

УДК 004.424
ББК 32.372
К94

Куслейка Д.

К94 Визуализация данных при помощи дашбордов и отчетов в Excel / пер. с англ. А. Ю. Гинько. – М.: ДМК Пресс, 2021. – 338 с.: ил.

ISBN 978-5-97060-966-8

В этой книге поэтапно рассматриваются средства Excel, связанные с визуализацией данных, – от элементарных операций с рабочими листами до построения полноценных диаграмм и дашбордов. Вы научитесь анализировать огромные массивы данных и представлять их в сжатом и понятном виде, обеспечивать их визуальную привлекательность, автоматизировать анализ и отчеты.

Издание будет полезно пользователям Excel любого уровня, менеджерам, специалистам по финансам и бизнес-аналитике, руководителям отделов и всем, кому безразличны данные и их представление.

УДК 004.424
ББК 32.372

All rights reserved. This Translation publish under license with the original publisher John Wiley & Sons, Inc.

Все права защищены. Любая часть этой книги не может быть воспроизведена в какой бы то ни было форме и какими бы то ни было средствами без письменного разрешения владельцев авторских прав.

ISBN 978-1-119-69872-2 (англ.)
ISBN 978-5-97060-966-8 (рус.)

© John Wiley & Sons, Inc., 2021
© Перевод, оформление, издание,
ДМК Пресс, 2021



Содержание

От издательства	11
Об авторе	12
О техническом редакторе	13
Благодарности	14
Введение	15
Часть I Отображение данных на дашборде	18
Глава 1 Основы дашбордов	19
Когда необходимо использовать дашборды	19
Что такое дашборд?	21
Ключевые показатели эффективности	22
Пользовательские требования	22
Типы конечных пользователей	23
Сбор данных	24
Сводные таблицы	24
Функция ПОЛУЧИТЬ_ДААННЫЕ_СВОДНОЙ_ТАБЛИЦЫ (GETPIVOTDATA)	28
Функции рабочего листа	30
Функция ВПР (VLOOKUP)	30
Функция ПРОСМОТРХ (XLOOKUP)	31
Функции ИНДЕКС (INDEX) и ПОИСКПОЗ (MATCH)	32
Функция СУММПРОИЗВ (SUMPRODUCT)	33
Формулы массива	36
Умные таблицы	37
Структурированные ссылки на таблицы	39
Текст по столбцам	41
Удаление дубликатов	43
Построение дашбордов	45
Организация элементов	46
Разнообразие элементов	47

Отображение трендов	48
Форматирование дашбордов	50
Форматирование чисел.....	54
Глава 2 Практические примеры дашбордов	56
Отслеживание процесса	56
Практический пример: отслеживание процесса написания ПО	56
Планирование и составление макета.....	57
Сбор данных.....	59
Построение визуальных элементов	60
Размещение элементов на дашборде.....	72
Отображение ключевых показателей эффективности	73
Практический пример: отдел по работе с персоналом	73
Планирование и составление макета.....	73
Сбор данных.....	74
Построение визуальных элементов	75
Размещение элементов на дашборде.....	86
Отображение финансовых показателей.....	89
Практический пример: финансовая информация и показатели	89
Планирование и составление макета.....	90
Сбор данных.....	90
Построение визуальных элементов	92
Размещение элементов на дашборде.....	101
Глава 3 Организация данных для дашбордов	104
Разделение слоев данных	104
Слой источника	106
Слой подготовки и анализа.....	108
Слой представления	109
Работа с внешними данными	109
Power Query против Power Pivot	109
Текстовые файлы	110
Файлы Excel	115
Базы данных Access	123
Базы данных SQL Server.....	129
Преобразование данных при помощи Power Query	132
Управление столбцами и строками	135
Преобразование столбцов	138
Преобразование типов данных.....	138
Преобразование чисел.....	140
Разделение столбцов	141
Часть II Основы визуализации данных	145
Глава 4 Основы эффективной визуализации	146
Создание эффективной визуализации.....	146
Держитесь в рамках одного экрана	146

Повысьте привлекательность дашборда	148
Расскажите свою историю быстро	148
История должна быть согласована с данными.....	152
Выберите правильный тип диаграммы	152
Управление цветом.....	154
Как использовать цвет	154
Изменение цвета точек данных с изменением значений	155
Использование контраста для выделения данных	155
Группирование данных при помощи цвета	156
Советы по использованию цвета	157
Используйте пустое пространство.....	157
Используйте простую цветовую палитру	159
Используйте цвета, согласующиеся с данными	159
Используйте контраст	159
Используйте неинформационные пиксели при необходимости.....	160
Равнение на текст	160
Шрифты.....	160
Легенды	161
Оси	163
Подписи данных	163
Демонстрация аналитических выводов (инсайтов) на графиках	164
Сравнение	165
Анализ состава	166
Взаимосвязи	168
Глава 5 Неграфические визуализации	170
Основы пользовательских числовых форматов.....	170
Четыре секции форматирования	170
Специальные символы	172
Заместитель цифры	172
Запятые и точки.....	172
Текст	173
Символ подчеркивания	174
Звездочка.....	174
Экранирование специальных символов.....	175
Финансовый формат чисел.....	175
Форматы даты и времени.....	177
Условные пользовательские числовые форматы.....	178
Использование значков	179
Цветовые шкалы	179
Гистограммы	184
Наборы значков	187
Создание спарклайнов	189
Типы спарклайнов.....	189
Создание спарклайнов.....	190
Группы спарклайнов	191

Настройка спарклайнов.....	192
Изменение источника данных	192
Изменение цвета и толщины линий спарклайнов.....	193
Настройка осей	195
Глава 6 Использование фигур для создания инфографики	197
Работа с фигурами	197
Вставка фигур.....	197
Настройка фигур.....	200
Встраивание данных в фигуры	202
Создание баннера	203
Создание закладки	206
Работа с несколькими фигурами.....	208
Создание простых графиков с применением фигур.....	210
Создание пользовательской инфографики.....	212
Добавление других иллюстраций.....	214
Часть III Рассказываем историю с помощью визуализации	219
Глава 7 Визуализация сравнения показателей эффективности	220
Отдельные показатели	220
Гистограмма.....	222
Практический пример: продажи по кварталам	225
Пулевая диаграмма.....	226
Практический пример: расходы в сравнении с планом.....	227
Гистограмма с группировкой.....	231
Практический пример: брак на производстве.....	232
Воронка.....	233
Практический пример: конверсия продаж.....	234
Точечная диаграмма	236
Практический пример: температура и продажи	238
Пузырьковая диаграмма	241
Практический пример: ипотека	241
Точечная гистограмма.....	244
Практический пример: объемы производства	244
Глава 8 Визуализация частей от целого.....	255
Круговая диаграмма	255
Кольцевая диаграмма	257
Практический пример: продажи по регионам	258
Вафельная диаграмма	260
Практический пример: привилегии для сотрудников.....	261
Диаграмма солнечные лучи	265
Практический пример: анализ времени производственного процесса	267
Гистограммы.....	268
Практический пример: средний чек в ресторане.....	271

Диаграмма дерево.....	272
Практический пример: виды страхования.....	274
Каскадная диаграмма.....	275
Практический пример: чистая прибыль.....	277
Глава 9 Визуализация изменений с течением времени.....	281
Линейный график.....	281
Практический пример: продажи товаров по категориям.....	284
Столбчатая диаграмма с отклонениями.....	289
Практический пример: продажи домов по месяцам.....	289
Комбинированная диаграмма.....	297
Практический пример: выручка от грузовых перевозок и мили.....	298
Линейный график с отклонениями.....	300
Практический пример: отслеживание выручки по кварталам....	301
Смежные ящики с усами.....	305
Практический пример: зарплаты по отделам.....	307
Анимированная диаграмма.....	309
Сводные диаграммы.....	309
Промежуточные данные с формулами.....	313
Анимация диаграммы при помощи макроса.....	316
Автоматизация диаграмм.....	319
Манипулирование объектами диаграмм.....	319
Создание панельных диаграмм.....	324
Предметный указатель.....	334



Об авторе

Дик Куслейка (Dick Kusleika) работает с пакетом Microsoft Office на протяжении более 20 лет. В течение 12 лет подряд удостоивался статуса Microsoft MVP. Перу Дика принадлежит несколько книг по Excel и Access.



О техническом редакторе

Даг Холланд (Doug Holland) – инженер-программист и разработчик архитектуры в Microsoft Corporation, обладающий степенью магистра Оксфордского университета в области программной инженерии. Перед тем как присоединиться к компании Microsoft, Даг получил статус Microsoft MVP и Intel Black Belt Developer.



Введение

В настоящее время бизнес хранит и обрабатывает огромные массивы данных – больше, чем когда бы то ни было. И это касается не только крупных компаний. Мелкие и средние предприятия сегодня также обрабатывают беспрецедентное количество информации. Руководство компаний использует собранные данные для принятия важных решений, но сложность состоит в том, что с такими объемами информации в исходном виде им просто не справиться. В связи с этим основной задачей *бизнес-аналитики* (Business Intelligence – BI) является преобразование сырых необработанных данных в вид, пригодный для анализа.

Сама бизнес-аналитика появилась не вчера, но в последнее время значительно повысилось количество и качество инструментов в этой области. А это вкупе с возросшей доступностью данных привело к тому, что сегодня, похоже, только ленивый не строит дашборды.

Excel постепенно становится флагманом в мире бизнес-аналитики, если еще не стал. Компания Microsoft в последние годы вложила немало ресурсов в развитие инструментов бизнес-аналитики, встроенных в Excel и существующих отдельно. За это время появилось целое семейство продуктов Power, включающее в себя Power Query, PowerPivot и Power BI, а также значительно обогатился выбор типов встроенных диаграмм, который еще пару версий назад был далеко не таким разнообразным.

Постепенно некогда исключительно специализированные инструменты вошли в состав Excel и стали доступны всем и каждому. В прошлом вам понадобился бы целый серьезный проект для сбора данных, а без специальных инструментов о создании дашбордов не могло быть и речи. Сейчас же все эти средства наверняка уже установлены на вашем компьютере и просто ждут своего часа. И основой, объединяющей в себе все эти полезные инструменты, является именно Excel – программа, с которой вы работаете на ежедневной основе вне зависимости от масштаба вашего бизнеса.

Возможно, вы сами хотите начать строить дашборды, но считаете, что у вас для этого недостаточно знаний и умений. Или руководство компании поставило вам такую задачу. В этой книге мы пройдем поэтапно по всем средствам Excel, связанным с визуализацией данных, – от фигур через условное форматирование и до полноценных диаграмм. Попутно мы будем разбирать

реальные практические примеры, и вы сможете увидеть, как конкретные бизнес-требования превращаются в графики и дашборды.

Организация книги

Содержимое данной книги поделено на три части. В первой части я расскажу о дашбордах в целом, а также мы рассмотрим три реальных примера формирования дашбордов. Во второй части мы сосредоточимся на том, как можно выжать максимум возможного из отдельных компонентов, составляющих дашборды, и поговорим о некоторых неграфических элементах визуализации. Третья часть будет посвящена конкретным типам диаграмм с рассмотрением реальных практических примеров применения каждого из них.

Глава 1 «Основы дашбордов» В данной главе мы поговорим о самых базовых фундаментальных вещах, касающихся проектирования дашбордов, включая основания для их применения и простейшие принципы их представления и форматирования.

Глава 2 «Практические примеры дашбордов» В этой главе мы рассмотрим три практических примера. В каждом из них будут подробно описаны бизнес-требования к дашбордам и принципы их построения.

Глава 3 «Организация данных для дашбордов» Эта глава будет целиком посвящена данным. В ней мы досконально разберем особенности организации данных по слоям. Также рассмотрим несколько типов внешних источников данных и посмотрим, как информацию из них можно перенести в Excel.

Глава 4 «Основы эффективной визуализации» Данную главу будет полезно почитать новичкам в мире визуализации. Здесь мы узнаем, что делает визуализацию по-настоящему эффективной, как использовать такие элементы, как цвет и текст, а также как выбирать тип диаграммы для данных, которые вам необходимо представить конечному пользователю.

Глава 5 «Неграфические визуализации» Не все компоненты дашборда представляют собой диаграммы. В данной главе мы подробно остановимся на особенностях визуализации в Excel и погрузимся в мир пользовательского форматирования.

Глава 6 «Использование фигур для создания инфографики» В этой главе мы поговорим о применении элементарных фигур в Excel. Также вы увидите, как можно оригинально компоновать данные при помощи фигур.

Глава 7 «Визуализация сравнения показателей эффективности» Данная глава будет посвящена типам диаграмм, идеально подходящим для сравнения различных показателей эффективности. Вместе с тем мы рассмотрим практические примеры применения каждого из этих типов визуального представления информации.

Глава 8 «Визуализация частей от целого» В этой главе мы подробно рассмотрим варианты представления данных для ситуаций, когда показатель разбивается на составляющие. Так же, как и раньше, все размышления будут подкреплены конкретными практическими примерами.

Глава 9 «Визуализация изменений с течением времени» Здесь мы коснемся типов диаграмм, подходящих для представления динамики показателей во времени. В качестве дополнения мы рассмотрим способы управления графиками посредством языка программирования Visual Basic for Applications (VBA).

Сопроводительные файлы

Все рабочие книги и вспомогательные файлы для работы с примерами из этой книги собраны по адресу www.wiley.com/go/datavizwithexcel или на сайте издательства dmkpress.com со страницы книги.

Отображение данных на дашборде

Глава 1. Основы дашбордов

Глава 2. Практические примеры дашбордов

Глава 3. Организация данных для дашбордов

Основы дашбордов

В этой главе:

- когда необходимо использовать дашборды;
- пользовательские требования;
- сбор данных;
- построение дашбордов;
- форматирование дашбордов.

Дашборды еще никогда не были столь популярны. Сегодня в нашем распоряжении невиданное ранее количество информации и поражающие воображение средства ее визуализации. По своей сути *дашборд* (dashboard) является коллекцией диаграмм. Но это гораздо больше, чем просто коллекция. Разместите на странице разрозненные графики и диаграммы. Что перед вами? Дашборд? Фактически да, но он не слишком хорош. Создание действительно полезных дашбордов требует определенной подготовки, знаний и умений. В этой вводной главе мы познакомимся с понятием дашборда, его концепцией, а также пройдемся по навыкам, которыми необходимо обладать, чтобы полноценно работать с ним.

Когда необходимо использовать дашборды

Дашборды используются для представления данных. При этом сами данные могут проходить разные стадии: изначально мы имеем дело с сырыми необработанными данными, следующая стадия – их агрегирование, а далее анализ и презентация. Стадия, на которой находятся ваши данные, зависит от их источника и предназначения. Существует немало способов агрегировать данные и бесконечное количество методов их анализа и представления. К примеру, счет можно назвать агрегацией строк в табличной части счета, а отчет о продажах – агрегацией счетов. Таким образом, для счета исходными данными являются его строки, а для отчета о продажах – счета. На рис. 1.1 показаны данные на разных стадиях.

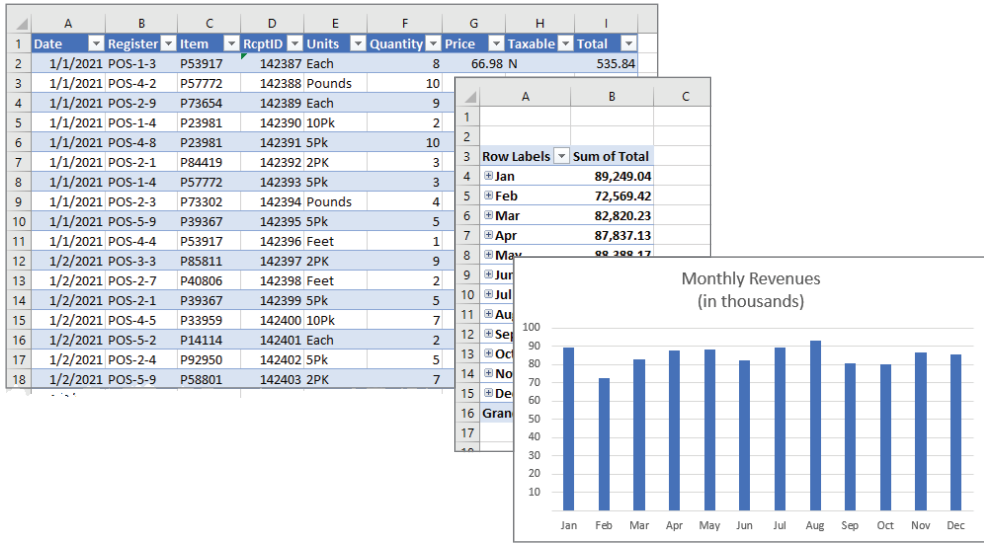


Рис. 1.1. Исходные данные, агрегированные, проанализированные и представленные

Исходными, или *сырыми, данными* (raw data) называют информацию, не прошедшую стадию обработки. Это могут быть транзакции из оперативной системы учета, данные о продажах из магазина или сведения с измерительных приборов, такие как температура или уровень жидкости. Если вашим исходным сырьем являются необработанные данные, вам придется выполнить их агрегирование и, возможно, предварительный анализ перед помещением на дашборд.

Примечание Рабочая книга, графики из которой показаны в этой главе, называется *Chapter1Figures.xlsx* и располагается на сайте с сопроводительными файлами.

Агрегированные данные (aggregated data) определенным образом группируются и суммируются. Например, в отчете о произведенных изделиях за месяц суммируются объемы производства по неделям или дням. А ежедневное производство может складываться из смен. Зачастую построение дашбордов начинается именно с агрегированных данных.

Дашборды призваны рассказывать истории о данных. *Анализ данных* (analyzing data), в свою очередь, служит для определения того, какие истории получаются и какие из них достойны того, чтобы быть рассказанными. Анализ – это нечто большее, чем просто извлечение выводов из данных. Он также включает в себя исследование природы данных и определение вопросов, поднимаемых этими данными. Во время анализа данных зачастую приходится делать шаг назад и выполнять агрегирование иным способом.

Заключительная стадия данных – это стадия *представления* (presentation), – именно здесь обитают дашборды. Процесс построения дашборда может начинаться на любой стадии. Если исходные данные для дашборда вы получаете от аналитика, скорее всего, история, которую они должны рассказывать, уже определена, а значит, вам останется лишь эффективно донести

ее до пользователя графически. Если же вы начинаете с необработанных данных, вам необходимо будет сначала агрегировать их, а затем проанализировать, чтобы определиться с этими историями.

Дашборды постоянно развиваются. Если раньше это были исключительно статические визуализации, рассказывающие одну историю, то теперь они включают в себя сложнейшие средства бизнес-аналитики, позволяющие рассказывать сразу несколько историй или дать пользователю самому сделать для себя важные выводы. С появлением мощного пакета Microsoft Power BI и его дальнейшей интеграции с Excel посредством Power Pivot и Power Query средства самостоятельного анализа данных стали гораздо более доступными и фактически вышли на первый план.

Ссылка Об инструментах Power Pivot, Power Query и Power BI мы будем говорить в главе 3.

Что такое дашборд?

Дашбордом называется один или несколько визуальных элементов, рассказывающих историю о лежащих в их основе данных. *Отчет*, в котором агрегированы данные, нельзя назвать дашбордом, поскольку он не рассказывает историю. Такой тип представления информации называется таблицей или отчетом – зачастую эти термины означают одно и то же. Дашборд же обязательно должен содержать визуальные элементы, а не просто список с данными.

Рассказываемая история является важнейшим аспектом любого дашборда. Как мы уже говорили, именно на стадии анализа данных определяются важные аспекты истории. На дашбордах чаще всего отображаются *ключевые показатели эффективности* (key performance indicator – KPI). Эти показатели сами по себе могут быть историями, и в следующем разделе мы поговорим о них подробнее. Распространенной ошибкой при построении дашборда является начало истории с вывода. У пользователя может быть заранее сложившееся мнение о том, что именно должен рассказать дашборд. Но именно данные должны лежать в основе истории, а не наоборот. Постарайтесь преобразовать вывод в вопрос. Если кто-то хочет, чтобы ваш дашборд показывал снижение продаж по причине плохой погоды, вы должны задаться вопросом о том, как среднесуточная температура коррелирует с дневными продажами, или сравнить продажи в дождливые и солнечные дни.

Данные, лежащие в основе дашборда, всегда с чем-то связаны. А с чем именно – зависит от того, кто просматривает дашборд. К примеру, рядовой менеджер отдела по работе с персоналом может запросить данные, связанные с удержанием сотрудников, такие как коэффициенты приема на работу и увольнения с работы, количество добровольных отставок и выходов на пенсию. В то же время управляющий отдела может заинтересоваться более обобщенной информацией об увольнениях, а также данными о затратах на заработную плату и льготах. Руководство компании хочет располагать сведениями о сотрудниках на уровне общих финансов, бухгалтерии и юридических аспектов.

Ключевые показатели эффективности

Как определяются ключевые показатели эффективности и их польза, выходит далеко за рамки этой книги. В руководстве компаний эти показатели устанавливаются на основании знаний о деятельности предприятия. Если у вас коммерческая организация, чистая прибыль будет мерилom вашего успеха, и вам не нужно анализировать какие-то данные, чтобы понять это. Показатели эффективности отличаются в разных компаниях, но некоторые из них используются чаще остальных в своих областях. К примеру, руководству финансовых отделов прежде всего будут интересны чистая прибыль, свободный денежный поток и оборотный капитал. А на производстве на первый план могут выходить объем этого самого производства и коэффициент использования линии.

Пользовательские требования

Никогда не приступайте к разработке дашборда без четкого плана действий. Здесь как при строительстве дома в отсутствие плана – все может рухнуть, и придется возводить здание заново. Чтобы разработать план, поговорите с конечными пользователями дашборда. Существует по крайней мере три типа пользователей, которых вам стоит опросить при составлении плана: непосредственно заказчика дашборда, человека, поставляющего исходные данные, и конечного пользователя. При этом все они могут быть представлены в одном лице, и даже в вашем собственном.

Получите максимум информации от *заказчика дашборда*. Если у него есть общая идея того, что хочется получить, постарайтесь углубиться в детали, чтобы картина стала наиболее полной. Как я уже упоминал ранее, у заказчика в сознании может быть даже итоговый вывод дашборда. Сделайте все, чтобы превратить этот вывод в один или несколько вопросов.

Что касается *источника данных*, то этот аспект зачастую незаслуженно упускается из виду. На этапе разработки плана обязательно узнайте, откуда берутся исходные данные для дашборда, а также проходят ли они предварительную агрегацию или анализ. В зависимости от проекта вы можете стремиться к получению максимально сырых исходных данных, насколько это возможно, чтобы можно было изменить направление анализа в любой момент. Детальные данные всегда можно агрегировать как-то иначе, тогда как разагрегировать их будет почти невозможно.

Определите, поступают ли исходные данные изнутри организации или извне, кто занимается поддержкой этих данных и как часто они обновляются. К примеру, финансовые данные из бухгалтерской системы могут быть доступны раз в месяц или в квартал, тогда как информацию с точки продажи можно запрашивать практически в реальном времени.

Если у вас в наличии нет исходных данных для дашборда, ваш проект может раздвоиться: одна часть будет добывать данные, а другая заниматься построением дашборда. Более того, может оказаться так, что данные не просто

недоступны в настоящий момент, а их не существует вовсе. Например, если на предприятии не ведется учет брака на линии, получить исторические сведения из этой области будет просто невозможно. В этом случае вам необходимо инициировать сбор нужной вам информации, хотя это и приведет к задержке выпуска итогового дашборда. И чем раньше вы обговорите все эти моменты с заказчиком, тем более реалистичными будут его ожидания.

Типы конечных пользователей

Конечных пользователей можно разбить на типы по тому, как они будут использовать дашборд. *Наблюдатели* (monitor) используют дашборды для контроля состояния компании или проекта в тот или иной момент времени. Находясь в автомобиле, вы постоянно наблюдаете за своим дашбордом в виде приборной панели, контролируя скорость, уровень топлива и индикаторы диагностики. *Решители* (decider) применяют дашборды с целью определиться, стоит ли принимать то или иное решение. А *менеджеры по производству* (production manager) могут опираться на дашборд по продажам и использованию линий для решения о том, стоит ли запускать третью смену.

Планировщики (planner) являются представителями руководства компании и занимаются определением вектора развития организации. Они следят за более общими показателями, а решения принимают политического характера. При этом планировщики могут воспользоваться информацией об оперативных результатах деятельности подразделения с целью распределения ресурсов на ближайшие пять лет. *Презентаторы* (presenter) используют дашборды для представления информации. Дашборд, показанный на собрании акционеров, может позволить им увидеть информацию, которой они не располагают на ежедневной основе.

Разумеется, все эти категории пользователей могут пересекаться между собой. К примеру, наблюдатели могут принимать определенные решения в зависимости от показателей. А акционеры – сменить руководство компании, если им не понравятся те или иные цифры. Полезно знать свою аудиторию, а иногда и аудиторию своей аудитории, чтобы предоставлять информацию в правильном виде.

Также важно определиться с тем, как часто будет использоваться дашборд. Если ваш дашборд посвящен влиянию проведения Олимпийских игр на финансовое благополучие города-организатора, разумеется, можно предположить, что он будет построен один раз. Для таких «одноразовых» дашбордов нет нужды беспокоиться о хранении данных и эффективности доступа к ним.

Если дашборд планируется обновлять периодически или он должен работать в реальном времени, вам придется позаботиться о беспрепятственном доступе к актуальным исходным данным. Кроме того, в этом случае стоит задуматься и об автоматизации. Дашборд, работающий в реальном времени, должен быть полностью автоматизирован, тогда как ежегодные дашборды могут быть автоматизированы лишь частично.

С самого начала работы над проектом ведите его *документацию*. Для этого вовсе не обязательно устанавливать какое-то специализированное про-

граммное обеспечение, достаточно будет обычного текстового файла или таблицы. Фиксируйте всю информацию о пользователях дашборда и данных, которые уже исследовали. Также документируйте весь путь исходных данных от хранилища до таблицы. Даже если ваш дашборд будет частично или полностью автоматизирован, отражайте весь поток данных документально, чтобы вы или кто-то другой могли внести необходимые изменения при необходимости. Представьте, что вы создаете дашборд не сегодня, а через год, когда забудутся все детали, и попытайтесь ответить на все вопросы, которые у вас возникнут.

Планируйте каждую стадию разработки дашборда. По готовности плана отправьте его заказчиком для обсуждения деталей и после каждого внесения изменений проводите новое согласование. Ожидая ответ, вы можете заниматься следующей стадией проекта, но никогда не принимайте важных решений без согласования с заказчиком. Нет ничего хуже, чем сдать финальную версию дашборда, содержащую недочеты, которые можно было устранить на ранних стадиях.

Наконец, запланируйте *ревизию дашборда* после его внедрения. Если вы посылаете дашборд пользователям еженедельно, постарайтесь встретиться с ними через несколько месяцев, чтобы убедиться, что дашборд по-прежнему отвечает всем их требованиям. Это особенно важно, если от вас требуются какие-то действия по публикации дашборда. У меня была ситуация, когда я ежемесячно создавал дашборды, не заметив, что механизм их публикации сломался. Прошло какое-то время, но никто так и не поинтересовался, почему дашборд не обновляется. Оказалось, что за это время требования в компании изменились, и этот дашборд просто утратил актуальность.

Сбор данных

Для пользователя Excel работа с данными является самой приятной частью процесса построения дашборда. Как вы, разумеется, знаете, Excel обладает очень богатым инструментарием для агрегации и манипуляции данными, включая сотни приемов и функций для сортировки, разделения данных и удаления дубликатов. В следующих разделах мы обсудим эти средства более подробно.

Ссылка Если исходные данные для вашего дашборда располагаются за пределами Excel, первым делом необходимо их получить. Работу с внешними данными мы детально будем рассматривать в главе 3.

Сводные таблицы

Одним из наиболее мощных инструментов Excel является *сводная таблица* (PivotTable). Сводная таблица представляет собой отчет, в котором выполнены фильтрация, сортировка и агрегация исходных данных по вашим требованиям. Сводные таблицы являются полностью интерактивными – это

означает, что вы можете по желанию переносить поля в них в нужные вам секции для контроля над тем, как будут агрегироваться ваши данные.

Ссылка Диаграммы, созданные на основании сводных таблиц, называются *сводными диаграммами* (PivotChart). Сводные диаграммы от обычных также отличаются своей интерактивностью. Подробно о сводных диаграммах мы будем говорить в главе 9.

Все, что вы можете делать со сводными таблицами, вы можете делать и с *формулами рабочего листа*. При этом сводная таблица выполнит агрегирование данных гораздо быстрее, чем это сделают формулы. Любые изменения в сводных таблицах, включая добавление, удаление или перестановку столбцов, будут выполняться практически мгновенно. Переписывать формулы под изменившиеся требования будет дольше. На рис. 1.2 показан фрагмент исходных данных на листе Excel, а на рис. 1.3 – сводная таблица на их основании.

	A	B	C	D	E
1	State	Region	Month	Qtr	Sales
2	California	West	Jan	Qtr-1	1,118
3	California	West	Feb	Qtr-1	1,960
4	California	West	Mar	Qtr-1	1,252
5	California	West	Apr	Qtr-2	1,271
6	California	West	May	Qtr-2	1,557
7	California	West	Jun	Qtr-2	1,679
8	Washington	West	Jan	Qtr-1	1,247
9	Washington	West	Feb	Qtr-1	1,238
10	Washington	West	Mar	Qtr-1	1,028
11	Washington	West	Apr	Qtr-2	1,345
12	Washington	West	May	Qtr-2	1,784
13	Washington	West	Jun	Qtr-2	1,574
14	Oregon	West	Jan	Qtr-1	1,460
15	Oregon	West	Feb	Qtr-1	1,954
16	Oregon	West	Mar	Qtr-1	1,726
17	Oregon	West	Apr	Qtr-2	1,461
18	Oregon	West	May	Qtr-2	1,764
19	Oregon	West	Jun	Qtr-2	1,144

Рис. 1.2. Исходные данные для сводной таблицы

Sum of Sales		Column Labels		
Row Labels	Qtr-1	Qtr-2	Grand Total	
▣ Central	21,025	22,176	43,201	
Illinois	4,243	3,623	7,866	
Kansas	4,776	4,007	8,783	
Kentucky	3,342	4,626	7,968	
Missouri	4,669	4,556	9,225	
Oklahoma	3,995	5,364	9,359	
▣ East	17,267	17,864	35,131	
Florida	4,722	4,630	9,352	
Massachusetts	3,442	4,155	7,597	
New Jersey	4,365	4,316	8,681	
New York	4,738	4,763	9,501	
▣ West	16,778	18,242	35,020	
Arizona	3,795	4,663	8,458	
California	4,330	4,507	8,837	
Oregon	5,140	4,369	9,509	
Washington	3,513	4,703	8,216	
Grand Total	55,070	58,282	113,352	

Рис. 1.3. Сводная таблица

Сводная таблица, показанная на рис. 1.3, была создана путем перетаскивания поля *Qtr* в *область столбцов* (Columns), полей *Region* и *State* – в *область строк* (Rows), а поля *Sales* – в *область значений* (Values). По умолчанию Excel суммирует поле *Sales*, являющееся числовым. Для нечисловых данных Excel применяет агрегацию по количеству. На рис. 1.4 показана панель *Поля сводной таблицы* (PivotTable Fields) с заполненными областями.

На панели с полями можно выделить следующие области:

- *сервис* (кнопка с шестеренкой). Здесь вы можете изменить расположение панели и отсортировать поля по именам. Если у вас много полей, возможно, вы предпочтете макет, предполагающий компактное размещение большего количества элементов;

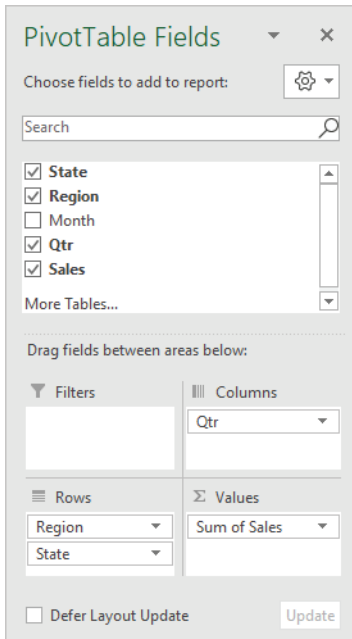


Рис. 1.4. Панель
Поля сводной таблицы

- *поиск* (Search). Осуществляя ввод в этом поле, вы сможете ограничить список полей для выбора;
- *фильтры* (Filters). Если перенести поле в эту область, Excel разместит над сводной таблицей выпадающий список с его элементами. Пользователь сможет выбрать нужные ему фильтры для отображения в сводной таблице;
- *столбцы* (Columns). Поля, помещенные в эту область, будут располагаться в сводной таблице в виде колонок. Для реализации иерархии в столбцах вы можете разместить здесь несколько полей. В этом случае поле, стоящее выше по иерархии, будет объединять в себе нижестоящие поля, и значения в таблице под ним будут относиться к этому родительскому элементу. На рис. 1.5 показано, как можно поле *Month* вложить в поле *Qtr*. Значения *Jan*, *Feb* и *Mar* относятся к *Qtr-1*, так что только они располагаются под этим элементом, соответствующим первому кварталу;

Column Labels									
Qtr-1			Qtr-1 Total	Qtr-2			Qtr-2 Total	Grand Total	
Jan	Feb	Mar		Apr	May	Jun			
Sum of Sales	18,579	19,194	17,297	55,070	20,754	18,771	18,757	58,282	113,352

Рис. 1.5. Вложенные поля в области столбцов

- *строки* (Rows). Поля, расположенные в этой области, будут располагаться в сводной таблице в виде строк. Как и в случае со столбцами, вы можете вкладывать поля одно в другое в этой области, что также при-

ведет к образованию иерархии. В сводной таблице, изображенной на рис. 1.3, поле *State* вложено в поле *Region* по строкам;

- *значения* (Values). Значения полей, помещенных в эту область, будут агрегироваться в теле сводной таблицы. Наиболее часто используемыми функциями агрегирования являются *Сумма* (Sum) и *Количество* (Count). Но есть и другие способы агрегирования значений, включая вычисление *среднего значения* (Average), *минимума* (Min) и *максимума* (Max). Агрегации показываются в сводной таблице на пересечении строк и столбцов. Например, на рис. 1.3 на пересечении значений *Qtr-1* и *Missouri* отображены только суммы, относящиеся к штату Миссури за первый квартал года.

Также Excel добавляет в сводную таблицу *общие итоги* и *промежуточные*, если в строках или столбцах есть вложенные поля. Эти итоги могут быть показаны или скрыты по вашему желанию. Чтобы отключить общие итоги, щелкните мышью в любом месте сводной таблицы и в разделе **Работа со сводными таблицами** (PivotTable Tools) на вкладке **Анализ** (Analyze) нажмите на кнопку **Параметры** (Options), чтобы открылось диалоговое окно **Параметры сводной таблицы** (PivotTable Options), показанное на рис. 1.6. Снимите флажки **Показывать общие итоги для строк** (Show Grand Totals For Rows) и **Показывать общие итоги для столбцов** (Show Grand Totals For Columns) на вкладке **Итоги и фильтры** (Totals & Filters).

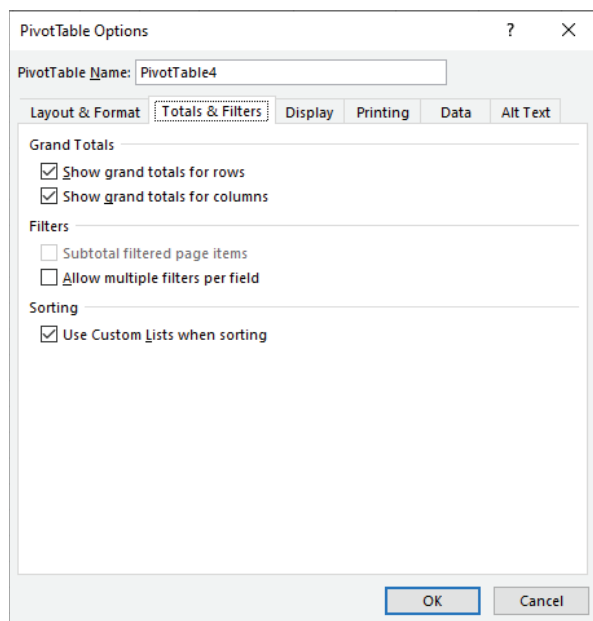


Рис. 1.6. Диалоговое окно **Параметры сводной таблицы**

Если вы хотите скрыть промежуточные итоги для вложенных полей, щелкните мышью по одному из полей в соответствующей области и выберите

в контекстном меню пункт **Параметры поля** (Field Settings). В открывшемся диалоговом окне **Параметры поля** (Field Settings), показанном на рис. 1.7, выберите один из следующих переключателей:

- **автоматические** (Automatic): промежуточные итоги будут отображаться с тем же типом агрегирования (сумма, количество и т. д.), что и дочернее поле;
- **нет** (None): промежуточные итоги для поля выводиться не будут;
- **другие** (Custom): промежуточные итоги будут отображаться с тем же или другим типом агрегирования по сравнению с дочерним полем.

На рис. 1.8 показана сводная таблица со скрытыми общими и промежуточными итогами.

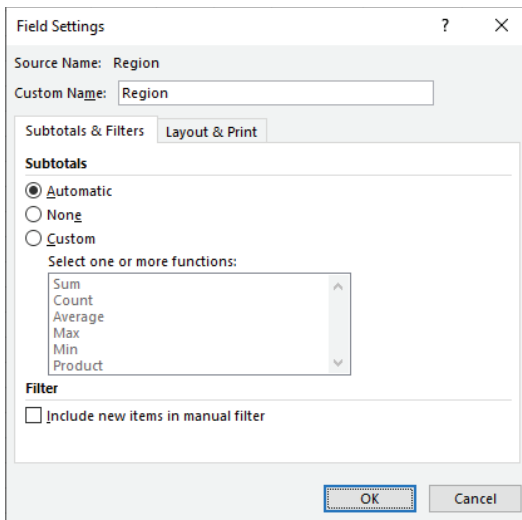


Рис. 1.7. Диалоговое окно **Параметры поля**

Sum of Sales		Column Labels	
Row Labels	Qtr-1	Qtr-2	
Central			
Illinois	4,243	3,623	
Kansas	4,776	4,007	
Kentucky	3,342	4,626	
Missouri	4,669	4,556	
Oklahoma	3,995	5,364	
East			
Florida	4,722	4,630	
Massachusetts	3,442	4,155	
New Jersey	4,365	4,316	
New York	4,738	4,763	
West			
Arizona	3,795	4,663	
California	4,330	4,507	
Oregon	5,140	4,369	
Washington	3,513	4,703	

Рис. 1.8. Сводная таблица без общих и промежуточных итогов

Функция **ПОЛУЧИТЬ.ДААННЫЕ.СВОДНОЙ.ТАБЛИЦЫ** (GETPIVOTDATA)

Сводные таблицы динамичны по своей природе. Это означает, что данные в них могут менять свое расположение в зависимости от действий пользователя, что затрудняет обращение к ним в формулах на рабочем листе. К счастью, в Excel есть очень полезная функция **ПОЛУЧИТЬ.ДААННЫЕ.СВОДНОЙ.ТАБЛИЦЫ** (GETPIVOTDATA) для гарантии обращения к нужным вам данным анализа.

К примеру, если вам необходимо получить информацию о продажах в штате Калифорния за первый квартал, вы можете воспользоваться этой функцией следующим образом¹:

¹ Здесь и далее мы будем приводить формулы для русской и английской версий Excel. – Прим. перев.

```
=ПОЛУЧИТЬ.ДААННЫЕ.СВОДНОЙ.ТАБЛИЦЫ("Sales";$I$3;"State";"California";"Region";"West";"Qtr";"Qtr-1")
```

```
=GETPIVOTDATA("Sales", $I$3, "State", "California", "Region", "West", "Qtr", "Qtr-1")
```

Если источник данных пополнится новой информацией и нужная нам цифра переместится в сводной таблице в другое место, наша формула по-прежнему будет указывать на нее. Первый аргумент формулы ссылается на поле в области значений, которое вы хотите вернуть. Следующим передается аргумент с адресом верхней левой ячейки сводной таблицы с данными. Дальнейшие шесть аргументов представляют собой пары значений полей для искомого пересечения.

Эти пары значений могут следовать в любом порядке, но внутри каждой пары первым должно значиться имя поля, а вторым – его значение. Пара аргументов "State" и "California" предполагает, что мы должны взять строку, в которой для поля *State* указано значение *California*. После учета всех перечисленных пар аргументов мы сузим возвращаемое значение до единственной нужной нам цифры.

Также функция *ПОЛУЧИТЬ.ДААННЫЕ.СВОДНОЙ.ТАБЛИЦЫ* может быть использована для получения итоговых значений в сводных таблицах. Например, чтобы получить общий итог по значению столбца *Qtr-1*, используйте следующую запись:

```
=ПОЛУЧИТЬ.ДААННЫЕ.СВОДНОЙ.ТАБЛИЦЫ("Sales";$I$3;"Qtr";"Qtr-1")
```

```
=GETPIVOTDATA("Sales", $I$3, "Qtr", "Qtr-1")
```

А общий итог по всей сводной таблице можно получить так:

```
=ПОЛУЧИТЬ.ДААННЫЕ.СВОДНОЙ.ТАБЛИЦЫ("Sales";$I$3)
```

```
=GETPIVOTDATA("Sales", $I$3)
```

Как вы понимаете, каждую формулу с использованием функции *ПОЛУЧИТЬ.ДААННЫЕ.СВОДНОЙ.ТАБЛИЦЫ* со всеми ее аргументами писать вручную может быть весьма утомительно. К счастью, Excel вычисляет эту формулу за вас каждый раз, когда вы щелкаете по ячейкам сводной таблицы. Просто введите в строку формул знак равенства и щелкните по нужной вам ячейке, и функция будет заполнена всеми нужными вам аргументами. Если вы хотите обращаться к ячейкам сводной таблицы без использования функции *ПОЛУЧИТЬ.ДААННЫЕ.СВОДНОЙ.ТАБЛИЦЫ*, вы можете просто указывать в формуле адрес ячейки. Чтобы изменить *адресацию данных в сводных таблицах*, принятую по умолчанию, нажмите на вкладку **Файл** (File) и выберите пункт **Параметры** (Options). Перейдите в раздел **Формулы** (Formulas) и снимите флажок **Использовать функции GetPivotData для ссылок в сводной таблице** (Use GetPivotData Functions For PivotTable References).

Предупреждение В отличие от формул, сводные таблицы автоматически не обновляются при изменении исходных данных. Для принудительного обновления нажмите в разделе Работа со сводными таблицами (PivotTable Tools) на вкладке Анализ (Analyze) кнопку Обновить (Refresh).

Функции рабочего листа

Excel располагает сотнями *функций рабочего листа* (worksheet function) для манипулирования данными, которые можно использовать на дашбордах. На практике вы будете использовать далеко не все из них. Но есть функции, которые при формировании исходных данных для дашбордов применяются чаще остальных. В следующих разделах мы перечислим их и кратко пройдемся по их функционалу.

Функция ВПР (VLOOKUP)

Функция *ВПР* (VLOOKUP) используется для поиска значения в одном столбце таблицы и возвращения результата из другого столбца из той же строки. Синтаксис функции *ВПР* приведен ниже:

ВПР(lookup_value; table_array; col_index_num; range_lookup)

VLOOKUP(lookup_value, table_array, col_index_num, range_lookup)

На рис. 1.9 показан пример применения функции ВПР.

	A	B	C	D	E	F	G	H	I	J	K
1											
2						Salesperson	January	February	March	April	May
3						Rachel Thomas	2,076.08	2,741.45	1,540.96	2,142.28	1,863.27
4		Hailey Porter	4,274.40			Payton Brooks	3,109.17	2,136.32	4,774.18	4,675.77	3,519.00
5						Ava Mills	1,619.69	1,104.68	1,566.24	1,797.64	4,233.14
6						Jayden Cruz	2,207.33	4,083.45	4,205.58	3,921.56	4,830.75
7						Brady Bailey	3,547.31	4,787.54	3,557.96	1,545.73	3,228.25
8						Hailey Porter	1,598.16	1,935.84	1,977.14	4,274.40	2,700.35
9						Cameron Cook	4,801.98	4,326.55	2,342.97	1,532.29	1,619.69
10						Arianna Gardner	2,855.90	1,785.04	3,266.20	4,989.87	2,140.13
11						Ashley Austin	4,704.37	4,084.04	3,752.84	3,889.23	3,777.19
12						Camila Phillips	1,935.64	2,482.00	3,077.91	3,772.08	2,081.66
13						Julia Sims	4,439.92	1,640.12	1,323.81	1,865.97	1,692.28
14						Robert Wallace	4,752.36	2,357.20	1,464.17	4,888.25	4,723.95

Рис. 1.9. Функция ВПР в действии

В данном примере мы видим таблицу с суммами комиссий по продавцам и месяцам, расположенную в диапазоне *F3:R52*. Приведенная формула вычисляет комиссию за апрель по продавцу из ячейки *B4*. При этом имя продавца должно находиться в первом столбце указанного диапазона поиска. По сути, формула проходит по столбцу *F*, пока не найдет искомое имя, после чего возвращает значение из пятого по счету столбца в найденной строке.

Аргумент *range_lookup* установлен в значение *FALSE*, чтобы производился четкий поиск. Если присвоить этому аргументу значение *TRUE*, Excel будет считать, что данные в первом столбце отсортированы, и найдет ближайшее к искомому значению, если *четкий поиск* не даст результата. В некоторых ситуациях *нечеткий поиск* может пригодиться, но это бывает не слишком

часто, – в большинстве случаев значение аргумента *range_lookup* должно быть установлено в *FALSE*.

Функция ПРОСМОТРХ (XLOOKUP)

Функция *ПРОСМОТРХ* (XLOOKUP) появилась в Excel относительно недавно. Она работает подобно функции *ВПР* (VLOOKUP), но не требует, чтобы искомое значение обязательно располагалось в первом столбце указанного диапазона. Ниже приведен синтаксис функции *ПРОСМОТРХ*:

```
ПРОСМОТРХ(lookup_value; lookup_array; return_array; if_not_found; match_mode; search_mode)
XLOOKUP(lookup_value, lookup_array, return_array, if_not_found, match_mode, search_mode)
```

Первый аргумент у функций *ПРОСМОТРХ* и *ВПР* одинаковый. Но вместо того чтобы передавать целую таблицу, предполагая, что искомое значение располагается в первом ее столбце, мы дадим функции *ПРОСМОТРХ* диапазон для поиска и диапазон для возвращаемого значения.

Функция *ПРОСМОТРХ* также включает в себя аргумент *if_not_found*, позволяющий передать значение, которое будет возвращено в случае безуспешного поиска. Если не задать этот аргумент, в случае неудачи функция *ПРОСМОТРХ*, подобно *ВПР*, вернет ошибку #Н/Д (#N/A).

Аргумент *match_mode* похож на аргумент *range_lookup* функции *ВПР*. Но вместо логических значений *TRUE* и *FALSE* необходимо передать число, указывающее Excel, какой тип поиска применять. Значение этого аргумента по умолчанию равно нулю и означает четкий поиск.

Заключительный аргумент *search_mode* позволяет указать Excel, как именно выполнять поиск. Если вы имеете дело с большим набором данных, использование этого аргумента может положительно сказаться на скорости выполнения операции. Но в большинстве случаев аргумент *search_mode* можно опустить.

На рис. 1.10 показано использование этой функции с целью поиска имени продавца с максимальной суммой комиссии за апрель.

	A	B	C	D	E	F	G	H	I	J	K
1											
2						Salesperson	January	February	March	April	May
3						Rachel Thomas	2,076.08	2,741.45	1,540.96	2,142.28	1,863.27
4		Arianna Gardner	4,989.87			Payton Brooks	3,109.17	2,136.32	4,774.18	4,675.77	3,519.00
5						Ava Mills	1,619.69	1,104.68	1,566.24	1,797.64	4,233.14
6						Jayden Cruz	2,207.33	4,083.45	4,205.58	3,921.56	4,830.75
7						Brady Bailey	3,547.31	4,787.54	3,557.96	1,545.73	3,228.25
8						Hailey Porter	1,598.16	1,935.84	1,977.14	4,274.40	2,700.35
9						Cameron Cook	4,801.98	4,326.55	2,342.97	1,532.29	1,619.69
10						Arianna Gardner	2,855.90	1,785.04	3,266.20	4,989.87	2,140.13
11						Ashley Austin	4,704.37	4,084.04	3,752.84	3,889.23	3,777.19
12						Camila Phillips	1,935.64	2,482.00	3,077.91	3,772.08	2,081.66
13						Julia Sims	4,439.92	1,640.12	1,323.81	1,865.97	1,692.28
14						Robert Wallace	4,752.36	2,357.20	1,464.17	4,888.25	4,723.95

Рис. 1.10. Функция *ПРОСМОТРХ*

Вместо того чтобы осуществлять поиск значения в первой колонке диапазона, формула ищет указанную сумму в пятой колонке и возвращает соответствующее имя продавца из первой. А с учетом того, что в ячейке *C4* мы ввели формулу `=МАКС(J3:J52)` (`=MAX(J3:J52)`), на выходе получим имя продавца с максимальной месячной комиссией, как и ожидали. Функция *ПРОСМОТРХ* проходит по столбцу *J*, ищет нужную нам сумму комиссии и возвращает имя соответствующего продавца из колонки *F*. Поскольку мы знаем, что искомое значение в столбце *J* точно присутствует, нам нет необходимости использовать дополнительные аргументы.

Функции *ИНДЕКС (INDEX)* и *ПОИСКПОЗ (MATCH)*

Если вы используете версию Excel, в которой не представлена относительно новая функция *ПРОСМОТРХ*, вы можете добиться похожих результатов путем сочетания двух полезных функций *ИНДЕКС (INDEX)* и *ПОИСКПОЗ (MATCH)*. Функция *ИНДЕКС* возвращает значение из переданного диапазона на основании номеров строки и столбца. Для диапазона, состоящего из единственного столбца, достаточно указать только смещение по строкам. Функция *ПОИСКПОЗ (MATCH)* в связке с *ИНДЕКС* позволяет вычислить смещение по строкам. На рис. 1.11 показано совместное применение функций *ИНДЕКС (INDEX)* и *ПОИСКПОЗ (MATCH)* для получения результата, аналогичного функции *ПРОСМОТРХ* из предыдущего примера.

	A	B	C	D	E	F	G	H	I	J	K
1											
2						Salesperson	January	February	March	April	May
3						Rachel Thomas	2,076.08	2,741.45	1,540.96	2,142.28	1,863.27
4		Arianna Gardner	4,989.87			Payton Brooks	3,109.17	2,136.32	4,774.18	4,675.77	3,519.00
5						Ava Mills	1,619.69	1,104.68	1,566.24	1,797.64	4,233.14
6						Jayden Cruz	2,207.33	4,083.45	4,205.58	3,921.56	4,830.75
7						Brady Bailey	3,547.31	4,787.54	3,557.96	1,545.73	3,228.25
8						Hailey Porter	1,598.16	1,935.84	1,977.14	4,274.40	2,700.35
9						Cameron Cook	4,801.98	4,326.55	2,342.97	1,532.29	1,619.69
10						Arianna Gardner	2,855.90	1,785.04	3,266.20	4,989.87	2,140.13
11						Ashley Austin	4,704.37	4,084.04	3,752.84	3,889.23	3,777.19
12						Camila Phillips	1,935.64	2,482.00	3,077.91	3,772.08	2,081.66
13						Julia Sims	4,439.92	1,640.12	1,323.81	1,865.97	1,692.28
14						Robert Wallace	4,752.36	2,357.20	1,464.17	4,888.25	4,723.95

Рис. 1.11. Функции *ИНДЕКС (INDEX)* и *ПОИСКПОЗ (MATCH)*

Базовый синтаксис функции *ПОИСКПОЗ* следующий:

`ПОИСКПОЗ(lookup_value; lookup_array; match_type)`

`MATCH(lookup_value, lookup_array, match_type)`

Функция *ПОИСКПОЗ* работает подобно *ВПР*, за исключением того, что не возвращает значение из другой колонки. Вместо этого она возвращает позицию значения из аргумента *lookup_value* в диапазоне, заданном аргументом *lookup_array*. Нулевое значение, переданное в качестве третьего параметра

match_type, указывает на четкий поиск, как в случае со значением *FALSE* для функции *ВПР*. В данном примере функция *ПОИСКПОЗ* возвращает значение 8, поскольку сумма, указанная в ячейке *C4*, располагается на восьмой строке в диапазоне поиска *J3:J52*.

После этого итоговая формула приобретет следующий вид:

```
=ИНДЕКС(F3:F52;8)
```

```
=INDEX(F3:F52,8)
```

Теперь рассмотрим синтаксис функции *ИНДЕКС*:

```
ИНДЕКС(reference; row_number; column_number)
```

```
INDEX(reference, row_number, column_number)
```

Поскольку в качестве первого аргумента у нас используется единственный столбец (*F3:F52*), для поиска нужной ячейки достаточно будет указать только аргумент *row_number*.

Функция СУММПРОИЗВ (SUMPRODUCT)

Функция *СУММПРОИЗВ* (*SUMPRODUCT*) является одной из наиболее мощных функций рабочего листа в Excel. Она предназначена для перемножения двух и более диапазонов значений и подсчета общей суммы полученных произведений. Простой пример использования функции *СУММПРОИЗВ* показан на рис. 1.12.

	A	B	C	D	E	F	G
1	Quantity	Price					
2	2	9.07					
3	3	4.46					
4	10	8.43					
5	9	4.53					
6							
7	Invoice Total:		156.59				
8							

Рис. 1.12. Функция *СУММПРОИЗВ*

В данном примере есть два столбца, представляющих количество проданных товаров и их цену. Применение функции *СУММПРОИЗВ* позволит перемножить соответствующие количество и цену построчно и суммировать полученные результаты. Такой же итог мы могли бы получить, если бы заполнили столбец *C* формулой $=A2*B2$ с протягиванием вниз, после чего вычислили бы общую сумму при помощи формулы $=СУММ(C2:C5)$ ($=SUM(C2:C5)$).

Функция *СУММПРОИЗВ* работает по принципу формул массива, о которых мы будем говорить в следующем разделе, и именно в этом состоит ее главная сила. С функцией *СУММПРОИЗВ* вы можете отфильтровывать ненужные значения, умножая их на ноль, чтобы они не оказывали влияния на итоговый

расчет. На рис. 1.13 показано, как можно использовать эту функцию для расчета итоговой суммы по одному товару за один месяц.

L6												
=SUMPRODUCT((\$C\$2:\$C\$3651=K6)*(MONTH(\$A\$2:\$A\$3651)=4)*(\$I\$2:\$I\$3651))												
	A	B	C	D	E	F	G	H	I	J	K	L
1	Date	Register	Item	RcptID	Units	Quantity	Price	Taxable	Total			
2	1/1/2021	POS-1-3	P53917	142387	Each	8	66.98	N	535.84			
3	1/1/2021	POS-4-2	P57772	142388	Pounds	10	75.43	N	754.30			
4	1/1/2021	POS-2-9	P73654	142389	Each	9	93.00	Y	837.00			
5	1/1/2021	POS-1-4	P23981	142390	10Pk	2	51.03	N	102.06			
6	1/1/2021	POS-4-8	P23981	142391	5Pk	10	57.56	N	575.60		P23981	5,000.32
7	1/1/2021	POS-2-1	P84419	142392	2PK	3	44.53	N	133.59			
8	1/1/2021	POS-1-4	P57772	142393	5Pk	3	59.39	Y	178.17			
9	1/1/2021	POS-2-3	P73302	142394	Pounds	4	69.06	N	276.24			
10	1/1/2021	POS-5-9	P39367	142395	5Pk	5	61.18	N	305.90			
11	1/1/2021	POS-4-4	P53917	142396	Feet	1	19.67	N	19.67			

Рис. 1.13. Фильтрация при помощи функции СУММПРОИЗВ

Приведенный здесь пример функции *СУММПРОИЗВ* можно условно разбить на три секции: две секции со сравнением и одну – с результатом. Магия функции *СУММПРОИЗВ* состоит в том, что, когда операция сравнения возвращает *TRUE*, при умножении это логическое значение интерпретируется как единица, а *FALSE* – как ноль. В первой секции ($\$C\$2:\$C\$3651=K6$) производится сравнение каждого значения в столбце *C* со значением в ячейке *K6* и возвращается *TRUE* или *FALSE*. Во второй секции ($\text{MONTH}(\$A\$2:\$A\$3651)=4$) сравнивается месяц каждого значения в колонке *A* со значением 4 и также возвращается *TRUE* или *FALSE*.

В заключительной секции ($\$I\$2:\$I\3651) никаких сравнений нет. Здесь мы просто получаем длинный список значений из столбца *I*. Каждая из этих секций возвращает массив из 3650 значений. При этом первые два массива содержат значения *TRUE* и *FALSE*, а последний массив – числа из столбца *I*.

После этого функция *СУММПРОИЗВ* перемножает значения массивов построчно. К примеру, для первой строки в массиве, соответствующей второй строке с данными, функция производит вычисление вида $FALSE * FALSE * 535,84$, поскольку значение в ячейке *C2* не равно значению в ячейке *K6*, и месяц из ячейки *A2* не равен четырем. Как мы уже знаем, логическое значение *FALSE* в математических операциях интерпретируется как ноль, а умножение любого числа на ноль дает ноль. Таким образом, результирующее значение по строке также будет равно нулю.

Для данных в строке 6 значений в столбце *C* совпадают с искомым значением в ячейке *K6*, но месяц не совпадает. В итоге получаем математическое выражение $TRUE * FALSE * 575,60$. Логическое значение *TRUE* будет интерпретировано как единица, а *FALSE* – как ноль. Результатом снова будет ноль. И только в строке 924 в первый раз нам встретится ситуация с двумя совпадениями в колонках *C* и *A*. В итоге формула приобретет вид $TRUE * TRUE * 47,12$, что в результате даст 47,12.

Таким образом, Excel будет возвращать ноль всякий раз, когда оба сравнения будут проваливаться. И только когда сравнения будут возвращать *TRUE*,

результат будет отличным от нуля. При этом совершенно не важно, будет ли у вас две секции сравнения или 20, Excel будет воспринимать каждое совпадение как единицу и соответствующим образом вычислять результат.

В этом отношении функция *СУММПРОИЗВ* похожа на сводные таблицы. В сводных таблицах происходит агрегирование исходных значений в местах пересечения значений в строках и столбцах. Преимущество сводных таблиц перед функциями состоит в том, что в них легче вносить изменения. В то же время результаты формульных вычислений будут обновляться автоматически при изменении исходных данных, тогда как сводную таблицу необходимо обновлять вручную.

Еще одно сходство между функцией *СУММПРОИЗВ* и сводными таблицами заключается в возможности агрегировать значения не по сумме, а по количеству элементов. Чтобы посчитать количество совпадений при помощи функции *СУММПРОИЗВ*, нужно включить в нее только секции сравнения. К примеру, чтобы подсчитать количество строк в нашей таблице по определенному товару за апрель, можно вызвать следующую функцию:

```
=СУММПРОИЗВ(($C$2:$C$3651=K6)*(МЕСЯЦ($A$2:$A$3651)=4))
=SUMPRODUCT(($C$2:$C$3651=K6)*(MONTH($A$2:$A$3651)=4))
```

Как видите, эта формула очень похожа на предыдущую, за тем лишь исключением, что в ней не присутствует указание на столбец *I*. Теперь в случае успешного сравнения в обоих столбцах Excel перемножит две единицы, что даст в результате единицу. В данном примере строк, удовлетворяющих нашему условию, будет ровно 18. В итоге будет возвращена сумма для такого количества единиц, что приведет к результату 18.

Также вы можете объединять в функции *СУММПРОИЗВ* сравнения по одному и тому же столбцу при помощи сложения. К примеру, если вам нужно найти все строки за апрель, в которых присутствует либо товар P23981, либо P73302, вы можете использовать следующую формулу:

```
=СУММПРОИЗВ(((C$2:C$3651="P23981")+($C$2:$C$3651="P73302"))*(МЕСЯЦ($A$2:$A$3651)=4)*($I$2:$I$3651))
=SUMPRODUCT(((C$2:C$3651="P23981")+($C$2:$C$3651="P73302"))*(MONTH($A$2:$A$3651)=4)*($I$2:$I$3651))
```

В данном примере присутствует две секции сравнения, ссылающихся на столбец *C*. Как вы понимаете, оба этих условия никогда не дадут результат *TRUE*, поскольку в ячейке не может быть указано два товара одновременно. Таким образом, остаются варианты математических выражений $0 + 0$, $0 + 1$ или $1 + 0$, что приведет к возвращению единицы, если ячейка ссылается на какой-либо из указанных товаров, и нуля, если ни один из искомых товаров в ячейке не присутствует. Поскольку оператор умножения в Excel обладает более высоким приоритетом по сравнению со сложением, мы должны заключить в скобки наше сравнение, чтобы оно было выполнено в первую очередь. Используйте эту технику со сложением только в случае обращения к одному и тому же столбцу. Если вы укажете разные столбцы, для каких-то строк могут быть возвращены два логических значения *TRUE*, что приведет к выполнению математического выражения $1 + 1 = 2$, а это нарушит все ваши расчеты.

Формулы массива

Функция *СУММПРОИЗВ*, которую мы обсуждали в предыдущем разделе, является особым случаем. Она просто суммирует результаты, но не умеет применять другие виды агрегации. *Формулы массива* используют те же вычисления, что и функция *СУММПРОИЗВ*, но с возможностью указать тип агрегации, применяемый к фильтруемым значениям. Особенностью работы с формулами массива является то, что при их вводе необходимо нажимать сочетание клавиш **Ctrl+Shift+Enter**, а не просто **Enter**, как в случае с той же функцией *СУММПРОИЗВ* и другими функциями рабочего листа.

Например, если вы хотите применительно к тому же набору данных вернуть максимальную сумму из столбца *I* по товару P23981 за апрель, то можете воспользоваться функцией *МАКС* (MAX) совместно с формулой массива, как показано ниже и на рис. 1.14:

```
=МАКС((С$С2:С$3651=K6)*(МЕСЯЦ($A$2:$A$3651)=4)*($I$2:$I$3651))
=MAX((C$2:C$3651=K6)*(MONTH($A$2:$A$3651)=4)*($I$2:$I$3651))
```

На рис. 1.14 видно, что наше выражение в строке формул оказалось заключено в фигурные скобки. При этом мы не вводили эти скобки. Они автоматически появились при нажатии сочетания клавиш **Ctrl+Shift+Enter** вместо **Enter**. Если написанная вами формула массива возвращает неправильный результат, первым делом необходимо проверить наличие фигурных скобок в строке формул. В случае их отсутствия нажмите на клавишу **F2**, чтобы перейти в режим редактирования, после чего снова нажмите сочетание клавиш **Ctrl+Shift+Enter**.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Date	Register	Item	RcptID	Units	Quantity	Price	Taxable	Total			
2	1/1/2021	POS-1-3	P53917	142387	Each	8	66.98	N	535.84			
3	1/1/2021	POS-4-2	P57772	142388	Pounds	10	75.43	N	754.30			
4	1/1/2021	POS-2-9	P73654	142389	Each	9	93.00	Y	837.00			
5	1/1/2021	POS-1-4	P23981	142390	10Pk	2	51.03	N	102.06			
6	1/1/2021	POS-4-8	P23981	142391	5Pk	10	57.56	N	575.60		P23981	797.70
7	1/1/2021	POS-2-1	P84419	142392	2PK	3	44.53	N	133.59			
8	1/1/2021	POS-1-4	P57772	142393	5Pk	3	59.39	Y	178.17			
9	1/1/2021	POS-2-3	P73302	142394	Pounds	4	69.06	N	276.24			

Рис. 1.14. Формула массива для нахождения максимальной суммы по условию

Как и в случае с функцией *СУММПРОИЗВ*, эта формула массива содержит две секции сравнения и одну секцию со значениями. После перемножения результатов выполнения трех вычислений по тем же правилам, что и раньше, будет рассчитана не сумма, а максимальное значение. Фактически если в этой формуле поменять функцию *МАКС* на *СУММ* (SUM), вы получите тот же результат, что и в примере с функцией *СУММПРОИЗВ*.

Нахождение минимального значения в списке отфильтрованных элементов требует использования несколько иной техники. Поскольку значения,

которые вы исключаете из выбора, приравниваются к нулю, минимальное значение в списке из положительных чисел всегда будет равняться нулю. К счастью, функция *МИН* (MIN) игнорирует текстовые выражения, и мы можем использовать это свойство для включения в выборку только ненулевых значений. В следующей формуле мы вычислим минимальное значение по товару P23981 за апрель:

```
=МИН(ЕСЛИ((С$2:С$3651=K15)*(МЕСЯЦ($A$2:$A$3651)=4)=0;"";$I$2:$I$3651))
=MIN(IF((С$2:С$3651=K15)*(MONTH($A$2:$A$3651)=4)=0,"",$I$2:$I$3651))
```

Две условные секции мы разместили здесь внутри условного аргумента функции *ЕСЛИ* (IF). Если в результате перемножения результатов сравнения мы получим ноль, в список фильтруемых значений будет вставлена пустая строка (две двойные кавычки). В противном случае будет вставлено содержимое столбца *I*. Поскольку функция *МИН* игнорирует текстовые выражения, все значения, полученные в результате неудачных сравнений, будут отброшены, а минимальное значение будет вычислено на основе совпадений, завершившихся успешно.

Умные таблицы

Умные таблицы (table) представляют собой особую область с данными на рабочем листе. Эти области можно сравнить с таблицами баз данных, но они обладают лишь некоторыми их свойствами. Двумя главными преимуществами применения умных таблиц является то, что формат и формулы в них распространяются на все строки определенного столбца, и ссылки на таблицу автоматически адаптируются под ее изменяющиеся размеры. На рис. 1.15 показана небольшая умная таблица и диаграмма, построенная на ее основании.

Если в таблице изменится количество строк или столбцов, диаграмма будет перерисована автоматически. Таким образом, график не привязан жестко к диапазону *A1:C6*, а ссылается на таблицу с именем *Table1* и следит за изменениями ее размеров.

Первое, что нужно сделать, – это преобразовать диапазон на рабочем листе Excel в таблицу. Для этого выделите диапазон и на вкладке **Вставка** (Insert) нажмите на кнопку **Таблица** (Table). В результате появится диалоговое окно **Создать таблицу** (Create Table), показанное на рис. 1.16.

В открывшемся окне вы видите ссылку на диапазон создаваемой таблицы, а также можете указать, является ли первая строка в указанном диапазоне строкой с заголовками. Нажмите на кнопку **ОК**, убедившись, что все заполнено правильно. После этого выделите созданную таблицу на рабочем листе и на вкладке **Вставка** (Insert) нажмите на кнопку **Рекомендуемые диаграммы** (Recommended Charts) и выберите вариант **Гистограмма с группировкой** (Clustered Column) для вставки соответствующего графика. Вы можете изменить заголовок диаграммы, как сделал я, но для данного примера это вовсе не обязательно.

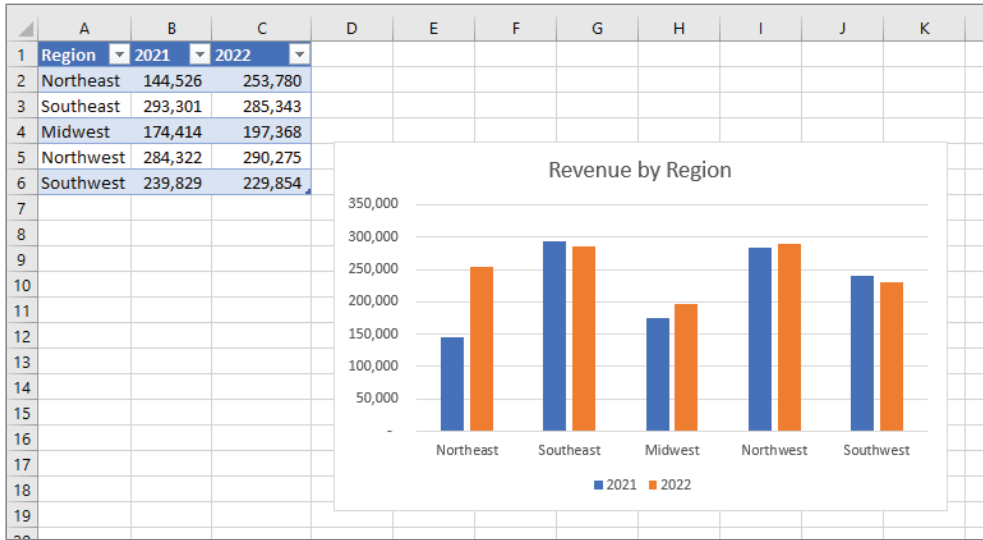


Рис. 1.15. Диаграмма на базе умной таблицы

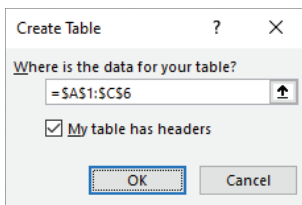


Рис. 1.16. Диалоговое окно создания таблицы

Теперь представьте, что вам необходимо добавить в таблицу данные по следующему году и вывести регион *Texas* из состава *Southwest*. Работая с диапазонами в Excel, вам пришлось бы создавать новый ряд для нового года и изменить существующие ряды под расширенный диапазон. Но, поскольку вы используете умные таблицы, то можете просто добавить новые данные, и диаграмма автоматически подстроится.

Чтобы это сделать, встаньте в ячейку *D1* и введите номер года 2023. На рис. 1.17 видно, что Excel автоматически добавил новую колонку в таблицу и создал ряд на диаграмме.

Теперь введите слово *Texas* в ячейку *A7*. Excel автоматически добавит новую строку в таблицу, в которую вы сможете ввести нужные данные. На рис. 1.18 показаны обновленная таблица и график.

Excel сам добавил новый столбец на диаграмму для нового региона, а ряд 2023 автоматически заполнился данными из таблицы. При этом значения, введенные для штата *Texas*, приняли тот же формат, что и остальные данные в существующих столбцах таблицы. Что касается нового столбца 2023, то данные в нем не были отформатированы, и вам придется вручную выполнить эту операцию. На диаграмме вам не пришлось ничего менять вручную. Все, что вы сделали, – это добавили новую информацию в таблицу.

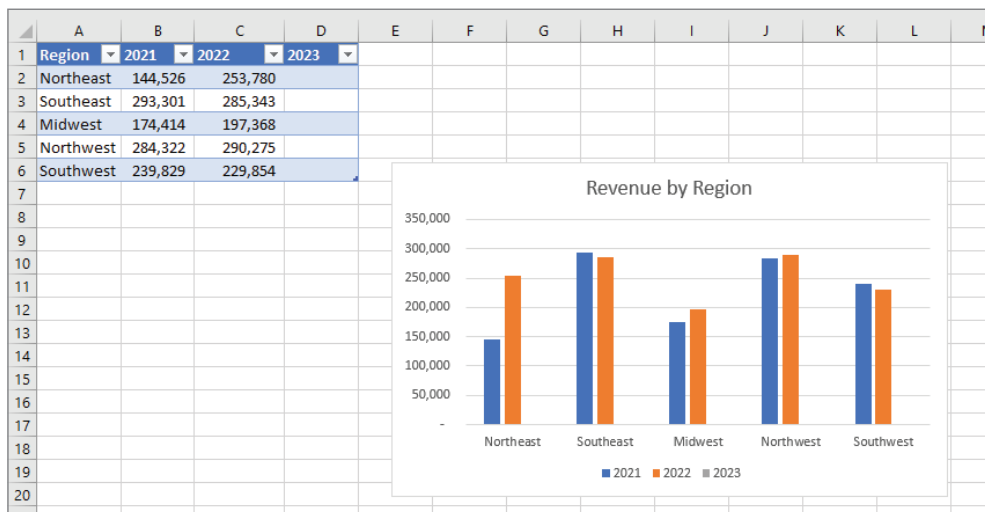


Рис. 1.17. Таблица расширяется под новые данные

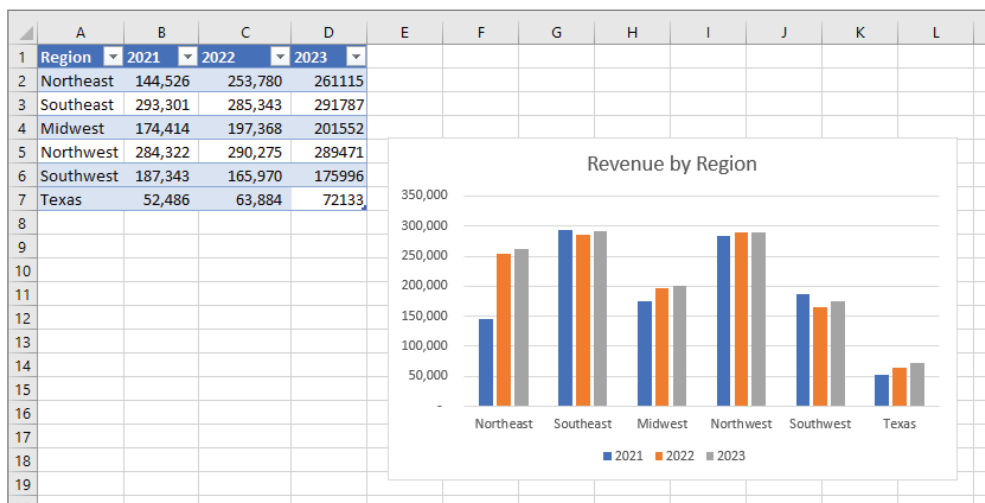


Рис. 1.18. Диаграмма адаптировалась под новую таблицу

Структурированные ссылки на таблицы

Как мы уже видели, диаграмма из предыдущего примера автоматически подстраивается под данные в таблице при ее расширении. Формулы, использующие при обращении к таблицам *структурированные ссылки* (structured table referencing), будут учитывать все изменения, происходящие в источнике. Давайте для этого примера изменим название нашей умной таблицы на *tblRevenue*. Для этого щелкните по таблице и на вкладке **Конструктор таблиц** (Table Tools Design) измените имя в поле **Имя таблицы** (Table Name).

Чтобы создать свою первую формулу с использованием структурированных ссылок, щелкните по свободной ячейке на рабочем листе, начните вводить формулу, например так: =СУММ((SUM), и щелкните по любой ячейке из диапазона B2:B7. Стоит отметить, что если вы начнете вводить данные в соседнюю с таблицей ячейку, Excel решит, что вы собираетесь ее расширить, и будет действовать соответственно. Так что лучше делать это в ячейке, отстоящей от таблицы. По умолчанию Excel преобразует вашу формулу для использования структурированных ссылок, как показано ниже:

```
=СУММ(tblRevenue[2021])
```

```
=SUM(tblRevenue[2021])
```

Как видите, вместо обращения к диапазону B2:B7 Excel преобразовал аргумент функции в tblRevenue[2021], ссылаясь на столбец 2021 в таблице с именем tblRevenue. Если вы готовите исходные данные для дашбордов, то можете активно использовать структурированные ссылки, чтобы ваши формулы всегда покрывали все ваши данные.

При этом вовсе не обязательно указывать на данные мышью, если вы хотите воспользоваться структурированными ссылками. Вы вполне можете вводить их и вручную. Начните ввод в ячейку: =СУММ(tbl. На рис. 1.19 показан выпадающий список, который предложит вам Excel на выбор.

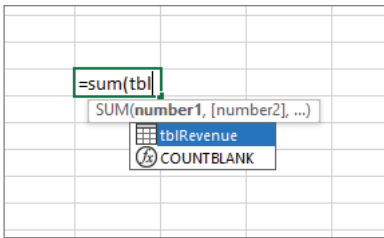


Рис. 1.19. Выпадающий список с вариантами формул в Excel

Если в вашей рабочей книге больше одной таблицы и вы соблюдаете определенные правила их именования (например, начинаете их названия с *tbl*), то этот метод позволит вам получить полный список доступных таблиц на выбор. В нашем случае таблица только одна и она автоматически выделена в списке, поскольку стоит первой. Выделив нужную таблицу в списке, вы можете нажать на клавишу **Tab**, чтобы завершить ввод автоматически, – без необходимости вводить все имя таблицы вручную. После ввода имени таблицы откройте квадратную скобку, чтобы раскрыть выпадающий список опций, доступных в этой таблице. На рис. 1.20 показан список опций для таблицы *tblRevenue*.

Первый вариант в списке – одиночный знак @. При использовании формул внутри таблиц этот символ указывает Excel, что необходимо вернуть значение из той же строки, в которой введена формула. К примеру, ссылка [*@2021*] вернет значение в столбце с именем 2021, расположенное в той же строке. В то же время ссылка [*2021*] вернет весь столбец.

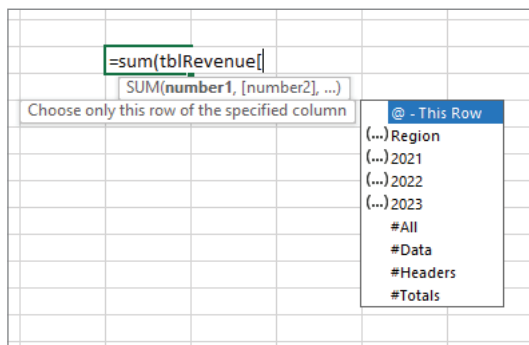


Рис. 1.20. Опции для завершения ввода формулы

Следующие четыре варианта представляют собой адресацию столбцов в таблице. Excel осведомлен о структуре таблиц и автоматически генерирует именованные диапазоны для столбцов. Последние четыре опции служат для обращения к *особым областям* (special area) таблицы. Ссылка *#Все* (*#All*) позволяет обратиться ко всей таблице в целом, включая заголовки и итоги. *#Данные* (*#Data*) обращается к тем же данным, что и *#Все*, за исключением заголовков и итогов. Ссылки *#Заголовки* (*#Headers*) и *#Итоги* (*#Totals*) служат для обращения к заголовкам и итогам соответственно. Вы можете перемещаться по выпадающему списку при помощи клавиш курсора и делать выбор, нажимая на клавишу **Tab**.

К примеру, если вы переместитесь на пункт *#Данные*, нажмете на клавишу **Tab** и введете закрывающую квадратную и круглую скобки, вы получите следующую формулу, суммирующую все данные в таблице:

```
=СУММ(tblRevenue[#Данные])
```

```
=SUM(tblRevenue[#Data])
```

Текст по столбцам

Когда мы имеем дело с исходными данными за пределами Excel, то часто сталкиваемся с тем, что информация собрана по строкам, а не аккуратно разбита на столбцы. Встроенный в Excel инструмент под названием *Текст по столбцам* (Text to Columns) поможет вам привести подобные данные в вид, удобный для анализа. На рис. 1.21 показан пример исходных данных, вставленных в колонку *A* на рабочем листе. Это, по сути, столбчатые данные, разделенные запятыми и представленные в виде одной колонки.

Для разделения этих данных на столбцы необходимо выполнить следующие действия.

1. Выделите информацию и на вкладке **Данные** (Data) и нажмите на кнопку **Текст по столбцам** (Text to Columns).
2. В открывшемся диалоговом окне **Мастер распределения текста по столбцам** (Convert Text to Columns Wizard), показанном на рис. 1.22, выберите вариант **С разделителями** (Delimited) и нажмите на кнопку **Далее** (Next).

	A	
1	January,69,47,38,97,36	
2	February,14,47,62,90,67	
3	March,42,78,83,94,82	
4	April,72,15,65,61,21	
5	May,56,70,84,22,33	
6	June,37,38,97,39,34	
7	July,18,96,76,91,88	
8	August,13,59,79,43,80	
9	September,70,36,33,77,69	
10	October,29,58,25,99,87	
11	November,56,83,88,39,24	
12	December,93,99,52,33,97	

Рис. 1.21. Исходные данные, собранные вместе

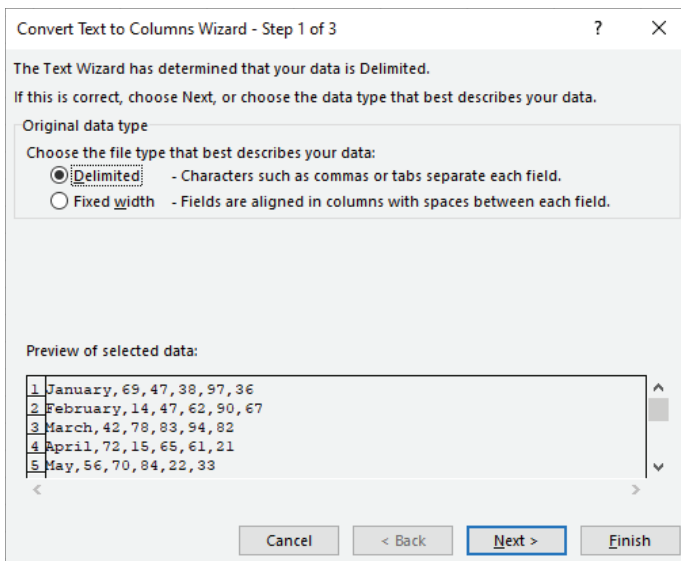


Рис. 1.22. Мастер преобразования текста в столбцы

- Снимите флажок **Знак табуляции** (Tab) и установите флажок **Запятая** (Comma). В нижней части окна Excel покажет предварительный вариант разделения данных на столбцы, как видно по рис. 1.23. Нажмите на кнопку **Далее** (Next).
- На заключительном шаге вы можете отформатировать столбцы. Обычно Excel нормально справляется с конвертированием чисел и дат. Если в исходных данных есть числа с лидирующими нулями, Excel избавится от этих нулей, если не форматировать соответствующий столбец как текст. На рис. 1.24 показано, как данные были разбиты на столбцы.

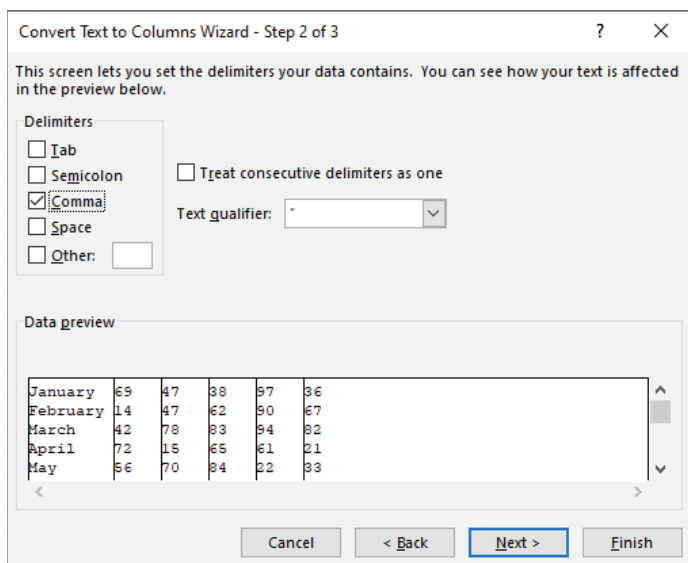


Рис. 1.23. Распределение исходных данных с разделителями

	A	B	C	D	E	F
1	January	69	47	38	97	36
2	February	14	47	€2	90	€7
3	March	42	78	83	94	82
4	April	72	15	€5	€1	21
5	May	56	70	84	22	33
6	June	37	38	97	39	34
7	July	18	96	76	91	88
8	August	13	59	79	43	80
9	September	70	36	33	77	69
10	October	29	58	25	99	87
11	November	56	83	88	39	24
12	December	93	99	52	33	97

Рис. 1.24. Данные, разделенные по колонкам

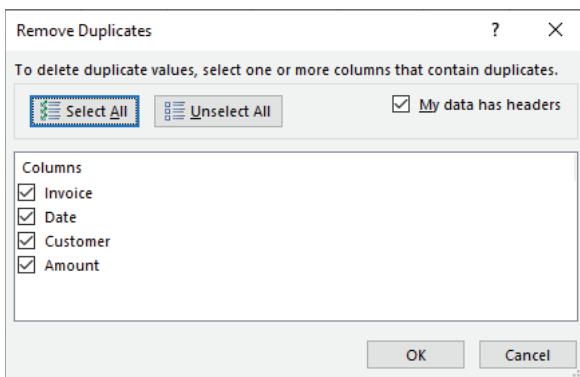
Удаление дубликатов

Еще одной распространенной проблемой при работе с сырыми данными являются дубликаты. Встроенный инструмент Excel *Удалить дубликаты* (Remove Duplicates) позволяет быстро и легко избавиться от повторов в данных без задействования сводных таблиц и сложных формул. На рис. 1.25 показаны данные с дубликатами.

	A	B	C	D
1	Invoice	Date	Customer	Amount
2	IN001000	12/1/2021	Atlantic Northern	847.97
3	IN001000	12/1/2021	Atlantic Northern	847.97
4	IN001003	12/21/2021	Big T Burgers and Fries	763.62
5	IN001003	12/21/2021	Big T Burgers and Fries	763.62
6	IN001006	12/25/2021	Sixty Second Avenue	516.08
7	IN001007	12/4/2021	Mainway Toys	706.85
8	IN001008	12/19/2021	Carrys Candles	726.68
9	IN001013	12/2/2021	The Legitimate Businessmens Club	666.51
10	IN001014	12/14/2021	Zevo Toys	780.58
11	IN001016	12/15/2021	North Central Positronics	370.02
12	IN001022	12/25/2021	Fake Brothers	142.65
13	IN001022	12/25/2021	Fake Brothers	142.65
14	IN001022	12/25/2021	Fake Brothers	142.65
15	IN001023	12/14/2021	General Services Corporation	931.21
16	IN001024	12/3/2021	General Products	918.33
17	IN001026	12/6/2021	Taco Grande	303.72
18	IN001032	12/10/2021	Videlectrix	939.56
19	IN001036	12/15/2021	Wernham Hogg	909.72
20	IN001037	12/30/2021	Sample, inc	563.13
21	IN001040	12/24/2021	Taco Grande	472.36
22	IN001040	12/24/2021	Taco Grande	472.36
23	IN001042	12/7/2021	Roxxon	909.71
24	IN001046	12/1/2021	LuthorCorp	121.00
25	IN001048	12/15/2021	Videlectrix	897.77

Рис. 1.25. Данные с дублирующимися значениями

Чтобы исключить повторяющиеся значения, перейдите на вкладку **Данные** (Data) и нажмите на кнопку **Удалить дубликаты** (Remove Duplicates). Это приведет к открытию одноименного диалогового окна, показанного на рис. 1.26.

Рис. 1.26. Диалоговое окно **Удалить дубликаты**

В диалоговом окне показаны все столбцы в исходных данных, и вы можете сами выбрать, в каких из них удалять дубликаты. В примере, представленном на рисунках, повторяющиеся записи присутствуют во всех столбцах, так что

выберем их все. В других случаях может понадобиться определять дубликаты всего в одном или двух столбцах. Чтобы избежать удаления важных данных, постарайтесь задействовать как можно больше столбцов. При нажатии на кнопку **ОК** Excel отобразит сообщение, показанное на рис. 1.27, с информацией о количестве удаленных строк.

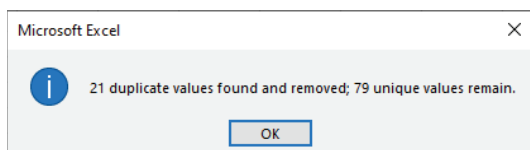


Рис. 1.27. Данные о количестве удаленных строк

Нажмите на кнопку **ОК**, чтобы закрыть окно, и в выбранном диапазоне останутся только уникальные записи. Вы всегда можете отменить выполненное действие, если увидите, что допустили какие-то ошибки.

Построение дашбордов

Если вы обычный пользователь Excel вроде меня, вам наверняка должна доставлять удовольствие работа с цифрами и сводными таблицами. Возможно, вы не слишком комфортно чувствуете себя в разработке визуальных элементов. К счастью, сейчас есть очень много примеров визуального отображения информации на различных сайтах и в новостных статьях. *Создание дашбордов* – довольно древняя наука, и за это время люди выработали вполне четкие правила, придерживаясь которых, вы сможете создавать приятные глазу визуализации. А когда освоите их все, вероятно, превратитесь в ярого критика любых дашбордов, которые будут встречаться вам на пути.

Первое правило дашборда гласит: не отклоняйся от общей мысли! Если вы тщательно провели этап планирования, значит, у вас уже есть понимание того, с какими данными вы имеете дело и что нужно конечным пользователям. Каждый новый визуальный элемент на вашем дашборде должен стать продолжением истории, которую вы пытаетесь донести до зрителя. Если вы можете рассказать свою историю при помощи всего одного простого графика, не стоит перегружать дашборд лишними элементами. Не существует какого-то минимального количества элементов, при котором ваш визуальный отчет может называться дашбордом.

Не позволяйте имеющимся у вас данным определять, какие элементы будут размещены на дашборде. С большой вероятностью в вашем распоряжении будет больше исходных данных, чем вам необходимо. Когда у вас есть под рукой хорошо агрегированные и причесанные данные, трудно удержаться от соблазна воткнуть их на свободное место дашборда, – просто потому, что это не требует сил. Но силы потребуются пользователю дашборда, чтобы не сбиться с общей мысли, заключенной в дашборде.

Разумеется, это не означает, что каждый элемент на дашборде должен быть абсолютным хитом. В любом визуальном отчете есть более и менее значимые элементы. Диаграммы, дополняющие общий сюжет, но не являющиеся критически важными для истории в целом, могут быть включены в дашборд, если не добавят лишнего хаоса. Со временем вы начнете чувствовать баланс между полнотой информации, представленной на дашборде, и его простотой.

Организация элементов

Существует большое количество мнений и информации о том, как *организовывать элементы* на странице, – особенно это касается компоновки веб-страниц. Полагаю, есть три основополагающих принципа, охватывающих все эти мнения, – это размер, цвет и размещение.

Самый большой визуальный элемент на дашборде будет привлекать больше всего внимания. Таким образом, вы должны сделать более заметными элементы, несущие действительно важную информацию. Конечно, он не должен быть единственной доминантой на странице, но по его размеру пользователь должен понять, что представленная информация особенно важна для принятия решения. Вы можете использовать пустое пространство вокруг элемента для визуального увеличения его размера. На рис. 1.28 показан дашборд с тремя элементами, один из которых выделен при помощи размера.

Визуальные элементы в ярком контрастном исполнении будут восприниматься пользователем как более важные по сравнению с диаграммами, выполненными в мягких пастельных тонах. В то же время, если на вашем дашборде все элементы будут контрастными, страница будет выглядеть нелепо из-за своей аляповатости, – лучше с этим не переусердствовать. Конечно, я не выступаю в защиту серых и скучных отчетов. Но внимание пользователя лучше привлекать незначительными и редкими отхождениями от общей цветовой гаммы. Так или иначе, в целом вам необходимо использовать ограниченное количество цветов на дашборде, при этом цвета должны идеально сочетаться друг с другом. Но это не говорит о том, что вы не можете выделить важный элемент, немного отойдя от принятых правил.

Расположение визуальных элементов на странице должно отвечать следующему общепринятому требованию: наиболее важные элементы должны располагаться в верхнем левом углу дашборда и по мере снижения значимости смещаться к правому нижнему. Вспомните, как вы читаете текст на странице. Вы начинаете с левого верхнего угла и постепенно двигаетесь ниже. Таким образом, именно этот угол в отчетах привлекает наибольшее внимание зрителя. Согласно еще одному шаблону, необходимо размещать наиболее важный элемент визуализации по центру дашборда и окружать его менее значимыми. Этот шаблон может быть действенным, если один элемент на вашем дашборде существенно превосходит остальные по значимости.

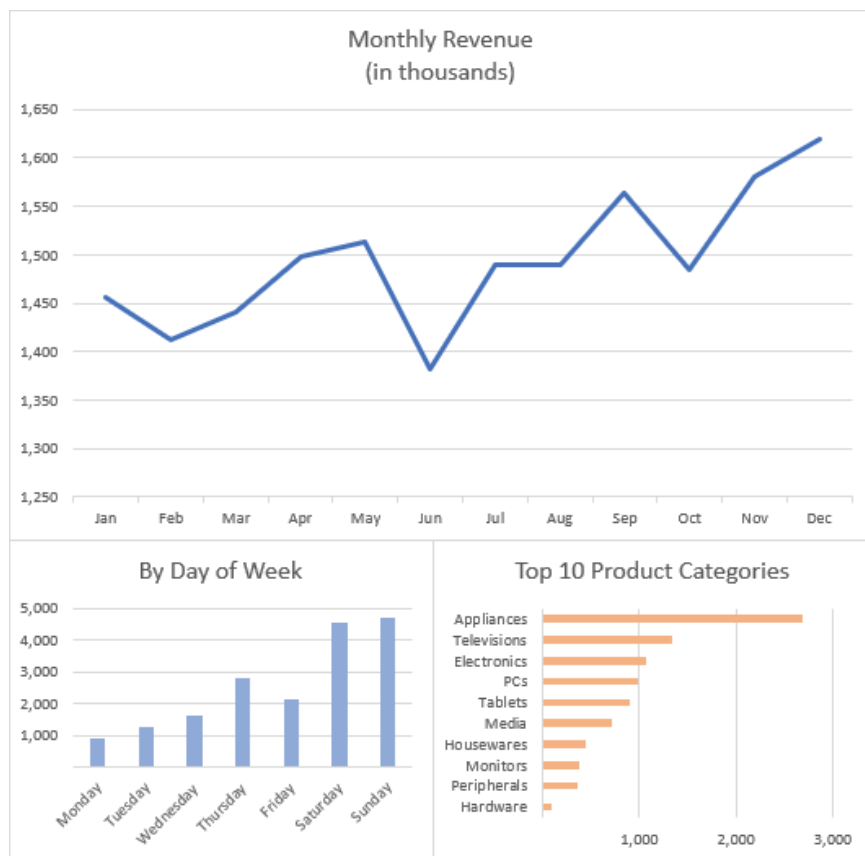


Рис. 1.28. Дашборд с элементами разных размеров

Разнообразие элементов

Обычно дашборд состоит из визуальных элементов разных типов. Если в отчете будут присутствовать только линейные диаграммы, пользователю придется сильно потрудиться, чтобы хоть что-то в них понять. При этом есть одно исключение из этого правила, и работает оно тогда, когда для каждого элемента на странице используются одни и те же исходные данные, различающиеся одним измерением. К примеру, это могут быть продажи товаров по регионам. В этом случае похожие графики на странице дадут пользователю понять, что в основе лежат единые данные. Если же исходные данные для элементов отличаются более чем на одно измерение, – например, выводится информация в штуках и долларах на шкале времени, необходимо задействовать разные типы диаграмм. На рис. 1.29 показан дашборд, в котором используются восемь элементов одного типа, отличающихся единственным измерением.

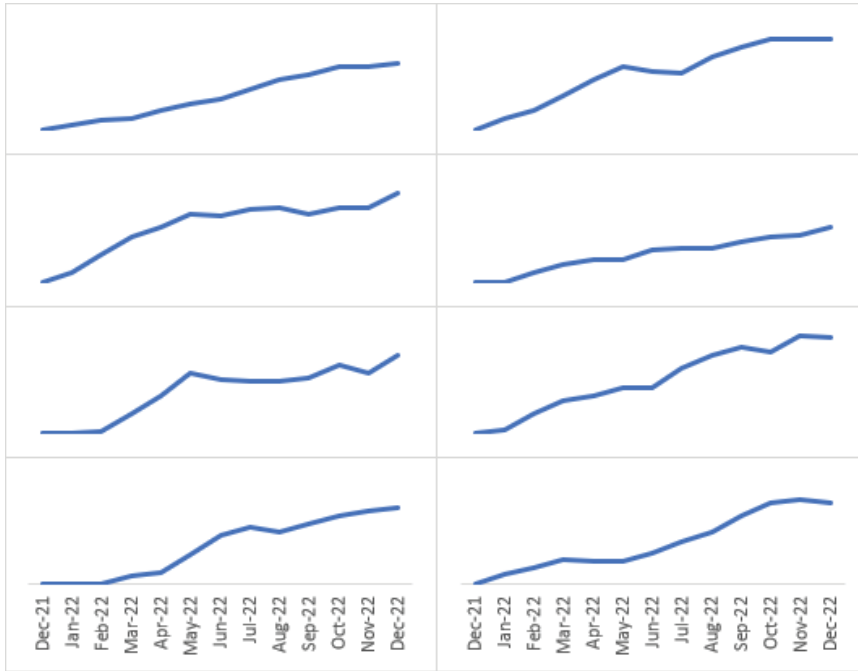


Рис. 1.29. Дашборд, состоящий из однотипных визуальных элементов

В третьей части книги мы подробнее поговорим о выборе типов визуализации для дашборда и о том, что эти типы должны в полной мере отвечать требованиям данных и сюжету рассказываемой истории. Но случаются и наложения. К примеру, линейные графики идеально подходят для демонстрации изменений с течением времени, тогда как столбчатые диаграммы наиболее эффективны в случае с дискретными периодами, такими как месяцы.

Это не означает, что все элементы дашборда должны быть обязательно разных типов. Но лучше не размещать однотипные диаграммы вплотную друг к другу и не делать их одного цвета и размера. Также здесь полезно вспомнить правило, о котором мы уже говорили, – не размещать в отчете данные просто потому, что они у вас есть. Не включайте в дашборд диаграммы, исходя из их популярности или потому что только недавно узнали об их существовании. Элементы, выведенные в отчете, должны четко следовать вашему замыслу и помогать рассказывать историю.

Отображение трендов

Если вы хотите показать на графике *линию тренда* (trendline), вы всегда можете это сделать. Но вы должны помнить, что некоторым диаграммам линия тренда не добавляет никакой полезной информации. На рис. 1.30 показан график с пунктирной линией тренда. Но тенденция здесь легко различима и без вспомогательной линии, так что она тут просто не нужна.

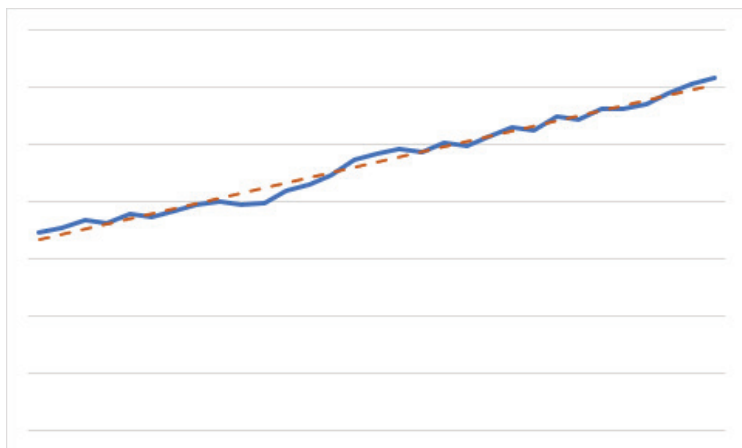


Рис. 1.30. Линия тренда, не помогающая считыванию информации

В других случаях линия тренда может оказаться очень полезной. Если в ваших данных наблюдается изменчивость, затрудняющая определение общей тенденции, и линия тренда хорошо согласуется с канвой общего рассказа, смело добавляйте ее. На рис. 1.31 показан график дневных продаж. Периодические спады во время выходных существенно затрудняют отслеживание роста продаж, и линия тренда здесь подойдет как нельзя лучше.

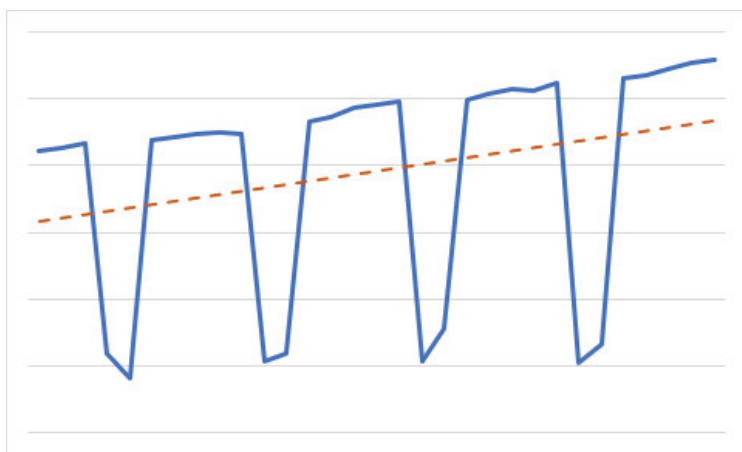


Рис. 1.31. Линия тренда на графике с периодической изменчивостью данных

Чтобы добавить линию тренда на график, щелкните правой кнопкой мыши по ряду и выберите пункт **Добавить линию тренда** (Add Trendline). По умолчанию будет добавлен *линейный тренд* (linear trendline), а справа появится панель **Формат линии тренда** (Format Trendline), показанная на рис. 1.32.

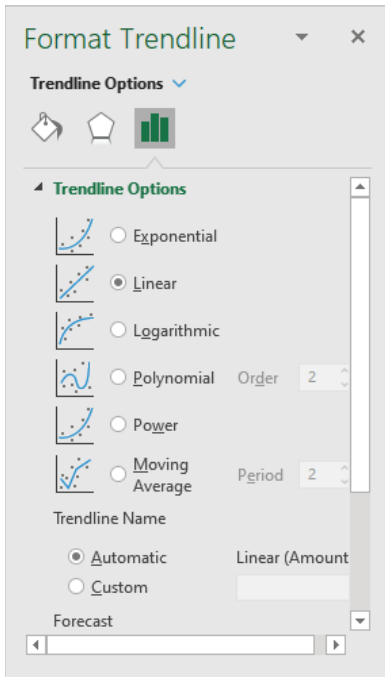


Рис. 1.32. Панель
Формат линии тренда

Вы можете указать на выбор одного из следующих *типов линии тренда*:

- *экспоненциальный* (Exponential): идеально подходит для восходящих или нисходящих данных нарастающими темпами;
- *линейный* (Linear): лучший выбор для простых линейных данных;
- *логарифмический* (Logarithmic): подходит для данных, которые растут или убывают, после чего стабилизируются;
- *полиномиальный* (Polynomial): идеален для данных с высокими колебаниями;
- *степенной* (Power): подходит для восходящих или нисходящих данных постоянными темпами;
- *скользящее среднее* (Moving Average): применяется к наборам данных с постоянными перепадами, которые необходимо сгладить.

Форматирование дашбордов

Форматирование (formatting) может отличать просто хороший дашборд от великолепного. К счастью, Excel предлагает богатый выбор опций форматирования для большинства поддерживаемых типов диаграмм. Но, используя те же самые опции, можно также испортить дашборд, уведя его в сторону от задуманной истории.

С форматированием нужно следовать тому же простому правилу и делать его максимально простым, насколько это возможно. Перебор с форматированием будет отвлекать пользователя от правильного интерпретирования данных и приведет к снижению эффективности диаграмм. На рис. 1.33 показана диаграмма с чрезмерным форматированием, а на рис. 1.34 – та же диаграмма, но в стиле минимализма. И если первая показывает, скорее, возможности Excel в отношении форматирования графиков, вторая ставит во главу угла сами данные.

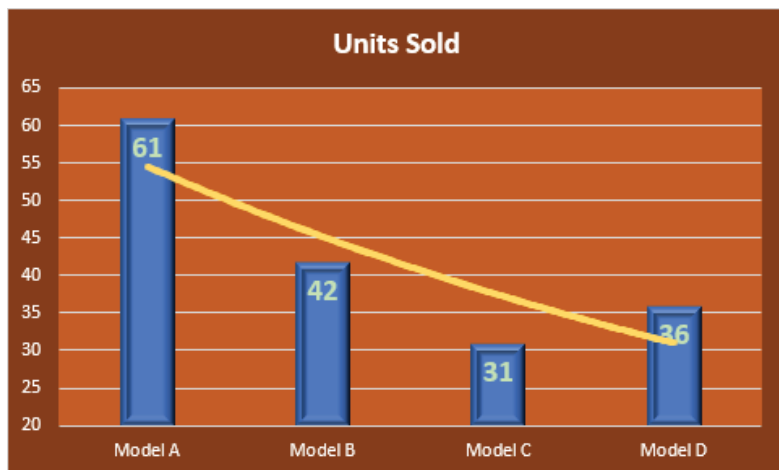


Рис. 1.33. Излишнее форматирование отвлекает от сути визуализации

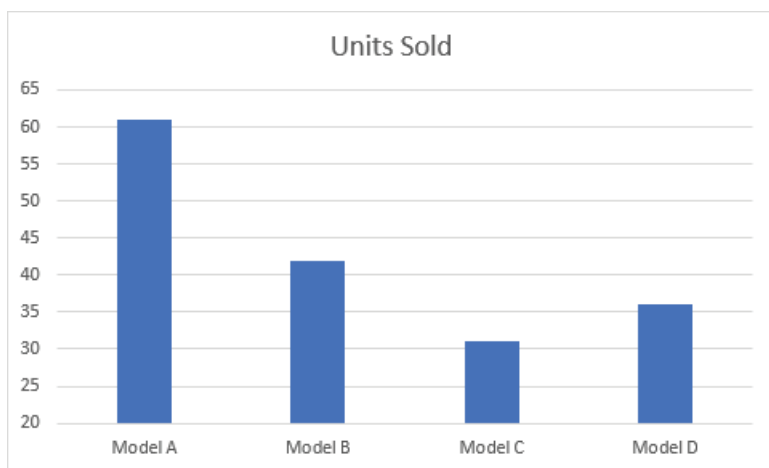


Рис. 1.34. Простое форматирование не мешает сконцентрироваться на сути

Главным врагом при форматировании визуальных элементов в Excel является *трехмерность* (3D). Применение трехмерных эффектов почти всегда

отвлекает зрителя, тем самым затрудняя донесение до него важной информации. Например, при использовании трехмерности столбики в первом ряду будут более объемными, а стоящие позади будут частично заслонены. Графики 3D выглядят привлекательно, но обычно мешают конечному пользователю правильно интерпретировать данные и делать верные выводы.

Ниже перечислим другие опции форматирования, которые применительно к диаграммам стоит использовать с осторожностью или не использовать вовсе:

- *цвет заливки* (background color): для фоновой заливки диаграмм лучше всего подходит белый или светлые цвета. Использование черного цвета или темных оттенков приведет к снижению контраста между значимыми элементами диаграммы и фоном. Заливка никогда не несет информационной нагрузки, поэтому не стоит ее особенно выделять;
- *цвет элементов* (foreground color): не используйте много цветов при оформлении элементов диаграмм. Выделение цветом необходимо применять точно – например, для привлечения внимания к важным точкам данных. *Цветовая палитра* (color pallet) в Excel, показанная на рис. 1.35, помогает выбрать гармоничную цветовую гамму путем предложения различных уровней светлого и темного тонов. В галерее тем Excel, представленной на рис. 1.36, есть заготовленные сочетания хорошо подходящих друг другу цветов;



Рис. 1.35. Цветовая палитра Excel

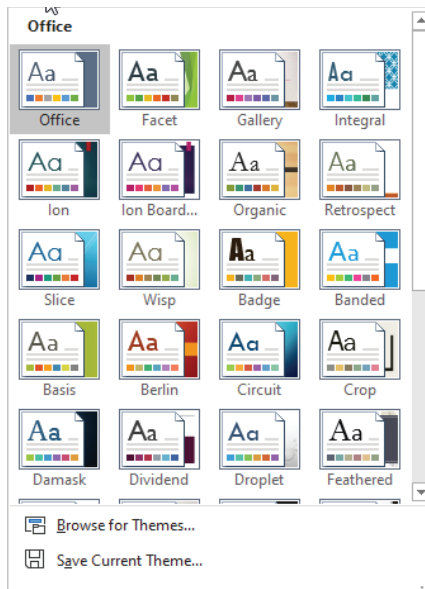


Рис. 1.36. Галерея тем в Excel

- *градиенты* (gradient): градиентная заливка позволяет закрашивать элементы на диаграмме с присутствием цветовых переходов. В таком подходе нет ни малейшего смысла, и его стоит избегать;

- *рисунки* (pictures): этот тип заливки позволяет окрашивать элементы диаграммы при помощи повторяющихся изображений. Это не только не имеет смысла, но и значительно снижает эффективность элементов визуализации. Так что изображения также не стоит использовать при оформлении диаграмм;
- *эффекты* (effects): на панели **Формат ряда данных** (Format Data Series) для большинства типов диаграмм присутствует отдельная вкладка **Эффекты** (Effects), на которой можно настроить такие особенности форматирования, как *тень* (Shadow), *свечение* (Glow) и *сглаживание* (Bevel). Если использовать эти эффекты с умом, они могут добавить графику привлекательности. Но зачастую разработчики ими злоупотребляют, что делает диаграммы трудночитаемыми. В основном вы не должны использовать в работе эту вкладку – обращайтесь к ней только в случае крайней необходимости;
- *границы* (borders): границами называются линии, обрамляющие *область построения* (plot area) на диаграмме. В этих линиях почти никогда нет нужды. Оси сами по себе обрамляют область построения. Но даже если по тем или иным причинам вы скроете оси, в границах необходимость может не возникнуть. Решив все же воспользоваться границами, выбирайте для них более светлые тона, чтобы элементы данных выделялись на их фоне;
- *линии сетки* (gridlines): на большинстве типов диаграмм Excel по умолчанию включает линии сетки. Однако зачастую они не несут никакой смысловой нагрузки. Если на вашей диаграмме присутствует важная точка данных вдалеке от линий сетки, вы можете снабдить ее меткой или выделить цветом. Сделав выбор в пользу линий сетки, избегайте контрастных цветов, чтобы они не перетягивали на себя внимание зрителя;
- *подписи данных* (data label): если каждая точка данных на диаграмме будет сопровождаться меткой или подписью, это будет таблица, которую будет очень трудно читать. Используйте подписи только для важных отметок на графике, например для минимума и максимума. Подписи также могут быть использованы для описания неожиданных значений на графике. На рис. 1.37 показан график, на котором подпись использована для объяснения резкого спада производства;
- *оси* (axes): в большинстве случаев оси используются на графике для предоставления дополнительного контекста и масштаба. Но иногда роль осей полностью нивелируется – например, когда на диаграмме присутствуют большие колебания. Не оставляйте оси на графике только потому, что Excel добавил их по умолчанию. Подумайте о том, как оси вписываются в вашу историю, и помогут ли они пользователю правильно интерпретировать информацию.

Избегайте добавления картинок на ваши диаграммы. Диаграммы уже сами по себе являются картинками, и нет смысла сопровождать их другими изображениями. Это может отвлечь конечного пользователя от важной информации.

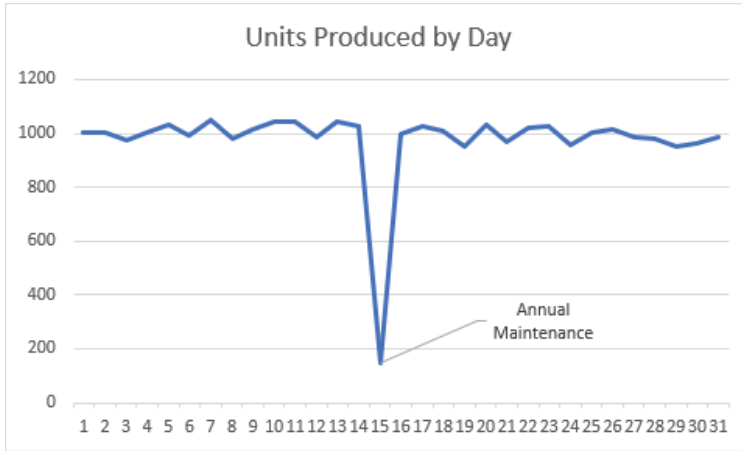


Рис. 1.37. Подписи данных могут быть использованы для неожиданных спадов на графике

Форматирование чисел

Ваши диаграммы почти всегда будут содержать в себе числовые данные, будь то на осях или в подписях к точкам на графике. По умолчанию Excel использует форматирование из источника данных, но это не всегда то, что нам нужно. Всегда используйте запятые, чтобы облегчить чтение числовых данных.

Форматирование чисел на графиках должно выполняться с учетом их специфики. К примеру, если вы выводите информацию о доходах в десятках миллионов долларов, не стоит показывать центы. В Excel используется очень гибкая система форматирования, позволяющая настроить отображение значений до мелочей. К примеру, вы можете показывать число 17,483,262 как 17.4 и использовать заголовок или подписи на осях, чтобы дать понять пользователю, что речь идет о миллионах. На диаграмме, показанной на рис. 1.38, используется эта техника.

Чтобы настроить форматирование чисел на оси, перейдите в секцию **Число** (Number) на панели **Формат оси** (Format Axis), как показано на рис. 1.39. Нули указывают на расположение чисел. Если справа от запятой числа нет, соответствующая тройка чисел, представляющая сотни или тысячи, будет скрыта.

В табл. 1.1 показаны некоторые способы форматирования больших чисел.

Вы можете использовать знак доллара для отображения валюты и буквы вроде К и М для тысяч и миллионов, чтобы обозначать масштаб чисел. Применяйте этот прием, если у вас несколько чисел на диаграмме, – к примеру, для выделения важных точек данных. На осях знак валюты или буквы после чисел использовать не следует – вместо этого масштаб графика лучше указать в заголовке диаграммы или в подписи к оси.



Рис. 1.38. Форматирование больших чисел на оси значений

Рис. 1.39. Секция Число на панели Формат оси

Таблица 1.1. Пользовательское форматирование больших чисел

17.5	0,,.0
17	0,,
\$17.5	\$0.0,,
\$17	\$0,,
17.5M	0.0,,»M»
17M	0,,»M»
\$17.5M	\$0,,.0»M»
\$17M	\$0,,»M»
17,483K	0,000,,»K»
\$17,483K	\$0,000,,»K»