

Оглавление

Об авторах.....	12
О редакторе.....	13
Предисловие от издательства.....	14
Предисловие	15
ЧАСТЬ I ВВЕДЕНИЕ	19
Глава 1.1 DAX в бизнес-аналитике	20
Пятислойная модель бизнес-аналитики	21
Аналитика предприятия и аналитика конечных пользователей	22
Для чего используется DAX и где его найти	24
Excel.....	25
Power BI	25
SQL Server Analysis Services	25
Azure Analysis Services.....	26
Инструменты для разработки моделей и написания выражений DAX.....	26
Сделано при помощи DAX: визуальные элементы и интерактивные отчеты	27
Подходы к разработке решений.....	30
Использование Power BI для ускорения разработки аналитических решений.....	31
Мы не знаем точно, что нам нужно	32
Наши данные некорректны	33
Цикл цифровой трансформации.....	33
Заключение.....	35
Глава 1.2 Разработка модели	36
Колоночное хранение данных.....	36
Реляционные базы данных.....	36
Колоночные базы данных.....	37
Типы данных и кодирование.....	38
Связи	41
Данные в Excel	41
Реляционные базы данных.....	42
Реляционная модель Power BI	43
Свойства связей.....	44
Активные и неактивные связи.....	45
Направление фильтрации	46
Кратность.....	49

Оптимальная структура модели данных.....	50
Схемы «звезда» и «снежинка»	50
Проблемы, присущие схеме «звезда»	51
Концепции реляционных моделей, которых нужно избегать при моделировании в Power BI	52
Независимые измерения.....	52
Только одна таблица фактов.....	54
Хранилище данных как единый источник истины	55
Использование связей типа «многие ко многим»	56
Производительность модели и использование памяти	57
Заключение.....	59

Глава 1.3 Применение DAX..... 60

Вычисляемые столбцы.....	60
Вычисляемые таблицы.....	62
Меры.....	64
Фильтры безопасности	66
Запросы DAX.....	66
Таблицы дат	68
Создание таблицы дат.....	69
Практические советы по использованию языка DAX.....	71
Первым делом меры!	71
Создавайте явные меры.....	71
Используйте базовые меры в качестве строительных блоков....	71
Прячьте от пользователей ненужные им элементы модели....	72
Не смешивайте данные и меры – создайте для мер отдельную таблицу.....	72
Типы таблиц	74
Заключение.....	75

Глава 1.4 Контексты и фильтры 76

Введение в контексты DAX	77
Контекст строки.....	77
Контекст запроса	79
Контекст фильтра	81
Обнаружение фильтров	82
Сравнение контекста запроса и фильтра с контекстом строки.....	83
Фильтрация в DAX: использование функции CALCULATE	83
Шаг 1. Настройка контекста фильтра.....	85
Шаг 2. Удаление существующих фильтров	87
Шаг 3. Применение новых фильтров	87
Шаг 4. Вычисление выражения в новом контексте.....	88
Удаление фильтров при помощи функций группы ALL.....	90
Функции логики операций со временем	93
Изменение поведения связей	96

Табличные функции в DAX.....	98
Табличные агрегации.....	98
Использование виртуальных таблиц.....	100
Контекст в табличных функциях.....	102
Производительность при использовании табличных функций.....	105
Фильтрация при помощи табличных функций	106
Использование функции CALCULATETABLE	107
Фильтры и таблицы.....	108
Использование функции TREATAS.....	112
Переменные в DAX.....	113
Заключение.....	116

ЧАСТЬ II ПРИМЕНЕНИЕ DAX НА РЕАЛЬНЫХ ПРИМЕРАХ..... 117

Глава 2.1 Безопасность в DAX 118

Знакомство с безопасностью на уровне строк (RLS).....	118
Роли безопасности	119
Фильтры безопасности в DAX.....	121
Динамическая безопасность на уровне строк.....	122
Особенности моделирования данных с RLS.....	125
Проверка ролей безопасности.....	130
Тестирование отчетов с динамическим подключением	132
Модель для имперсонализации	133
Добавление таблицы rImpersonation в модель данных ...	134
Добавление тестовой роли безопасности.....	134
Собираем все вместе	136
Иерархическая безопасность с применением функций PATH... ..	137
Иерархические таблицы	137
Знакомство с функциями группы PATH	138
PATH	138
PATHCONTAINS.....	139
PATHLENGTH	139
PATHITEM.....	139
PATHITEMREVERSE	139
Использование функций PATH совместно с RLS.....	140
Продвинутая навигация по иерархии в RLS.....	140
Безопасность атрибутов.....	143
Применение безопасности атрибутов	144
Безопасность на уровне объектов и ее ограничения.....	144
Динамическая защита атрибутов: безопасность на уровне значений.....	145
Безопасность на уровне значений: моделирование данных	146
Безопасность на уровне значений: фильтры безопасности ..	149
Безопасность на уровне значений: сложные сценарии	150

Разработка моделей данных с безопасностью на уровне значений.....	153
Защита уровней агрегации.....	153
Меры не могут быть защищены, а таблицы фактов – могут.....	154
Ограничение гранулярности таблицы фактов.....	154
Защита уровней агрегации с помощью составных моделей данных.....	155
Сочетание защиты уровней агрегации с безопасностью на уровне значений.....	158
Защита уровня агрегации как атрибута.....	161
Заключение.....	165

Глава 2.2 Динамическое изменение визуализации 166

Рассматриваемый пример.....	167
Динамические меры.....	169
Базовые меры показателей.....	169
Создание вспомогательной таблицы.....	171
Создание динамической меры.....	172
Динамический выбор показателя и используемой даты.....	174
Динамические метки.....	179
Описание сценария.....	180
Создание вспомогательной таблицы.....	180
Создание меры DAX с использованием динамических меток.....	182
Комбинирование динамических меток с динамическими вычислениями.....	184
Заключение.....	186

Глава 2.3 Альтернативные календари 187

Недельный и григорианский календари.....	187
Что такое недельный календарь?.....	188
Номера недель.....	188
Периоды.....	190
Кварталы.....	190
Годы.....	191
Создание недельного календаря.....	191
Настройка дат.....	191
Определение даты начала года.....	193
Определение даты конца года.....	194
Создание дополнительных столбцов в таблице дат.....	196
Анализ с применением логики операций со временем для недельного календаря.....	199
Модель Power BI.....	199
Расчет нарастающих итогов с начала года.....	201
Расчет приростов продаж.....	203
Расчет скользящих средних продаж за отчетный год.....	207
Поддержание актуальности отчетов.....	211

Таблица Date Selection	211
Создание вариантов выбора в таблице	214
Использование вариантов выбора дат в мерах.....	217
Заключение.....	219

Глава 2.4 Использование механизма AutoExist 220

Модель данных Power BI	220
Как Power BI визуализирует данные из модели	222
Фильтры и контекст	222
Как использование мер способно изменить поведение визуального элемента.....	224
Знакомство с запросами DAX визуальных элементов.....	226
Что такое AutoExist и как он работает	229
Использование нескольких фильтров в визуальном элементе.....	229
Как AutoExist позволяет оптимизировать вычисление формул DAX	231
Пример: пропущенные рабочие дни	233
Условия задачи	234
Структура модели.....	234
Анализ сумм заказов.....	235
Расширение таблицы Calendar	238
Анализ рабочих дней	240
Куда делся мой рабочий день?	241
Как решить проблему пропущенных рабочих дней	242
Источник проблемы	243
Изменение структуры модели для обхода механизма AutoExist	244
Всегда учитывайте контекст.....	244
Исправляем ошибку с рабочими днями	248
Оптимизация производительности отчетов с AutoExist.....	250
Гранулярность в таблицах фактов.....	250
Фильтрация в нескольких таблицах фактов.....	251
Оптимизация структуры модели данных.....	254
Оптимизация визуальных элементов	255
Заклучение.....	257

Глава 2.5 Взаимодействие внутри компании 258

Моделирование процесса ведения продаж в компании QuantoVikes.....	258
Процесс ведения продаж	259
Модель данных Power BI	260
Взаимоотношения между филиалами одной компании.....	262
Представление с филиалами против сводного представления	263
Сопоставление внутренних продаж и покупок	264
Визуализация внутренних движений компании.....	268

Будущие продажи.....	273
Анализ одиночных продаж.....	273
Анализ долгосрочных продаж.....	276
Анализ старых заказов.....	278
Анализ текущих заказов.....	282
Оптимизация расчета для текущих заказов.....	288
Анализ излишков по счетам.....	290
Анализ своевременных заказов.....	294
Коррекция для запаздывающих заказов.....	296
Дальнейшая оптимизация.....	299
Тестирование сложных вычислений.....	301
Заключение.....	303

Глава 2.6 Исследуем будущее: прогнозирование и будущая стоимость..... 305

Финансовые расчеты.....	306
Приведенная стоимость и чистая приведенная стоимость....	307
Внутренняя норма доходности.....	308
Финансовые функции DAX.....	308
Бизнес-сценарий и модель данных.....	311
Подбор приемлемой ставки и индексов.....	314
Вычисление будущей стоимости.....	316
Первоначальные инвестиции и остаточная стоимость.....	316
Нерегулярные денежные потоки.....	317
Повторяющиеся денежные потоки.....	318
Положительные и отрицательные денежные потоки.....	321
Вычисление чистой приведенной стоимости.....	322
Вычисление внутренней нормы доходности.....	326
Расчет аренды, покрывающей издержки.....	328
Поиск суммы покрывающей издержки аренды с помощью аппроксимации.....	329
Оптимизация аппроксимационного подхода.....	333
Заклучение.....	338

Глава 2.7 Анализ состояния запасов..... 339

Моделирование данных, ориентированных на статусы.....	339
Гранулярность складских запасов.....	344
Базовые складские расчеты.....	346
Целевой уровень запасов.....	349
Прогнозирование складских запасов.....	355
Два типа прогнозирования.....	355
Прогнозирование продаж как инструмент предсказания изменений запасов.....	356
Экстраполяция как инструмент предсказания изменений запасов.....	360
Вычисление долгосрочных складских запасов.....	366

Использование целевого уровня запасов на основе прогноза	370
Использование линейной регрессии при экстраполяции товарных запасов.....	372
Заключение.....	376

Глава 2.8 Планирование кадровых ресурсов 377

Модель данных Power BI для планирования кадровых ресурсов	377
Расчет показателей продажи.....	380
Оптимизация расчета показателей продажи.....	383
Вычисление требуемых объемов ресурсов для реализованных проектов	385
Работа с итоговыми значениями	389
Оптимизация вычисления трудовых ресурсов	390
Оптимизация модели данных Power BI	393
Уровни агрегации.....	396
Заключение.....	397

Предметный указатель 399

Об авторах

Мишель Розема (Michiel Rozema) – один из ведущих специалистов в области Power BI из Нидерландов. Он обладает степенью магистра по математике и работает в сфере информационных технологий в роли управляющего и консультанта уже более 25 лет. В течение восьми лет Мишель возглавлял отдел аналитики в Microsoft Netherlands, запустив в стране платформу Power BI. Является автором двух книг на голландском языке, посвященных Power Pivot и Power BI. Также состоит в числе основателей официальной группы голландских пользователей Power BI и организаторов Power BI Summer School. Выступает на различных конференциях по Power BI. В 2019 году Мишель впервые стал обладателем статуса Microsoft MVP и вместе с соавтором книги Хенком Влотманом, также обладающим статусом MVP, основал компанию Quanto, специализирующуюся на консультациях в области Power BI.

Хенк Влотман (Henk Vlootman) – старший бизнес-консультант в области Power Platform, Power BI и Excel. Является обладателем статуса Microsoft MVP с 2013 года и сооснователем официальной группы голландских пользователей Power BI и Power BI Summer School. Выступает в роли спикера на многочисленных конференциях, посвященных Power BI, по всему миру. Хенк написал две книги по Excel и две по Power Pivot/Power BI. Карьеру начал в 1992 году, основав собственную компанию, после чего добился успеха в роли консультанта по Excel. В настоящее время вместе с соавтором книги Мишелем Роземой управляет компанией Quanto, специализирующейся на Power BI.

О редакторе

Грег Деклер (Greg Deckler) – обладатель статуса MVP в категории Data Platform и активный член сообщества информационных технологий в Колумбусе, штат Огайо. Является основателем пользовательской группы Columbus Azure ML and Power BI User Group (CAMLPUUG) и спикером многочисленных конференций в США. Будучи активным блогером и членом сообщества, много помогает новичкам в области Power BI. Имеет в своем активе более 180 загрузок в Power BI Quick Measures Gallery и более 5000 ответов на вопросы пользователей. Грег является заместителем председателя отдела облачных технологий в региональной компании Fusion Alliance. Его перу принадлежат три книги, посвященные Power BI: *Learn Power BI*, *DAX Cookbook* и *Power BI Cookbook, второе издание*. Также он разработал внешний инструмент для Power BI Desktop под названием *Microsoft Hates Greg's Quick Measures* и записывает видео по Power BI на своем канале в YouTube.

Я бы хотел поблагодарить моего сына, семью и все сообщество Power BI за их поддержку.

Предисловие 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.dmkipress.com, зайдя на страницу книги и оставив комментарий в разделе «Отзывы и рецензии». Также можно послать письмо главному редактору по адресу dmkipress@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://dmkipress.com/authors/publish_book/ или напишите в издательство по адресу dmkipress@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 – возможно, ошибку в основном тексте или программном коде, – мы будем очень благодарны, если вы сообщите нам о ней. Сделав это, вы избавите других читателей от недопонимания и поможете нам улучшить последующие издания этой книги.

Если вы найдете какие-либо ошибки в коде, пожалуйста, сообщите о них главному редактору по адресу dmkipress@gmail.com, и мы исправим это в следующих тиражах.

Нарушение авторских прав

Пиратство в интернете по-прежнему остается насущной проблемой. Издательство «ДМК Пресс» очень серьезно относится 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 dmkipress@gmail.com.

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

Предисловие

Прочитав эту книгу, вы сможете вывести свои аналитические навыки работы с платформой Power BI на новый качественный уровень. Вы узнаете об истинной мощи языка запросов DAX и научитесь строить на его основе эффективные решения в практической области.

Для кого эта книга

Если вы – аналитик со знанием DAX в Power BI или других аналитических инструментов Microsoft, эта книга поможет вам выйти на новый уровень владения языком DAX и начать использовать аналитические модели гораздо более эффективно.

Эта книга не для новичков – практический опыт работы с DAX вам все-таки понадобится.

Структура книги

Глава 1.1. DAX в бизнес-аналитике. В этой главе мы рассмотрим сферу бизнес-аналитики в целом и аналитические модели в современных решениях, которым отводится центральная роль в этой области. Модели Power BI идеально подходят для применения в этой среде, и не в последнюю очередь благодаря мощи языка DAX.

Глава 1.2. Разработка модели. Во второй главе книги мы поговорим о фундаментальных концепциях, лежащих в основе моделей Power BI. Вы узнаете, чем модели данных в Power BI принципиально отличаются от других систем управления данными и как выглядит оптимальная модель в этой платформе.

Глава 1.3. Применение DAX. В этой главе мы пройдемся по концепциям, реализуемым в Power BI с помощью языка DAX, в числе которых – вычисляемые столбцы и таблицы, меры, правила безопасности и запросы. Мы также дадим несколько полезных советов по работе с DAX.

Глава 1.4. Контексты и фильтры. Здесь мы поговорим о контексте строки, контексте запроса и контексте фильтра, а также обсудим, какую роль играют контексты при вычислении формул DAX. Кроме того, мы посмотрим, как можно преобразовывать контексты при помощи функции CALCULATE, добавляя и удаляя фильтры из существующих контекстов. В дополнение мы познакомимся с функциями логики операций со временем и табличными функциями DAX, узнаем о связи между таблицами и фильтрами, а также поговорим о переменных DAX.

Все эти главы будут посвящены базовым концепциям моделирования данных и языка DAX, чтобы подготовить вас к более глубокому погружению в тему.

Во второй части книги мы рассмотрим более продвинутые приемы работы с DAX на примерах из реальной практики, с которыми мы сталкивались и продолжаем сталкиваться в своей работе.

Глава 2.1. Безопасность в DAX. В этой главе мы поговорим о концепциях безопасности в моделях Power BI и роли DAX в этой области. Мы подробно обсудим механизм безопасности на уровне строк, роли безопасности, иерархии, атрибуты и уровни агрегации в процессе моделирования данных с применением DAX.

Глава 2.2. Динамическое изменение визуализации. Здесь мы рассмотрим использование таблиц-помощников и функции SWITCH для перехвата пользовательского ввода. Мы продемонстрируем богатые возможности динамического изменения привязки данных с помощью DAX с целью создания высокодинамичных визуальных элементов. В зависимости от ваших намерений таблица-помощник может содержать всего несколько строк с параметрами или большой список значений на основе других данных из модели Power BI.

Глава 2.3. Альтернативные календари. В этой главе мы поговорим о реализации логики операций со временем в случаях, когда используемый вами календарь отличается от стандартного григорианского, применение которого подразумевается по умолчанию в моделях Power BI. В конце главы мы предложим более гибкую альтернативу фильтру Power BI с относительными датами, которая будет поддерживать работу с нестандартными календарями.

Глава 2.4. Использование механизма AutoExist. Здесь мы поговорим о том, *какие именно* вычисления производятся при наполнении визуального элемента данными из модели Power BI. Понимание принципов работы механизма AutoExist позволит вам лучше разобраться в том, почему иногда данные в визуальных элементах не соответствуют вашим ожиданиям. Кроме того, вы узнаете, что именно этому полезному механизму мы обязаны быстротой формирования отчетов с визуальными элементами, в которых задействовано множество столбцов из разных таблиц.

Глава 2.5. Взаимодействие внутри компании. В этой главе обсуждаются такие распространенные бизнес-сценарии, как организация взаимодействия в рамках одной компании и консолидированные представления, а также выставление счетов по открытым продажам. Мы узнаем, как отслеживать состояние контекста, строить меры на DAX и проводить углубленный анализ.

Глава 2.6. Исследуем будущее: прогнозирование и будущая стоимость. В данной главе мы поговорим о финансовых показателях для анализа прогноза инвестиций. Мы обсудим такие распространенные метрики, как будущая стоимость, приведенная стоимость, чистая приведенная стоимость и внутренняя норма доходности, а также функции DAX для их расчета: XNPV и XIRR. Кроме того, мы познакомимся с параметрами What-if и научимся использовать их в сложных вычислениях.

Глава 2.7. Анализ состояния запасов. Здесь мы рассмотрим способы анализа складских запасов, хотя все указанные приемы могут быть с легкостью применены к анализу любых данных, связанных со статусами. Мы обсудим варианты моделирования таких данных, научимся рассчитывать состояние запасов на любой момент времени и сравнивать фактические значения с целевыми. Также мы заглянем в будущее и научимся строить прогнозы в DAX, в том числе при помощи линейной регрессии.

Глава 2.8. Планирование кадровых ресурсов. В заключительной главе книги мы поговорим об анализе потребности организации в человеческих ресурсах (выраженных в штатных единицах) при подготовке к проектам. С технической точки зрения мы на этом примере научимся работать с несколькими таблицами фактов, к которым нужно будет обращаться одновременно для получения требуемого результата. При этом нашей целью будет не только извлечение корректных цифр, но и оптимизация этого процесса.

Как извлечь максимум из книги

Как мы уже упоминали, для комфортного чтения этой книги вы должны обладать определенным практическим опытом работы с DAX и желанием развивать его и узнать что-то новое о применении этого языка запросов в сложных нестандартных сценариях. Таким образом, мы не будем подробно рассказывать о том, как работают функции DAX и какие аргументы они принимают, если речь не идет об исключительных случаях, поскольку всю эту вводную информацию можно легко найти в документации.

Все модели данных, обсуждаемые в этой книге, можно загрузить из хранилища GitHub по адресу <https://github.com/PacktPublishing/Extreme-DAX> и со страницы книги. Кроме того, в каждой главе мы будем давать ссылку на конкретный файл или файлы PBIX, с которыми будем работать. Мы рекомендуем вам параллельно с чтением разбирать указанные файлы самостоятельно.

Для открытия файлов вы можете использовать последнюю версию Power BI Desktop. Примеры всех визуализаций и запросов, показанные в этой книге, реализованы в сопроводительных файлах.

Сопроводительные файлы

Сопроводительные файлы к книге можно загрузить со страницы на сайте издательства и из хранилища GitHub по адресу <https://github.com/PacktPublishing/Extreme-DAX>.

Цветные изображения

По следующей ссылке вы можете скачать в виде PDF все рисунки и диаграммы, использованные в книге: https://static.packt-cdn.com/downloads/9781801078511_ColorImages.pdf.

Условные обозначения

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

Код в тексте: так в тексте книги мы будем выделять код на языке DAX. Пример: «Применение неактивных связей позволит вам избежать многократного включения выражения ALL('Calendar') в код DAX».

Блоки кода будут выделены следующим образом:

```
Total Sales =  
CALCULATE(  
    SUM(fProjectSales[Budget]),  
    USERELATIONSHIP(fProjectSales[StartDate], 'Calendar'[Date])  
)
```

Для привлечения внимания к важной части текста мы будем использовать курсив. Пример: «Обратите внимание, что связь между таблицами fProjectSales и Calendar мы сделали *неактивной*. Причина в том, что в большинстве вычислений мы не будем группировать проекты по дате начала, а вместо этого будем распределять результаты по целому временному интервалу, *начинающемуся* с этой даты».

Новые термины, важные слова и текст, который вы видите на экране, будут подсвечены жирным шрифтом, например: «Раздел **Workloads** содержит настройки, связанные с производительностью».



Предупреждения и важные примечания будут выводиться так.

ЧАСТЬ I

Введение

DAX в бизнес-аналитике

Нет сомнений в том, что в наше время одним из главных активов любой организации является информация. Мы прекрасно ощущаем это на себе, когда компании выстраиваются в очередь в погоне за нашими личными данными. И это происходит не от того, что мы им интересны как личности (хотя, разумеется, среди нас есть и выдающиеся люди), просто в совокупности данные о нас всех позволяют компаниям делать важные выводы и тем самым повышать эффективность своего бизнеса.

И это верно не только применительно к коммерческим структурам. Общественные организации, клиники и университеты также активно ведут сбор информации в целях улучшения качества своей деятельности. Так или иначе, сегодня именно информация является двигателем прогресса и инноваций.

В то же время процесс преобразования исходных данных в информацию, а затем в аналитические сведения может быть весьма сложным и утомительным. Он включает в себя сбор данных из самых разных источников, исследование скрытых структур и зависимостей в них с учетом требуемого контекста. Именно поэтому задачи по анализу данных традиционно отдаются на откуп отделам информационных технологий, хотя это далеко не всегда бывает оптимально, поскольку речь зачастую идет о *бизнес-аналитике*, и знание специфики ведения бизнеса здесь крайне важно.

Эта книга посвящена одному из мощнейших средств аналитики на сегодняшний день – **DAX** (Data Analysis eXpressions – выражения для анализа данных). Мы предполагаем, что у вас уже есть определенный опыт работы с этим языком запросов и вы хотите улучшить свои навыки. А для этого очень важно понимать, для чего DAX подходит идеально, а для чего – нет. Кроме того, вы должны выработать привычку использовать DAX везде, где он может проявить себя наилучшим образом.

В этой главе мы рассмотрим основные концепции, которые помогут вам заложить эти основы. Темы, которые будут рассмотрены в этой главе:

- пятислойная модель бизнес-аналитики;
- аналитика предприятия и аналитика конечных пользователей;
- для чего используется DAX и где его найти;
- инструменты для разработки моделей и написания выражений DAX;
- сделано при помощи DAX: визуальные элементы и интерактивные отчеты;
- подходы к разработке решений;
- цикл цифровой трансформации.

Пятислойная модель бизнес-аналитики

Чтобы можно было обсуждать аналитику структурированно и всеобъемлюще, мы разработали простую схему, описывающую основные компоненты и ответственности в рамках аналитического решения. И имя для схемы, представленной на рис. 1.1.1, выбрано тоже простое – *пятислойная модель бизнес-аналитики*.

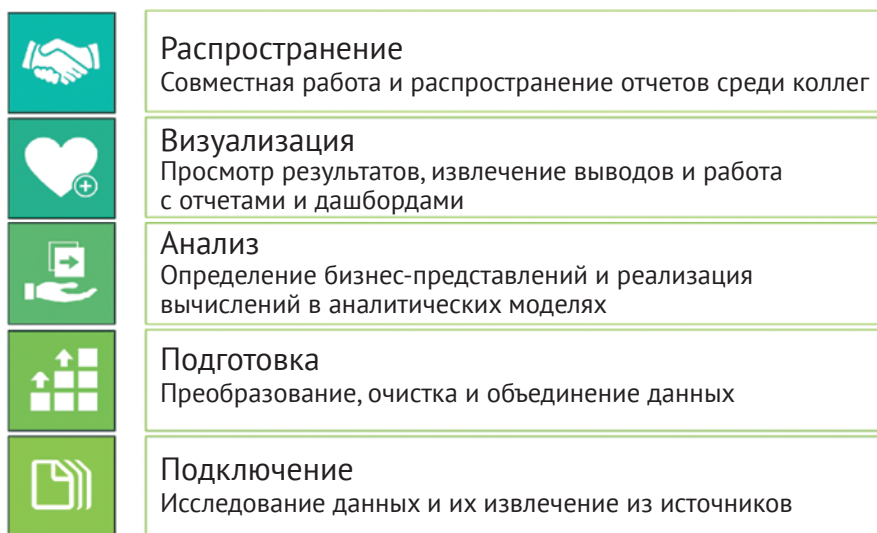


Рис. 1.1.1. Пятислойная модель бизнес-аналитики

Первый и нижний слой – **Подключение** – знаменует собой отправную точку для аналитики: если вам нужно проанализировать данные, для начала их необходимо как-то получить. Источниками при этом могут быть файлы Excel, текстовые файлы, полноценные базы данных или серверы в интернете.

Обычно исходные данные поставляются не в самом лучшем виде для анализа, особенно если они происходят из разных источников. И для исправления этой неприятности служит второй слой на нашей схеме – **Подготовка**. Этот важный процесс может включать в себя различные действия, такие как преобразование типов, трансформация данных, восстановление исторических сведений или их сопоставление на основе ключевых атрибутов.

Получение чистых и опрятных наборов данных на этапах подключения и подготовки может отнять очень много сил и времени. Создание *хранилища данных* (data warehouse), – а именно этим занимаются в отделе ИТ при подготовке данных, – зачастую выливается в разработку сопутствующих проектов, которые длятся годами. Но самое грустное, что к моменту готовности хранилища все понимают, что мир давно ушел вперед и оно утратило свою актуальность.

После приведения исходных данных в формат, готовый для обработки, приходит черед третьего слоя под названием **Анализ**. Именно здесь бьется сердце всего аналитического решения в целом. Построение аналитических моделей позволяет осуществлять срез и фильтрацию данных, рассчитывать агрегаты всех видов и добавлять вычисления для получения специфических выводов.

На слое **Визуализация** происходит создание отчетов и дашбордов на основе созданных ранее аналитических моделей. Мы назвали этот слой **Визуализация**, а не просто **Вывод**, поскольку для нас принципиально не просто предоставить пользователю результаты, а отобразить их таким образом, чтобы он мог сделать важные выводы и сконцентрировать внимание на ключевых показателях. Классические отчеты со страницами детализации не очень подходят для извлечения важных сведений, и пользователю нередко приходится полностью выгружать отчет в Excel и там агрегировать нужные ему показатели. Действительно эффективная визуализация данных позволяет пользователю точно выделить для себя наиболее важные составляющие и при необходимости погрузиться в детальные сведения прямо из отчета.

На верхнем слое – **Распространение** – выполняется осуществление доступа пользователей к отчетам и дашбордам по заранее установленным правилам.

Строите ли вы свое аналитическое решение в Excel, используете ли Power BI, разрабатываете собственную систему бизнес-аналитики уровня предприятия или вовсе не используете никакие системы автоматизации, вы так или иначе будете проходить сквозь все пять слоев. При должной организации процесса разработки аналитического решения все слои четко отделены друг от друга, и все преобразования данных выполняются в нужном слое. Это позволяет, в частности, избежать многократного дублирования логики. К тому же четкое разграничение между слоями поможет легко и безболезненно справиться с возникшими изменениями – к примеру, при смене источников данных.

Как вы увидите далее в этой главе, DAX живет именно в слое **Анализ** и тесно связан с соседними слоями **Подготовка** и **Визуализация**. Визуализации мы посвятим отдельный раздел, а пока нам необходимо разобраться, *кто и что делает в бизнес-аналитике*.

Аналитика предприятия и аналитика конечных пользователей

Компании в наше время становятся все более зависимы от данных. В частности, они нередко оценивают свою деятельность, обращаясь к дашбордам с *ключевыми показателями эффективности* (key performance indicators – KPI). Обычно эти дашборды строятся, исходя из очень высоких стандартов – в высших эшелонах организаций тщательно обсуждают долгосрочную стратегию и бизнес-процессы, способы измерения KPI и варианты их отслеживания. В результате утвержденные дашборды с выводом этих ключевых показателей могут долгое время не меняться, а их созданием и поддержкой занимается центральный отдел информационных технологий или специализированный центр бизнес-аналитики.

Также следствием зависимости компаний от данных является тот факт, что *каждое* принимаемое решение основывается исключительно на полученных цифрах. В результате появляется необходимость в более динамичной аналитике, способной отвечать на оперативные, не запланированные заранее запросы. Такой вид аналитики получил название *аналитика самообслуживания* (self-service BI), что предполагает возможность построения полноценных бизнес-решений без вмешательства отдела ИТ. Взглянув еще раз на нашу схему пя-

тислойной модели бизнес-аналитики, вы можете понять, что это практически невозможно. Мало у кого из пользователей есть необходимые знания, навыки и время, чтобы обслуживать все пять слоев. В идеале конечные пользователи и отдел информационных технологий должны дополнять друг друга, занимаясь приоритетными для них задачами.

Мы различаем два вида аналитики: *аналитику предприятия* (enterprise BI) и *аналитику конечных пользователей* (end-user BI). В первом случае вся аналитическая нагрузка по созданию и поддержке решений ложится на плечи отдела ИТ, а в качестве технической базы для обслуживания большого количества пользователей используются масштабируемые серверные системы и облачные платформы. При этом большое внимание уделяется качеству данных, а все сопутствующие процессы хорошо выстроены и отлажены.

Что касается аналитики конечных пользователей, то здесь, как ясно из названия, ключевую роль играют бизнес-пользователи. Для качественного выполнения своей повседневной работы им необходимо быстро получать те или иные данные из общей системы, и они это делают. Многие из них пытаются строить собственные аналитические решения в Excel, однако сложность структуры документов в Excel в конечном счете приводит к еще большим временным затратам на поддержку актуальности данных и масштабирование решения. К тому же Excel не предназначен для обработки сколько-нибудь больших данных. Зная это, компания Microsoft расширила возможности Excel за счет Power Query и Power Pivot как средств подготовки и анализа данных, что позволило превратить этот инструмент в подобие полноценной аналитической платформы для конечного пользователя. В результате совокупность этих инструментов вылилась в продукт, который мы сегодня называем **Power BI**.

Прелесть аналитической платформы Microsoft состоит в том, что в рамках нее инструменты аналитики предприятия и аналитики конечного пользователя могут работать совместно без каких-либо сложностей. И движущей силой этого процесса выступает технология, лежащая в основе Power BI. По сути, с ее появлением мы получили возможность сочетать ранее разрозненные ветви аналитики в одной архитектуре с разными уровнями возможностей самообслуживания, что показано на рис. 1.1.2.



Рис. 1.1.2. Интеграция аналитики предприятия и аналитики конечного пользователя

Наша пятислойная модель идеально подходит для демонстрации возможностей пользователей в рамках системы самообслуживания. Пользователи, нуждающиеся в собственной аналитике, могут интегрироваться в общую аналитическую систему на любом уровне. К примеру, они могут ограничиться созданием собственных визуализаций на базе существующих (корпоративных) аналитических моделей или строить свои аналитические модели на основе централизованно подготовленных наборов данных. Наконец, опытные бизнес-пользователи могут самостоятельно собирать и подготавливать данные, объединять их с корпоративными данными и создавать свои модели и отчеты. Более того, в своей системе они могут использовать объекты, созданные другими пользователями.

Очевидно, что при движении вниз сквозь описанные выше слои значительно повышается и уровень сложности выполняемых пользователем операций. И речь идет не только о компетенциях, которыми, безусловно, он должен обладать, но и об ответственности в виде следования корпоративным стандартам и инструкциям. Централизованные отделы ИТ и бизнес-аналитики должны всячески способствовать тому, чтобы избранные пользователи могли строить собственные аналитические решения. При правильной организации это ведет к созданию среды, в которой все пользователи оказываются в выигрыше от доступа к нужной им аналитической информации. Мы называем это *коллективной аналитикой* (collective analytics).

Power BI предлагает возможности по реализации коллективной аналитики. Одним из таких инструментов является DAX, который способен наделить пользователей возможностью не только создавать свои собственные аналитические решения, но и успешно встраиваться в разработку решений корпоративного уровня. О последней возможности мы поговорим в одном из следующих разделов, а пока нужно напомнить вам, что такое DAX и где его найти.

Для чего используется DAX и где его найти

В аналитических решениях, построенных на базе платформы Microsoft, DAX присутствует на слое **Анализ**. DAX живет внутри аналитических моделей в виде языка написания формул для определения вычислений и прочей логики. По сути, модели и DAX являются двумя сторонами одной и той же монеты: дизайн модели непосредственно влияет на сложность кода DAX, а ваши навыки в области DAX определяют вид и сложность модели (о ключевых концепциях моделирования данных мы будем говорить в главе 1.2).

Мощь языка DAX кроется в его потрясающих возможностях в плане агрегации данных. В состав DAX входит множество функций и конструкций, помогающих выстроить агрегацию идеальным образом для получения необходимой аналитики. В прошлом далеко не все типы агрегаций можно было рассчитывать напрямую – вместо этого они должны были быть реализованы на этапе подготовки данных. К примеру, накопительные итоги с начала года могут быть рассчитаны в DAX при помощи одной функции, тогда как в Excel или традиционных инструментах отчетности вам понадобится настроить немало метрик, показывающих, какие транзакции принадлежат к периоду накопления.

Конец ознакомительного фрагмента.

Приобрести книгу можно

в интернет-магазине

«Электронный универс»

e-Univers.ru