билетов на аттракционы: (ценность  $i$  аттракционов) – (цена  $i$  билетов на аттракционы).
- Определите максимальную потребительскую выгоду.
- Определите количество приобретенных единиц. Если максимальная потребительская выгода отрицательная, это количество равно нулю. В противном случае используйте функцию ПОИСКПОЗ (MATCH) и найдите количество единиц, приносящее максимальную выгоду.
- С помощью функции ВПР (VLOOKUP) найдите выручку, соответствующую количеству приобретенных единиц.
- Вычислите прибыль как (выручка – затраты).
- Используйте двунаправленную таблицу данных для определения фиксированного взноса и цены для билета на аттракцион, которые приносят максимальную прибыль.

Рассмотрим лист **two-part tariff** в файле **Nlp.xlsx** и рис. 90.4.

Сначала я назначил ячейке **F2** имя **фикс\_взнос**, а ячейке **F3** — имя **цена\_аттр**. В ячейках **F2** и **F3** я ввел пробные значения для фиксированного взноса и цены за билет на аттракцион. Затем я определил ценность, которую клиент придает каждому аттракциону, скопировав формулу  $=10,25-(D6/2)$  из ячейки **E6** в **E7:E25**. Оказалось, что клиент оценивает первый аттракцион в \$9,75, второй аттракцион — в \$9,25 и т.д.

Для вычисления общей ценности первых  $i$  аттракционов я скопировал формулу  $=СУММ(\$E\$6:E6)$  из **F6** в **F7:F25**. Эта формула суммирует все значения в столбце **E** из текущей строки и над ней. Скопировав формулу  $=\text{фикс\_взнос}+\text{цена\_аттр}*D6$  из **G6** в **G7:G25**, я вычислил цену за  $i$  аттракционов. Например, цена пяти билетов на аттракционы — \$68,50.

Напомню, что потребительская выгода для  $i$  аттракционов равна (ценность  $i$  аттракционов) – (стоимость  $i$  билетов на аттракционы). Скопировав формулу  $=F6-G6$  из ячейки **H6** в ячейки **H7:H25**, я вычислил потребительскую выгоду от покупки билетов на любое количество аттракционов. Например, потребительская выгода от билетов на пять аттракционов равна  $-\$24,75$ , что является следствием большого фиксированного взноса.

В ячейке **H4** я вычислил максимальную потребительскую выгоду по формуле  $=\text{МАКС}(H6:H25)$ . Помните, что если максимальная потребительская выгода отрицательная, у вас ничего не купят. В противном случае клиент приобретет то количество единиц предлагаемой продукции, которое дает максимальную потребительскую прибыль. Следовательно, я ввел в ячейку **I1** формулу  $=\text{ЕСЛИ}(H4>=0;\text{ПОИСКПОЗ}(H4;H6:H25;0);0)$  и узнал количество приобретенных билетов на аттракционы (в данном случае, 15). Обратите внимание, что функция **ПОИСКПОЗ (MATCH)** воз-



	C	D	E	F	G	H	I	J
1						Units bought	15	
2	cost		fixed fee	\$ 56.00		Revenue	\$ 93.50	
3	\$ 2.00		price per ride	\$ 2.50		Prod Cost	\$ 30.00	
4					Max surplus	0.25		
5		Unit	Value	Cum Value	Price paid	Surplus		Profit
6		1	9.75	9.75	58.5	-48.75		\$ 63.50
7		2	9.25	19	61	-42		
8		3	8.75	27.75	63.5	-35.75		
9		4	8.25	36	66	-30		
10		5	7.75	43.75	68.5	-24.75		
11		6	7.25	51	71	-20		
12		7	6.75	57.75	73.5	-15.75		
13		8	6.25	64	76	-12		
14		9	5.75	69.75	78.5	-8.75		
15		10	5.25	75	81	-6		
16		11	4.75	79.75	83.5	-3.75		
17		12	4.25	84	86	-2		
18		13	3.75	87.75	88.5	-0.75		
19		14	3.25	91	91	0		
20		15	2.75	93.75	93.5	0.25		
21		16	2.25	96	96	0		
22		17	1.75	97.75	98.5	-0.75		
23		18	1.25	99	101	-2		
24		19	0.75	99.75	103.5	-3.75		
25		20	0.25	100	106	-6		

**Рис. 90.4.** Определение оптимального двухчастного тарифа

вращает число строк, на которое необходимо сдвинуться вниз в диапазоне H6:H24, чтобы найти первое совпадение со значением максимальной выгоды.

Теперь я присвою диапазону D5:G25 имя **lookup**. Тогда в четвертом столбце этого диапазона я смогу найти общую выручку на основе количества приобретенных билетов на аттракционы (которое уже вычислено в ячейке I1). Общая выручка вычисляется в ячейке I2 по формуле **=ЕСЛИ(I1=0;0;ВПР(I1;lookup;4))**. Обратите внимание, что если билеты не были приобретены, то выручка отсутствует. В ячейке I3 по формуле **=I1\*C3** я вычислил общую себестоимость катания на аттракционах для приобретенного количества билетов. В ячейке J6 по формуле **=I2-I3** я вычислил прибыль как выручку минус расходы.

Теперь я могу использовать двунаправленную таблицу данных, чтобы определить комбинацию фиксированного взноса и цены билета на аттракцион, максимизирующую прибыль. Таблица данных показана на рис. 90.5. (Многие строки и столбцы скрыты.) При настройке таблицы данных я изменял фиксированный взнос от \$10,00 до \$60,00 (ячейки K10:K60), а цену билета — от \$0,50 до \$5,00 (ячейки L9:BE9). Прибыль пересчитана в ячейке K9 по формуле **=J6**.

Я выделил табличный диапазон (ячейки K9:BE60) и затем на вкладке **Данные (Data)** в группе **Прогноз (Forecast)** в списке **Анализ "что, если" (What If Analysis)** выбрал **Таблица данных (Data Table)**. В поле **Подставлять значения по строкам в (Column Input Cell)** я указал ячейку F2 (фиксированный взнос), а в поле **Подставлять значения по столбцам в (Row Input Cell)** — ячейку F3 (цена билета на аттракцион). После нажа-

тия ОК в диалоговом окне Таблица данных (Table) вычислится прибыль по каждой комбинации фиксированного взноса и цены билета на аттракцион, представленной в таблице.

	J	K	L	M	N	AE	AF	BB	BC	BD	BE
5	Profit				Max Profit						
6	\$ 63.50				\$ 63.50						
7											
8			Unit	cost							
9		\$ 63.50	0.5	0.6	0.7	2.4	2.5	4.7	4.8	4.9	5
10		10	-18.5	-16.6	-14.7	16	17.5	39.7	38	39	40
11		11	-17.5	-15.6	-13.7	17	18.5	40.7	39	40	41
12		12	-16.5	-14.6	-12.7	18	19.5	41.7	40	41	42
13		13	-15.5	-13.6	-11.7	19	20.5	42.7	41	42	43
14		14	-14.5	-12.6	-10.7	20	21.5	43.7	42	43	44
15		15	-13.5	-11.6	-9.7	21	22.5	44.7	43	44	45
16		16	-12.5	-10.6	-8.7	22	23.5	45.7	44	45	46
17		17	-11.5	-9.6	-7.7	23	24.5	46.7	45	46	47
18		18	-10.5	-8.6	-6.7	24	25.5	47.7	46	47	48
19		19	-9.5	-7.6	-5.7	25	26.5	48.7	47	48	49
20		20	-8.5	-6.6	-4.7	26	27.5	49.7	48	49	50
21		21	-7.5	-5.6	-3.7	27	28.5	50.7	49	50	51
22		22	-6.5	-4.6	-2.7	28	29.5	51.7	50	51	52
23		23	-5.5	-3.6	-1.7	29	30.5	52.7	51	52	53
24		24	-4.5	-2.6	-0.7	30	31.5	53.7	52	53	54
25		25	-3.5	-1.6	0.3	31	32.5	54.7	53	54	55
49	Fixed	49	20.5	22.4	24.3	55	56.5	0	0	0	0
50	cost	50	21.5	23.4	25.3	56	57.5	0	0	0	0
51		51	22.5	24.4	26.3	57	58.5	0	0	0	0
52		52	23.5	25.4	27.3	58	59.5	0	0	0	0
53		53	24.5	26.4	28.3	59	60.5	0	0	0	0
54		54	25.5	27.4	29.3	60	61.5	0	0	0	0
55		55	26.5	28.4	30.3	61	62.5	0	0	0	0
56		56	27.5	29.4	31.3	62	63.5	0	0	0	0
57		57	28.5	30.4	32.3	63	0	0	0	0	0
58		58	29.5	31.4	33.3	0	0	0	0	0	0
59		59	30.5	32.4	34.3	0	0	0	0	0	0
60		60	31.5	33.4	35.3	0	0	0	0	0	0

**Рис. 90.5.** Двухнаправленная таблица для вычисления оптимального двухчастного тарифа

Для выделения цветом двухчастного тарифа, максимизирующего прибыль, я применил Условное форматирование, выделив диапазон L10:BE60. На вкладке Главная (Home) в группе Стили (Styles) я щелкнул на Условное форматирование (Conditional Formatting). Далее выбрал Правила отбора первых и последних значений (Top/Bottom Rules) и Первые 10 элементов (Top 10 Items). Для форматирования только максимальной прибыли замените в диалоговом окне 10 на 1. Здесь фиксированный взнос \$56,00 и цена билета на аттракцион \$2,50 дают прибыль \$63,50, что почти в два раза превышает прибыль при линейном ценообразовании. Фиксированный взнос \$59,00 и цена билета \$2,30 также дают прибыль \$63,50.

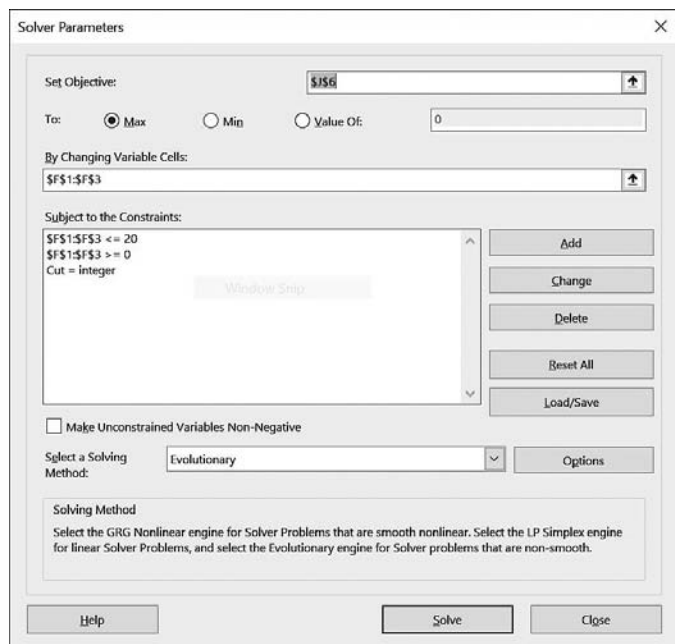
Поскольку план ценообразования со скидкой на количество включает выбор трех переменных (отсечка, высокая и низкая цена), невозможно определить план со скидкой на количество, максимизирующий прибыль, с помощью таблицы данных. Можно подумать, что для определения стратегии ценообразования со скидкой на количество, максимизирующей прибыль, следует взять модель поиска решения (с изменяющимися ячейками, в которых находятся значения отсечки, высокой и низкой цены). До Excel 2010 Поиск решения (Solver) не всегда верно находил корректные решения, если значение в целевой ячейке вычислялось по формулам,

содержащим функции ЕСЛИ (IF). В Microsoft Excel 2019 эта проблема полностью устранена. В файле Qd.xlsx (рис. 90.6) я использовал эволюционный метод поиска решения, чтобы найти план со скидкой на количество, максимизирующий прибыль. Предположим, что все единицы продукции, приобретенные до отсечки (отсечка), проданы по высокой цене (высок\_цена), а остальные единицы — по низкой цене (низк\_цена). Единственное изменение внесено в столбец G. После присвоения ячейкам F1:F3 имен в ячейках E1:E3 я вычислил сумму, которую клиент уплатит за любое количество единиц продукции, скопировав формулу =ЕСЛИ(D6<=отсечка; D6\*высок\_цена; высок\_цена\*отсечка+(D6-отсечка)\*низк\_цена) из G6 в G7:G25. Диалоговое окно Параметры поиска решения (Solver Parameters) с настройками показано на рис. 90.7. Как я говорил в главе 36 «Расположение складов по методу ОПГ с несколькими начальными точками и согласно эволюционному поиску решения», я использую здесь эволюционный метод поиска решения, поскольку в модели присутствуют функции ЕСЛИ, включающие изменяемые ячейки. Параметр Скорость изменения (Mutation Rate) эволюционного метода решения также был изменен на 0,5. Обратите внимание на ограничение, в котором указано, что отсечка должна быть целым числом. Для всех изменяемых ячеек я указал верхнюю границу — \$20. Разумеется, если при поиске решения цена приблизится к \$20, следует сделать ограничение для изменяемых ячеек менее строгим.

Поиск решения дает максимальную прибыль \$63,97, полученную при цене \$18,48 на первые четыре приобретенные единицы продукции и цене \$1,84 на остальные единицы продукции. Клиенты купят 16 единиц продукции. Если бы поиск решения выполнялся дольше, был бы найден истинный максимум прибыли, составляющий \$64,00.

	C	D	E	F	G	H	I	J
1			Cut	4		Units bought	16	
2	cost		HP	\$ 18.48		Revenue	\$ 95.97	
3	\$ 2.00		LP	\$ 1.84		Prod Cost	\$ 32.00	
4					Max surplus	0.03156491		
5		Unit	Value	Cum Value	Price paid	Surplus		Profit
6		1	9.75	9.75	18.48110027	-8.731100274		\$ 63.97
7		2	9.25	19	36.96220055	-17.96220055		
8		3	8.75	27.75	55.44330082	-27.69330082		
9		4	8.25	36	73.9244011	-37.9244011		
10		5	7.75	43.75	75.76140393	-32.01140393		
11		6	7.25	51	77.59840676	-26.59840676		
12		7	6.75	57.75	79.4354096	-21.6854096		
13		8	6.25	64	81.27241243	-17.27241243		
14		9	5.75	69.75	83.10941526	-13.35941526		
15		10	5.25	75	84.94641809	-9.946418094		
16		11	4.75	79.75	86.78342093	-7.033420926		
17		12	4.25	84	88.62042376	-4.620423759		
18		13	3.75	87.75	90.45742659	-2.707426592		
19		14	3.25	91	92.29442942	-1.294429425		
20		15	2.75	93.75	94.13143226	-0.381432257		
21		16	2.25	96	95.96843509	0.03156491		
22		17	1.75	97.75	97.80543792	-0.055437923		
23		18	1.25	99	99.64244076	-0.642440755		
24		19	0.75	99.75	101.4794436	-1.729443588		
25		20	0.25	100	103.3164464	-3.316446421		

**Рис. 90.6.** Поиск плана со скидкой на количество, максимизирующего прибыль



**Рис. 90.7.** Настройки Поиска решения для плана со скидкой на количество, максимизирующего прибыль

## Задания

Вы — владелец небольшого загородного клуба, у которого три типа клиентов. Их оценки для каждой игры в гольф в течение месяца приведены в таблице:

Номер игры	Клиенты типа 1	Клиенты типа 2	Клиенты типа 3
1	\$60	\$50	\$40
2	\$50	\$45	\$30
3	\$40	\$30	\$20
4	\$30	\$15	\$10
5	\$20	\$0	\$0
6	\$10	\$0	\$0

1. Найдите двухчастный тариф, максимизирующий прибыль.
2. Предположим, вы хотите предложить чистое объединение в набор. Например, член клуба может сыграть до пяти игр в гольф за \$60 в месяц. Член клуба не имеет возможности выбрать какой-либо другой вариант, кроме этого. Какой чистый набор максимизирует прибыль?

## ГЛАВА 91

# Формулы массива и функции, возвращающие массив

### Обсуждаемые вопросы

- Что такое формула массива?
- Как интерпретировать формулы, такие как  $=(D2:D7)*(E2:E7)$  и  $=СУММ(D2:D7*E2:E7)$ ?
- Имеется список имен в столбце. Имена в списке часто меняются. Существует ли простой способ перенести имена в одну строку так, чтобы изменения в исходном столбце отражались бы в этой новой строке?
- Имеется список ежемесячных доходов по акциям. Существует ли способ определить количество значений доходов, попадающих в интервалы от  $-30$  до  $-20\%$ , от  $-10$  до  $0\%$  и т. д., которые обновлялись бы автоматически при изменении исходных данных?
- Как написать формулу, суммирующую вторые цифры из списка целых чисел?
- Существует ли способ просмотра двух списков имен и поиска имен, встречающихся в обоих списках?
- Как написать формулу, вычисляющую среднее значение для чисел в списке, которые больше или равны медиане списка?
- У небольшой компании декоративной косметики имеется база данных по продажам, в которой указан продавец, количество проданных единиц продукции и сумма в долларах для каждой сделки. Мне известно, что для обобщения этих данных подойдут статистические функции или функции СЧЁТЕСЛИМН (COUNTIFS), СУММЕСЛИМН (SUMIFS) и СРЗНАЧЕСЛИМН (AVERAGEIFS). Но могу ли я использовать также функции, возвращающие массив, чтобы обобщить эти данные и ответить на вопросы типа: сколько единиц косметики продал продавец, или сколько помады было продано, или сколько единиц косметики было продано определенным продавцом?
- Что такое массив констант и для чего он нужен?
- Как изменить формулу массива?
- Как оценить тренд и сезонность выручки для магазина игрушек по заданной квартальной выручке?
- Как вычислить медианный размер сделки в каждой стране по заданному списку сделок в различных странах?

- Имеются данные о торговых сделках с разбивкой по продавцам, продуктам и местоположению. Как определить стандартное отклонение проданных единиц продукции для каждого сочетания «продавец — продукт — регион»?
- Как применить функцию СУММПРОИЗВ (SUMPRODUCT) взамен условного счета, условного суммирования и условного определения среднего наравне с формулами массивов?

## Ответы на вопросы

### ? Что такое формула массива?

В Microsoft Excel *формулы массива* часто предоставляют наименее затратный или наиболее эффективный подход к выполнению сложных вычислений. Формула массива возвращает результат либо в одну ячейку, либо в диапазон ячеек. Формулы массива выполняют операции на двух или более наборах значений, называемых *аргументами-массивами*. Все аргументы-массивы, указанные в формуле, должны содержать одинаковое число строк и столбцов.

При вводе формулы массива сначала вы должны выделить диапазон, в который Excel должен поместить результаты вычисления формулы. Затем после ввода формулы в первую ячейку выделенного диапазона необходимо нажать **Ctrl+Shift+Enter**. Если этого не сделать, то полученные результаты будут неправильными или бессмысленными. Процесс ввода формулы массива и последующее нажатие **Ctrl+Shift+Enter** можно назвать *вводом массива* в формулу.

Excel также содержит целый ряд функций, возвращающих массив. В главе 44 «Обобщение данных с помощью описательной статистики» мы рассмотрели функцию МОДА.НСК (MODE.MULT), возвращающую массив. Еще две такие функции, ЛИНЕЙН (LINEST) и ТЕНДЕНЦИЯ (TREND), встречались нам в главах 59 «Введение во множественную регрессию» и 60 «Включение качественных факторов во множественную регрессию».

Для использования как формулы массива, так и функции, возвращающей массив, сначала необходимо выделить диапазон, в который должны быть записаны результаты. Затем после ввода функции или формулы в первую ячейку выделенного диапазона необходимо нажать **Ctrl+Shift+Enter**. В этой главе я расскажу еще о трех полезных функциях, возвращающих массив: ТРАНСП (TRANSPOSE), ЧАСТОТА (FREQUENCY) и ЛГРФПРИБЛ (LOGEST).

Как вы далее убедитесь, удалить какую-либо часть диапазона ячеек, содержащего результаты, вычисленные по формуле массива, невозможно. Кроме того, формулу массива нельзя скопировать в диапазон, содержащий и пустые ячейки, и формулы массива. Например, если в ячейке C10 находится формула массива, то нельзя скопировать ее в диапазон C10:J15, поскольку диапазон содержит пустые ячейки и формулу массива в ячейке C10. Чтобы обойти эту проблему, скопируйте формулу из C10 в D10:J10 в строке 10, а затем скопируйте содержимое C10:J10 в C11:J15. (Тем самым вы сначала создаете верхний столбец данных.)

Наилучший способ научиться работе с функциями и формулами массива — рассмотреть конкретные примеры, так что приступим.

**❓ Как интерпретировать формулы, такие как  $\text{=(D2:D7)*(E2:E7)}$  и  $\text{=СУММ(D2:D7*E2:E7)}$ ?**

На листе **Total Wages** в файле **Arrays.xlsx** указано количество отработанных часов и почасовая ставка для шести работников (рис. 91.1).

	C	D	E	F	G	H
				total owed each person	grand total owed	
1		Hours	wage rate			
2	John	3	\$ 6.00	\$ 18.00	\$ 295.00	$\text{=(D2:D7*E2:E7)}$
3	Jack	4	\$ 7.00	\$ 28.00		$\text{=(D2:D7*E2:E7)}$
4	Jill	5	\$ 8.00	\$ 40.00		$\text{=(D2:D7*E2:E7)}$
5	Jane	8	\$ 9.00	\$ 72.00		$\text{=(D2:D7*E2:E7)}$
6	Jean	6	\$ 10.00	\$ 60.00		$\text{=(D2:D7*E2:E7)}$
7	Jocelyn	7	\$ 11.00	\$ 77.00		$\text{=(D2:D7*E2:E7)}$

**Рис. 91.1.** Вычисление заработной платы по формуле массива

Допустим, вам нужно вычислить общую сумму заработной платы для каждого работника. Для этого достаточно скопировать формулу  $\text{=D3*E3}$  из F2 в F3:F7. В таком подходе нет никакой ошибки, но формула массива предоставляет более элегантное решение. Начните с выделения диапазона F2:F7, в котором будет рассчитана общая сумма зарплаты для каждого работника. Затем введите формулу  $\text{=(D2:D7*E2:E7)}$  и нажмите **Ctrl+Shift+Enter**. Общая заработная плата каждого работника будет вычислена корректно. Кроме того, в строке формул появится формула  $\text{={D2:D7*E2:E7}}$ . Фигурные скобки — способ, которым Excel дает вам понять, что вы создали формулу массива. (Вводить фигурные скобки в начале и конце формулы не нужно, их Excel вставляет автоматически. Тем не менее, чтобы обозначить, что это формула массива и ее ввод завершается нажатием **Ctrl+Shift+Enter**, в тексте я заключил ее в фигурные скобки.)

Чтобы увидеть, как работает эта формула, щелкните в строке формул, выделите D2:D7 в формуле и нажмите F9. В строке появится выражение  $\{3;4;5;8;6;7\}$ , так Excel представляет диапазон ячеек D2:D7 в виде массива. Теперь выделите в строке формул E2:E7 и снова нажмите F9. В строке появится выражение  $\{6;7;8;9;10;11\}$  — массив, соответствующий диапазону E2:E7. Звездочка (\*) указывает на перемножение соответствующих элементов в каждом массиве. Поскольку перемножаемые диапазоны включают по шесть ячеек каждый, Excel создает массивы из шести элементов, и поскольку был выделен диапазон из шести ячеек, зарплата каждого работника отображается в своей собственной ячейке. Если бы был выделен диапазон только из пяти ячеек, шестой элемент массива не отобразился бы.



Предположим, что вы хотите вычислить итоговую сумму зарплат всех работников. Это можно сделать, например, с помощью формулы `=СУММПРОИЗВ(D2:D7;E2:E7)`. Но, как и в предыдущем случае, для вычисления итоговой суммы попробуем создать формулу массива. Начните с выделения одной ячейки (я выбрал G2), в которую будет записан результат. Затем в ячейку G2 введите формулу `{=СУММ(D2:D7*E2:E7)}`. После нажатия **Ctrl+Shift+Enter** появится результат  $(3)(6) + (4)(7) + (5)(8) + (8)(9) + (6)(10) + (7)(11) = 295$ . Чтобы увидеть, как работает формула, выделите часть формулы D2:D7\*E2:E7 в строке формул и нажмите F9. В строке появится выражение `СУММ({18;28;40;83;60;77})`, которое показывает, что Excel создал массив из шести элементов, с первым элементом  $3*6(18)$ , вторым элементом  $4*7(28)$  и т. д. до последнего элемента  $7*11(77)$ . Затем значения элементов массива суммируются, и в результате получается \$295,00.

**?** Имеется список имен в столбце. Имена в списке часто меняются. Существует ли простой способ перенести имена в одну строку так, чтобы изменения в исходном столбце отражались бы в этой новой строке?

В файле `Arrays.xlsx` на листе `Transpose` (рис. 91.2) содержится список имен в ячейках A4:A8. Задача состоит в переносе этих имен в одну строку (диапазон ячеек C3:G3). Если известно, что исходный список имен никогда не будет изменяться, можно для решения этой задачи скопировать диапазон ячеек и установить флажок **транспонировать (Transpose)** в диалоговом окне **Специальная вставка (Paste Special)** (см. также главу 14 «Команда Специальная вставка»). К сожалению, при использовании специальной вставки, если имена в столбце A изменятся, в строке 3 эти изменения отражаться не будут. В такой ситуации вам понадобится функция **ТРАНСП (TRANSPOSE)**.

	A	B	C	D	E	F	G
3			Julie	Jason	Jack	Jill	Jane
4	Julie		<code>{=TRANSPOSE(A4:A8)}</code>	<code>{=TRANSPOSE(A4:A8)}</code>	<code>{=TRANSPOSE(A4:A8)}</code>	<code>{=TRANSPOSE(A4:A8)}</code>	<code>{=TRANSPOSE(A4:A8)}</code>
5	Jason						
6	Jack						
7	Jill						
8	Jane						

**Рис. 91.2.** Применение функции ТРАНСП

Функция **ТРАНСП** — это функция, возвращающая массив, которая преобразует строки выбранного диапазона в столбцы и наоборот. В этом примере перед использованием функции **ТРАНСП** выделите диапазон C3:G3, в который должен быть помещен транспонированный список имен. Затем в ячейку C3 введите формулу массива `{=ТРАНСП(A4:A8)}`. Теперь список имен находится в одной строке. Важно, что теперь при изменении каких-либо имен в диапазоне A4:A8 соответствующие имена в транспонированном диапазоне также изменятся.

**?** Имеется список ежемесячных доходов по акциям. Существует ли способ определить количество значений доходов, попадающих в интервалы от -30



до  $-20\%$ , от  $-10$  до  $0\%$  и т. д., которые обновлялись бы автоматически при изменении исходных данных?

Для решения этой задачи подойдет функция **ЧАСТОТА** (**FREQUENCY**). Эта функция вычисляет, сколько значений в массиве (называемом массивом данных) попадает в заданные интервалы значений (указанные в массиве интервалов). Функция **ЧАСТОТА** имеет синтаксис: **ЧАСТОТА(массив\_данных;массив\_интервалов)**.

Работа функции **ЧАСТОТА** проиллюстрирована в файле **Arrays.xlsx** на листе **Frequency** (рис. 91.3). На нем в диапазоне ячеек **A4:A77** приведены ежемесячные доходы по фиктивным акциям.

В ячейках **A1:B1** и **A2:B2** (с помощью функций **МИН** и **МАКС**) я вычислил, что все доходы находятся в интервале от  $-43\%$  до  $53\%$ . На основе этой информации я задал массив интервалов в ячейках **C7:C17**, начиная с  $-0,4$  и заканчивая  $0,6$ . Затем я выделил диапазон **D7:D18**, так как хочу, чтобы результаты работы функции **ЧАСТОТА** были помещены в него. В этом диапазоне в ячейке **D7** будет подсчитано число точек данных не больше  $-0,4$ . В ячейке **D8** будет подсчитано число точек данных больше  $-0,4$  и не больше  $-0,3$  и т. д. В ячейке **D17** будет подсчитано число всех точек данных больше  $0,5$  и не больше  $0,6$  и, наконец, в ячейке **D18** — число всех точек данных больше  $0,6$ .

	A	B	C	D	E
1	min	-43.84%			
2	max	52.56%			
3	Returns				
4	43.81%			Total	
5	-8.30%			74	
6	-25.12%		Bin values		
7	-43.84%		-0.4	1 {=FREQUENCY(A4:A77,C7:C17)}	
8	-8.64%		-0.3	1 {=FREQUENCY(A4:A77,C7:C17)}	
9	49.98%		-0.2	2 {=FREQUENCY(A4:A77,C7:C17)}	
10	-1.19%		-0.1	5 {=FREQUENCY(A4:A77,C7:C17)}	
11	46.74%		0	13 {=FREQUENCY(A4:A77,C7:C17)}	
12	31.94%		0.1	11 {=FREQUENCY(A4:A77,C7:C17)}	
13	-35.34%		0.2	13 {=FREQUENCY(A4:A77,C7:C17)}	
14	29.28%		0.3	12 {=FREQUENCY(A4:A77,C7:C17)}	
15	-1.10%		0.4	11 {=FREQUENCY(A4:A77,C7:C17)}	
16	-10.67%		0.5	4 {=FREQUENCY(A4:A77,C7:C17)}	
17	-12.77%		0.6	1 {=FREQUENCY(A4:A77,C7:C17)}	
18	19.17%			0 {=FREQUENCY(A4:A77,C7:C17)}	
19	25.06%				
20	19.03%				

**Рис. 91.3.** Применение функции **ЧАСТОТА**

В ячейку **D7** я ввел формулу **=ЧАСТОТА(A4:A77;C7:C17)** и нажал **Ctrl+Shift+Enter**. Эта формула говорит Excel посчитать число точек данных из **A4:A77** (массива данных), лежащих в каждом из интервалов, определенных в **C7:C17**. Результаты показыва-

ют, что есть одна точка данных больше, чем  $-0,4$ , и меньше либо равна  $-0,3$ . Тринадцать точек данных больше, чем  $0,1$ , и меньше либо равны  $0,2$ . Если в массиве данных изменить какие-либо точки данных, результаты, возвращенные функцией **ЧАСТОТА** в ячейках D7:D17, отразят эти изменения. Если отобразить исходные данные для функции **ЧАСТОТА** в виде таблицы, функция будет учитывать все данные, которые будут добавляться к таблице.

### ❓ Как написать формулу, суммирующую вторые цифры из списка целых чисел?

В файле **Arrays.xlsx** на листе **Sum Up 2nd Digit** в ячейках A4:A10 находится список из семи целых чисел (рис. 91.4). Требуется просуммировать вторые цифры каждого числа с помощью одной формулы. Эту сумму можно было бы получить, скопировав формулу **=ЗНАЧЕН(ПСТР(A4;2;1))** из B4 в B5:B10. По этой формуле вычисляется (как числовое значение) второй символ в ячейке A4. Затем можно сложить значения в диапазоне B4:B10 и получить в общей сложности 27.

	A	B	C	D
1		Total		
2		27		
3				
4	140	4		
5	85	5		
6	76	6		
7	1610	6	27	Error trapped
8	302	0	{=SUM(VALUE(MID(A4:A10;2;1)))}	27
9	434	3		{=SUM(IF(LEN(A4:A11)>=2,VALUE(MID(A4:A11;2;1)),0))}
10	13	3		
11	1	#VALUE!		
12				

**Рис. 91.4.** Суммирование вторых цифр в наборе целых чисел

Функция, возвращающая массив, намного упрощает этот процесс. Просто выделите ячейку C7 и введите формулу массива **{=СУММ(ЗНАЧЕН(ПСТР(A4:A10;2;1)))}**. По этой формуле будет вычислен правильный результат — 27.

Для просмотра работы функции выделите часть формулы **ПСТР(A4:A10;2;1)** в строке формул и нажмите F9. Появится выражение **{"4";"5";"6";"6";"0";"3";"3"}**. Эта строка значений показывает, что Excel создал в ячейках A4:A10 массив, состоящий из вторых цифр (рассматриваемых как текст). Часть формулы с функцией **ЗНАЧЕН** (**VALUE**) преобразует эти текстовые строки в числовые значения, которые суммируются частью формулы с функцией **СУММ**.

Обратите внимание, что в ячейку A11 введено число из одной цифры (1). Поскольку у этого числа нет второй цифры, часть формулы с функцией **ПСТР** (**MID**) возвращает сообщение об ошибке **#ЗНАЧ!**. Каким образом следует изменить эту формулу массива для учета целых чисел, состоящих из одной цифры? Просто введите в ячейку E8 формулу массива **{СУММ(ЕСЛИ(ДЛСТР(A4:A11)>=2;**

ЗНАЧЕН(ПСТР(A4:A11;2;1));0)). По этой формуле любое целое число из одной цифры заменяется значением 0, и в результате получается правильная сумма.

### ❓ Существует ли способ просмотра двух списков имен и поиска имен, встречающихся в обоих списках?

В файле Arrays.xlsx на листе Matching Names хранятся два списка имен (в столбцах D и E), как показано на рис. 91.5. Требуется определить, какие имена из списка 1 встречаются также и в списке 2. Для решения задачи выделите диапазон C5:C28 и введите формулу массива `{=ПОИСКПОЗ(D5:D28;E5:E28;0)}` в ячейку C5. Эта формула относится ко всему диапазону C5:C28. В ячейке C5 формула проверяет, имеет ли имя в ячейке D5 совпадения в столбце E. Если таковое существует, возвращаются позиция первого совпадения в диапазоне E5:E28. Если такового не существует, возвращается сообщение об ошибке #Н/Д (недоступно). Аналогично в ячейке C6 формула проверяет, имеет ли совпадение второе имя из списка 1. Как видно из рисунка, например, имя Artest не встречается в списке 2, а имя Harrington встречается (первое совпадение находится во второй ячейке диапазона E5:E28).

	B	C	D	E	F	G
4			List 1	List 2		
5	NO	#N/A	Artest	BMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
6	NO	#N/A	Artest	Harrington	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
7	YES	2	Harrington	BMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
8	NO	#N/A	Artest	Harrington	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
9	NO	#N/A	Artest	Harrington	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
10	NO	#N/A	Artest	BMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
11	YES	2	Harrington	BMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
12	NO	#N/A	Artest	Mercer	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
13	NO	#N/A	Artest	Harrington	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
14	NO	#N/A	Artest	Harrington	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
15	YES	8	Mercer	BMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
16	NO	#N/A	Artest	Mercer	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
17	NO	#N/A	O'Neal	RMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
18	NO	#N/A	O'Neal	BMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
19	NO	#N/A	O'Neal	RMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
20	NO	#N/A	O'Neal	RMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
21	YES	13	RMiller	BMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
22	NO	#N/A	O'Neal	RMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
23	NO	#N/A	O'Neal	RMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
24	YES	13	RMiller	BMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
25	NO	#N/A	O'Neal	Mercer	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
26	NO	#N/A	O'Neal	BMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
27	NO	#N/A	O'Neal	RMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>
28	NO	#N/A	O'Neal	BMiller	<code>{=IF(ISERROR(C5:C28),"NO","YES")}</code>	<code>{=MATCH(D5:D28,E5:E28,0)}</code>

Рис. 91.5. Поиск совпадений в двух списках

Чтобы записать в ячейку значение Да для каждого имени из списка 1, встречающегося в списке 2, и Нет для противоположного случая, выделите диапазон ячеек B5:B28 и введите в ячейку B5 формулу массива `{=ЕСЛИ(ЕОШИБКА(C5:C28);"Нет";"Да")}`. По этой формуле для каждой ячейки в C5:C28, содержащей сообщение об ошибке #Н/Д, записывается Нет и для всех ячеек с числовыми значениями за-

писывается **Да**. Обратите внимание, что формула `=ЕОШИБКА(х)` дает в результате **ИСТИНА**, если в *х* имеется ошибка, и **ЛОЖЬ** в противном случае.

**?** Как написать формулу, вычисляющую среднее значение для чисел в списке, которые больше или равны медиане списка?

В файле *Arrays.xlsx* на листе *Average Those > Median* (рис. 91.6) диапазон *D5:D785* (с именем *Prices*) содержит список цен. Требуется получить среднее значение для всех цен, которые не уступают медиане цены. В ячейке *F2* я вычислил медиану по формуле `=МЕДИАНА(prices)`. В ячейке *F3* я вычислил среднее значение чисел, которые больше или равны медиане, по формуле `=СУММЕСЛИ(prices;">="&F2; prices)/СЧЁТЕСЛИ(prices;">="&F2)`. Эта формула суммирует все цены, по меньшей мере, не уступающие значению медианы (243), и затем делит сумму на число цен, не уступающих значению медианы. Среднее значение для всех цен, которые не меньше медианы цены, составляет \$324,30.

	D	E	F	G
2		median	243	=MEDIAN(prices)
3		answer	324.2977	with formula
4	Price		↓	=SUMIF(prices,">="&F2,prices)/COUNTIF(prices,">="&F2)
5	224			
6	321		324.2977	with array
7	133			{=AVERAGE(IF(prices>=MEDIAN(prices),prices,""))}
8	310			
9	370			
10	223			
11	380			
12	253			
13	211			
14	248			
15	146			
16	334			
17	393			
18	295			
19	398			
20	166			
21	162			
22	340			

**Рис. 91.6.** Среднее значение для цен, которые не меньше медианы цены

Более простой подход — выделить ячейку *F6* и ввести формулу массива `{=СРЗНАЧ(ЕСЛИ(prices>=МЕДИАНА(prices);prices;""))}`. Эта формула создает массив, содержащий цену, если цена больше или равна медиане цены, или пробел в противном случае. Вычисление среднего значения для этого массива дает требуемый результат.

**?** У небольшой компании декоративной косметики имеется база данных по продажам, в которой указан продавец, количество проданных единиц про-

дукции и сумма в долларах для каждой сделки. Мне известно, что для обобщения этих данных подойдут статистические функции или функции СЧЁТЕСЛИМН (COUNTIFS), СУММЕСЛИМН (SUMIFS) и СРЗНАЧЕСЛИМН (AVERAGEIFS). Но могу ли я использовать также функции, возвращающие массив, чтобы обобщить эти данные и ответить на вопросы типа: сколько единиц косметики продал продавец, или сколько помады было продано, или сколько единиц косметики было продано определенным продавцом?

В файле Makeuparray.xlsx хранится список 1900 торговых сделок компании декоративной косметики. Для каждой сделки указаны номер сделки, продавец, дата сделки, проданный продукт, количество проданных единиц продукции и объем продаж в долларах. Часть данных представлена на рис. 91.7.

	I	J	K	L	M	N
4	Trans Number	Name	Date	Product	Units	Dollars
5	1	Betsy	4/1/2004	lip gloss	45	\$ 137.20
6	2	Hallagan	3/10/2004	foundation	50	\$ 152.01
7	3	Ashley	2/25/2005	lipstick	9	\$ 28.72
8	4	Hallagan	5/22/2006	lip gloss	55	\$ 167.08
9	5	Zaret	6/17/2004	lip gloss	43	\$ 130.60
10	6	Colleen	11/27/2005	eye liner	58	\$ 175.99
11	7	Cristina	3/21/2004	eye liner	8	\$ 25.80
12	8	Colleen	12/17/2006	lip gloss	72	\$ 217.84
13	9	Ashley	7/5/2006	eye liner	75	\$ 226.64
14	10	Betsy	8/7/2006	lip gloss	24	\$ 73.50
15	11	Ashley	11/29/2004	mascara	43	\$ 130.84
16	12	Ashley	11/18/2004	lip gloss	23	\$ 71.03
17	13	Emilee	8/31/2005	lip gloss	49	\$ 149.59
18	14	Hallagan	1/1/2005	eye liner	18	\$ 56.47
19	15	Zaret	9/20/2006	foundation	-8	\$ (21.99)
20	16	Emilee	4/12/2004	mascara	45	\$ 137.39
21	17	Colleen	4/30/2006	mascara	66	\$ 199.65
22	18	Jen	8/31/2005	lip gloss	88	\$ 265.19
23	19	Jen	10/27/2004	eye liner	78	\$ 236.15
24	20	Zaret	11/27/2005	lip gloss	57	\$ 173.12
25	21	Zaret	6/2/2006	mascara	12	\$ 38.08

**Рис. 91.7.** База данных компании декоративной косметики

Эти данные можно обобщить с помощью статистических функций для баз данных, описанных в главе 50 «Обработка данных с помощью статистических функций для баз данных», или с помощью функций СЧЁТЕСЛИМН (COUNTIFS) и СУММЕСЛИМН (SUMIFS). (См. главы 20 и 21.) Как будет показано в этом разделе, простую и эффективную альтернативу этим функциям составляют функции, возвращающие массив.

**О Сколько единиц косметики продала Джен?** На этот вопрос легко ответить с помощью функции СУММЕСЛИ (SUMIF). На указанном листе диапазоне J5:J1904 присвойте имя **Имя**, а диапазону M5:M1904 — имя **Количество**. В ячейку E7 введите формулу **=СУММЕСЛИ(Имя;"Джен";Количество)** для суммирования

всех единиц продукции, проданных Джен (Jen). Всего Джен продала 9537 единиц косметики. Также ответ на этот вопрос можно получить, если ввести формулу массива  $\{=\text{СУММ}(\text{ЕСЛИ}(\text{J5:J1904}=\text{"Jen"};\text{M5:M1904};0))\}$  в ячейку E6. Эта формула создает массив, содержащий количество проданных единиц косметики в сделках Джен и 0 для всех остальных сделок. Таким образом, суммирование элементов этого массива также дает в результате количество единиц товара, проданных Джен, — 9537 (рис. 91.8).

	A	B	C	D	E	F	G
5					units sold by jen	units lipstick sold by jen	units sold by Jen or lipstick
6				Array function	9537	1299	17061
7				other functions	9537	1299	17061
8							
9					Name	Product	
10					Jen	lipstick	
11							
12					Name	Product	
13					Jen		
14						lipstick	
15							
16		eye liner	foundation	lip gloss	lipstick	mascara	
17	Ashley	1920	1373	1985	1066	2172	
18	Betsy	1987	2726	1857	1305	1582	
19	Cici	1960	2031	1701	1035	2317	
20	Colleen	1107	2242	1831	765	2215	
21	Cristina	1770	1729	1734	788	1790	
22	Emilee	2490	1803	1725	720	1545	
23	Hallagan	2288	2387	1840	1045	1873	
24	Jen	2302	1883	1792	1299	2261	
25	Zaret	2715	2117	1868	800	1268	

**Рис. 91.8.** Суммирование данных с помощью формул массива

- **Сколько помады продала Джен?** Для решения этой задачи необходим критерий обработки только двух столбцов (Имя и Продукт). Ответ можно получить в ячейке F7 по формуле  $=\text{БДСУММ}(\text{J4:N1904};4;\text{E9:F10})$  со статистической функцией для баз данных. Формула показывает, что Джен продала 1299 штук помады. На этот вопрос можно также ответить с помощью формулы массива  $\{=\text{СУММ}((\text{J5:J1904}=\text{"Jen"})*(\text{L5:L1904}=\text{"lipstick"})*\text{M5:M1904})\}$ , введенной в ячейку F6.

Для понимания этой формулы необходимо немного знать булеву алгебру. Для части формулы  $(\text{J5:J1904}=\text{"Jen"})$  создается булев массив. Для каждой ячейки в J5:J1904, содержащей Jen, в массив включается значение ИСТИНА, а для каждой ячейки в J5:J1904, которая не содержит Jen, в массив включается значение ЛОЖЬ. Аналогично, для части формулы  $(\text{L5:L1904}=\text{"lipstick"})$  также создается булев массив, в котором значения ИСТИНА соответствуют каждой ячейке в диапазоне, содержащей lipstick, а значение ЛОЖЬ соответствует каждой ячейке, не содержащей lipstick. При перемножении булевых массивов создается еще один массив по следующим правилам:



- ИСТИНА\*ИСТИНА = 1;
- ИСТИНА\*ЛОЖЬ = 0;
- ЛОЖЬ\*ИСТИНА = 0;
- ЛОЖЬ\*ЛОЖЬ = 0.

Коротко говоря, перемножение булевых массивов имитирует оператор И. При умножении произведения булевых массивов на значения в диапазоне M5:M1904 создается новый массив. В любой строке, где Джен продавала помаду этот массив содержит количество проданной помады. Во всех остальных строках этого массива содержится значение 0. Сумма элементов этого массива дает общее количество помады, проданной Джен (1299 шт.).

- **Сколько единиц продукции было продано Джен и сколько всего было продано помады?** В ячейке G7 я использовал статистическую функцию для баз данных =БДСУММ(J4:N1904;4;E12:F14) и вычислил, что единицы продукции, проданные Джен, и проданная помада в сумме составили 17 061 штуку. В ячейке G6 я вычислил сумму для количества продукции, проданной Джен, и количества проданной помады по формуле массива {СУММ(ЕСЛИ((J5:J1904="jen")+(L5:L1904="lipstick");1;0)\*M5:M1904)}.

Опять же часть формулы (J5:J1904="jen")+(L5:L1904="lipstick") создает два булевых массива. Первый массив содержит значения ИСТИНА тогда и только тогда, когда продавцом является Джен (в формуле регистр символов не учитывается). Второй массив содержит значения ИСТИНА тогда и только тогда, когда проданный продукт является помадой (lipstick). Булевы массивы складываются по следующим правилам:

- ЛОЖЬ + ИСТИНА = 1;
- ИСТИНА + ИСТИНА = 1;
- ИСТИНА + ЛОЖЬ = 1;
- ЛОЖЬ + ЛОЖЬ = 0.

Коротко говоря, сложение булевых массивов имитирует оператор ИЛИ. Таким образом, эта формула создает массив, в котором для каждой строки, где Джен является продавцом или помада является проданным продуктом, количество проданных единиц продукции умножается на 1. Во всех остальных строках количество проданных единиц продукции умножается на 0. Результат тот же, что и со статистической функцией для баз данных (17 061).

- **Как просуммировать количество единиц каждого продукта, проданного каждым продавцом?** Формулы массива легко дают ответы на такие вопросы. Сначала составьте список продавцов в диапазоне ячеек A17:A25 и список названий продуктов в диапазоне ячеек B16:F16. Затем в ячейку B17 введите формулу массива {СУММ((J\$5:J\$1904=\$A17)\*(\$L\$5:\$L\$1904=B\$16)\*\$M\$5:\$M\$1904)}.

Эта формула вычисляет только количество подводки для глаз ( **eyeliner**), проданной Эшли (**Ashley**), — 1920 штук. Скопировав эту формулу в **C17:F17**, я вычислю поштучное количество каждого товара, проданного Эшли. Затем я скопирую формулу из **C17:F17** в **C18:C25** и вычислю поштучное количество каждого товара, проданного каждым продавцом. Обратите внимание, что я добавил знак доллара к **A** в ссылке на ячейку **A17** для получения имени соответствующего продавца и знак доллара к **16** в ссылке на ячейку **B16** для получения соответствующего продукта.

## ПРИМЕЧАНИЕ

Внимательные читатели могут спросить, почему бы просто не выделить формулу в **B17** и попытаться скопировать ее за один шаг? Напомню, что невозможно скопировать формулу массива в диапазон, содержащий и пустые ячейки, и формулы массива. Вот почему необходимо сначала скопировать формулу из **B17** в **C17:F17**, а затем уже перетянуть ее вниз для завершения таблицы.

Чтобы лучше понять эти формулы, нажмите **Вычислить формулу (Evaluate Formula)** и повторно щелкайте по **Вычислить (Evaluate)**, чтобы понять, как Excel рассчитывает по этой формуле конечный результат.

## ? Что такое массив констант и для чего он нужен?

Можно создать собственные массивы и применять их в формулах массива. Для этого заключите значения элементов массива в фигурные скобки **{ }**. Кроме того, текст необходимо указывать в кавычках ( **" "** ). В качестве элементов массива можно также включать логические значения **ИСТИНА** или **ЛОЖЬ**. В массивы констант нельзя включать формулы или символы, такие как знак доллара или запятая.

В качестве примера использования массива констант рассмотрим лист **Creating Powers** в файле **Arrays.xlsx** (рис. 91.9).

	C	D	E	F
3	<b>Sales</b>	<b>Sales^2</b>	<b>Sales^3</b>	<b>Sales^4</b>
4	2	4	8	16
5	4	16	64	256
6	8	64	512	4096
7	10	100	1000	10000
8	14	196	2744	38416
9	20	400	8000	160000

**Рис. 91.9.** Возведение объемов продаж во вторую, третью и четвертую степень

На этом листе даны объемы продаж за шесть месяцев, и для каждого месяца требуется создать вторую, третью и четвертую степень объемов продаж. Сначала выделите диапазон **D4:F9**, в который будут помещены результаты вычислений. Введите формулу массива **=C4:C9^{2,3,4}** в ячейку **D4**. По этой формуле в диапазоне



D4:D9 каждое значение из диапазона C4:C9 возводится в квадрат. В диапазоне ячеек E4:E9 по этой формуле каждое значение из C4:C9 возводится в куб. Наконец, в диапазоне ячеек F4:F9 каждое число из диапазона C4:C9 возводится в четвертую степень. Массив констант {2, 3, 4} необходим для перебора значений степени.

### ? Как изменить формулу массива?

Допустим, формула массива создает результаты во многих ячейках, а вам необходимо изменить, переместить или удалить результаты. Отдельный элемент массива изменить невозможно. Чтобы изменить формулу массива, сначала выделите все ячейки диапазона массива. Затем выделите одну ячейку в массиве. Нажав F2, вы сможете внести изменения в эту ячейку. После внесения изменений нажмите Ctrl+Shift+Enter для применения изменений. Теперь весь массив обновится.

### ? Как оценить тренд и сезонность выручки для магазина игрушек по заданной квартальной выручке?

В файле Toysrustrend.xlsx (рис. 91.10) содержатся данные о квартальной выручке (в миллионах долларов) магазина игрушек за 1997–2002 гг. Я хочу оценить квартальную тенденцию выручки, а также сезонность, связанную с каждым кварталом (первый квартал: январь — март; второй квартал: апрель — июнь; третий квартал: июль — сентябрь; четвертый квартал: октябрь — декабрь). Например, тенденция 1% за квартал означает, что продажи повысились на 1% за квартал. Индекс сезонности 0,80 для первого квартала, например, означает, что продажи за первый квартал составляют приблизительно 80% от продаж среднего квартала.

	C	D	E	F	G	H	I	J	K	L	M	N	O
2								mean quarter	0.582197	q4	172%		
3													
4					1	2	3	index	80%	73%	75%		
5	Year	Quarter	Sales	Quarter#	Q1 dummy	Q2 dummy	Q3 dummy	Forecast	q3	q2	q1	trend	const
6	1997	1	1646	1	1	0	0	1853.104665	0.467772	0.43	0.435	1.0086	4219.6
7	1997	2	1738	2	0	1	0	1826.702203					
8	1997	3	1883	3	0	0	1	2025.018084					
9	1997	4	4868	4	0	0	0	4366.196216					
10	1998	1	1924	5	1	0	0	1917.500329					
11	1998	2	1989	6	0	1	0	1890.180377					
12	1998	3	2142	7	0	0	1	2095.387765					
13	1998	4	4383	8	0	0	0	4517.922188					
14	1999	1	2043	9	1	0	0	1984.133752					
15	1999	2	2020	10	0	1	0	1955.864428					
16	1999	3	2171	11	0	0	1	2168.202803					
17	1999	4	4338	12	0	0	0	4674.920659					
18	2000	1	2186	13	1	0	0	2053.082697					
19	2000	2	2204	14	0	1	0	2023.83101					
20	2000	3	2465	15	0	0	1	2243.548175					
21	2000	4	5027	16	0	0	0	4837.374851					
22	2001	1	2319	17	1	0	0	2124.427627					
23	2001	2	1994	18	0	1	0	2094.15944					
24	2001	3	2220	19	0	0	1	2321.51181					
25	2001	4	4799	20	0	0	0	5005.474351					
26	2002	1	2061	21	1	0	0	2198.251805					
27	2002	2	2021	22	0	1	0	2166.931793					

Рис. 91.10. Оценка тренда и сезонности для выручки магазина игрушек

В решении задачи нам поможет функция ЛГРФПРИБЛ (LOGEST). Допустим, вы хотите предсказать переменную  $y$  по независимым переменным  $x_1, x_2, \dots, x_n$ , и предполагаете, что для некоторых значений  $a, b_1, b_2, \dots, b_n$  отношение между  $y$  и  $x_1, x_2, \dots, x_n$  задается уравнением  $y = a(b_1)^{x_1}(b_2)^{x_2}\dots(b_n)^{x_n}$ . (Назовем его уравнением 1.)

Функция ЛГРФПРИБЛ (LOGEST) используется для определений значений  $a, b_1, b_2, \dots, b_n$ , наилучшим образом приближающих это уравнение к наблюдаемым данным. Для использования функции ЛГРФПРИБЛ в целях вычисления тренда и сезонности нужно иметь в виду следующее:

- $y$  равен квартальной выручке;
- $x_1$  равен номеру квартала. (В хронологическом порядке: текущий квартал — это первый квартал, следующий квартал — второй квартал и т. д.);
- $x_2$  равен 1, если квартал является первым кварталом года, и 0 в противном случае;
- $x_3$  равен 1, если квартал является вторым кварталом года, и 0 в противном случае;
- $x_4$  равен 1, если квартал является третьим кварталом года, и 0 в противном случае.

Необходимо выбрать один квартал для исключения из модели. (Я произвольно выбрал четвертый квартал.) Этот подход аналогичен подходу с фиктивными переменными из главы 60. Теперь модель оценки выглядит следующим образом:  $y = a(b_1)^{x_1}(b_2)^{x_2}\dots(b_3)^{x_3}(b_4)^{x_4}$ . После того как функция ЛГРФПРИБЛ определит значения  $a, b_1, b_2, b_3$  и  $b_4$ , больше всего соответствующие набору данных, эти значения можно будет интерпретировать следующим образом:

- $a$  — константа для масштабирования прогнозов;
- $b_1$  — константа, представляющая среднее поквартальное процентное увеличение продаж в магазине игрушек;
- $b_2$  — константа, измеряющая отношение продаж в первом квартале к продажам в опущенном (четвертом) квартале;
- $b_3$  — константа, измеряющая отношение продаж во втором квартале к продажам в опущенном квартале;
- $b_4$  — константа, измеряющая отношение продаж в третьем квартале к продажам в опущенном квартале.

Сначала в ячейках G6:I27 я создал фиктивные переменные для кварталов 1–3, скопировав формулу =ЕСЛИ(\$D6=G\$4;1;0) из G6 в G6:I27. Напомню, что четвертый квартал также определен в Excel, поскольку для четвертого квартала все три фиктивных переменных имеют значение 0, и поэтому нет необходимости заводить для него фиктивную переменную.

Теперь я выделил диапазон K6:O6, потому что хочу, чтобы рассчитанные функцией ЛГРФПРИБЛ коэффициенты были помещены в него. Константа  $a$  займет самую

правую ячейку, за ней последуют коэффициенты в соответствии с независимыми переменными. Таким образом, рядом с константой окажется коэффициент тренда, затем коэффициент для квартала 1 и т. д.

Функция ЛГРФПРИБЛ имеет синтаксис: ЛГРФПРИБЛ(известные\_значения\_у; известные\_значения\_х; ИСТИНА; ИСТИНА). После ввода в ячейку К6 формулы массива {=ЛГРФПРИБЛ(Е6:Е27;F6:I27;ИСТИНА;ИСТИНА)} будут рассчитаны коэффициенты, показанные на рис. 91.10. Уравнение прогноза квартальной выручки (в млн) имеет вид:

$$4219,57 \times 1,0086^{\text{номер\_кварт}} \times 0,435^{K1\_макет} \times 0,426^{K2\_макет} \times 0,468^{K3\_макет}.$$

В первом квартале коэффициент K1\_макет равен 1, а коэффициенты для второго и третьего кварталов равны 0. (Напомним, что любое число, возведенное в нулевую степень, равно 1.) Таким образом, прогнозируемая выручка для первого квартала составляет  $4219,57 \times 1,00860086^{\text{номер\_кварт}} \times (0,435)$ .

Во втором квартале коэффициенты K1\_макет и K3\_макет равны 0, а коэффициент K2\_макет равен 1. Прогнозируемая выручка для этого квартала равна  $4219,57 \times 1,0086^{\text{номер\_кварт}} \times 0,426$ . В третьем квартале коэффициенты K1\_макет и K2\_макет равны 0, а коэффициент K3\_макет равен 1. Прогнозируемая выручка для этого квартала составляет  $4219,57 \times 1,0086^{\text{номер\_кварт}} \times 0,468$ . Наконец, в четвертом квартале коэффициенты K1\_макет, K2\_макет и K3\_макет равны 0. Прогнозируемая выручка для этого квартала составляет  $4219,57 \times 1,0086^{\text{номер\_кварт}}$ .

Таким образом, по оценкам квартальная тенденция к росту доходов составляет 0,9% (около 3,6% в год). После корректировки тренда получаем:

- выручка за квартал 1 в среднем составляет 43,5% от выручки за квартал 4;
- выручка за квартал 2 в среднем составляет 42,6% от выручки за квартал 4;
- выручка за квартал 3 в среднем составляет 46,8% от выручки за квартал 4.

Для вычисления индекса сезонности по каждому кварталу присвойте опущенному кварталу значение 1 и найдите весовой коэффициент для среднего квартала по следующей формуле (см. ячейку K2 на рис. 91.10):

$$\frac{0,435 + 0,426 + 0,468 + 1}{4} = 0,582.$$

Затем можно вычислить относительный индекс сезонности для кварталов 1–3, скопировав формулу =K6/\$K\$2 из K4 в L4:M4. Сезонность для квартала 4 вычисляется в ячейке M2 по формуле =1/K2. После корректировки тренда получаем:

- объем продаж в квартале 1 составляет 75% от объема продаж в типичном квартале;
- объем продаж в квартале 2 составляет 73% от объема продаж в типичном квартале;

- объем продаж в квартале 3 составляет 80% от объема продаж в типичном квартале;
- объем продаж в квартале 4 составляет 172% от объема продаж в типичном квартале.

Пусть вам необходимо подготовить прогноз на каждый квартал в соответствии с подобранным уравнением (уравнение 1). Вы можете использовать для этого функцию РОСТ (GROWTH), возвращающую массив. Ее синтаксис: РОСТ(известные\_значения\_y; известные\_значения\_x;новые\_значения\_x;ИСТИНА). Эта функция возвращает прогнозы для новых значений  $x$ , когда уравнение 1 подогнано к данным, содержащимся в диапазонах с известными значениями  $y$  и  $x$ . Выделите диапазон J6:J27 и введите в ячейку J6 формулу массива {=РОСТ(E6:E27;F6:I27;F6:I27;ИСТИНА)}. Формула сгенерирует прогнозы по уравнению 1 для выручки в каждом квартале. Например, прогноз на квартал 4 1997 г. по уравнению 1 составил \$4,366 млрд.

### ❓ Как вычислить медианный размер сделки в каждой стране по заданному списку сделок в различных странах?

В файле Medians.xlsx (рис. 91.11) содержатся данные о выручке по сделкам компании во Франции, США и Канаде. Здесь необходимо вычислить медианный размер сделки в каждой стране. Предположим, например, что вам нужно вычислить медианный размер сделки в США. Простой способ это сделать — создать массив, содержащий только данные по США, заменив остальные данные пустым значением. Далее вычислите медиану этого нового массива. Присвойте имя Страна столбцу C и имя Выручка столбцу D, затем в ячейку G5 введите формулу массива {=МЕДИАНА(ЕСЛИ(Страна=F5;Выручка;""))} для замещения выручки в каждой строке, содержащей сделку, совершенную не в США, пустым значением и для вычисления

	C	D	E	F	G
1	Country	Revenue			
2	US	5919			
3	Canada	4005			Median
4	US	6456			Revenue
5	France	8328		US	6376.5
6	Canada	9426		Canada	6326
7	US	5929		France	7403
8	France	7746			
9	Canada	9292			
10	US	8839			
11	France	7403			
12	Canada	3911			
13	US	7458			
14	Canada	8094			
15	France	4727			
16	Canada	5675			

**Рис. 91.11.** Вычисление медианного размера сделок в каждой стране

медианного размера сделок в США (\$6376,50). Формула, скопированная из G5 в G6:G7, вычисляет медианный размер сделки для Канады и Франции.

**❓ Имеются данные о торговых сделках с разбивкой по продавцам, продуктам и местоположению. Как определить стандартное отклонение проданных единиц продукции для каждого сочетания «продавец — продукт — регион»?**

Файл *Stdevif.xlsx* содержит 456 сделок по продаже косметики. (См. рис. 91.12.) Для каждой сделки известны продавец, количество проданных единиц продукции, продукт и местоположение. Необходимо определить для каждого сочетания «продавец — продукт — регион» стандартное отклонение проданных единиц продукции для соответствующих транзакций. Сначала на вкладке **Формулы (Formulas)** в группе **Определенные имена (Defined Names)** выберите **Создать из выделенного (Create From Selection)** для присвоения имен столбцам данных. Затем скопируйте данные на новый лист (Sheet2) и выберите для создания всех уникальных комбинаций «продавец — продукт — регион» инструмент **Удалить дубликаты (Remove Duplicates)** на вкладке **Данные (Data)** в группе **Работа с данными (Data Tools)**. Вставьте эти комбинации в диапазон ячеек L6:N51 (снова на листе Sheet1). Затем скопируйте формулу массива  $\{=\text{СТАНДОТКЛОН}(\text{ЕСЛИ}(\text{Имя}=\text{L7}; \text{ЕСЛИ}(\text{Продукт}=\text{M7}; \text{ЕСЛИ}(\text{Регион}=\text{N7}; \text{Кол\_во}))))\}$  из ячейки O7 в O8:O51. Эта формула сначала создаст массив из единиц (если Имя в ячейке равно Ashley) и нулей (если Имя в ячейке не равно Ashley). Затем массив из единиц и нулей для продукта mascara (тушь) и наконец — массив из единиц и нулей для региона East. Затем эти массивы перемножаются, в результате давая массив, соответствующий строкам, в которых Ashley продавала mascara в регионе East. Этот последний массив из единиц и нулей умножается на столбец Кол\_во (Units), и вы получаете все объемы продаж туши в сделках Ashley в регионе East. Функция **СТАНДОТКЛОН** возвращает стандартное отклонение (29,2) проданных единиц во всех сделках Ashley по продаже туши в регионе East. Копирование этой формулы в ячейки O8:O51 дает все требуемые стандартные отклонения. Чтобы лучше понять эту формулу вы можете применить инструмент **Вычислить формулу (Evaluate Formula)** на вкладке **Формулы (Formulas)** в группе **Зависимости формул (Formula Auditing)**.

**❓ Как применить функцию СУММПРОИЗВ (SUMPRODUCT) взамен условного счета, условного суммирования и условного определения среднего наравне с формулами массивов?**

Функцию **СУММПРОИЗВ** мы использовали с главы 30 «Поиск решения при определении оптимального ассортимента продукции» для решения многих задач на оптимизацию. Как показано в файле *Sumproductstricks.xlsx*, функция **СУММПРОИЗВ** может применяться и для имитации многих расчетов, выполняемых с помощью массивов, условного счета и условного суммирования. Как показано на рис. 91.13, допустим, имеются объемы продаж компьютерных чипов по разным странам.

Продemonстрируем теперь на нескольких примерах скрытую мощь функции **СУММПРОИЗВ**.

	G	H	I	J	K	L	M	N	O
4	Name	Product	Units	Location					
5	Ashley	mascara	43	east					
6	Ashley	mascara	93	east		Name	Product	Location	Std dev
7	Ashley	lip gloss	63	east		Ashley	mascara	east	29.20046
8	Ashley	mascara	19	east		Ashley	lip gloss	east	26.32319
9	Ashley	eye liner	41	east		Ashley	eye liner	east	23.8838
10	Ashley	foundatio	84	east		Ashley	foundatio	east	28.5482
11	Ashley	lipstick	-8	east		Ashley	lipstick	east	28.74022
12	Ashley	eye liner	76	east		Betsy	lip gloss	east	26.83129
13	Ashley	lip gloss	31	east		Betsy	lipstick	east	26.21341
14	Ashley	foundatio	8	east		Betsy	mascara	east	22.94559
15	Ashley	eye liner	81	east		Betsy	foundatio	east	28.26438
16	Ashley	foundatio	12	east		Betsy	eye liner	east	36.03074
17	Ashley	lip gloss	50	east		Cici	lip gloss	east	35.35887
18	Ashley	eye liner	39	east		Cici	mascara	east	34.82632
19	Ashley	mascara	17	east		Cici	lipstick	east	25.92939
20	Ashley	lip gloss	32	east		Cici	eye liner	east	19.77361
21	Ashley	lipstick	71	east		Cici	foundatio	east	22.94816
22	Ashley	lip gloss	40	east		Colleen	eye liner	east	39.11266
23	Ashley	eye liner	92	east		Colleen	lipstick	east	23.16679

**Рис. 91.12.** Вычисление стандартных отклонений на срезах данных

	A	B	C	D	E
1	Month	Product	Country	Revenue	Budget
2	January	Chip 1	US	\$4,000.00	\$5,454.00
3	January	Chip 1	Canada	\$3,424.00	\$5,341.00
4	January	Chip 1	US	\$8,324.00	\$1,232.00
5	January	Chip 1	France	\$5,555.00	\$3,424.00
6	January	Chip 1	Canada	\$5,341.00	\$8,324.00
7	January	Chip 1	US	\$1,232.00	\$5,555.00
8	January	Chip 1	France	\$3,424.00	\$5,341.00
9	January	Chip 1	Canada	\$8,324.00	\$1,232.00
10	January	Chip 1	US	\$5,555.00	\$3,424.00
11	January	Chip 1	France	\$5,341.00	\$8,324.00
12	January	Chip 1	Canada	\$1,232.00	\$5,555.00
13	January	Chip 1	US	\$3,424.00	\$5,341.00
14	January	Chip 1	Canada	\$8,383.00	\$5,454.00
15	January	Chip 1	France	\$8,324.00	\$1,232.00
16	January	Chip 1	Canada	\$5,555.00	\$3,424.00
17	January	Chip 1	US	\$5,341.00	\$8,324.00
18	January	Chip 1	France	\$1,232.00	\$5,555.00
19	January	Chip 1	France	\$3,523.00	\$9,295.00
20	February	Chip 2	Canada	\$5,555.00	\$3,424.00
21	February	Chip 2	US	\$5,454.00	\$4,000.00
22	February	Chip 2	US	\$5,341.00	\$8,324.00
23	February	Chip 2	France	\$1,232.00	\$5,555.00

**Рис. 91.13.** Объемы продаж компьютерных чипов



**❓ Для скольких сделок фактические доходы оказались меньше, чем заложенные в бюджет (прогнозные)?**

В ячейке G3 видно (рис. 91.14), что по формуле `=СУММПРОИЗВ(--(D2:D208<E2:E208))` 113 сделок принесли доход меньше заложенного в бюджет. Чтобы лучше понять эту формулу, поместите курсор в ячейку G3 и нажмите **Вычислить формулу (Evaluate Formula)** на вкладке **Формулы (Formulas)**. Вначале вы увидите массив записей со значениями **ИСТИНА** и **ЛОЖЬ**, показанный на рис. 91.15. Например, истинно то, что доход меньше заложенного в бюджет для первых двух сделок, но не для третьей и четвертой. Символ `--` преобразовывает записи **ИСТИНА** в единицы, а **ЛОЖЬ** — в нули, что и приводит к конечному результату: 113!

**❓ У скольких сделок доход меньше заложенного в бюджет, с разбиением по типу чипа?**

Для ответа на этот вопрос скопируем из G6 в G7:G8 формулу `= СУММПРОИЗВ(--(D2:D208<E2:E208); --(B2:B208=F6))`. Она возвращает значение **ИСТИНА** для каждой строки данных тогда и только тогда, если чип — это чип № 1, а доход в строке меньше заложенного в бюджет. Опять же результаты **ИСТИНА** превращаются в единицы, а **ЛОЖЬ** — в нули. И мы видим, что для 39 сделок с чипом № 1 доход меньше спрогнозированного.

**❓ Сколько сделок происходит в конкретных странах?**

Скопировав из G11 в G12:G13 формулу `=СУММПРОИЗВ(--($C$2:$C$208=F11))`, мы рассчитаем число сделок в каждой стране. Фрагмент формулы `$C$2:$C$208` возвращает значение **ИСТИНА** в каждой строке, где страной является Франция, и значение **ЛОЖЬ** по другим строкам. Затем превращаем результаты **ИСТИНА** в единицы, **ЛОЖЬ** — в нули, и получаем, что 75 сделок имели место во Франции.

**❓ Каков общий доход от каждого чипа?**

Скопировав формулу `= СУММПРОИЗВ(--($B$2:$B$208=F16);$D$2:$D$208)` из G16 в G17:G18, мы рассчитаем общий доход от каждого чипа. Фрагмент формулы `--($B$2:$B$208=F16)` возвращает значение **ИСТИНА** для строк по чипу № 1. Символ `--` преобразует результат **ИСТИНА** в единицы, которые перемножаются с доходом по каждой строке, находящимся в столбце D.

**❓ Рассчитайте общий доход для каждой комбинации страны и чипа.**

Скопировав из H22 в H23:H30 формулу `= СУММПРОИЗВ(--($B$2:$B$208=F22);--($C$2:$C$208=G22);$D$2:$D$208)`, мы рассчитаем доход по каждой комбинации чипа и страны. Фрагмент `--($B$2:$B$208=F22)` возвращает единицу по каждой строке, где фигурирует сделка с чипом № 1; тогда как фрагмент `--($C$2:$C$208=G22)` возвращает единицу по каждой строке, где страна — Франция. Перемножив эти два массива со столбцом D и просуммировав результат, получаем общий доход по чипу № 1 во Франции.





	O	P	Q	R	S	T	U
10	Revenue by Country						
11	France	\$368,873.00	=SUMPRODUCT(--(\$C\$2:\$C\$208=O11),\$D\$2:\$D\$208)				
12	US	\$327,946.00	=SUMPRODUCT(--(\$C\$2:\$C\$208=O12),\$D\$2:\$D\$208)				
13	Canada	\$329,459.00	=SUMPRODUCT(--(\$C\$2:\$C\$208=O13),\$D\$2:\$D\$208)				

Рис. 91.16. Расчет дохода по странам

## Задания

Данные к заданиям 1–5 содержатся в файле Chapter91data.xlsx.

1. На листе **Duplicate** находятся два списка имен. Подсчитайте с помощью формулы массива число имен, встречающихся в обоих списках.
2. На листе **Find Errors** находятся некоторые вычисления. Подсчитайте с помощью формулы массива число ячеек с ошибками. Подсказка: вставьте в формулу массива функцию **ЕОШИБКА (ISERROR)**.
3. На листе **Sales** находятся данные о продажах в магазине игрушек за 48 месяцев. Создайте формулу массива, суммирующую (начиная с третьего месяца) объемы продаж для каждого пятого месяца. Подсказка: воспользуйтесь функцией **ОСТАТ (MOD)**. Функция **ОСТАТ(число;делитель)** возвращает остаток от деления числа на делитель. Например, формула **=ОСТАТ(7;5)** дает в результате 2.
4. Также вычислите на листе **Sales** с помощью функции, возвращающей массив, третью, пятую и седьмую степени для объемов продаж в каждом месяце.
5. На листе **Product** находятся данные о продажах с апреля по август для продуктов 1–7. Объемы продаж указаны в одном столбце, по месяцам. Преобразуйте данные таким образом, чтобы объемы продаж для каждого месяца были указаны в одной строке и изменения в исходных данных отражались бы в новых преобразованных данных.
6. На основе данных из файла **Historicalinvest.xlsx** подсчитайте количество лет, в которые доходы по акциям, облигациям и казначейским векселям составляли бы от –20 до –15%, от –15 до –10% и т. д.
7. Матрица  $m \times n$  представляет собой прямоугольную таблицу чисел, содержащую  $m$  строк и  $n$  столбцов. Например, матрица  $3 \times 3$ :

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

Рассмотрим две матрицы,  $A$  и  $B$ . Предположим, что число столбцов в матрице  $A$  равно числу строк в матрице  $B$ . Тогда матрицу  $A$  можно умножить на

матрицу  $B$ . (Произведение записывается как  $AB$ .) Элемент в строке  $i$  и столбце  $j$  матрицы  $AB$  вычисляется путем применения функции СУММПРОИЗВ (SUMPRODUCT) к строке  $i$  матрицы  $A$  и столбцу  $j$  матрицы  $B$ . Матрица  $AB$  будет иметь столько же строк, сколько в матрице  $A$ , и столько же столбцов, сколько в матрице  $B$ . Функция МУМНОЖ (MMULT) — это функция, возвращающая массив, с помощью которой можно перемножать матрицы. Умножьте с помощью функции МУМНОЖ следующие матрицы:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}.$$

8. В квадратной матрице одинаковое число строк и столбцов. Пусть дана квадратная матрица  $A$ . Предположим, что существует такая матрица  $B$ , что матрица  $AB$  представляет собой матрицу, в которой все диагональные элементы равны 1, а все другие элементы равны 0. Тогда можно сказать, что  $B$  — это обратная матрица к матрице  $A$ . Функция МОБР (MINVERSE) вычисляет обратную матрицу к квадратной матрице. С помощью функции МОБР найдите обратные матрицы к матрицам  $A$  и  $B$  из задания 7.
9. Вы вложили часть  $f_i$  своих денег в инвестицию  $i$  ( $i = 1, 2, \dots, n$ ). Предположим, что стандартное отклонение годовой процентной доходности инвестиции  $i$  равно  $s_i$  и корреляция между годовой процентной доходностью инвестиции  $i$  и инвестиции  $j$  равна  $r_{ij}$ . Найдите дисперсию и стандартное отклонение годовой процентной доходности портфеля. Это просто сделать с помощью перемножения матриц. Создайте следующие три матрицы:
  - матрица 1 — это матрица  $1 \times n$ ,  $i$ -й элемент которой равен  $s_i f_i$ ;
  - матрица 2 — это матрица  $n \times n$ , в строке  $i$  и столбце  $j$  которой находится элемент  $r_{ij}$ ;
  - матрица 3 — это матрица  $n \times 1$ ,  $i$ -й элемент которой равен  $s_i f_i$ .

Дисперсия годовой процентной доходности портфеля — это просто произведение (матрица 1)\*(матрица 2)\*(матрица 3). Файл *Historicalinvest.xlsx* содержит годовые доходы от акций, облигаций и казначейских векселей. Используйте функции МУМНОЖ и ТРАНСП, чтобы вычислить (на основе указанных исторических данных) дисперсию и стандартное отклонение портфеля, в котором инвестиции распределены следующим образом: 50% в акции, 25% в облигации и 25% в казначейские векселя.

Данные к заданиям 10–13 находятся в файле *Makeupdb.xlsx*.

10. На какую сумму продала Джен блеск для губ (lip gloss)?
11. Каково среднее количество помады (lipstick), проданной Джен в регионе Восток (East)?

12. На какую сумму продала косметику Эмили и было продано косметики в регионе Восток?
13. На какую сумму продали блеск для губ Коллин и Зарет в регионе Восток?
14. На основе данных в файле **Chapter91data.xlsx** оцените тренд и сезонность квартальной выручки компаний Ford и GM.
15. В примере с магазином игрушек (см. файл **Toysrustrend.xlsx**) на основе данных за 1999–2001 гг. подготовьте прогноз для квартальной выручки в 2002 г.
16. В файле **Lillydata.xlsx** содержатся данные исследования рынка, которые использовались при разработке нового лекарства от давления. Пятнадцать экспертов (шесть из компании Lilly и девять из других компаний — см. столбец N) сравнивали пять сценариев для четырех потенциальных продуктов Lilly. Пятый вариант в каждом сценарии заключался в том, что предпочтение отдавалось препарату конкурента, а не четырем препаратам из списка компании Lilly.

Например, в первом сценарии во втором варианте рассматривался препарат компании Lilly, снижающий давление на 18 пунктов, дающий побочные эффекты с вероятностью 14% и стоивший \$16.

В диапазоне **I6:N21** содержатся данные о выборе, сделанном каждым экспертом для каждого из пяти сценариев. Например, первый эксперт (из Lilly) выбрал препарат конкурента в сценарии 1 и первый в списке препарат в сценарии 2. Исходя из этой информации, решите следующие задачи.

- Введите формулу, которую можно скопировать из **I2** в **I2:M5** и которая вычисляет в **I2:M5** цену для каждого сценария и варианта.
- Введите в ячейку **I23** формулу массива, которую можно скопировать в **I23:I32** и затем в **J23:M32** и которая вычисляет для каждого сценария частоту каждого ответа (1–5) с разбивкой по месту работы эксперта (в Lilly и не в Lilly). Таким образом, в первом сценарии один эксперт из Lilly выбрал ответ 1, три эксперта выбрали ответ 2 и два эксперта выбрали ответ 5.

17. В файле **Arrayexam1data.xlsx** содержатся данные о продажах и даты продаж для различных компаний. Задача состоит в разбивке продаж на квартальной основе с помощью формул массива.

Подведите итоги (только с помощью формул массива) по компаниям и по кварталам (рис. 91.17).

Например, ячейка **L7** должна содержать продажи компании ACS за квартал 1 (с 1 января по 31 марта) и т. д. Проверьте ответы по сводной таблице.

18. Объясните, почему формула массива **{=СУММ(1/СЧЁТЕСЛИ(Info;Info))}** вычисляет число уникальных записей в диапазоне **Info**. Примените эту формулу к данным в файле **Unique.xlsx** и убедитесь, что она возвращает число уникальных записей.

	C	D	E	F	G	H	I
1							
2	Date	Num	Name	Memo	Clr	Split	Amount
3	01/01/2004	1053544-03	ACS			2000 - Accounts Payable	54.63
4	01/01/2004	1052976-01	ACS			2000 - Accounts Payable	98.75
5	01/01/2004	1053544-01	ACS			2000 - Accounts Payable	224.96
6	01/01/2004	1053544-02	ACS			2000 - Accounts Payable	224.94
7	01/02/2004	4818189	Cardinal Health	67210		2000 - Accounts Payable	135.00
8	01/03/2004	4820762	Cardinal Health	67210		2000 - Accounts Payable	55,817.37
9	01/03/2004	4820860	Cardinal Health	67210		2000 - Accounts Payable	78.35
10	01/05/2004	47786	ActSys Medical, Inc.	INT014		2000 - Accounts Payable	3,455.00
11	01/05/2004	4822189	Cardinal Health	67210		2000 - Accounts Payable	650.81
12	01/05/2004	4822835	Cardinal Health	67210		2000 - Accounts Payable	6,950.36
13	02/26/2004	49426	ActSys Medical, Inc.	INT014		2000 - Accounts Payable	1,596.00
14	02/27/2004	4936605	Cardinal Health	67210		2000 - Accounts Payable	18,879.05
15	02/27/2004	4936809	Cardinal Health	67210		2000 - Accounts Payable	741,427.08
16	02/27/2004	1056106-02	ACS			2000 - Accounts Payable	55.69
17	02/28/2004	4938525	Cardinal Health	67210		2000 - Accounts Payable	35,812.86
18	04/30/2004	5063491	Cardinal Health	67210		2000 - Accounts Payable	13,466.25
19	04/30/2004	5063060	Cardinal Health	67210		2000 - Accounts Payable	24,092.33
20	05/01/2004	5065649	Cardinal Health	67210		2000 - Accounts Payable	25,148.30
21	05/04/2004	90021858	BioMed Plus, Inc.			2000 - Accounts Payable	1,107.00
22	05/05/2004	5073200	Cardinal Health	67210		2000 - Accounts Payable	47,485.66
23	05/05/2004	451499-0	Cardinal Health	67210		2000 - Accounts Payable	(2,336.24
24	05/05/2004	5072364	Cardinal Health	67210		2000 - Accounts Payable	3,764.50
25	05/05/2004	1881410 RI	Briggs Corporation	200977		2000 - Accounts Payable	147.49
26	05/06/2004	5076324	Cardinal Health	67210		2000 - Accounts Payable	51,598.62

**Рис. 91.17.** Формат ответа для задания 17

19. В файле Salaries.xlsx содержатся данные о зарплатах игроков НБА. Напишите формулу массива, которая суммирует четыре самых больших зарплаты игроков. Подсказка: используйте массив констант {1, 2, 3, 4} в сочетании с функцией **НАИБОЛЬШИЙ** (**LARGE**).

# Запись макросов

### Обсуждаемые вопросы

- Что такое макрос?
- Как используется вкладка Разработчик (Developer) при записи и запуске макросов?
- Как записать макрос?
- Как записать макрос, применяющий требуемый формат к произвольному диапазону ячеек?
- Как запустить макрос?
- Как при записи макросов использовать относительные ссылки?
- Как записать макрос, фильтрующий сводную таблицу так, чтобы отображались только данные о 20 основных клиентах?

## Ответы на вопросы

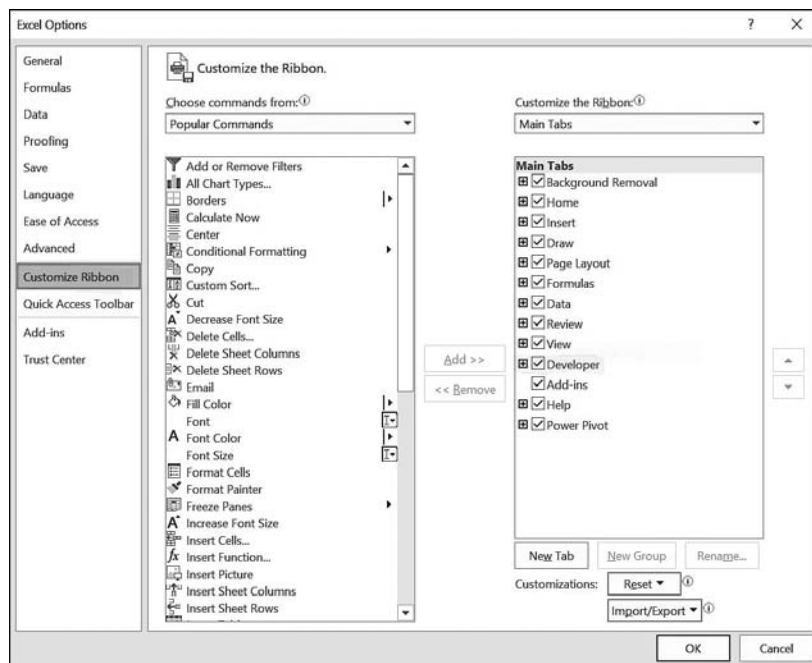
### ❓ Что такое макрос?

*Макросы* позволяют автоматизировать выполнение в Excel повторяющихся действий. Например, для применения определенного формата к диапазону ячеек вы можете создать макрос, выполняющий это задание по форматированию одним нажатием кнопки. Это может быть, к примеру, набор заголовков, который вы часто используете в своих листах. После создания соответствующего макроса вы сможете вставлять эти заголовки в любое место листа одним нажатием кнопки. В этой главе будет показано, насколько просто записывать макросы. Вы можете достичь многого, записывая макросы, но есть круг макросов, которые нельзя записать таким образом, — они создаются с помощью языка программирования Excel VBA (Visual Basic для приложений). Читателям, интересующимся VBA, рекомендую книгу Excel 2016 Power Programming with VBA (Wiley, 2016).

### ❓ Как используется вкладка Разработчик (Developer) при записи и запуске макросов?

Для создания макросов (как для записи, так и для написания макросов на VBA) и для их запуска вам необходимо наличие на ленте вкладки **Разработчик (Developer)**.

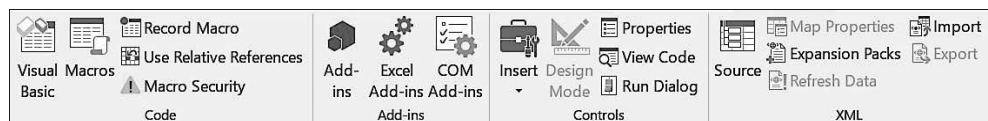
Для размещения на ленте вкладки Разработчик (если ее там еще нет) перейдите на вкладку Файл (File) и выберите Параметры (Options). Затем выберите в левой панели раздел Настроить ленту (Customize Ribbon) и установите флажок Разработчик (Developer) в разделе Основные вкладки, как показано на рис. 92.1. Затем нажмите ОК в диалоговом окне Параметры Excel (Excel Options).



**Рис. 92.1.** Помещение вкладки Разработчик на ленту

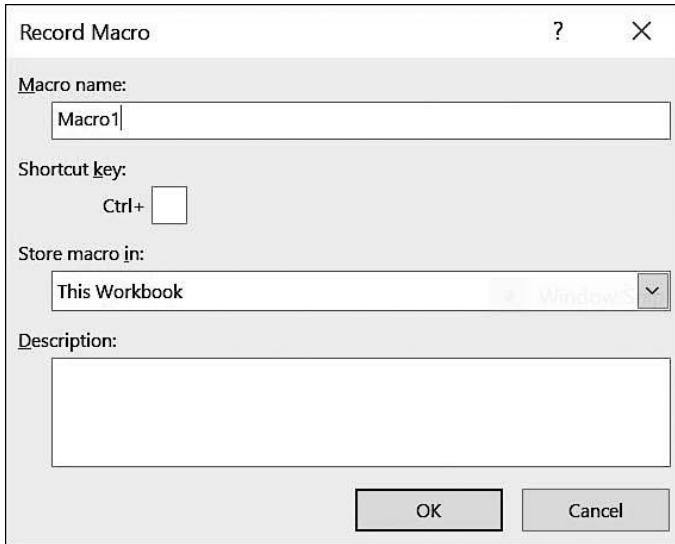
### ❓ Как записать макрос?

Для записи макроса перейдите на вкладку Разработчик (Developer). Отобразятся опции, представленные на рис. 92.2.



**Рис. 92.2.** Вкладка Разработчик

Для начала записи макроса щелкните Запись макроса (Record Macro) в группе Код (Code). Откроется диалоговое окно Запись макроса (Record Macro), представленное на рис. 92.3.



**Рис. 92.3.** Диалоговое окно Запись макроса

Рассмотрим подробнее диалоговое окно **Запись макроса (Record Macro)**.

- **Имя макроса (Macro Name).** Excel присваивает макросу имя по умолчанию (например, *Macro1*). Возможно, вам потребуется более информативное имя для вашего макроса. Например, если макрос меняет начертание в выделенных ячейках на курсив и цвет шрифта на красный, вы можете назвать макрос *красный курсив*.
- **Сочетание клавиш (Shortcut Key).** Как вскоре будет показано, есть много способов запустить макрос. Задавая *сочетание клавиш*, вы определяете последовательность клавиш, которую можно использовать для запуска макроса. Например, если ввести **Shift+F**, то нажатие последовательности **Ctrl+Shift+F** запустит макрос на выполнение. Однако следует внимательно подходить к выбору последовательности клавиш. Например, так как **Ctrl+B** меняет начертание выделенного текста на полужирное, нельзя использовать для макроса это сочетание клавиш. Задание сочетания клавиш не является обязательным.
- **Сохранить в (Store Macro In).** По умолчанию макрос сохраняется в текущей книге. Затем этот макрос становится доступен при каждом открытии книги. Если вам нужно, чтобы данный макрос был всегда доступен при работе в Excel, выберите вариант **Личная книга макросов (Personal Macro Workbook)**.
- **Описание (Description).** Заполнение этого поля необязательно, но может оказаться полезным, особенно если вы используете свой файл Excel совместно с кем-то еще.

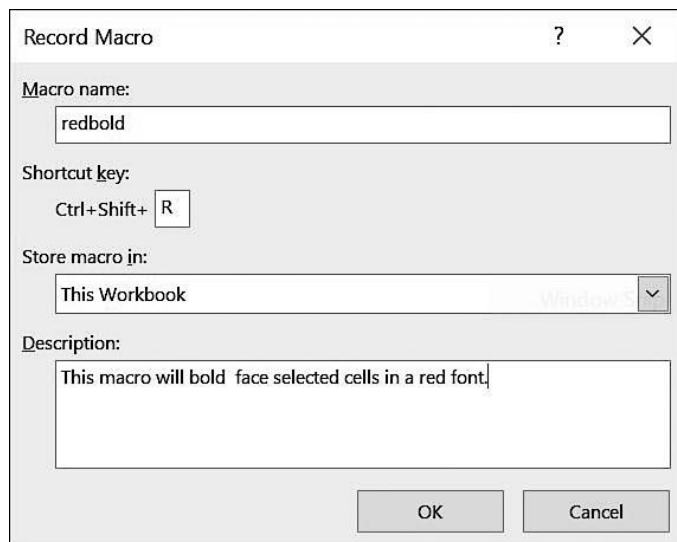
После заполнения диалогового окна **Запись макроса (Record Macro)** можно приступить к записи макроса. Нажмите **OK** и выполните задания, которые требуется

выполнять при работе этого макроса. Когда вы закончите, выберите **Остановить запись (Stop Recording)** на вкладке **Разработчик (Developer)** в группе **Код (Code)**. (Пункт меню **Запись макроса (Record Macro)** изменится на **Остановить запись (Stop Recording)**.) Кроме того, остановить запись можно, нажав маленький серый прямоугольник, отображающийся в левой нижней части экрана (под закладками листов книги).

### ❓ Как записать макрос, применяющий требуемый формат к произвольному диапазону ячеек?

Предположим, вам в работе часто требуется выделять числовые данные полужирным красным начертанием. Не составляет труда создать макрос, автоматизирующий эту рутинную работу. (Готовый результат вы можете увидеть в файле **Redbold.xlsm**.) До записи макроса введите какие-нибудь данные (например, числа в диапазоне ячеек **F3:F6**), которые затем будет нужно отформатировать. Выберите диапазон ячеек и выполните следующие шаги:

1. На вкладке **Разработчик (Developer)** в группе **Код (Code)** выберите **Запись макроса (Record Macro)** и заполните открывшееся диалоговое окно, как показано на рис. 92.4. Обратите внимание, что задана комбинация клавиш **Ctrl+Shift+R**, которая будет запускать макрос. Нажмите **OK**, чтобы приступить к записи макроса.



**Рис. 92.4.** Параметры макроса с именем **redbold**

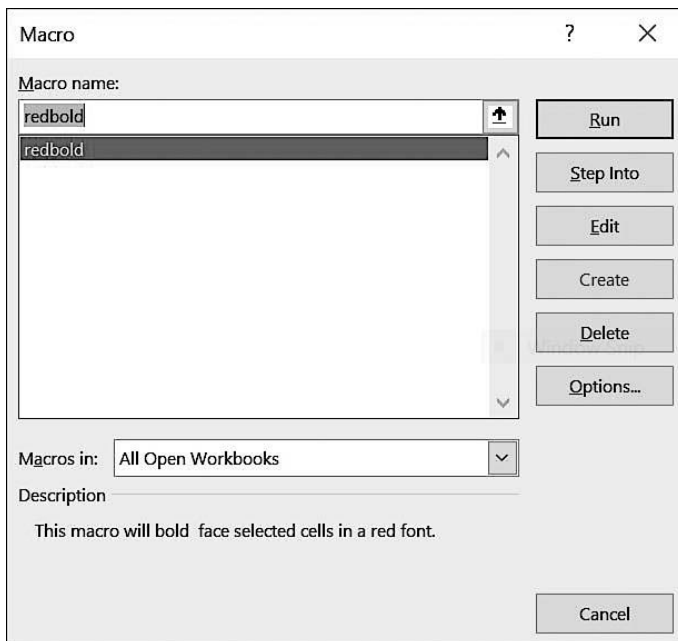
2. Нажмите **Ctrl+1** (или щелкните правой кнопкой мыши и выберите **Формат ячеек (Format Cells)**) для перехода в диалоговое окно **Формат ячеек (Format Cells)**.



Выберите в диалоговом окне вкладку **Шрифт (Font)** и далее **Полужирный (Bold)** в списке **Начертание (Font Style)** и красный цвет в разделе **Цвет (Color)**.

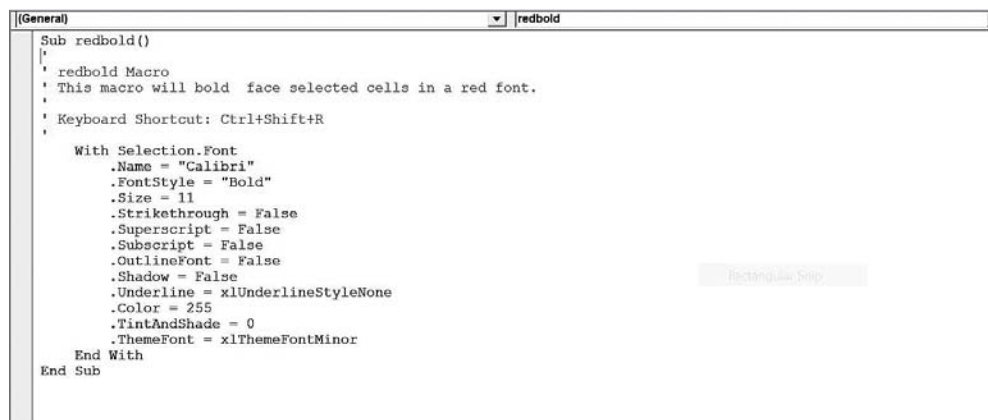
3. Щелкните **Остановить запись (Stop Recording)**.
4. Так как книга содержит макрос, при сохранении файла укажите тип файла **Книга Excel с поддержкой макросов (Excel Macro-Enabled workbook)**. Имя получившегося файла — **Redbold.xlsm**. Расширение **.xlsm** указывает, что файл содержит книгу с поддержкой макросов.
5. Для запуска макроса можно использовать сочетание клавиш **Ctrl+Shift+R**.
6. Можно проверить работу макроса, введя, например, данные в ячейки **J5:J7**, затем выделить их и запустить макрос снова. Ячейки **J5:J7** станут отображаться полужирным начертанием красного цвета. Можно провести этот тест на других диапазонах ячеек.

Кроме того, макрос можно запустить, перейдя на вкладку **Разработчик (Developer)** и выбрав **Макросы (Macros)** в группе **Код (Code)**. Откроется список доступных макросов, как показано на рис. 92.5. Если выбрать макрос **redbold** и щелкнуть на **Выполнить (Run)**, начнется выполнение макроса.



**Рис. 92.5.** Список доступных макросов в открытых книгах

При записи макроса записывается соответствующий код VBA. При выборе **Изменить (Edit)** в диалоговом окне **Макрос (Macro)** вы увидите этот код VBA (рис. 92.6).



**Рис. 92.6.** Записанный код макроса для макроса redbold

Заметьте, что все, что требовалось, — это выбрать красный цвет шрифта (инструкция `.Color=255`) и выбрать полужирное начертание (инструкция `.FontStyle="Bold"`). Записанный код макроса неэффективен, так как содержит много избыточных инструкций. Далее будет показано, как можно отредактировать макрос.

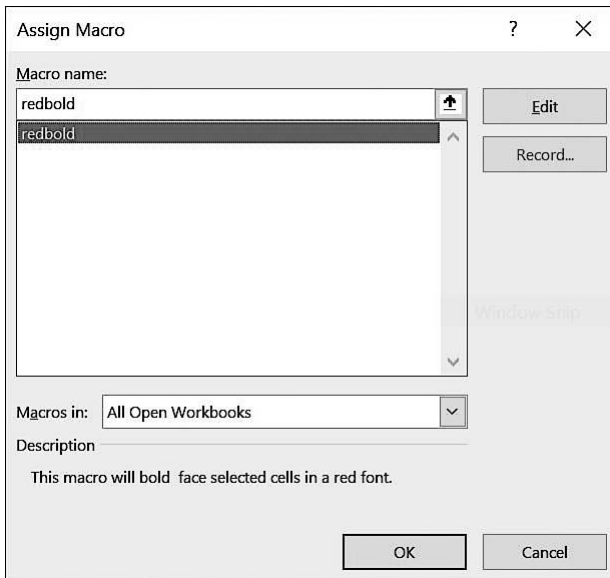
### ? Как запустить макрос?

Мы уже рассмотрели запуск макроса с помощью комбинации клавиш и выбора макроса в диалоговом окне **Макрос (Macro)**. В этом разделе мы рассмотрим запуск макроса через элемент управления — кнопку (см. главу 27 «Счетчики, полосы прокрутки, переключатели, флажки, группы и поля со списками») или через размещение макроса в панели быстрого доступа (которая находится слева над лентой). Для запуска макроса нажатием кнопки выберите **Вставить (Insert)** в группе **Элементы управления (Controls)** на вкладке **Разработчик (Developer)**, далее выберите форму элемента управления **Кнопка (Button)** (первый элемент списка) из раздела **Элементы управления (Form Controls)**. Разместите кнопку, и откроется диалоговое окно, представленное на рис. 92.7. В этом диалоговом окне вы можете назначить макрос redbold этой кнопке. Выберите нужный макрос и нажмите **OK** в окне **Назначить макрос объекту (Assign Macro)**.

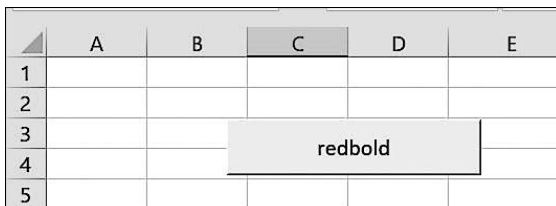
Чтобы изменить имя кнопки, соответствующей макросу `redbold`, щелкните правой кнопкой мыши на кнопке и выберите **Изменить текст (Edit Text)**. Замените текст на redbold. Для подтверждения внесения изменений щелкните мышью на произвольной ячейке листа (рис. 92.8). Проверьте работу кнопки: введите некоторые новые данные, выберите их и щелкните по кнопке. Новые данные будут выделены полужирным начертанием, а текст станет красного цвета.

Для размещения макроса в панели быстрого доступа выполните следующие шаги:

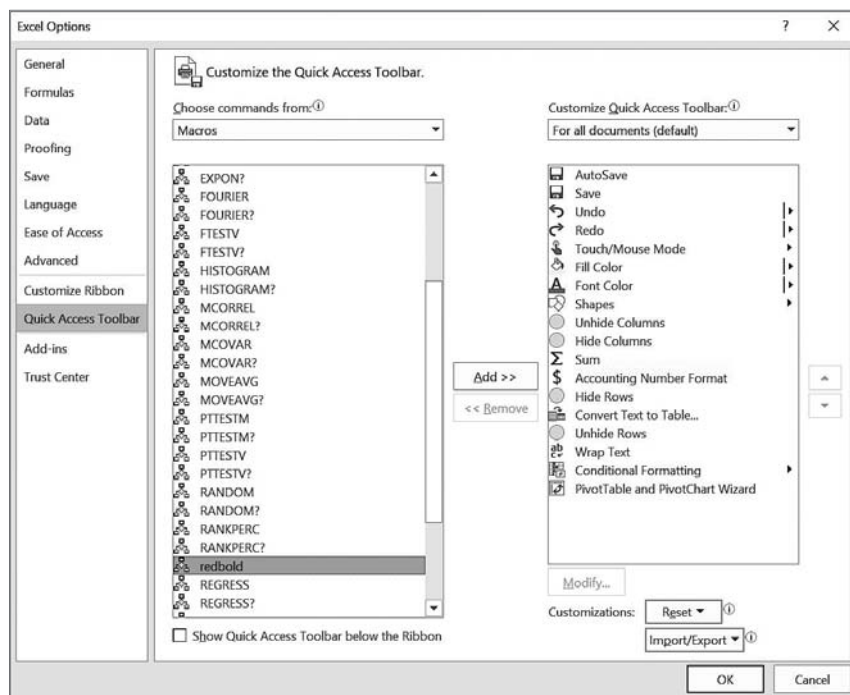
1. Щелкните правой кнопкой мыши на панели быстрого доступа и выберите **Настройка панели быстрого доступа (Customize Quick Access Toolbar)**.
2. В диалоговом окне **Параметры Excel (Excel Options)**, представленном на рис. 92.9, выберите **Макросы (Macros)** в раскрывающемся списке **Выбрать команды из: (Choose Commands From)**.
3. Выберите макрос **redbold** и щелкните кнопку **Добавить>> (Add)**.
4. Щелкнув кнопку **Изменить... (Modify)** (справа внизу), вы можете настроить расположение значка, используемого для запуска макроса, чтобы он отображался последним в панели быстрого доступа. Нажмите **ОК** в диалоговом окне **Изменение кнопки (Modify Button)** и в диалоговом окне **Параметры Excel (Excel Options)**.
5. Если вы выделите диапазон ячеек, щелчок на значке, который вы добавили на панель быстрого доступа, запустит на выполнение макрос **redbold**.



**Рис. 92.7.** Назначение макроса кнопке



**Рис. 92.8.** Кнопка запуска макроса redbold



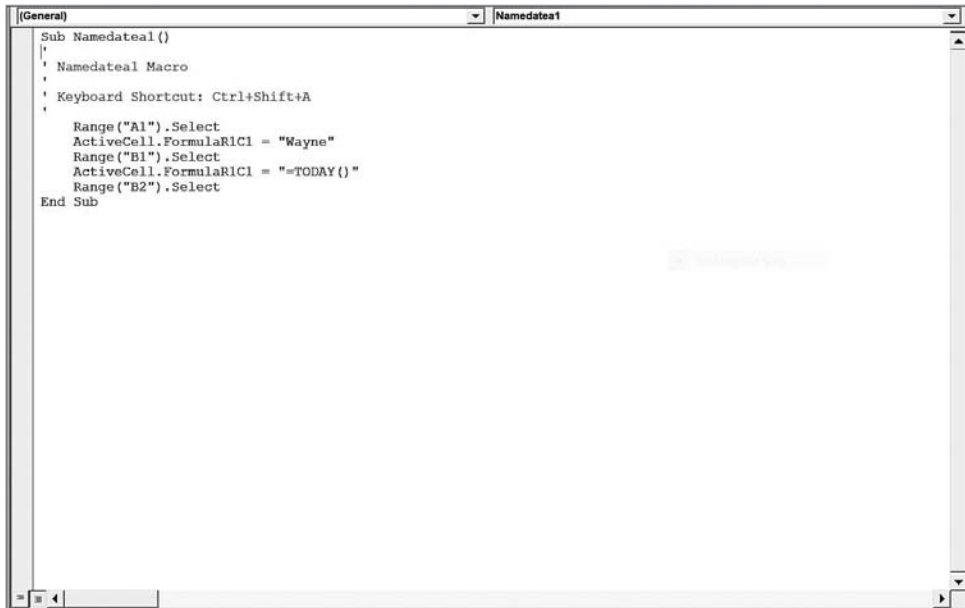
**Рис. 92.9.** Выбор макроса redbold для добавления в панель быстрого доступа

### ❓ Как при записи макросов использовать относительные ссылки?

Предположим, вам требуется ввести свое имя в заданную ячейку и текущую дату в соседнюю справа от нее ячейку.

Сначала создадим макрос, всегда помещающий ваше имя в ячейку A1, а текущую дату — в ячейку B1. (См. файл Namedatemacro.xlsm.)

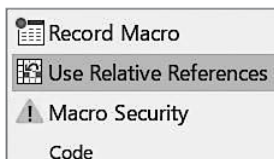
1. На вкладке Разработчик (Developer) выберите Запись макроса (Record Macro), назовите макрос *Namedatea1* и присвойте ему сочетание клавиш Ctrl+Shift+A. Нажмите OK, чтобы закрыть диалоговое окно Запись макроса.
2. Перейдите в ячейку A1 и введите свое имя, затем перейдите в ячейку B1 и введите формулу =TODAY(), которая всегда возвращает текущую дату.
3. Нажмите Остановить запись (Stop Recording).
4. Для проверки работы макроса перейдите в какую-либо ячейку и введите сочетание клавиш Ctrl+Shift+A. Вы увидите, что введенное вами имя появится в ячейке A1, а текущая дата в ячейке B1. Теперь выберите Изменить (Edit) в диалоговом окне Макрос (Macro), и вы увидите код VBA, приведенный на рис. 92.10.



**Рис. 92.10.** Макрос, помещающий имя в ячейку A1 и текущую дату в ячейку B1

Просмотрев код, вы увидите, что сначала Excel выбирает ячейку A1. Она становится активной, и Excel вводит *Wayne* в текущую активную ячейку (A1). Потом он перемещается в ячейку B1 и вводит в нее формулу =TODAY().

Так как при записи макроса не было изменено перемещение курсора, Excel затем перемещает курсор на одну ячейку вниз, в ячейку B2. Если удалить из макроса строку Range("B2").Select и перезапустить выполнение макроса, вы увидите, что макрос завершит работу в ячейке B1, а не B2. Этот простой пример показывает возможности редактирования кода макроса. Теперь предположим, что вам нужно иметь возможность выбрать любую ячейку, чтобы ввести в нее свое имя и чтобы текущая дата отображалась в соседней справа от нее ячейке. Это означает, что макрос должен содержать относительные ссылки. Для этого выберите опцию **Относительные ссылки (Use Relative References)** в группе **Код (Code)** на вкладке **Разработчик (Developer)**, как показано на рис. 92.11.



**Рис. 92.11.** Включение режима относительных ссылок

После включения режима относительных ссылок макрос не будет содержать явных ссылок на ячейки, кроме тех, которые были выбраны до записи макроса. При выполнении этого макроса все вводы будут выполняться относительно текущей активной ячейки. Для создания макроса, который пишет ваше имя в выделенной ячейке и помещает текущую дату в соседней с ней справа ячейке, выполните следующие действия. (См. файл *Namedaterelative.xlsm*.)

1. Откройте новую книгу и переместитесь в произвольную ячейку (например, H5).
2. Перейдите на вкладку Разработчик (Developer), включите режим Относительные ссылки (Use Relative References) в группе Код (Code).
3. Выберите Запись макроса (Record Macro) и присвойте макросу имя *Имядата-относит.* Щелкните ОК.
4. Введите свое имя в активной ячейке (в H5).
5. Переместитесь на одну ячейку вправо, в I5, и введите формулу =TODAY().
6. Дважды щелкните на заголовке столбца I, чтобы его ширина вмещала дату.
7. Щелкните Остановить запись (Stop Recording).
8. Создайте кнопку для запуска макроса, как было описано ранее.
9. Перейдите в некоторую ячейку листа и запустите макрос. Вы увидите свое имя в выбранной ячейке и текущую дату в соседней с ней справа. В завершение выполнения макроса курсор переместится на одну ячейку вниз от даты. Если вам нужно, чтобы этого не происходило, отредактируйте макрос, удалив инструкцию, перемещающую курсор от ячейки с датой вниз.

Код VBA для этого макроса приведен на рис. 92.12.



**Рис. 92.12.** Код VBA, помещающий в произвольную ячейку ваше имя и рядом текущую дату

Далее приведены описания всех использованных в макросе инструкций:

- **Sub Namedaterelative()**. Макрос начинается с инструкции Sub, содержащей имя макроса.
- **' Namedaterelative Macro**. Инструкции, начинающиеся с одиночной прямой кавычки ('), являются комментариями и не исполняются. Данная инструкция — комментарий, определяющий имя макроса.
- **ActiveCell.FormulaR1C1 = «Wayne»**. Эта инструкция вставляет текст Wayne<sup>1</sup> в выбранную перед запуском макроса ячейку.
- **ActiveCell.Offset(0, 1).Range(«A1»).Select**. Эта инструкция перемещает курсор на одну ячейку вправо от текущей ячейки. Обратите внимание на использование Offset, сходное с использованием функции СМЕЩ (OFFSET), о которой рассказывалось в главе 22.
- **ActiveCell.FormulaR1C1 = «=TODAY()»**. Эта инструкция вводит текущую дату в активную ячейку, которая находится на один столбец правее ячейки, содержащей ваше имя.
- **ActiveCell.Offset(1, 0).Range(«A1»).Select**. Эта инструкция перемещает курсор на одну ячейку вниз от ячейки с текущей датой. Если удалить эту инструкцию из макроса, активной ячейкой после завершения работы макроса будет ячейка, содержащая текущую дату.
- **ActiveCell.Columns(«A:A»).EntireColumn.EntireColumn.AutoFit**. Инструкция подбирает ширину столбца, содержащего дату, чтобы дата отображалась полностью.
- **End Sub**. Инструкция End Sub отмечает конец макроса.

❓ **Как записать макрос, фильтрующий сводную таблицу так, чтобы отображались только данные о 20 основных клиентах?**

Файл Ptcustomerstemp.xlsx (в папке Template) содержит сводную таблицу из главы 45 «Сводные таблицы и срезы для описания данных», в которой собраны продажи по 100 клиентам. Мы создадим макрос, который при щелчке кнопкой отфильтрует сводную таблицу и выведет данные о 20 основных клиентах. Выполните следующие шаги:

1. На вкладке Разработчик (Developer) в группе Код (Code) выберите Запись макроса (Record Macro). В открывшемся диалоговом окне присвойте макросу имя Topcustomers. Щелкните ОК, чтобы начать запись.
2. Перейдите в ячейку A4 и, нажав на стрелку, открывающую меню, выберите Фильтры по значению (Value filters).
3. Выберите Первые 10 (Top 10). Затем в диалоговом окне Фильтр "Первые 10" (Top 10 Filter (Customer)) измените значение 10 на 20, чтобы задать фильтр на отображение первых 20 клиентов. Щелкните ОК.

<sup>1</sup> В вашем макросе в этой инструкции будет ваш текст. — *Примеч. пер.*

4. Нажмите **Остановить запись (Stop Recording)**.
5. Создайте кнопку для запуска макроса. При запуске макроса будут отображаться только данные по первым 20 клиентам. Для проверки работы готового макроса откройте файл **Ptcustomers.xlsm** из папки **Practice Files**, сбросьте значение фильтра в ячейке **A4** и нажмите кнопку **Top20**.

Вы можете теперь записать макрос, восстанавливающий исходную сводную таблицу. (См. задание 4.)

## Задания

1. Игроков команды МБА зовут Уэйн, Тейлор, Бритни и Кэти. Создайте макрос, вводящий имена игроков в ячейки **A1:D1**.
2. Измените макрос из задания 1 так, чтобы имена игроков вводились в одной строке, начиная с произвольно заданной ячейки.
3. Создайте макрос, размещающий толстую красную границу вокруг любого набора выделенных ячеек.
4. Для примера сводной таблицы из файла **Ptcustomers.xlsm** создайте макрос, сбрасывающий любые фильтры на данные по клиентам.
5. Создайте макрос, который будет форматировать начертание любого выделенного диапазона ячеек как полужирный курсив.
6. Сочетание клавиш **Ctrl+A** выделяет все ячейки на текущем листе. Используйте этот факт при создании макроса, который гарантированно отформатирует все ячейки листа начертанием полужирный курсив.
7. Часто требуется зафиксировать результаты в некотором диапазоне, вычисленные по формулам, и заменить формулы значениями. Создайте макрос, который, основываясь на выделенной ячейке, будет находить наибольший диапазон ячеек, содержащий текущую ячейку, и преобразовывать все формулы в значения. Подсказка: **Ctrl+Shift+\*** выделит требуемый диапазон. Протестируйте макрос, введя в лист набор функций **=СЛЧИС()** (**RAND**).
8. Создайте макрос, который выделяет желтым цветом столбец данных целиком. Этот макрос должен работать для любого количества строк в столбце. Подсказка: используйте сочетание клавиш **Ctrl+Shift+↓**.



## ГЛАВА 93

# Продвинутый анализ чувствительности

Обсуждаемый вопрос

- В электронной таблице много исходных параметров, поэтому двунаправленная таблица данных в этом случае неприемлема для анализа чувствительности. Как же мне все-таки провести анализ чувствительности?

## Ответ на вопрос

❓ В электронной таблице много исходных параметров, поэтому двунаправленная таблица данных в этом случае неприемлема для анализа чувствительности. Как же мне все-таки провести анализ чувствительности?

В файле Spiderplottemplate.xlsx (рис. 93.1) содержится расчет чистой приведенной стоимости (ЧПС/NPV) после уплаты налогов от нового продукта на основе базовых исходных параметров в ячейках D6:D13. Нужно построить график, показы-

	A	B	C	D	E	F
2						
3				Change	2	
4				percentchange	0	
5			used	base		
6	1	taxrate	0.4	0.4		
7	2	Year1sales	12000	12000		
8	3	Sales growth	0.05	0.05		
9	4	Year1price	7.5 \$	7.50		
10	5	Year1cost	6 \$	6.00		
11	6	intrate	0.15	0.15		
12	7	costgrowth	0.05	0.05		
13	8	pricegrowth	0.03	0.03		
14	Year	1	2	3	4	5
15	Unit Sales	12000	12600	13230	13891.5	14586.075
16	unit price	\$ 7.50	\$ 7.73	\$ 7.96	\$ 8.20	\$ 8.44
17	unit cost	\$ 6.00	\$ 6.30	\$ 6.62	\$ 6.95	\$ 7.29
18	Revenues	\$ 90,000.00	\$ 97,335.00	\$ 105,267.80	\$ 113,847.13	\$ 123,125.67
19	Costs	\$ 72,000.00	\$ 79,380.00	\$ 87,516.45	\$ 96,486.89	\$ 106,376.79
20	Before Tax Profits	\$ 18,000.00	\$ 17,955.00	\$ 17,751.35	\$ 17,360.24	\$ 16,748.88
21	Tax	\$ 7,200.00	\$ 7,182.00	\$ 7,100.54	\$ 6,944.10	\$ 6,699.55
22	Aftertax Profits	\$ 10,800.00	\$ 10,773.00	\$ 10,650.81	\$ 10,416.15	\$ 10,049.33
23						
24	NPV	\$35,492.08				

Рис. 93.1. ЧПС нового продукта

вающий изменение ЧПС, когда все исходные параметры, за исключением одного, остаются на уровне своих базовых значений, а значение оставшегося исходного параметра изменяется в диапазоне от  $-40\%$  до  $+40\%$  от его базового значения.

Хитрость в том, чтобы все-таки использовать двуправленную таблицу, указав ячейку E4 в поле Подставлять значения по строкам (Row input cell), что позволяет определить процентное изменение исходного параметра, и ячейку E3 в поле Подставлять значения по столбцам (Column input cell) для выбора исходного параметра (на основании меток в диапазоне ячеек A6:A13).

Скопировав из C6 в C7:C13 следующую формулу  $=\text{ЕСЛИ}(\text{A6}=\text{Change}; \text{D6}*(1+\text{percentage}); \text{D6})$ , мы можем быть уверены, что таблица данных, представленная на рис. 93.2, рассчитывает ЧПС, когда значение каждого исходного параметра варьирует в пределах  $40\%$  от его базового значения. Например, в ячейке R26 мы видим, что если рост базовых продаж сократится на  $20\%$ , а все другие исходные параметры останутся на базовом уровне, то ЧПС упадет до 34925,57 доллара.

	O	P	Q	R	S	T	U	V	W	X
23		-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4
24	taxrate	44956.63	42590.49	40224.35	37858.22	35492.08	33125.94	30759.8	28393.66	26027.52
25	Year1sales	21295.25	24844.45	28393.66	31942.87	35492.08	39041.29	42590.49	46139.7	49688.91
26	Sales growth	34368.87	34646	34925.57	35207.59	35492.08	35779.06	36068.54	36360.54	36655.08
27	Year1price	-47879.9	-27036.9	-6193.89	14649.09	35492.08	56335.06	77178.05	98021.03	118864
28	Year1cost	104667.2	87373.41	70079.63	52785.86	35492.08	18198.3	904.523	-16389.3	-33683
29	intrate	41110.49	39577.06	38133.97	36774.38	35492.08	34281.39	33137.14	32054.62	31029.49
30	costgrowth	41685.97	40159.19	38618.02	37062.35	35492.08	33907.1	32307.32	30692.62	29062.91
31	pricegrowth	30984.39	32101.76	33225.48	34355.58	35492.08	36635.01	37784.4	38940.28	40102.67

Рис. 93.2. Таблица данных для анализа чувствительности

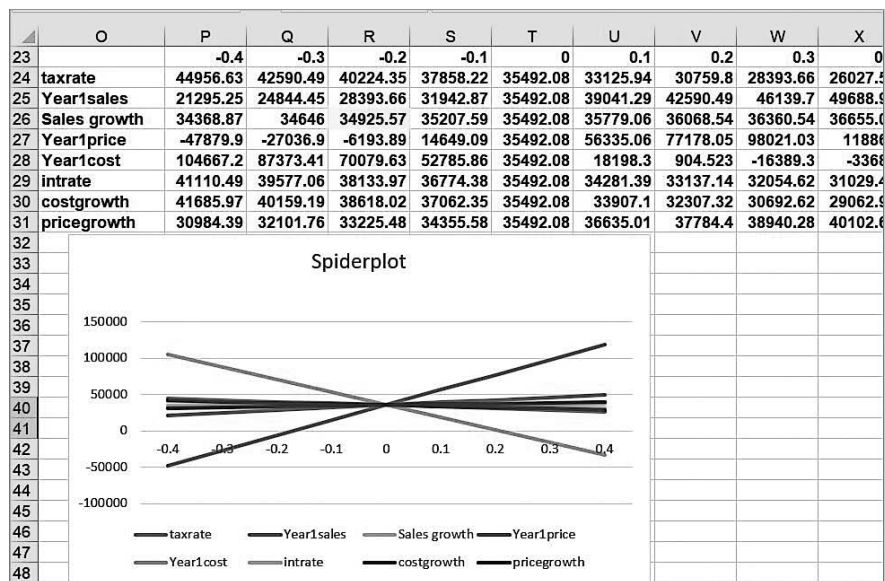


Рис. 93.3. «Паукограмма»

Таблица данных с ячейкой **E4** для подстановки значений по строкам и **E3** для подстановки значений по столбцам отражает изменение ЧПС при изменении каждого исходного параметра в диапазоне 40% от его базового значения.

Скопировав результаты из таблицы данных в диапазон **P23:X31** и имена исходных параметров в диапазон **O24:O31**, мы можем построить линейную диаграмму, комбинирующую все данные из таблицы, представленную на рис. 93.3 (также называемую «паукограммой», spiderplot).

## Задания

1. Измените пример, приведенный в главе, чтобы создать «паукограмму», отражающую изменения значения каждого исходного параметра в диапазоне 80% от базового значения (с шагом в 20%).
2. В файле **Customers.xlsx** содержится расчет числа клиентов на конец десятого года на основании трех входных параметров: число клиентов на начало года, ежегодное количество новых клиентов и коэффициент оттока (учитывающий уход постоянных клиентов). Постройте «паукограмму», показывающую, как варьирует число клиентов на конец десятого года при изменении каждого исходного параметра в диапазоне до 40% от его базового уровня.

# Список вопросов

## Стр. Глава 1. Основы моделирования в электронных таблицах

- 22 Как эффективно определить понедельные выплаты всех моих сотрудников?
- 23 Как эффективно определить, сколько пекарня должна каждому из своих поставщиков?
- 26 Как оценить, какое количество посетителей будет иметь новый фитнес-клуб через 10 лет?
- 28 Как правильно задавать порядок выполнения операций в Excel?
- 29 Как определить влияние на прибыль изменения в цене и себестоимости единицы товара для кофейни по соседству?

## Глава 2. Имена диапазонов

- 36 Необходимо вычислить общий объем продаж в штатах Аризона, Калифорния, Монтана, Нью-Йорк и Нью-Джерси. Можно ли для вычисления общего объема продаж воспользоваться формулой  $AZ+CA+MT+NY+NJ$  вместо формулы  $\text{СУММ}(A21:A25)$  и получить правильный ответ?
- 37 Для чего нужна формула  $\text{СРЗНАЧ}(A:A)$ ?
- 38 Чем различаются имена с областью действия «книга» и «лист»?
- 39 Мне начинают нравиться имена диапазонов. Я стал определять имена диапазонов во многих книгах, которые я создаю в офисе. Однако эти имена не появляются в формулах. Как добиться отображения недавно созданных имен диапазонов в ранее созданных формулах?
- 40 Каким образом можно вставить список имен всех диапазонов (и представляемых ими ячеек) в лист?
- 40 Предполагаемый годовой доход вычисляется как кратный прошлогоднему доходу . Может ли формула выглядеть как  $(1+\text{прирост})*\text{предыдущий\_год}$ ?
- 41 Для каждого дня недели даны почасовая оплата и количество отработанных часов. Можно ли вычислить итоговую сумму оплаты за каждый день по формуле  $\text{почасовая\_оплата}*\text{часы}$ ?

## Глава 3. Функции поиска

- 44 Как создать формулу для вычисления налоговых ставок на основе дохода?
- 46 Как посмотреть цену продукта по идентификатору продукта?
- 47 Допустим, что цена продукта изменяется со временем. Известна дата продажи продукта. Как создать формулу для вычисления цены продукта?

## Глава 4. Функция Индекс

- 51 У меня есть список расстояний между различными городами США. Как должна выглядеть функция, возвращающая расстояние между Сиэтлом и Майами?
- 52 Можно ли создать формулу со ссылкой на весь столбец, содержащий сведения о расстоянии между Сиэтлом и другими городами?

**Глава 5. Функция ПОИСКПОЗ**

- 56 Как создать формулу, возвращающую объем продаж какого-либо продукта за определенный месяц, если даны ежемесячные продажи для нескольких продуктов? Например, сколько продукта 2 было продано за июнь?
- 57 Как создать формулу, определяющую бейсболиста с самой высокой зарплатой, если дан список зарплат игроков? А бейсболиста с пятой по величине зарплатой?
- 58 Как создать формулу, определяющую период окупаемости для первоначальных инвестиционных вложений, если известны ежегодные денежные потоки от реализации инвестиционного проекта?

**Глава 6. Текстовые функции и инструмент Мгновенное заполнение**

- 66 У меня есть лист, каждая ячейка которого содержит описание продукта, идентификатор продукта и цену продукта. Как поместить все описания в столбец А, все идентификаторы в столбец В и все цены в столбец С?
- 68 Каждый день я получаю данные о совокупных продажах в США, которые вычисляются в одной ячейке как сумма продаж на востоке, севере и юге. Как мне извлечь продажи на востоке, севере и юге в отдельные ячейки?
- 71 В конце каждого семестра студенты оценивают мою преподавательскую деятельность по шкале от 1 до 7 баллов. Мне известно, сколько студентов выставили тот или иной балл. Как создать гистограмму оценки качества преподавания?
- 71 Числовые данные были загружены из интернета или из базы данных. При попытке использовать их в вычислениях появляется ошибка #ЗНАЧ!. Как решить эту проблему?
- 72 Текстовые функции — вещь хорошая, но есть ли простой способ (без них) извлечь имена и фамилии из данных, создать список адресов электронной почты из списка имен или выполнить другие стандартные операции с текстовыми данными?
- 74 Что такое символы Unicode?
- 75 Чем новая функция ОБЪЕДИНИТЬ лучше старой функции СЦЕПИТЬ (или &)?
- 75 Мне нравятся текстовые функции, а как пользоваться функцией Excel ТЕКСТ?

**Глава 7. Даты и функции работы с датами**

- 81 Когда я ввожу даты в Excel, то часто получаю число, например 37625, а не дату вида 04.01.2003. Что делать?
- 82 Как с помощью формулы автоматически отобразить сегодняшнюю дату?
- 82 Как определить дату через 50 рабочих дней после другой даты? Как исключить праздничные дни?
- 84 Как определить количество рабочих дней между двумя датами?
- 85 На листе Excel введено 500 дат. Какие формулы нужны, чтобы извлечь из каждой даты месяц, год, день месяца и день недели?
- 85 У меня есть значения года, месяца и дня. Существует ли простой способ восстановления фактической даты?
- 85 У меня бизнес по купле-продаже автомобилей. Для некоторых автомобилей известны даты покупки и продажи. Каким образом можно определить, сколько месяцев эти автомобили пробыли на стоянке?
- 86 Как поместить в лист статичную (неизменяемую) дату?

**Глава 8. Оценка инвестиций по чистой приведенной стоимости**

- 90 Что такое чистая приведенная стоимость (ЧПС)?
- 91 Как работать с функцией ЧПС в Excel?

- 92 Как рассчитать ЧПС, если денежные потоки приходят в начале или середине года?
- 92 Как рассчитать ЧПС, если денежные потоки приходят через неравные промежутки времени?

### **Глава 9. Внутренняя ставка доходности**

- 97 Как вычислить внутреннюю ставку доходности денежных потоков?
- 97 Всегда ли проект имеет единственное значение ВСД?
- 99 Существуют ли условия, гарантирующие единственное значение ВСД проекта?
- 99 Если единственное значение ВСД имеют два проекта, как можно использовать значение ВСД для этих проектов?
- 100 Как вычислить ВСД для нерегулярно поступающих денежных потоков?
- 100 Что такое модифицированная внутренняя ставка доходности (МВСД) и как она рассчитывается?

### **Глава 10. Еще несколько финансовых функций Excel**

- 104 Вы покупаете копировальный аппарат. Как лучше поступить: заплатить \$11 000 сразу или выплачивать по \$3000 в год в течение пяти лет?
- 106 В конце каждого года в течение следующих 40 лет я собираюсь инвестировать по \$2000 в год вплоть до выхода на пенсию и зарабатывать 8% в год на своих инвестициях. Какую сумму я получу при выходе на пенсию?
- 108 Я занял \$10 000 на срок 10 месяцев под 8% годовых. Каковы мои ежемесячные платежи? Каковы ежемесячные выплаты по основной сумме и по процентам?
- 111 Я собираюсь занять \$80 000 и вносить ежемесячные платежи в течение 10 лет. Максимальный ежемесячный платеж, который я могу себе позволить, составляет \$1000. На какую максимальную процентную ставку я могу рассчитывать?
- 112 Если я возьму займы \$100 000 под 8% годовых и буду вносить платежи в размере \$10 000 в год, через сколько лет я выплачу кредит?
- 112 Я дипломированный бухгалтер и часто работаю со сложными методами расчета амортизации оборудования. Имеются ли в Excel функции для расчета амортизации?

### **Глава 11. Циклические ссылки**

- 119 Я часто получаю в Excel сообщение о циклической ссылке. Означает ли это, что я сделал ошибку?
- 120 Как исправить циклическую ссылку?

### **Глава 12. Функции ЕСЛИ, ЕСЛИОШИБКА, ЕСЛИМН, ВЫБОР и ПЕРЕКЛЮЧ**

- 126 При заказе 500 единиц продукции необходимо заплатить \$3,00 за единицу продукции. При заказе от 501 до 1200 единиц цена за единицу падает до \$2,70. При заказе от 1201 до 2000 единиц стоимость единицы продукции составляет \$2,30. При заказе свыше 2000 единиц — \$2,00. Как написать формулу, выражающую затраты на приобретение как функцию количества приобретенных единиц продукции?
- 127 Я купил 100 акций по цене \$55 за акцию. Для хеджирования риска снижения стоимости акций я купил 60 шестимесячных европейских пут-опционов. Каждый опцион имеет цену исполнения \$45 и стоит \$5 долларов. Как разработать таблицу, в которой указана процентная доходность моего инвестиционного портфеля через шесть месяцев для ряда возможных будущих цен?
- 129 Многие аналитики фондового рынка считают что торговые правила для торговли по скользящим средним могут опережать рынок. Общее рекомендуемое правило для торговли по скользящим средним гласит: покупай, когда цена акции превысит средний показатель за последние 15 меся-

цев, и продавай, когда цена акции упадет ниже среднего показателя за последние 15 месяцев. Как это торговое правило может быть сформулировано для биржевого индекса Standard&Poor's 500 (S&P)?

- 132 В игре в кости брошены две кубика. Если сумма чисел при первом броске равна 2, 3 или 12 — вы проиграли. Если сумма чисел при первом броске равна 7 или 11 — вы выиграли. Во всех остальных случаях игра продолжается. Как создать формулу, определяющую состояние игры после первого броска?
- 132 В большинстве гипотетических финансовых отчетов в качестве средства уравнивания активов и обязательств используются денежные средства. Я знаю, что более реалистично использовать для этого долгосрочные обязательства. Как настроить гипотетический отчет с долгосрочными обязательствами в качестве средства уравнивания?
- 137 При копировании формулы с функцией ВПР для вычисления зарплаты сотрудников я получаю множество сообщений об ошибке #Н/Д. Затем при расчете средней зарплаты сотрудников я не могу получить числовой ответ из-за ошибки #Н/Д. Можно ли заменить ошибки #Н/Д пробелом и рассчитать, таким образом, среднюю зарплату?
- 140 Лист содержит ежеквартальные доходы компании Walmart. Как вычислить доход за каждый год и внести его в строку, содержащую продажи в первом квартале соответствующего года?
- 141 Формула с функцией ЕСЛИ может оказаться довольно длинной. Сколько вложенных функций ЕСЛИ можно поместить в ячейку? Каково максимально допустимое количество символов в формуле Excel?
- 141 Вложенные функции ЕСЛИ могут оказаться слишком длинными и сложными. Как новая функция ЕСЛИМН облегчает написание вложенных функций ЕСЛИ?
- 142 Как работает функция ВЫБОР?
- 144 Как работает новая функция ПЕРЕКЛЮЧ?

### **Глава 13. Время и функции времени**

- 151 Как указать время в Excel?
- 152 Как указать время и дату в одной ячейке?
- 152 Как в Excel обрабатываются значения времени?
- 153 Как отобразить на листе текущее время?
- 154 Как создать значения времени с помощью функции ВРЕМЯ?
- 154 Как использовать функцию ВРЕМЗНАЧ для преобразования строки из текстового формата в формат времени?
- 154 Как извлечь часы, минуты и секунды из заданного значения времени?
- 154 Как определить, сколько часов отработал сотрудник, если известно время начала и окончания работы?
- 155 Я суммирую общее время, отработанное сотрудником, но у меня никогда не получается больше 24 часов. Где ошибка?
- 155 Как создать последовательность равномерно распределенных временных интервалов?
- 156 Как поместить в книгу статичное временное значение?

### **Глава 14. Команда Специальная вставка**

- 158 Как переместить результаты расчетов (но не формулы) в другую часть листа?
- 160 В столбце имеется список имен. Как поместить список в строку вместо столбца?
- 161 Я скачал с сайта в Excel процентные ставки по казначейским векселям США. В полученных данных отображается 5, если процентная ставка составляет 5%, 8 — если процентная ставка состав-

ляет 8%, и т. д. Как быстро разделить все результаты на 100, чтобы, например, пятипроцентная ставка отображалась как 0,05?

### **Глава 15. Трехмерные формулы и гиперссылки**

- 165 Существует ли простой способ создания многостраничной книги, в которой все листы имеют одинаковую структуру? Можно ли создать формулы со ссылками на ячейки, расположенные на разных листах?
- 167 Я работаю с 200-страничной книгой. Как организовать простую навигацию по страницам?

### **Глава 16. Инструменты проверки зависимостей и надстройка Inquire**

- 173 Я только что получил лист с 5000 строк, на котором вычисляется чистая приведенная стоимость (ЧПС) нового автомобиля. На этом листе мой финансовый аналитик сделал предположение о годовом проценте роста цены продукта. Каких ячеек на листе касается это предположение?
- 175 Мне кажется, что мой финансовый аналитик сделал ошибку при вычислении прибыли до уплаты налогов за первый год. Какие ячейки на модели листа использовались в этом расчете?
- 177 Как инструменты проверки зависимостей работают с данными на нескольких листах или в нескольких книгах?
- 178 Что такое надстройка Inquire и как ее установить?
- 178 Как применить надстройку Inquire для сравнения книг?
- 179 Как применить надстройку Inquire для анализа структуры книги?
- 179 Как применить надстройку Inquire для анализа связей между листами и книгами?
- 180 Как применить надстройку Inquire для анализа влияющих и зависимых ячеек для конкретной ячейки?
- 180 Как с помощью Inquire очистить ячейку от избыточного форматирования?

### **Глава 17. Анализ чувствительности с помощью таблиц данных**

- 183 Я подумываю открыть магазин в местном торговом центре и продавать натуральный лимонад. Я хотел бы знать до открытия магазина, как прибыль, доход и переменные затраты будут зависеть от цены и себестоимости единицы продукции.
- 187 Я собираюсь построить новый дом. Сумма кредита (с периодом погашения 15 лет) зависит от цены, по которой я продам свой нынешний дом. Я также не уверен в годовой процентной ставке, которую я получу по завершении процесса продажи. Можно ли определить, как ежемесячные платежи будут зависеть от суммы кредита и годовой процентной ставки?
- 188 Крупная интернет-компания подумывает о покупке еще одного магазина. Текущий годовой доход от этого магазина составляет \$100 млн, а расходы — \$150 млн. Текущие прогнозы показывают, что доходы магазина ежегодно растут на 25%, а расходы — на 5%. Известно, однако, что прогнозы могут быть ошибочными, и хотелось бы знать с учетом различных предположений о годовом росте доходов и расходов, через сколько лет магазин начнет приносить прибыль.
- 190 Как создать диаграмму на основе таблицы данных?

### **Глава 18. Инструмент Подбор параметра**

- 196 Сколько стаканов лимонада в год по определенной цене должен продать магазин для достижения безубыточности?
- 196 Мы собираемся выплатить ипотеку за 15 лет. Годовая процентная ставка составляет 6%. Банк сообщил, что максимальный ежемесячный платеж для нас составляет \$2000. Какую сумму мы можем взять в кредит?



- 197 Мне всегда плохо давались сюжетные задачи по алгебре в средней школе. Может ли Excel упростить решение таких задач?

### **Глава 19. Анализ чувствительности с помощью Диспетчера сценариев**

- 201 Я хочу создать для компании наиболее благоприятный, наименее благоприятный и наиболее вероятный сценарии продаж модели автомобиля, изменяя значения объема продаж за год 1, годового роста продаж и розничной цены в год 1. В таблице данных для анализа чувствительности можно изменять только один или два исходных параметра, поэтому для решения этой задачи таблица данных не подходит. Существует ли в Excel другой инструмент, который при анализе чувствительности позволяет изменять более двух параметров?

### **Глава 20. Функции СЧЁТЕСЛИ, СЧЁТЕСЛИМН, СЧЁТ, СЧЁТЗ и СЧИТАТЬПУСТОТЫ**

- 209 Сколько песен исполнил каждый певец?
- 210 Сколько песен было исполнено не Эминемом?
- 210 Сколько песен длились не меньше четырех минут?
- 210 Сколько песен длились дольше, чем средняя продолжительность звучания песен из списка?
- 211 Сколько песен было исполнено певцами, фамилии которых начинаются с S?
- 211 Сколько песен было исполнено певцами, фамилии которых состоят из шести букв?
- 211 Сколько песен прозвучало после 15 июня 2005 г.?
- 211 Сколько песен прозвучало до 2009 г.?
- 211 Сколько было исполнено песен, длящихся ровно четыре минуты?
- 211 Сколько песен, длящихся ровно четыре минуты, было исполнено Брюсом Спрингстином?
- 212 Сколько песен, длящихся от трех до четырех минут, было исполнено Мадонной?
- 212 Как подсчитать в диапазоне количество ячеек, содержащих числа?
- 212 Как подсчитать количество пустых ячеек в диапазоне?
- 212 Как подсчитать количество непустых ячеек в диапазоне?

### **Глава 21. Функции СУММЕСЛИ, СРЗНАЧЕСЛИ, СУММЕСЛИМН, СРЗНАЧЕСЛИМН, МАКСЕСЛИ и МИНЕСЛИ**

- 217 Каков объем продаж в долларах для каждого продавца?
- 217 Сколько единиц товара было возвращено?
- 218 Каков объем продаж в долларах начиная с 2005 г.?
- 218 Сколько штук блеска для губ было продано? Какова выручка от продажи блеска для губ?
- 219 Каков объем продаж в долларах для всех продавцов, за исключением Джен?
- 219 Каково среднее количество единиц проданного товара по всем сделкам, совершенным конкретным продавцом?
- 219 Каков объем продаж помады в долларах у Джен?
- 219 Каково среднее количество проданной помады (в единицах товара) по всем продажам, совершенным Зарет?
- 219 Каково среднее количество проданной помады (в единицах товара) по всем сделкам, совершенным Зарет, для сделок по крайней мере с 50 единицами товара?
- 219 Каков объем продаж помады в долларах по всем сделкам, совершенным конкретным продавцом, для сделок на сумму свыше \$100? Каков объем продаж в долларах для сделок на сумму менее \$100?
- 220 Может ли Excel искать максимумы и минимумы по условию?

**Глава 22. Функция СМЕЩ**

- 224 Как создать ссылку на диапазон ячеек, который отстоит от ячейки или другого диапазона на заданное количество строк и столбцов?
- 225 Как можно выполнить операцию поиска при помощи крайнего правого столбца таблицы вместо крайнего левого?
- 226 Я часто скачиваю данные о продажах программных продуктов, сгруппированные по странам/регионам. Я должен отслеживать для Ирана выручку, затраты и количество проданных программных продуктов, но данные по Ирану не всегда находятся в одной и той же части листа. Можно ли создать формулу, в которой выручка, затраты и количество проданных программных продуктов выбираются всегда правильно?
- 227 Каждый разрабатываемый компанией препарат проходит три этапа разработки. У меня есть список затрат по месяцам на каждый препарат и я знаю продолжительность в месяцах каждого этапа разработки. Можно ли создать формулу, вычисляющую общие затраты для каждого препарата на каждом этапе разработки?
- 228 Я владелец небольшого видеопроката. На рабочем листе мой бухгалтер указал названия всех фильмов и количество копий на складе. К сожалению, он объединил информацию для каждого фильма в одной ячейке. Как мне переписать данные о количестве копий на складе для каждого фильма в отдельную ячейку?
- 230 Как работает в Excel инструмент Вычислить формулу?
- 231 Как написать формулу, всегда возвращающую последнее число в столбце?
- 231 Как создать диапазон, автоматически включающий новые данные?
- 233 Я ежемесячно составляю диаграммы по объемам продаж в штуках для продукта компании. Каждый месяц я загружаю новые данные. Хорошо бы диаграммы обновлялись автоматически. Есть ли простой способ добиться этого?

**Глава 23. Функция ДВССЫЛ**

- 239 Мои формулы на листе часто содержат ссылки на ячейки, или диапазоны ячеек, или на то и другое. Вместо замены ссылок в формулах можно ли поместить эти ссылки в собственные ячейки и быстро заменять в них ссылки без изменения базовых формул?
- 240 На каждом листе книги в ячейке D1 записываются ежемесячные продажи продукта. Можно ли написать и скопировать формулу для записи продаж продукта за каждый месяц на одном листе?
- 241 Предположим, что я суммирую значения в диапазоне A5:A10 по формуле =СУММ(A5:A10). Если я вставляю пустую строку где-либо между строками 5 и 10, моя формула будет автоматически обновлена до =СУММ(A5:A11). Каким образом следует составить формулу, которая при вставке пустой строки между строками 5 и 10 по-прежнему суммировала бы стоимости в исходном диапазоне A5:A10?
- 243 Как с помощью функции ДВССЫЛ обеспечить интерпретацию части формулы в качестве имени диапазона?
- 244 Книга содержит продажи по каждому продукту компании в разных странах и регионах, и каждый континент отображен на отдельном листе. Как объединить эти данные на одном листе?
- 245 Как составить список всех листов в книге?
- 246 Моя книга содержит множество листов. Можно ли легко составить оглавление с гиперссылками так, чтобы была связь со всеми листами в книге?

**Глава 24. Условное форматирование**

- 251 Как можно визуально отобразить, связаны ли последние температурные данные с глобальным потеплением?

- 253 Как работают правила выделения ячеек в условном форматировании?
- 254 Как можно проверить и настроить созданные правила?
- 256 Как работают гистограммы?
- 259 Как работают цветовые шкалы?
- 261 Как работают наборы значков?
- 264 Как применить цветовой код к ежемесячной доходности акций для отображения каждого удачного месяца одним цветом, а каждого неудачного — другим?
- 267 При поступлении данных о ежеквартальных доходах корпорации как можно выделить одним цветом кварталы, в которых доходы увеличились по сравнению с предыдущим кварталом, и другим цветом кварталы, в которых доходы снизились по сравнению с предыдущим кварталом?
- 270 Как в заданном списке дат выделить цветом даты, приходящиеся на выходные дни?
- 272 Наш тренер по баскетболу присвоил каждому игроку рейтинг от 1 до 10 на основе способностей игрока к игре в защите, нападении или в качестве центрового. Можно ли создать лист, визуально демонстрирующий способность каждого игрока играть на позиции, на которую его поставили?
- 273 Для чего нужны флажки «Остановить, если истина» (Stop If True) в диалоговом окне «Управление правилами» (Manage Rules)?
- 275 Как с помощью инструмента Формат по образцу скопировать условный формат?

## Глава 25. Сортировка в Excel

- 281 Как можно отсортировать данные торговых сделок сначала по продавцам, потом по продуктам, по количеству проданных единиц товара и в хронологическом порядке от самых старых до самых новых?
- 285 Мне всегда хотелось отсортировать данные на основе цвета ячеек или цвета шрифта. Возможно ли это в Excel 2016?
- 286 Мне нравятся большие наборы значков, описанные в главе 24. Можно ли отсортировать данные на основе значков в ячейке?
- 286 На листе имеется столбец, в котором указан месяц для каждой продажи. При сортировке по этому столбцу я получаю первым или первый месяц по алфавиту, или последний. Как можно отсортировать данные по этому столбцу в хронологическом порядке, то есть сначала получить сделки за январь, затем за февраль и т. д.?
- 287 Можно ли отсортировать данные без диалогового окна Сортировка (Sort)?
- 288 Мне часто требуется напечатать список городов, в которых моя компания имеет служебные подразделения. Имеется ли возможность создать свой список, а потом при вводе только первого города из списка и перетаскивании курсора вниз автоматически заполнить ячейки остальными городами из этого списка?

## Глава 26. Таблицы

- 291 Я указал на листе количество проданных единиц товара и общую выручку для каждого продавца, и теперь легко могу вычислить среднюю цену единицы товара для каждого продавца. Как создать формат, который при вводе новых данных копировался бы автоматически? Кроме того, существует ли простой способ автоматически скопировать формулы при добавлении новых данных?
- 296 Я ввел в книгу цены на природный газ за несколько лет и создал подходящую линейную диаграмму, отображающую ежемесячное изменение цен. Можно ли добиться, чтобы при добавлении новой информации о ценах диаграмма обновлялась автоматически?
- 297 Для каждой торговой сделки указаны продавец, дата, продукт, место продажи и объем продажи. Можно ли подсчитать, например, общий объем продаж помады на Востоке для Джен или Коллин?

- 300 Как срезы (добавлены в Microsoft Excel 2013) позволяют выполнить срез данных в таблице Excel?
- 302 Как сослаться на фрагмент таблицы в других частях листа?
- 304 Применяются ли условные форматы автоматически к новым данным, добавляемым в таблицу?

### **Глава 27. Счетчики, полосы прокрутки, переключатели, флажки, группы и поля со списками**

- 308 Мне необходимо запустить анализ чувствительности с множеством ключевых исходных данных, таких как объем продаж за первый год, годовой рост продаж, цена реализации в первый год и себестоимость единицы товара. Каким образом можно быстро изменить эти исходные данные и сразу отследить влияние такого изменения, например, на расчет ЧПС?
- 313 Как создать простой флажок, с помощью которого можно было бы включать и отключать условное форматирование?
- 316 Как создать лист, который позволит персоналу, занятому в цепочке поставок, одним нажатием кнопки назначить высокую, низкую или среднюю цену продукта?
- 317 Как предоставить пользователю простой способ ввода дня недели без необходимости набирать какой-либо текст?

### **Глава 28. Революция в аналитике**

- 320 Что такое аналитика?
- 320 Что такое прогнозная аналитика?
- 321 Что такое предписывающая аналитика?
- 322 Почему значение аналитики возрастает?
- 322 Насколько важна аналитика для вашей организации?
- 323 Что необходимо знать для проведения анализа?
- 323 Какие трудности могут возникнуть при осуществлении анализа?
- 325 Какие тенденции будут влиять на развитие аналитики?

### **Глава 30. Поиск решения при определении оптимального ассортимента продукции**

- 331 Как определить ежемесячный ассортимент продукции, максимизирующий прибыльность?
- 340 Всегда ли модель поиска решения имеет решение?
- 341 Что означает выдаваемый моделью поиска решения ответ «Значения ячейки целевой функции не сходятся»?

### **Глава 31. Поиск решения при планировании расписания работы сотрудников**

- 344 Каким спланировать рабочее расписание сотрудников для удовлетворения потребности в рабочей силе?

### **Глава 32. Поиск решения для задач транспортировки и распределения**

- 349 Как фармацевтическая компания может определить, где следует производить препараты и откуда их лучше отгружать заказчикам?

### **Глава 33. Поиск решения для бюджетирования капиталовложений**

- 355 Как с помощью инструмента Поиск решения компания может определить, какой из проектов ей следует принять к реализации?

**Глава 34. Поиск решения при финансовом планировании**

- 364 Можно ли с помощью инструмента Поиск решения проверить точность функции ПЛТ или определить ипотечные платежи с плавающей процентной ставкой?
- 366 Можно ли с помощью инструмента Поиск решения определить, сколько денег нужно отложить до выхода на пенсию?

**Глава 35. Поиск решения при оценке спортивных команд**

- 371 Как можно с помощью Excel распределить очки для команд НФЛ?

**Глава 36. Расположение складов по методу ОПГ с несколькими начальными точками и согласно эволюционному поиску решения**

- 381 В каком месте в США транспортная интернет-компания должна расположить единственный склад для минимизации суммарной протяженности транспортировки пакетов?
- 383 В каком месте в США транспортная интернет-компания должна расположить два склада для минимизации суммарной протяженности транспортировки пакетов?

**Глава 37. Штрафы и эволюционный поиск решения**

- 386 Что является ключом к успешному использованию эволюционного поиска решения?
- 387 Каким образом с помощью эволюционного поиска решения можно распределить 80 сотрудников финансового отдела Microsoft по четырем рабочим группам?

**Глава 38. Задача коммивояжера**

- 393 Как с помощью Excel решать задачи упорядочения?
- 394 Как с помощью Excel решить задачу коммивояжера?

**Глава 39. Импорт данных из текстового файла или документа**

- 398 Как импортировать данные из текстового файла в Excel для анализа?

**Глава 40. Инструмент Получить и преобразовать данные**

- 405 Как загрузить актуальный курс биткойна и сделать так, чтобы эти данные обновлялись каждый день?
- 409 Как загрузить цифры текущего населения городов США?

**Глава 41. Типы данных «Акции» и «География»**

- 411 Какие географические сведения можно использовать в формулах Excel?
- 414 Какие показатели деятельности корпораций можно использовать в формулах Excel?

**Глава 42. Проверка достоверности данных**

- 416 Я ввожу результаты профессиональных баскетбольных матчей в Excel. Я знаю, что команда набирает от 50 до 200 очков за игру. Один раз я ввел 1000 очков вместо 100, что привело к неправильным результатам анализа. Существует ли в Excel способ предотвращения ошибок такого типа?
- 418 Я ввожу даты и суммы деловых расходов на следующий год. В начале года я часто по ошибке ввожу в поле даты предыдущий год. Могут ли настройки Excel предотвратить ошибки такого типа?
- 419 Я ввожу длинный список чисел. Могу ли я получить в Excel предупреждение, если введу нечисловое значение?

- 420 Мой помощник при вводе многочисленных торговых сделок должен вводить сокращенные названия штатов. Можно ли создать список сокращенных названий штатов для сведения к минимуму ошибок ввода неправильных сокращений?

### **Глава 43. Обобщение данных на гистограммах и диаграммах Парето**

- 425 Говорят, одна картина стоит тысячи слов. Как с помощью Excel можно создать изображение (называемое гистограммой), которое обобщает значения в наборе данных?
- 432 Каковы самые распространенные типы гистограмм?
- 434 Что можно узнать, сравнивая гистограммы для разных наборов данных?
- 435 Как создать диаграмму Парето?

### **Глава 44. Обобщение данных с помощью описательной статистики**

- 440 Как определяется типичное значение набора данных?
- 444 Как измерить, насколько набор данных отклонился от своего типичного значения?
- 445 Как среднее значение и стандартное отклонение характеризуют набор данных?
- 447 Как с помощью описательной статистики сравнить наборы данных?
- 447 Как для заданной точки данных можно найти величину ее процентиля в наборе данных? Например, как найти девяностый перцентиль набора данных?
- 449 Как найти второе наибольшее или второе наименьшее число в наборе данных?
- 450 Как можно ранжировать числа в наборе данных?
- 450 Что такое урезанное среднее набора данных?
- 451 Существует ли быстрый способ получить для выбранного диапазона ячеек различные статистики, описывающие данные в этих ячейках?
- 451 Почему финансовые аналитики часто определяют средний доход по акциям с помощью среднего геометрического?
- 453 Как использовать ящики с усами (boxplots) для характеристики и сравнения наборов данных?

### **Глава 45. Сводные таблицы и срезы для описания данных**

- 461 Что такое сводная таблица?
- 462 Как с помощью сводной таблицы обобщить данные о продажах товаров в нескольких продуктовых магазинах?
- 466 Какие макеты сводных таблиц доступны в Microsoft Excel 2019?
- 467 Почему сводные таблицы называются сводными?
- 468 Как быстро изменить формат в сводной таблице?
- 468 Как свернуть и развернуть поля?
- 469 Как сортировать и фильтровать поля сводной таблицы?
- 473 Как обобщить данные в сводной таблице с помощью сводной диаграммы?
- 474 Для чего предназначена область ФИЛЬТРЫ в сводной таблице?
- 474 Как работают срезы сводной таблицы?
- 475 Как добавить пустые строки или скрыть промежуточные итоги в сводной таблице?
- 476 Как применить к сводной таблице Условное форматирование?
- 477 Как обновить вычисления при добавлении новых данных?

- 478 Я работаю в небольшой турфирме и должен разослать проспект о ней потенциальным клиентам. Мой бюджет ограничен, поэтому послать проспект необходимо тем людям, которые больше остальных тратят денег на путешествия. У меня есть случайная выборка из 925 человек, о которых известна следующая информация: возраст, пол и количество денег, потраченных на путешествия за прошлый год. Как с помощью этих данных определить влияние пола и возраста на расходы на путешествия? Какие выводы можно сделать о типе потенциального клиента, которому следует отправить проспект по почте?
- 483 Я провожу маркетинговое исследование автомобилей-универсалов. Мне необходимо определить, какие факторы влияют на вероятность покупки семьей автомобиля-универсала. У меня есть большая выборка семей с информацией об их размере (большая или нет) и доходе (высокий или низкий). Как определить влияние размера семьи и семейного дохода на вероятность покупки автомобиля-универсала?
- 485 Я работаю в компании, продающей микрочипы по всему миру. Ежемесячно я получаю данные о фактических и прогнозируемых объемах продаж чипа 1, чипа 2 и чипа 3 в Канаде, Франции и США. Я также получаю расхождение, или разницу, между фактической выручкой и выручкой, заложенной в бюджет. Для каждого месяца и каждой комбинации страны и продукта мне нужно отобразить следующие данные: фактическая выручка, выручка, заложенная в бюджет, фактическое расхождение, фактическая выручка как процент от годового дохода и расхождение как процент от выручки, заложенной в бюджет. Каким образом можно отобразить эту информацию?
- 488 Что такое вычисляемое поле?
- 490 Что использовать: фильтр отчета или срез?
- 491 Как сгруппировать элементы в сводной таблице?
- 492 Что такое вычисляемый объект?
- 494 Что такое детализация?
- 494 Мне часто приходится использовать конкретные данные в сводной таблице для вычисления прибыли, например апрельских продаж чипа 1 во Франции. К сожалению, при добавлении новых данных в сводную таблицу эти перемещаются. Существует ли в Excel функция, позволяющая всегда извлекать из сводной таблицы объем апрельских продаж чипа 1 во Франции?
- 497 Как применить временную шкалу для обобщения данных за разные периоды времени?
- 498 Как с помощью сводной таблицы получить нарастающий итог общего объема продаж для каждого периода года?
- 499 Как с помощью сводной таблицы сравнить объем продаж за месяц с объемом продаж за тот же месяц предыдущего года?
- 500 Как создать сводную таблицу на основе данных из нескольких местоположений?
- 503 Как создать сводную таблицу на основе уже имеющейся сводной таблицы?
- 504 Как использовать фильтр отчета для создания нескольких сводных таблиц?
- 505 Как быстро изменить настройки для сводной таблицы по умолчанию?

## Глава 46. Модель данных

- 509 Что представляет собой инструмент Модель данных и зачем он нужен?
- 510 Как добавить данные в модель данных?
- 511 Как удалить данные из модели данных?
- 511 Как создать связь в модели данных?
- 512 Как с помощью модели данных создать сводную таблицу?
- 513 Как добавить новые данные в модель данных?

- 513 Как с помощью модели данных создать новую сводную таблицу?
- 514 Как изменять и удалять связи?
- 514 Как работает функция Число различных элементов (DISTINCT COUNT)?

#### **Глава 47. Power Pivot**

- 518 Как загрузить данные в Power Pivot?
- 522 Как с помощью Power Pivot создать сводную таблицу?
- 523 Как использовать срезы с Power Pivot?
- 524 Что такое функции DAX и вычисляемые столбцы?
- 525 Как работает функция RELATED?
- 530 Как работает функция CALCULATE и что такое вычисляемая мера?

#### **Глава 48. 2D-картограммы и 3D-карты Power Map**

- 532 Как создать двумерную картограмму?
- 536 Как создать 3D-карту Power Map?
- 538 Как применить фильтр к карте Power Map?
- 538 Как изменить способ отображения данных на карте Power Map?
- 539 Как с помощью временной шкалы анимировать карту Power Map?
- 541 Можно ли обобщить данные с карты Power Map двумерной диаграммой?
- 542 Как создать карту Power Map, содержащую круговые диаграммы с подписями?
- 542 Как убедиться, что на карте Power Map правильно отображены геоточки?

#### **Глава 49. Спарклайны**

- 544 Как графически отобразить ежедневные расчеты с клиентами для каждого филиала банка в одной ячейке?
- 545 Как изменять спарклайны?
- 547 Как обобщить серию из побед и поражений команд НФЛ в одной ячейке?
- 548 Обновляются ли спарклайны автоматически при добавлении новых данных?

#### **Глава 50. Обработка данных с помощью статистических функций для баз данных**

- 551 На какую сумму продала блеск для губ Джен?
- 552 Каково среднее количество проданной помады для каждой сделки Джен в регионе Восток?
- 552 На какую общую сумму было продано косметики либо продавцом Эмили, либо в регионе Восток?
- 552 На какую сумму продали помаду Коллин и Зарет в регионе Восток?
- 552 Сколько сделок с помадой было совершено не в регионе Восток?
- 553 На какую сумму продала помаду Джен в 2004 г.?
- 553 Сколько единиц косметики было продано по цене не менее \$3,20?
- 554 Какова общая сумма выручки для каждого косметического товара, реализованного каждым продавцом?
- 555 Какие приемы применяются при создании диапазона условий?



- 555 У меня есть база со следующими данными для каждой сделки: выручка, дата продажи и идентификационный код продукта. Как быстро извлечь сумму выручки для сделки по заданной дате продажи и идентификационному коду?

### **Глава 51. Фильтрация данных и удаление дубликатов**

- 560 Как определить все сделки Джен по продаже помады в регионе Восток?
- 561 Как определить все сделки Кики и Коллин по продаже помады и туши в регионе Восток или Юг?
- 562 Как скопировать все сделки Кики и Коллин по продаже помады и туши в регионе Восток или Юг на другой лист?
- 562 Как очистить фильтры в столбце или базе данных?
- 562 Как определить все сделки по продаже более 90 единиц продукции и более чем на \$280?
- 564 Как определить все сделки, заключенные в 2005 или 2006 г.?
- 566 Как определить все сделки, заключенные в последние три месяца 2005 г. и в первые три месяца 2006 г.?
- 566 Как определить все сделки, заключенные продавцами, имена которых начинаются с С?
- 567 Как определить все сделки, для которых ячейка с названием продукта окрашена в красный цвет?
- 567 Как определить все сделки среди 30 сделок с максимальной выручкой, в которых продавцами были Халлаган или Джен?
- 568 Как быстро получить полный список продавцов?
- 569 Как просмотреть все комбинации «продавец — продукт — регион», встречающиеся в базе данных?
- 570 Если данные изменились, как повторно применить тот же самый фильтр?
- 570 Как извлечь все сделки по продаже основы под макияж, заключенные Эмили и Джен за первые шесть месяцев 2005 г., в которых средняя цена единицы продукции составила больше \$3,20?

### **Глава 52. Консолидация данных**

- 573 Компания продает продукцию в нескольких регионах США. В каждом регионе хранятся записи о количестве проданных единиц продукции в январе, феврале и марте. Существует ли простой способ создания мастер-книги, в которой постоянно объединяются продажи для всех регионов и проводится подсчет общей суммы выручки для каждого продукта, проданного в США в течение каждого месяца?

### **Глава 53. Создание промежуточных итогов**

- 578 Существует ли быстрый способ настроить лист для вычисления общей выручки и количества проданных единиц продукции по регионам?
- 581 Как можно при этом получить разбивку продаж по продавцам в каждом регионе?

### **Глава 54. Приемы работы с диаграммами**

- 584 Как создать комбинированную диаграмму?
- 586 Как создать вспомогательную ось?
- 587 Как обработать недостающие данные?
- 588 Как показать скрытые данные на диаграмме?
- 589 Как с помощью картинок сделать гистограмму интереснее?
- 591 Я отобразил данные о ежегодных продажах в виде гистограммы, но годы как метки столбиков не отображаются. Где ошибка?
- 592 Как включить метки данных и таблицу данных в диаграмму?

- 593 Как поместить в диаграмму метки данных на основе содержимого ячеек?
- 594 Как отследить эффективность отдела сбыта по продажам за определенный период?
- 596 Как создать ленточную диаграмму для проверки наличия запасов?
- 597 Как сохранить диаграмму в качестве шаблона?
- 597 Как с помощью диаграммы «термометр» отобразить прогресс?
- 598 Как создать на диаграмме динамические метки?
- 600 Как с помощью флажков указать, какие ряды данных требуется отобразить на диаграмме?
- 601 Как с помощью списка обеспечить выбор ряда данных для отображения на диаграмме?
- 602 Как создать диаграмму Ганта?
- 603 Как создать диаграмму на основе отсортированных данных?
- 604 Как создать гистограмму, которая при добавлении новых данных обновлялась бы автоматически?
- 605 Как добавить условные цвета в диаграмму?
- 606 Как с помощью диаграммы «водопад» отследить приближение объемов продаж к плану или разложить компоненты продажной цены?
- 608 Как с помощью функции ПОЛУЧИТЬ.ДААННЫЕ.СВОДНОЙ.ТАБЛИЦЫ и таблицы Excel создать динамические информационные панели?
- 611 Как вставить в диаграмму вертикальную линию для разделения производительности до и после слияния?
- 611 Как с помощью лепестковой диаграммы отобразить различия баскетболистов в силе, скорости и прыгучести?
- 612 Известно, что изменение двух переменных можно представить с помощью точечной диаграммы. А как представить изменение трех переменных с помощью пузырьковой диаграммы?
- 613 Как создать диаграмму «водопад»?
- 616 Как используются для обобщения иерархических данных диаграммы «дерево» и «солнечные лучи»?
- 619 Что такое диаграмма типа «воронка»?
- 619 Как отразить в диаграмме свежую динамику по интересующим меня акциям?

## **Глава 55. Оценка линейных зависимостей**

- 627 Как определить зависимость между ежемесячным производством и ежемесячными эксплуатационными расходами?
- 630 Насколько точно эта зависимость объясняет ежемесячные колебания эксплуатационных расходов предприятия?
- 630 Насколько точны мои прогнозы?
- 631 Какие функции при оценке линейной зависимости позволяют получить угловой коэффициент и начальную ординату прямой, точнее всего соответствующие данным?

## **Глава 56. Моделирование экспоненциального роста**

- 635 Как моделировать рост доходов компании с течением времени?

## **Глава 57. Степенная кривая**

- 640 Моя компания наращивает производство. Можно ли смоделировать зависимость между количеством произведенной продукции и временем, необходимым для производства одной единицы продукции?

**Глава 58. Представление зависимостей с помощью корреляции**

- 647 Каким образом связаны ежемесячные доходы от акций Microsoft, GE, Intel, GM и Cisco?

**Глава 59. Введение во множественную регрессию**

- 652 Завод выпускает три продукта. Как спрогнозировать эксплуатационные расходы завода на основе количества произведенной продукции?
- 657 Насколько точны мои прогнозы ежемесячных расходов на основе количества произведенной продукции?
- 657 Множественную регрессию можно запустить командой Анализ данных. Существует ли способ запустить регрессионный анализ без этой команды и поместить результаты регрессии на тот же лист, где находятся данные?

**Глава 60. Включение качественных факторов во множественную регрессию**

- 659 Как предсказать квартальные продажи автомобилей в США?
- 664 Как предсказать президентские выборы в США?
- 668 Какая функция Excel позволяет составить прогноз по уравнению множественной регрессии?

**Глава 61. Моделирование нелинейных характеристик и взаимосвязей**

- 670 Когда говорят, что независимая переменная оказывает нелинейное воздействие на зависимую переменную, что это означает?
- 671 Когда говорят, что между воздействиями двух независимых переменных на зависимую переменную существует взаимосвязь, что это означает?
- 671 Как выполнить проверку на наличие нелинейности и взаимосвязи в регрессии?

**Глава 62. Однофакторный дисперсионный анализ**

- 679 Владелец издательства, публикующего компьютерную литературу, хочет знать, влияет ли на продажи местоположение стенда с книгами в отделе компьютерной литературы в книжных магазинах. В частности, имеет ли значение, в какой зоне отдела компьютерной литературы (передней, задней или средней) находится стенд с книгами?
- 681 Если я определяю, имеют ли популяции значимо различные средние значения, почему метод называется дисперсионным анализом?
- 682 Как применить результаты однофакторного дисперсионного анализа для прогнозирования?

**Глава 63. Рандомизированные блоки и двухфакторный дисперсионный анализ**

- 685 Я пытаюсь проанализировать эффективность отдела продаж. Проблема в том, что кроме эффективности самого торгового представителя сумма его продаж зависит от региона, в который он назначен. Как можно включить в анализ распределение продавцов по регионам?
- 687 Как предсказать объем продаж на основе информации о торговых представителях и регионах? Насколько точными окажутся прогнозы?
- 688 Как определить, влияют ли на продажи видеониги изменение цены и расходов на рекламу? Как определить, значимо ли взаимодействие цены и рекламы?
- 690 Как интерпретировать влияние цены и рекламы на продажи при отсутствии значимого взаимодействия между ценой и рекламой?

**Глава 64. Скользящие средние для временных рядов**

- 694 Я пытаюсь проанализировать тенденцию роста ежеквартальных доходов Amazon.com с 1996 г. Объем продаж в четвертом квартале в США обычно больше (из-за Рождества), чем объем про-

даж в первом квартале следующего года. Эта закономерность затемняет тенденцию к росту продаж. Как отобразить тенденцию к росту доходов графически?

### **Глава 66. Метод прогнозирования «по отношению к скользящему среднему»**

- 705 Что такое тренд временного ряда?
- 706 Как определить индексы сезонности для временного ряда?
- 706 Существует ли простой способ включить тренд и сезонность в прогнозирование будущих продаж продукта?

### **Глава 67. Прогноз для особых случаев**

- 709 Как определить, влияют ли конкретные факторы на поток клиентов?
- 713 Как оценить точность прогноза?
- 716 Как проверить, являются ли ошибки прогноза случайными?

### **Глава 68. Введение в теорию вероятности**

- 718 Что такое эксперимент, выборочное пространство и событие?
- 719 Каковы аксиомы, которым должны удовлетворять вероятности наступления событий?
- 719 Что такое правило дополнений?
- 720 Что такое взаимоисключающие события?
- 720 Что такое правило сложения вероятностей?
- 721 Что такое независимые события?
- 722 Что такое условная вероятность?
- 722 Что такое формула полной вероятности?
- 723 Что такое теорема Байеса?

### **Глава 69. Введение в случайные величины**

- 728 Что такое случайная величина?
- 728 Что такое дискретная случайная величина?
- 729 Что такое среднее, дисперсия и стандартное отклонение случайной величины?
- 730 Что такое непрерывная случайная величина?
- 730 Что такое функция плотности вероятности?
- 731 Что такое независимые случайные величины?

### **Глава 70. Биномиальные, гипергеометрические и отрицательные биномиальные случайные величины**

- 734 Что такое биномиальная случайная величина?
- 735 Как с помощью функции БИНОМ.РАСП и БИНОМ.РАСП.ДИАП вычислить вероятности биномиального распределения?
- 736 Если одинаковое количество людей предпочитает кока-колу пепси-коле и наоборот, и у 100 человек спросили, предпочитают ли они кока-колу пепси-коле, какова вероятность того, что ровно 60 человек предпочитают кока-колу пепси-коле, и вероятность того, что от 40 до 60 человек предпочитают кока-колу пепси-коле?

- 736 Из всех лифтовых направляющих, производимых компанией, 3% считаются дефектными. Компания собирается отгрузить клиенту партию из 10 000 направляющих. Перед приемкой партии клиент случайным образом отбирает 100 направляющих и проверяет на наличие дефектов. Если дефекты обнаружатся не более чем у двух направляющих, клиент примет партию. Как определить вероятность приемки партии?
- 737 Авиакомпаниям не нравятся рейсы с незанятыми местами. Предположим, что в среднем 95% всех купивших билет регистрируются на рейс. Если авиакомпания продает 105 билетов на 100-местный рейс, какова вероятность того, что произойдет избыточное бронирование рейса?
- 737 В кафе Village Deli ежедневно приходит на ланч 1000 клиентов. В среднем 20% клиентов заказывают особый вегетарианский сэндвич. Эти сэндвичи готовят заранее. Сколько сэндвичей необходимо приготовить, если вероятность их нехватки должна быть равна 5%?
- 738 Что такое гипергеометрическая случайная величина?
- 739 Что такое отрицательная биномиальная случайная величина?

### **Глава 71. Пуассоновская и экспоненциальная случайные величины**

- 742 Что такое пуассоновская случайная величина?
- 743 Как вычислить распределение пуассоновской случайной величины?
- 744 Если количество клиентов, обращающихся в банк, описывается пуассоновской случайной величиной, какая случайная величина определяет время между обращениями?

### **Глава 72. Нормальная случайная величина и Z-оценка**

- 747 Каковы свойства нормальной случайной величины?
- 750 Как в Excel найти распределение нормальной случайной величины?
- 751 Как в Excel найти процентиля для нормальных случайных величин?
- 752 Почему нормальная случайная величина соответствует многим реальным ситуациям?
- 754 Что такое Z-оценка?

### **Глава 73. Распределение Вейбулла и бета-распределение: моделирование надежности механизмов и продолжительности работы**

- 758 Как оценить вероятность безотказной работы станка в течение не менее 20 часов?
- 761 Как оценить вероятность того, что продолжительность работ по монтажу перегородки из гипсокартона в здании составит более 200 часов?

### **Глава 74. Создание вероятностных высказываний на основе прогнозов**

- 765 На основании чего вычисляются прогнозы?
- 766 Фармацевтическая компания планирует продать в следующем году 60 млн единиц препарата. Какова вероятность того, что в следующем году будет продано более 65 млн единиц этого препарата?

### **Глава 75. Логарифмически нормальная случайная величина в моделировании курса акций**

- 768 Что такое логарифмически нормальная случайная величина?
- 769 По какой причине курс акций ведет себя подобно случайной величине с логарифмически нормальным распределением?
- 769 Как смоделировать будущий курс любых акций как логарифмически нормальную случайную величину?

- 771 Как вычислить вероятность того, что курс акций Microsoft превысит \$130 через шесть месяцев?
- 771 Как вычислить вероятность того, что курс акций Microsoft не превысит \$90 через шесть месяцев?
- 771 Как вычислить средний курс акций Microsoft через шесть месяцев?

#### **Глава 76. Импорт в Excel истории торгов (загрузка биржевых данных)**

- 773 Как просто импортировать ретроспективные биржевые данные в Excel?
- 775 Как рассчитать годовую доходность по моему портфелю акций?

#### **Глава 77. Введение в моделирование по методу Монте-Карло**

- 777 Где применяется моделирование по методу Монте-Карло?
- 778 Что происходит при вводе в ячейку формулы =СЛЧИС()?
- 779 Как моделируют значения дискретной случайной величины?
- 780 Как моделируют значения нормальной случайной величины?
- 781 Как компания — производитель поздравительных открыток может определить, сколько открыток следует произвести?

#### **Глава 78. Вычисление оптимальной цены предложения**

- 788 Как смоделировать случайную величину с биномиальным распределением?
- 789 Как определить, должна ли непрерывная случайная величина моделироваться как нормальная случайная величина?
- 790 Как с помощью моделирования определить оптимальную тендерную заявку для проекта строительства?

#### **Глава 79. Моделирование цен на акции и распределения средств между активами**

- 795 Я недавно купил 100 акций компании GE. Какова вероятность того, что в течение следующего года эта инвестиция принесет доход более 10%?
- 798 Я пытаюсь распределить свой инвестиционный портфель между акциями, казначейскими векселями и облигациями. Какое распределение средств между активами в рамках пятилетнего горизонта планирования принесет ожидаемый доход не менее 8% и минимизирует риск?

#### **Глава 80. Развлечения и игры: моделирование вероятностей для азартных игр и спортивных соревнований**

- 804 Какова вероятность выиграть в крэпс?
- 806 Какова вероятность получить «тройку» в пятикарточном дро-покере?
- 808 Какова вероятность выигрыша в Финале Четырех 2018 г. для каждой баскетбольной команды Национальной студенческой спортивной ассоциации?

#### **Глава 81. Анализ данных с помощью повторной выборки**

- 813 Девять партий продукции производятся при высокой температуре, а семь партий — при низкой. Какова вероятность того, что выход продукта выше при высокой температуре?

#### **Глава 82. Ценообразование опционов**

- 818 Что такое колл-опцион и пут-опцион?
- 819 В чем разница между американским и европейским опционами?

- 819 Как выглядят выигрыши по европейским колл- и пут-опционам, если рассматривать их как функцию курса акций на дату исполнения?
- 820 Какими параметрами определяется цена опциона?
- 821 Как оценить волатильность акций на основе исторических данных?
- 822 Как реализовать в Excel формулу Блэка — Шоулза?
- 824 Как изменение ключевых параметров изменяет цену колл- и пут-опционов?
- 825 Как оценить волатильность акций по формуле Блэка — Шоулза?
- 826 В мои точные формулы ценообразования опционов никто не должен вносить изменения. Как защитить формулы на листе от изменений?
- 828 Как с помощью ценообразования опционов компания может принимать лучшие инвестиционные решения?

### **Глава 83. Определение потребительской ценности**

- 833 Компания, работающая с кредитными картами, в настоящее время имеет коэффициент удержания клиентов, равный 80. Насколько возрастет прибыльность компании, если коэффициент удержания клиентов повысится до 90% и выше?
- 835 Междугородная телефонная компания предоставляет клиентам конкурентов бонус при смене компании. Насколько значительным должен быть бонус?

### **Глава 84. Оптимальный размер заказа в модели управления запасами**

- 839 Магазин электроники продает 10 000 смартфонов в год. Каждый раз при покупке смартфонов магазин несет расходы в размере \$10 за закупку. Магазин платит \$100 за каждый смартфон, и стоимость хранения смартфона на складе в течение года составляет \$20. Когда магазин закупает смартфоны, какого размера должна быть закупка?
- 841 Завод производит 10 000 серверов в год. Себестоимость производства каждого сервера составляет \$2000. Стоимость запуска производства партии серверов составляет \$200, а стоимость хранения сервера на складе в течение года — \$500. При необходимости завод может производить 25 000 серверов в год. Каким должен быть размер партии при производстве серверов?

### **Глава 85. Построение моделей управления запасами для неопределенного спроса**

- 845 При каком уровне запасов следует делать закупку, если необходимо минимизировать годовые затраты, связанные с дефицитом, хранением и закупками?
- 847 Что означает термин «95%-ный уровень обслуживания»?

### **Глава 86. Теория массового обслуживания (теория очередей)**

- 851 Какие факторы влияют на длину очереди и время, проведенное клиентами в очереди?
- 852 Какие условия должны выполняться, чтобы можно было анализировать среднее количество присутствующих клиентов или среднее время, проведенное клиентом в очереди?
- 853 Почему изменчивость снижает производительность системы очередей?
- 854 Можно ли быстро определить среднее время прохождения через систему контроля безопасности в аэропорту или среднее время ожидания в очереди в банке?

### **Глава 87. Оценка кривой спроса**

- 858 Что необходимо знать при назначении цены продукта?
- 859 Что такое эластичность спроса?

- 859 Существует ли простой способ оценить кривую спроса?
- 862 Каким образом кривая спроса показывает готовность покупателя платить за продукт?

### **Глава 88. Ценообразование продуктов с сопутствующими товарами**

- 863 Как покупка клиентами вместе с бритвенным станком лезвий способствует установлению максимально увеличивающей прибыль цены на бритву?

### **Глава 89. Ценообразование продуктов с помощью субъективно определяемого спроса**

- 868 Иногда ценовая эластичность продукта неизвестна. Или невозможно определить, какая из кривых спроса, линейная или степенная, является релевантной. Можно ли все равно оценить кривую спроса и воспользоваться инструментом Поиск решения, чтобы определить цену, максимизирующую прибыль?
- 869 Как небольшой аптеке определить цену на помаду, максимизирующую прибыль?

### **Глава 90. Нелинейное ценообразование**

- 874 Что такое линейное ценообразование?
- 874 Что такое нелинейное ценообразование?
- 876 Что такое объединение в набор и как оно может повысить прибыльность?
- 877 Как составить план нелинейного ценообразования с максимальной прибылью?

### **Глава 91. Формулы массива и функции, возвращающие массив**

- 884 Что такое формула массива?
- 885 Как интерпретировать формулы, такие как  $= (D2:D7) * (E2:E7)$  и  $= СУММ(D2:D7 * E2:E7)$ ?
- 886 Имеется список имен в столбце. Имена в списке часто меняются. Существует ли простой способ перенести имена в одну строку так, чтобы изменения в исходном столбце отражались бы в этой новой строке?
- 886 Имеется список ежемесячных доходов по акциям. Существует ли способ определить количество значений доходов, попадающих в интервалы от  $-30$  до  $-20\%$ , от  $-10$  до  $0\%$  и т. д., которые обновлялись бы автоматически при изменении исходных данных?
- 888 Как написать формулу, суммирующую вторые цифры из списка целых чисел?
- 889 Существует ли способ просмотра двух списков имен и поиска имен, встречающихся в обоих списках?
- 890 Как написать формулу, вычисляющую среднее значение для чисел в списке, которые больше или равны медиане списка?
- 890 У небольшой компании декоративной косметики имеется база данных по продажам, в которой указан продавец, количество проданных единиц продукции и сумма в долларах для каждой сделки. Мне известно, что для обобщения этих данных подойдут статистические функции или функции СЧЁТЕСЛИМН (COUNTIFS), СУММЕСЛИМН (SUMIFS) и СРЗНАЧЕСЛИМН (AVERAGEIFS). Но могу ли я использовать также функции, возвращающие массив, чтобы обобщить эти данные и ответить на вопросы типа: сколько единиц косметики продал продавец, или сколько помады было продано, или сколько единиц косметики было продано определенным продавцом?
- 894 Что такое массив констант и для чего он нужен?
- 895 Как изменить формулу массива?
- 895 Как оценить тренд и сезонность выручки для магазина игрушек по заданной квартальной выручке?



- 898 Как вычислить медианный размер сделки в каждой стране по заданному списку сделок в различных странах?
- 899 Имеются данные о торговых сделках с разбивкой по продавцам, продуктам и местоположению. Как определить стандартное отклонение проданных единиц продукции для каждого сочетания «продавец — продукт — регион»?
- 899 Как применить функцию СУММПРОИЗВ (SUMPRODUCT) взамен условного счета, условного суммирования и условного определения среднего наравне с формулами массивов?
- 901 Для скольких сделок фактические доходы оказались меньше, чем заложенные в бюджет (прогнозные)?
- 901 У скольких сделок доход меньше заложенного в бюджет, с разбиением по типу чипа?
- 901 Сколько сделок происходит в конкретных странах?
- 901 Каков общий доход от каждого чипа?
- 901 Рассчитайте общий доход для каждой комбинации страны и чипа.
- 902 Рассчитайте общий доход в каждой стране.

## **Глава 92. Запись макросов**

- 907 Что такое макрос?
- 907 Как используется вкладка Разработчик (Developer) при записи и запуске макросов?
- 908 Как записать макрос?
- 910 Как записать макрос, применяющий требуемый формат к произвольному диапазону ячеек?
- 912 Как запустить макрос?
- 914 Как при записи макросов использовать относительные ссылки?
- 917 Как записать макрос, фильтрующий сводную таблицу так, чтобы отображались только данные о 20 основных клиентах?

## **Глава 93. Продвинутый анализ чувствительности**

- 919 В электронной таблице много исходных параметров, поэтому двунаправленная таблица данных в этом случае неприемлема для анализа чувствительности. Как же мне все-таки провести анализ чувствительности?

*Уэйн Винстон*

**Бизнес-моделирование и анализ данных.  
Решение актуальных задач с помощью Microsoft Excel**

**6-е издание**

Перевела с английского *Ю. Бочина*

Заведующая редакцией	<i>Ю. Сергиенко</i>
Руководитель проекта	<i>Н. Римицан</i>
Ведущий редактор	<i>К. Тульцева</i>
Художественный редактор	<i>В. Мостипан</i>
Корректор	<i>М. Молчанова, Н. Сидорова, Г. Шкатова</i>
Верстка	<i>Л. Егорова</i>

Изготовлено в России. Изготовитель: ООО «Прогресс книга».  
Место нахождения и фактический адрес: 194044, Россия, г. Санкт-Петербург,  
Б. Сампсониевский пр., д. 29А, пом. 52. Тел.: +78127037373.

Дата изготовления: 12.2020. Наименование: книжная продукция. Срок годности: не ограничен.

Налоговая льгота — общероссийский классификатор продукции ОК 034-2014, 58.11.12 — Книги печатные профессиональные, технические и научные.

Импортер в Беларусь: ООО «ПИТЕР М», 220020, РБ, г. Минск, ул. Тимирязева, д. 121/3, к. 214, тел./факс: 208 80 01.

Подписано в печать 25.11.20. Формат 70×100/16. Бумага офсетная. Усл. п. л. 76,110. Тираж 1000. Заказ 0000.