

*Я посвящаю эту книгу
своей любимой жене Элике Мер (Elica Mehr)
и нашей драгоценной дочери Авине (Avina).*

Без их бесконечной поддержки все это было бы невозможно.

*Также я обязан своим успехом родителям,
которые всегда верили в меня и призывали следовать за своей мечтой*

Содержание

От издательства	13
Предисловие	14
Об авторе	15
О технических редакторах	16
Введение	17
Глава 1. Введение в моделирование данных в Power BI	23
Понятие слоев в Power BI Desktop	24
Слой подготовки данных (Power Query)	25
Слой модели данных	25
Вкладка Данные	26
Вкладка Модель данных	27
Слой визуализации данных	28
Вкладка Отчет	28
Поток данных в Power BI.....	29
Что означает моделирование данных в Power BI.....	30
Семантическая модель.....	31
Построение эффективной модели данных в Power BI	32
Схемы «звезда» (многомерное моделирование) и «снежинка»	34
Транзакционные модели против схемы «звезда»	34
Схема «снежинка».....	36
Понятие денормализации.....	36
Варианты лицензирования в Power BI.....	42
Максимальный размер набора данных.....	43
Добавочная загрузка данных	43
Группы вычислений	44
Общие наборы данных.....	45
Потоки данных Power BI	45
Итеративный подход к моделированию данных.....	45
Сбор информации от руководства	46
Подготовка данных на основе бизнес-логики.....	46
Моделирование данных.....	47
Проверка логики	47
Демонстрация бизнес-логики в базовой визуализации.....	47
Думай как профессиональный разработчик моделей данных	48
Заключение	48
Глава 2. DAX и моделирование данных	50
Понимание виртуальных таблиц.....	50

Создание вычисляемой таблицы.....	51
Использование виртуальных таблиц в мерах, часть 1	53
Использование виртуальных таблиц в мерах, часть 2	55
Визуальное представление виртуальных таблиц	56
Создание вычисляемых таблиц в Power BI Desktop	56
Использование DAX Studio.....	57
Связи в виртуальных таблицах.....	58
Логика операций со временем и моделирование данных	68
Определение валидности дат в измерении.....	68
Вычисления на основе сравнения периодов.....	76
Создание измерения дат при помощи DAX.....	84
Пометка календаря как таблицы дат.....	86
Создание измерения времени при помощи DAX	90
Заключение	92
Глава 3. Подготовка данных с помощью Power Query.....	95
Введение в язык формул M, используемый в Power Query.....	95
Power Query – регистрозависимый инструмент	96
Запросы	97
Выражения.....	97
Значения	97
Примитивные значения	97
Структурированные значения.....	98
Типы.....	102
Примитивные типы	102
Пользовательские типы.....	103
Введение в редактор Power Query.....	103
Панель Запросы.....	105
Таблицы	105
Настраиваемые функции	105
Параметры запросов.....	105
Константы.....	105
Группы.....	105
Панель Параметры запроса.....	106
Свойства.....	108
Область данных.....	109
Строка состояния	112
Расширенный редактор	113
Возможности Power Query в области моделирования данных	114
Качество столбца	115
Распределение столбцов.....	118
Профиль столбца.....	121
Параметры запросов	122
Настраиваемые функции	128
Рекурсивные функции	133
Заключение	135

Глава 4. Получение данных из различных источников	136
Получение данных из распространенных источников данных.....	136
Папка	137
CSV/Текст/TSV	142
Excel	148
Наборы данных Power BI.....	155
Потоки данных Power BI	159
SQL Server.....	160
SQL Server Analysis Services и Azure Analysis Services.....	162
SSAS многомерная/табличная	163
AAS.....	165
Канал OData	166
Сертификаты источников данных	169
Bronze	169
Silver.....	169
Gold/Platinum.....	170
Режимы подключения к данным	170
Импорт	171
Применение	171
Ограничения	171
DirectQuery.....	171
Применение	172
Ограничения	172
Подключение в режиме реального времени	172
Применение	173
Ограничения	173
Режимы хранения данных.....	173
Режимы хранения наборов данных	175
Заключение	177
Глава 5. Общие шаги по подготовке данных	178
Изменение типов данных	179
Разделение столбцов по разделителю	186
Объединение столбцов.....	189
Создание настраиваемого столбца.....	190
Создание столбца из примеров	193
Создание дубликата столбца	195
Фильтрация строк	197
Группирование данных.....	201
Добавление запросов.....	203
Объединение запросов.....	206
Создание дубликата запроса и ссылки на запрос	208
Замена значений.....	210
Извлечение чисел из текста.....	212
Работа с датой, временем и часовыми поясами	215
Заключение	218

Глава 6. Подготовка данных в Power Query для схемы

«звезда»	219
Выявление измерений и фактов.....	219
Количество таблиц в источнике данных	220
Связи между существующими таблицами.....	221
Наименьшая требуемая гранулярность полей с датой и временем	222
Определение измерений и фактов	223
Выявление возможных измерений	224
Выявление возможных фактов	225
Создание таблиц измерений.....	227
Geography	228
Sales Order	230
Product	233
Currency	236
Customer	237
Sales Demographic	238
Date	241
Time.....	245
Создание измерений Date и Time – Power Query против DAX.....	246
Создание таблиц фактов	247
Заключение	254

Глава 7. Эффективные методики подготовки данных

Общие рекомендации по подготовке данных.....	256
При работе с источником OData используйте частичную загрузку данных.....	256
Не забывайте о регистрозависимости Power Query	259
Помните о свертывании запросов и его влиянии на обновление данных.....	260
Понятие свертывания запросов	260
Свертывание запросов и режимы хранения DirectQuery и Dual.....	261
Свертывание запросов и источники данных	261
Индикация свертывания запросов	261
Рекомендации по выполнению свертывания запросов	263
Организируйте запросы в редакторе Power Query.....	267
Преобразование типов.....	268
Преобразование типов и влияние на моделирование данных.....	269
Включение преобразования типов в шаги	275
Изменение типов данных за один шаг.....	276
Оптимизация размера запросов.....	277
Избавьтесь от лишних строк и столбцов	277
Выполните агрегирование (группировку)	278
Отмените загрузку запросов.....	279
Соглашение о наименованиях	279
Заключение	280

Глава 8. Элементы моделирования данных	282
Моделирование данных в Power BI Desktop.....	282
Введение в таблицы.....	283
Свойства таблицы.....	283
Рекомендуемые таблицы.....	286
Вычисляемые таблицы.....	287
Введение в поля.....	292
Типы данных.....	292
Пользовательское форматирование.....	294
Столбцы.....	295
Вычисляемые столбцы.....	295
Группирование данных в столбцах и разделение их на ячейки.....	296
Свойства столбцов.....	300
Иерархии.....	304
Меры.....	305
Неявные меры.....	305
Явные меры.....	308
Текстовые меры.....	308
Использование связей.....	310
Первичные и внешние ключи.....	311
Управление составными ключами.....	311
Связь «один к одному».....	316
Связь «один ко многим».....	316
Связь «многие ко многим».....	316
Распространение фильтров.....	318
Двунаправленные связи.....	320
Заключение.....	323
Глава 9. Схема «звезда» и распространенные техники при моделировании данных	324
Работа со связями типа «многие ко многим».....	324
Связи «многие ко многим» с использованием таблицы-моста.....	327
Скрытие таблицы-моста.....	333
Повышенная бдительность при использовании двунаправленных связей.....	334
Работа с неактивными связями.....	337
Доступность таблицы по нескольким путям фильтра.....	337
Несколько прямых связей между двумя таблицами.....	339
Использование конфигурационных таблиц.....	341
Сегментирование.....	341
Динамическое условное форматирование с участием мер.....	342
Минусы создания вычисляемых столбцов.....	348
Организация модели данных.....	351
Скрытие второстепенных объектов.....	351
Скрытие неиспользуемых полей и таблиц.....	351
Скрытие ключевых полей.....	353
Скрытие неявных мер.....	354

Скрытие столбцов, использующихся в иерархиях, там, где это возможно	354
Создание таблиц мер.....	354
Рассуждения	356
Использование папок.....	357
Создание папки в нескольких таблицах в одно действие	357
Помещение меры в разные папки	359
Создание подпапок	359
Уменьшение размера модели путем отказа от автоматических таблиц с датами и временем	360
Заключение	362
Глава 10. Продвинутое моделирование данных	364
Использование агрегаций	364
Реализация агрегирования для источников, не поддерживающих DirectQuery.....	365
Реализация агрегации на уровне Date	366
Использование инструмента управления агрегированием.....	376
Управление агрегированием в Power BI Desktop для источников, поддерживающих DirectQuery, и больших данных	378
Проверка агрегирования	382
Добавочное обновление.....	387
Настройка добавочного обновления в Power BI Desktop	389
Проверка добавочного обновления	394
Иерархии типа родитель–потомок	396
Определение глубины иерархии	398
Создание уровней иерархии	400
Ролевые измерения	403
Использование групп вычислений.....	406
Требования	407
Терминология.....	407
Группы вычислений и логика операций со временем.....	408
Тестирование групп вычислений.....	414
Проблема с форматированием строк.....	415
Функции DAX для групп вычислений.....	417
Заключение	417
Глава 11. Безопасность на уровне строк	418
Безопасность на уровне строк при моделировании данных.....	419
Чем безопасность на уровне строк не является	419
Терминология безопасности на уровне строк.....	419
Роли	420
Правила.....	420
Проверка ролей.....	421
Назначение участникам ролей в службе Power BI	423
Назначение участникам ролей в Power BI Report Server.....	423

Реализация безопасности на уровне строк.....	425
Распространенные подходы в организации безопасности на уровне строк.....	426
Статическая безопасность на уровне строк.....	426
Создание ролей и определение правил	427
Проверка ролей.....	428
Публикация отчета в службе Power BI	429
Назначение участникам ролей.....	430
Проверка ролей в службе Power BI.....	432
Динамическая безопасность на уровне строк.....	432
Каждый пользователь имеет доступ только к своим данным	433
Менеджер может видеть данные подчиненных	436
Получение учетных данных пользователей из стороннего источника	442
Заключение	448

Глава 12. Дополнительные опции и возможности

моделирования данных	449
Медленно меняющиеся измерения	449
Медленно меняющиеся измерения типа 0 (SCD 0)	451
Медленно меняющиеся измерения типа 1 (SCD 1)	451
Медленно меняющиеся измерения типа 2 (SCD 2)	451
Безопасность на уровне объектов	455
Реализация безопасности на уровне объектов	455
Проверка ролей.....	458
Назначение участникам ролей в службе Power BI	460
Введение в потоки данных	462
Сценарии для использования потоков данных	462
Терминология потоков данных	463
Создание потока данных.....	464
Создание сущностей	467
Создание связанных сущностей из других потоков данных	471
Создание вычисляемых сущностей.....	474
Импорт и экспорт потоков данных.....	476
Составные модели.....	478
Новая терминология	479
Построение цепочек	479
Длина цепочки.....	479
Заключение	485
Предметный указатель.....	486

От издательства

Отзывы и пожелания

Мы всегда рады отзывам наших читателей. Расскажите нам, что вы думаете об этой книге – что понравилось или, может быть, не понравилось. Отзывы важны для нас, чтобы выпускать книги, которые будут для вас максимально полезны.

Вы можете написать отзыв на нашем сайте www.dmkpress.com, зайдя на страницу книги и оставив комментарий в разделе «Отзывы и рецензии». Также можно послать письмо главному редактору по адресу dmkpress@gmail.com; при этом укажите название книги в теме письма.

Если вы являетесь экспертом в какой-либо области и заинтересованы в написании новой книги, заполните форму на нашем сайте по адресу http://dmkpress.com/authors/publish_book/ или напишите в издательство по адресу dmkpress@gmail.com.

Список опечаток

Хотя мы приняли все возможные меры для того, чтобы обеспечить высокое качество наших текстов, ошибки все равно случаются. Если вы найдете ошибку в одной из наших книг, мы будем очень благодарны, если вы сообщите о ней главному редактору по адресу dmkpress@gmail.com. Сделав это, вы избавите других читателей от недопонимания и поможете нам улучшить последующие издания этой книги.

Нарушение авторских прав

Пиратство в интернете по-прежнему остается насущной проблемой. Издательства «ДМК Пресс» и Packt Publishing очень серьезно относятся к вопросам защиты авторских прав и лицензирования. Если вы столкнетесь в интернете с незаконной публикацией какой-либо из наших книг, пожалуйста, пришлите нам ссылку на интернет-ресурс, чтобы мы могли применить санкции.

Ссылку на подозрительные материалы можно прислать по адресу электронной почты dmkpress@gmail.com.

Мы высоко ценим любую помощь по защите наших авторов, благодаря которой мы можем предоставлять вам качественные материалы.

Предисловие

Я очень рад, что именно Сохейл, становление которого в статусе MVP я наблюдал лично, взялся за написание книги на тему экспертного моделирования данных в Power BI. Все студенты любят учиться на конкретных примерах, тогда как пробираться через дебри сухих документаций, извлекая важные сведения, им не очень-то интересно. Многие из них на этом этапе бросают обучение, так и не найдя реальных практических примеров из жизни, объясняющих стоящие перед ними задачи.

В книге Сохейла освещен обширный диапазон тем от начального до экспертного уровня в Power BI – инструменте, занимающем лидирующее место в иерархии продуктов для бизнес-аналитики. Power BI обладает потрясающей гибкостью, и его очень легко использовать. Данная книга подойдет как для начинающих аналитиков, так и для тех, кто проектирует системы, предназначенные для обслуживания тысяч пользователей. Я бы без колебаний посоветовал книгу Сохейла любому, интересующемуся бизнес-аналитикой, вне зависимости от уровня. Он пишет о Power BI в очень захватывающей манере, начиная повествование с элементарных строительных блоков, продолжая различными аспектами моделирования данных, включая описание схемы-звезды, управления двунаправленными связями, отношений «многие ко многим», вычисляемых таблиц, подготовки данных в Power Query, и заканчивая сложными темами, охватывающими RLS, OLS и составные модели данных. Кроме того, каждому студенту должен прийти по душе сугубо практический подход, использованный в этой книге.

*Кристиан Уэйд (Christian Wade),
руководитель группы проектов, Microsoft*

Об авторе

Сохейл Бахши (Soheil Bakhshi) является основателем сайта Data Vizioner и популярным практикующим консультантом по бизнес-аналитике. Сохейл обладает более чем 20-летним опытом работы с данными в области аналитики в Microsoft, включая работу с хранилищами данных и платформой Power BI. Обладая сертификатами MSCE и MCSA, он также может похвастаться статусом Microsoft MVP. Своими знаниями и страстью Сохейл делится на своем сайте www.biinsight.com, а также на конференциях по Power BI, проходящих по всему миру. В стремлении к простоте и эффективности Сохейл Бахши принял участие в разработке нескольких полезных инструментов, включая Power BI Exporter и Power BI Documenter.

О технических редакторах

Фелипе Вилела (Felipe Vilela) долгое время работал в области системных разработок, после чего более восьми лет назад переключился на бизнес-аналитику и хранилища данных, в основном с использованием продуктов от MicroStrategy. Фелипе работал со многими компаниями из Бразилии и США, внедряя и настраивая продукты от MicroStrategy. Параллельно он преподавал бизнес-аналитику и хранение данных на собственных курсах, а также на официальных курсах компании MicroStrategy. Фелипе ведет блог по адресу www.vilelamstr.com, кроме того, он был одним из разработчиков мобильного приложения для конференций MicroStrategy World 2016 и 2017. Обладает более чем тридцатью сертификатами MicroStrategy, включая сертификат МСЕР.

Никита Барсуков (Nikita Barsukov) является опытным специалистом по обработке и анализу данных, сосредоточенным на разработке комплексных аналитических решений. Никита родился и вырос в Украине, а обучение проходил в Финляндии и Швеции. Свой профессиональный путь он начал в области разработки программного обеспечения, но вскоре понял, что его предназначение – это анализ данных и проектирование аналитических инструментов, позволяющих людям лучше разбираться в своих данных. В настоящее время Никита Барсуков работает в Microsoft, где в составе команды разработчиков трудится над аналитическими решениями для Power Platform и Dynamics 365. Помимо работы Никита любит слушать подкасты и аудиокниги, играть в настольные игры с друзьями, а иногда и сам с собой, бегать, пить крафтовое пиво и читать книги. Он живет в Копенгагене с супругой и тремя детьми.

Ана Мария (Ana Maria) является консультантом и тренером по бизнес-аналитике, а также лауреатом звания Microsoft Data Platform MVP, партнером Microsoft Power BI и тренером LinkedIn Learning. Ана Мария находится в индустрии более 25 лет – в 1990-х она разрабатывала решения для FoxPro, а ныне работает, консультирует и преподает в области бизнес-аналитики. Она окончила факультет экономической информатики в Московском государственном университете управления, после чего получила степень магистра в испанском Университете Алькала. Ана Мария специализируется на работе с инструментами бизнес-аналитики от Microsoft, а также с SQL Server, Excel, Azure Machine Learning, R и Power BI. В качестве спикера, организатора или участника ее можно встретить на различных технических форумах и мероприятиях.

Введение

Microsoft Power BI является одним из наиболее популярных инструментов бизнес-аналитики на рынке программного обеспечения для настольных и облачных решений. Книга, которую вы держите в руках, может стать вашим проводником в мир моделирования данных в целом и применительно к Power BI. Вы узнаете, как подключаться к данным в различных источниках, объединять их при помощи связей и строить полноценные модели данных.

Из книги вы поймете, как использовать принципы моделирования данных и техники навигации для определения связей между сущностями и создания модели данных, после чего мы перейдем к вопросам определения новых метрик и выполнения пользовательских вычислений с использованием особенностей модели. С течением глав сложность и эффективность моделей данных будет увеличиваться, и вы научитесь использовать язык запросов DAX, а также новые техники моделирования. С помощью примеров мы покажем вам, как можно создавать новые или адаптировать существующие модели данных с учетом разнообразных бизнес-требований. Наконец, вы освоите применение относительно свежих продвинутых возможностей для оптимизации и расширения своих моделей данных, что позволит вам решать широкий спектр задач. К концу книги вы будете обладать всеми необходимыми знаниями для структурирования и обработки данных, поступающих из разных источников, и создания на их основе полноценных моделей данных, пригодных для построения отчетов и проведения полноценного анализа данных.

Для кого эта книга

Книга предназначена для пользователей систем бизнес-аналитики, а также специалистов и разработчиков в области анализа данных, желающих улучшить свое понимание техник моделирования данных с целью извлечь максимум возможного из Power BI. Наличие базовых знаний в области Power BI и понимание схемы данных «звезда» поможет вам в освоении тем, освещаемых в этой книге.

Структура книги

Глава 1 «Введение в моделирование данных в Power BI». В данной главе мы кратко опишем функционал программного продукта Power BI и расскажем о том, почему так важно уметь моделировать данные. Здесь мы также коснемся вопросов лицензирования Power BI, напрямую влияющих на возможности моделирования данных. Кроме того, мы познакомимся с поняти-

ем итеративного моделирования данных на примере реализации в Power BI.

Глава 2 «DAX и моделирование данных». В этой главе мы будем работать с языком запросов DAX не так много и глубоко, как в третьей и четвертой частях книги. Здесь мы сконцентрируемся на функционале языка, не самом очевидном для понимания, но очень важном с точки зрения моделирования данных. Начнем главу с краткого введения в DAX, после чего сразу перейдем к рассмотрению виртуальных таблиц и функций логики операций со временем, а также их применению в реальных сценариях.

Глава 3 «Подготовка данных с помощью Power Query». Здесь мы быстро пройдемся по базовому функционалу инструмента Power Query и способам его применения на практике. Мы сделаем особый упор на параметры запросов и пользовательские функции, а также рассмотрим несколько примеров, помогающих понять, как эти техники позволяют повысить гибкость и надежность создаваемых моделей данных.

Глава 4 «Получение данных из разных источников». В этой главе мы посмотрим на способы получения данных в Power BI из наиболее распространенных источников. После этого затронем тему сертификации источников данных, которая позволяет выстроить определенные ожидания по поводу типа получаемых данных из источника. Это бывает очень полезно при оценке усилий на проектирование модели данных. Также мы рассмотрим разные режимы подключения к данным.

Глава 5 «Общие шаги по подготовке данных». Здесь мы опишем на примерах наиболее распространенные действия, которые приходится выполнять при преобразовании данных, полученных из источника. В совокупности с уже приобретенными знаниями в более ранних главах описанные здесь шаги позволят вам в будущем проектировать высокоэффективные модели данных. Изучив все описанные возможности, вы сможете по своему усмотрению выбирать способ реализации своей модели данных.

Глава 6 «Подготовка данных в Power Query для схемы “звезда”». В этой главе мы подробно поговорим о вариантах подготовки запросов для создания модели данных типа «звезда» с рассмотрением реальных сценариев. Здесь мы будем активно использовать язык программирования M, встроенный в Power Query. С учетом всех полученных ранее знаний по предварительной подготовке данных в Power Query вам не должно составить труда разобраться с предложенными здесь примерами. Кроме того, вы научитесь создавать таблицы измерений и таблицы фактов, а также денормализовывать запросы при необходимости.

Глава 7 «Эффективные методики подготовки данных». Здесь мы поговорим о типичных шаблонах при преобразовании данных, полученных из разных источников. Использование этих шаблонов позволит вам повысить эффективность создаваемых моделей данных, которые будет легче поддерживать. Прочитав эту главу, вы сможете избегать распространенных ошибок при проектировании моделей данных, что значительно облегчит вам жизнь в будущем.

Глава 8 «Элементы моделирования данных». Эта глава будет посвящена составляющим компонентам моделей данных с точки зрения Power BI, рас-

смотренным на примерах. Здесь мы будем активно использовать DAX, так что базовое понимание этого языка запросов будет крайне желательным. При рассмотрении примеров мы будем иметь дело с полноценной моделью данных типа «звезда». В этой главе мы также коснемся темы особых таблиц, использование которых может позволить обогатить модель данных за счет добавления в нее сложной бизнес-логики.

Глава 9 «Схема “звезда” и распространенные техники при моделировании данных». Здесь мы поговорим о принятых нормах при проектировании моделей данных и постараемся сделать все, чтобы вы не допускали наиболее распространенных ошибок на этапе разработки модели. К примеру, проблему с типом данных в ключевом столбце, используемом в связи, бывает очень непросто обнаружить, тогда как предотвратить ее на этапе проектирования модели проще простого. Так что освоение распространенных техник позволит вам в конечном счете сэкономить драгоценные время и деньги.

Глава 10 «Продвинутые техники моделирования данных». В этой главе мы затронем тему особых приемов при моделировании данных, помогающих в решении поставленных бизнес-задач. Хороший специалист по моделям данных должен быть открыт всему новому. Вы в своей практике легко можете столкнуться с описанными в этой главе или очень похожими на них бизнес-требованиями. И здесь мы пытаемся донести до вас мысль о том, что вы всегда должны с готовностью принимать новые вызовы бизнеса и быть готовы применить все известные вам инновационные приемы для их решения.

Глава 11 «Безопасность на уровне строк». Здесь мы посмотрим, как реализуется в модели данных Power BI *безопасность на уровне строк* (row-level security – RLS). Работать с этой технологией бывает не так просто, и чтобы понять все ее тонкости, необходимо обладать глубокими знаниями в области моделирования данных и распространения фильтров. В данной главе мы поможем вам освоить эти концепции, что позволит вам в будущем проектировать эффективные и надежные модели.

Глава 12 «Дополнительные опции и возможности моделирования данных». В заключительной главе книги мы подробно поговорим о таких дополнительных концепциях моделирования данных, как *медленно меняющиеся измерения* (Slowly Changing Dimensions – SCD), *безопасность на уровне объектов* (Object-Level Security (OLS)), *потоки данных* (dataflows) и *составные модели* (composite model).

КАК ИЗВЛЕЧЬ МАКСИМУМ ИЗ КНИГИ

Вам необходимо загрузить и установить последнюю версию Power BI Desktop. Все выражения проходили проверку в мартовском релизе Power BI Desktop 2021 года и должны без проблем работать в более поздних версиях программы. В дополнение к Power BI Desktop желательно будет установить программы DAX Studio и Tabular Editor.

Программное/аппаратное обеспечение, используемое в книге	Требования к операционной системе
Power BI Desktop https://powerbi.microsoft.com/en-us/downloads/	<ul style="list-style-type: none"> – Windows 8.1 / Windows Server 2012 R2 или выше; – .NET 4.6.2 или выше; – Internet Explorer 11 или выше; – память (RAM): минимум 2 Гб, рекомендовано 4 Гб и более; – видеосистема: минимум 1440×900 или 1600×900 (16:9). Более низкие разрешения экрана, такие как 1024×768 или 1280×800, не поддерживаются, поскольку в этом случае некоторые элементы управления (например, закрытие стартовой заставки) не могут быть корректно отражены; – параметры экрана: если вы настроили дисплей таким образом, чтобы масштаб текста, приложений и других элементов мог превышать значение 100 %, вы можете не увидеть некоторых диалоговых окон в Power BI Desktop. Если у вас возникла такая проблема, откройте Параметры (Settings) ⇒ Система (System) ⇒ Дисплей (Display) и установите значение масштаба в 100 %; – процессор (CPU): 1 ГГц 64 бит (x64) или выше
DAX Studio https://daxstudio.org/downloads/	<ul style="list-style-type: none"> – Windows 7 или выше (рекомендуется Windows 10); – .NET Framework 4.7.1 или выше
Tabular Editor https://github.com/otykier/TabularEditor/releases/tag/2.16.0	<ul style="list-style-type: none"> – Windows Server 2019; – Windows Server 2016; – Windows 7 или выше (рекомендуется Windows 10); – клиентские библиотеки Azure Analysis Services «АМО» версии 18.4.0.5 или выше

ПРИМЕЧАНИЕ Начиная с 31 января 2021 года Power BI Desktop прекратил поддержку Windows 7.

Для проверки примеров из некоторых глав вам понадобится наличие аккаунта службы *Power BI* (Power BI Service). Вы можете зарегистрироваться в службе и приобрести лицензию в качестве отдельного пользователя. Подробнее почитать можно по адресу https://docs.microsoft.com/en-us/power-bi/fundamentals/service-self-service-signup-for-power-bi?WT.mc_id=5003466.

Если вы используете цифровую версию книги, мы советуем вам вводить весь код вручную или скачивать его с сайта www.dmkpress.com на странице с описанием данной книги. Это позволит вам избежать возможных ошибок при копировании и вставке кода.

При написании книги я предполагал, что вы знакомы с терминологией и базовыми принципами хранилищ данных и схемы «звезда». При этом в книге, когда это необходимо, приводятся термины с их кратким описанием.

ЗАГРУЗИТЕ СОПРОВОДИТЕЛЬНЫЕ ФАЙЛЫ

Скачать файлы с дополнительной информацией для книг издательства «ДМК Пресс» можно на сайте www.dmkpress.com на странице с описанием соответствующей книги.

УСЛОВНЫЕ ОБОЗНАЧЕНИЯ

На протяжении книги мы будем использовать следующие условные обозначения и шрифты.

Код в тексте: так в тексте книги мы будем обозначать код. Пример: «Таблица Customertable широкая и длинная».

Термины: так будут написаны важные термины, названия папок и файлов, пути, пользовательский ввод и прочее.

Блоки кода будут выделены следующим образом:

```
Sequential Numbers =  
SELECT COLUMNS(  
    GENERATESERIES(1, 20, 1)  
    , "ID"  
    , [Value]  
    )
```

Жирный шрифт: так будут выделяться новые термины, важные слова и текст, который вы видите на экране. Например, таким образом будут обозначаться пункты меню. Пример: «Нажмите на кнопку **Создать таблицу** (New table) на вкладке **Моделирование** (Modeling)».

СОВЕТЫ ИЛИ ВАЖНЫЕ ПРИМЕЧАНИЯ будут оформлены так.

Часть I

Моделирование данных в Power BI

В этой вводной части мы обсудим общие принципы моделирования данных при помощи Power BI. Мы будем предполагать, что вы уже знаете, что такое и для чего используются Power Query и DAX, а также понимаете базовые принципы схемы «звезда». В этой части вы узнаете, как применять виртуальные таблицы и функции логики операций со временем в DAX, кроме того, мы поговорим о принципах проектирования эффективных моделей данных на основе реальных сценариев.

Содержание этой части:

- глава 1 «Введение в моделирование данных в Power BI»;
- глава 2 «DAX и моделирование данных».

Глава 1

Введение в моделирование данных в Power BI

Power BI – это не просто инструмент для построения отчетов. Это полноценная платформа, предоставляющая богатейший спектр возможностей, – от подготовки исходных данных до их моделирования и визуализации. Кроме того, Power BI представляет собой целую экосистему, позволяющую пользователям вносить собственный вклад в аналитическую политику организации путем обмена наборами данных, отчетами и дашбордами, а также размещения в отчетах комментариев с мобильных устройств и их рассылки конкретным пользователям. Но все это возможно только при правильной настройке экосистемы Power BI. Даже самый красивый в мире отчет ровным счетом ничего не будет стоить, если он показывает неправильные цифры или на его формирование уходит много времени. Пользователи никогда не будут работать с таким отчетом.

Одним из важнейших факторов, влияющих на формирование эффективной экосистемы в Power BI, является правильность лежащих в ее основе данных. В реальных проектах вам зачастую приходится получать данные из различных источников. Но получение данных и их внедрение в систему – это только полдела. Самое главное – объединить эти данные в модель, позволяющую гарантировать целостность исходных сведений и их связь с бизнес-логикой.

В этой главе мы познакомим вас с таким понятием, как слои Power BI, и вместе посмотрим, как данные перемещаются между слоями, что помогает при эффективном решении потенциальных проблем. После этого мы поговорим о таком важнейшем аспекте платформы Power BI, как моделирование данных. Вы узнаете об ограничениях моделей данных и разных возможностях в зависимости от используемой лицензии. Наконец, мы познакомимся с понятием итеративного моделирования данных и его фазами.

Основные темы, которые мы рассмотрим в этой главе:

- понятие слоев в Power BI Desktop;
- что означает моделирование данных в Power BI;
- варианты лицензирования в Power BI;
- итеративный подход к моделированию данных.

ПОНЯТИЕ СЛОЕВ В POWER BI ДЕСКТОП

Как мы уже сказали ранее, *Power BI* – это не просто инструмент для формирования отчетов. Поскольку главным образом эта книга сконцентрирована на моделировании данных, нам бы не хотелось углубляться в детали Power BI как инструмента, но некоторые основные концепции без внимания мы оставить не можем. Говоря о моделировании данных в Power BI, мы фактически ссылаемся на программный продукт *Power BI Desktop*. Вы можете рассматривать Power BI Desktop как своеобразный аналог *Visual Studio* при разработке *табличной модели* (Tabular model) в *SQL Server Analysis Services* (SSAS). Power BI Desktop представляет собой бесплатный продукт от Microsoft, который можно загрузить по адресу <https://powerbi.microsoft.com/en-us/downloads/>. В этой книге мы будем подразумевать Power BI Desktop, когда говорим Power BI, если не указано иное.

На рис. 1.1 показан типичный процесс построения отчета в Power BI Desktop.

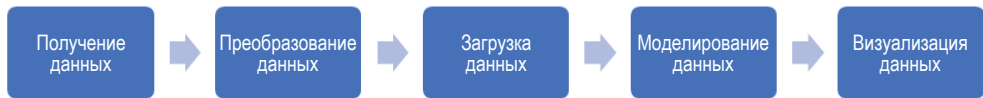


Рис. 1.1 ❖ Формирование нового отчета в Power BI

Для осуществления описанного на рис. 1.1 процесса мы используем различные *концептуальные слои* (conceptual layer) в Power BI. В Power BI Desktop эти слои отражены так, как показано на рис. 1.2.

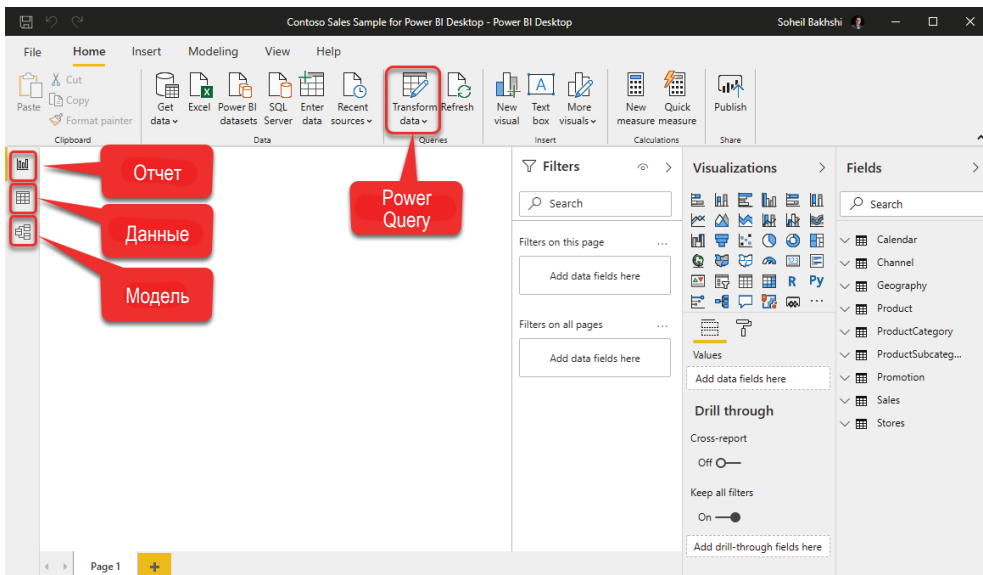


Рис. 1.2 ❖ Слои Power BI

Конец ознакомительного фрагмента.

Приобрести книгу можно

в интернет-магазине

«Электронный универс»

e-Univers.ru