

Содержание

<i>От издательства</i>	14
<i>Предисловие</i>	15
<i>Об авторах</i>	17
<i>О рецензентах</i>	19
<i>О переводчике</i>	20
<i>Введение</i>	21
Глава 1 Введение в язык M	26
История языка M	26
Кому следует изучать язык M?	28
Где и как применяется язык M?	29
Инструменты	29
Продукты и службы	30
Зачем изучать язык M?	35
Основы языка M	37
Выражение let	39
Особенности языка M	40
Формальная классификация	40
Неформальные особенности языка M	44
Заключение	46
Глава 2 Работа с Power Query/M	47
Технические требования	47
Краткий обзор Power Query Desktop	48
Краткий обзор	49
Ваш первый запрос	54
Параметры и настройки источника данных	57
Редактирование сгенерированного кода на языке M	64
Создание настраиваемых столбцов	66
Добавление столбца индекса	66
Добавление столбца из примеров	67
Математические операции	68
Добавление настраиваемых столбцов на основе кода на M	69
Использование расширенного редактора	71
Заключение	74
Глава 3 Извлечение и объединение данных	75
Технические требования	76
Доступ к файлам и папкам	76
Функция File.Contents	77

Текстовые и CSV-файлы	78
Файлы Excel	82
Работа с папками	85
Работа с файлами PDF	92
Работа с XML	94
Хранилище Azure	97
Дополнительные форматы файлов	100
Извлечение содержимого веб-страниц	101
Функции для работы с двоичными данными	104
Функции Lines	108
Доступ к базам данных и кубам	110
Функции Cube	112
Работа со стандартными протоколами данных	114
Дополнительные коннекторы	117
Популярные программные продукты	118
Функции для работы с данными идентификации	119
Комбинирование и объединение данных	120
Функция Table.Combine	120
Функции Table.NestedJoin и Table.Join	122
Функции Table.FuzzyNestedJoin и Table.FuzzyJoin	125
Заключение	126

Глава 4 Значения и выражения в языке M

Введение в типы значений	129
Двоичные значения	130
Значения календарного типа	132
Логические значения	139
Значения null	141
Числовые значения	143
Текстовые значения	145
Списки	146
Записи	147
Таблицы	149
Значения-функции	150
Значения-типы	151
Операторы	152
Выражения	155
Вложенные выражения let	157
Рекомендации по использованию выражений	160
Управляющие структуры	162
Перечисления	165
Заключение	168

Глава 5 Типы данных

Что такое типы данных?	169
Система типов	170
Столбцы со смешанными типами	171

Тип данных столбца против типа значения	173
Важность типов данных	175
Прозрачность и единообразие	175
Проверка допустимости данных	176
Другие причины использования явных типов данных	177
Типы данных, доступные в языке М	178
Примитивные типы	178
Пользовательские типы	187
Распознавание типов данных	194
Извлечение типов данных из источника	194
Преобразование типов данных	196
Преобразование типов значений	197
Преобразование типов столбцов	199
Избежание потери данных при преобразовании	201
Влияние выбранной локали	204
Фасеты	206
Утверждения типов	210
Приписывание типов	213
Что такое приписывание?	214
Функции, поддерживающие приписывание типов	216
Ошибки, связанные с приписыванием типов	219
Эквивалентность и совместимость типов	225
Эквивалентность типов	225
Согласованность типов	226
Заявление типов	227
Заключение	228

Глава 6 Структурированные значения

Введение в структурированные значения	229
Списки	230
Введение в списки	230
Операторы для работы со списками	232
Способы создания списков	235
Доступ к элементам списков	238
Распространенные операции со списками	240
Присваивание типов данных спискам	242
Записи	244
Знакомство с записями	244
Операторы для работы с записями	246
Способы создания записей	249
Доступ к полям записей	251
Распространенные операции с записями	253
Присваивание типов данных записям	257
Таблицы	259
Знакомство с таблицами	259
Операторы для работы с таблицами	260
Способы создания таблиц	263

Доступ к данным в таблицах	268
Распространенные операции с таблицами.....	272
Присваивание типов данных таблиц	273
Заключение	274
Глава 7 Концепции языка М	276
Технические требования.....	277
Введение в области видимости	277
Знакомство с глобальным окружением	281
Изучаем секции	282
Создание собственного глобального окружения	283
Введение в замыкания	284
Свертывание запросов.....	286
Управление метаданными	288
Заключение	295
Глава 8 Работа с вложенными структурами	296
Переходим к написанию кода вручную	296
Первые шаги.....	297
Преобразование значений в таблицах.....	303
Работа со списками	311
Преобразование списков.....	311
Извлечение элементов.....	314
Изменение размера списка	315
Фильтрация списка.....	320
Преобразование в список.....	321
Другие операции	324
Работа с записями	330
Преобразование записей.....	330
Извлечение значений полей	332
Изменение размера записей	333
Фильтрация записей	335
Преобразование в запись.....	337
Условный поиск или замена значений.....	341
Работа с таблицами	343
Преобразование таблиц.....	345
Извлечение значений ячеек	345
Изменение размера таблицы в длину.....	347
Изменение размера таблицы в ширину	349
Фильтрация таблиц.....	353
Преобразование в таблицу	358
Информация о таблицах	360
Работа со смешанными структурами	364
Список таблиц, списков или записей.....	364
Таблицы со списками, записями или таблицами внутри	367
Смешанные структуры	368
Заключение	375

Глава 9	Параметры и настраиваемые функции	376
	Параметры	376
	Введение в параметры	377
	Создание параметров	378
	Использование параметров в запросах	380
	Собираем все вместе	382
	Настраиваемые функции	388
	Что такое настраиваемые функции?	388
	Преобразование запросов в настраиваемые функции	390
	Вызов настраиваемых функций	396
	Выражение each	398
	Усовершенствование настраиваемых функций	402
	Обращение к именам столбцов и полей	408
	Отладка настраиваемых функций	410
	Область видимости функции	412
	Собираем все вместе	413
	Заключение	424
Глава 10	Работа с датами, временем и длительностями	426
	Технические требования	426
	Даты	426
	Таблица с календарем в языке M	433
	Другие форматы даты	435
	Дополнительные функции для работы с датами	439
	Время	441
	Создание таблицы времени	444
	Классификация смен	445
	Даты и время	445
	Часовые пояса	447
	Корректировка времени обновления данных	448
	Длительности	449
	Работа с длительностями	450
	Заклучение	451
Глава 11	Сравнение, замена, объединение и разделение	452
	Технические требования	452
	Ключевые концепции	453
	Вызов функции	453
	Распространенные ошибки	454
	Замыкания	454
	Функции высшего порядка	455
	Анонимные функции	457
	Упорядочивание значений	458
	Функции сравнения	458
	Comparer.Equal	459
	Comparer.Ordinal	460

Критерии сравнения	465
Числовые значения	466
Вычисление ключа сортировки	466
Списки с ключами и порядком сортировки	468
Настраиваемая функция сравнения с условной логикой	470
Настраиваемая функция сравнения с Value.Compare	470
Критерии равенства	472
Функции сравнения по умолчанию	472
Собственные функции сравнения	473
Выбор ключей	473
Комбинирование выбора ключей с функциями сравнения	474
Функции замены	474
Replacer.ReplaceText	476
Replacer.ReplaceValue	477
Собственные функции замены	481
Функции объединения	484
Combiner.CombineTextByDelimiter	486
Combiner.CombineTextByEachDelimiter	487
Combiner.CombineTextByLengths	488
Combiner.CombineTextByPositions	489
Combiner.CombineTextByRanges	490
Функции разделения	491
Splitter.SplitByNothing	493
Splitter.SplitTextByAnyDelimiter	494
Splitter.SplitTextByCharacterTransition	495
Splitter.SplitTextByDelimiter	496
Splitter.SplitTextByEachDelimiter	497
Splitter.SplitTextByLengths	498
Splitter.SplitTextByPositions	499
Splitter.SplitTextByRange	500
Splitter.SplitTextByRepeatedLengths	501
Splitter.SplitTextByWhitespace	501
Практические примеры	502
Удаление управляющих символов и лишних пробелов	503
Извлечение электронного адреса из строки	504
Разработка функции getEmail	505
Разделение объединенных значений в ячейках на строки	508
Множественная замена значений	512
Условное объединение строк	516
Заключение	520

Глава 12 Обработка ошибок и отладка

Технические требования	521
Что такое ошибка?	522
Сдерживание ошибок	522
Определение возникновения ошибок	523
Возбуждение ошибок	526

Выражение error	527
Оператор ... (многоточие)	529
Обработка ошибок.....	530
Работаем со значениями null при помощи оператора слияния	530
Опциональный доступ к элементу или полю.....	531
Перечисление MissingField.Type.....	532
Выражение if и условная логика.....	532
try	533
try и otherwise	533
try и catch	533
try и функция	534
Функции из стандартной библиотеки	534
Стратегии отладки.....	535
Распространенные ошибки.....	537
Синтаксические ошибки	537
Исправление ошибок – главный приоритет	538
DataSource.Error, невозможность обнаружить источник	539
Неизвестные или отсутствующие идентификаторы.....	540
Неизвестная функция	540
Ссылка на неизвестный столбец	541
Ссылка на неизвестное поле	542
Недостаточное количество элементов в перечислении	543
Ошибка Formula.Firewall	543
Expression.Error: ключ не соответствует ни одной строке в таблице	544
Expression.Error: ключ соответствует более чем одной строке в таблице	545
Expression.Error: вычисление привело к переполнению стека и не может быть продолжено	545
Собираем все вместе	546
Выбор столбцов	547
Разработка собственного решения	547
Отчет об ошибках уровня ячейки	551
Разработка собственного решения	551
Заключение	556

Глава 13 Итерации и рекурсия.....557

Введение в итерации	557
List.Transform	558
List.Accumulate.....	564
List.Generate.....	571
Рекурсия.....	599
Почему рекурсия так важна?.....	599
Рекурсия против итераций: краткое сравнение	600
Рекурсивные функции	600
Использование оператора @	601

Удаление последовательных пробелов.....	603
Рекурсия и быстроедействие.....	604
Заключение	605
Глава 14 Сложные шаблоны работы с данными	606
Распознавание шаблонов	607
Основы распознавания шаблонов	607
Извлечение фиксированных шаблонов.....	609
Объединение данных.....	630
Основы объединения данных	630
Извлечение, преобразование и объединение	632
Заключение	652
Глава 15 Оптимизация производительности.....	654
Использование памяти при выполнении запросов	654
Варианты ограничения памяти.....	655
Свертывание запросов	656
Свертывание запросов в действии	658
Выполнение запроса	660
Свертывание, несвертывание и частичное свертывание	660
Инструменты для определения свертываемости шагов запроса.....	661
Операции и их влияние на свертываемость запроса.....	666
Уровни конфиденциальности источников данных	668
Запросы на языке баз данных.....	668
Функции для предотвращения свертывания запросов	668
Стратегии для поддержки свертывания запросов.....	669
Файрвол формул	679
Что такое файрвол формул?	679
Введение в партиции	680
Основные принципы файрвола формул	680
Ошибка файрвола: ссылка на другие партиции.....	681
Ошибка файрвола: доступ к совместимым источникам данных	688
Оптимизация эффективности выполнения запросов.....	694
Фильтрация строк и удаление столбцов.....	694
Буферные и потоковые операции.....	695
Влияние буферизации и накопительные итоги	698
Источники данных.....	702
Советы по быстромудействию	704
Заключение	705
Глава 16 Расширения	706
Технические требования.....	706
Что такое расширение в Power Query?.....	706
Что можно сделать при помощи расширений?	709

Подготовка окружения	709
Установка Visual Studio Code	710
Установка Power Query SDK	710
Настройка Internet Information Services	710
Настройка Discord	712
Создание пользовательского коннектора.....	714
Создание проекта расширения.....	715
Настройка аутентификации	718
Настройка навигации и содержимого	728
Установка и использование пользовательского коннектора	736
Заключение	739
 <i>Предметный указатель</i>	 740

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. 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 электронной почты dmkpress@gmail.com.

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

Предисловие

Временами звезды сходятся так, что результат превосходит самые смелые ожидания. Именно так произошло и с книгой, которую вы держите в руках. Но я, кажется, опережаю события. Давайте вернемся в начало 2022 года, когда и началась история этой книги...

К тому времени я уже больше 30 лет профессионально занимался обработкой данных и специализировался на Excel и Power BI, параллельно проводя обучение. Я очень активно использовал в своей работе Power Query для очистки и преобразования данных перед их полноценным анализом и построением отчетов на их основе. При этом почти все преобразования я выполнял при помощи инструментов пользовательского интерфейса Power Query. Но каким бы удобным и продуманным ни был этот интерфейс, он не дает нам полной свободы действий, скрывая от нас большую часть из 700+ функций, доступных в языке M. И вот в начале 2022 года я решил, что пришла пора перестать работать с языком M на ощупь, а заняться полноценным его изучением. В то время это было непросто – несмотря на существование большого количества книг по Power Query, лишь малая доля повествования в них касалась языка M в чистом виде. В интернете на тот момент было множество ресурсов, посвященных языку M, но в основном они содержали решения узконаправленных задач, что также не очень подходило для нужд полноценного обучения.

Несмотря на все эти преграды, к концу 2022 года мои усилия по погружению в недра языка M начали давать определенные плоды. В частности, мои отчеты в Power BI стали работать намного быстрее, используемые в них выражения DAX значительно упростились, я стал тратить гораздо меньше времени на очистку данных и повторно использовать собственные настраиваемые функции с целью автоматизации рутинных процессов. Кроме того, я заметил, что понимание глубинных идей, лежащих в основе визуализации и преобразования сложных структур данных с помощью M, положительно сказалось на качестве моего кода в других языках программирования. В то же время я с горечью осознавал, что никто другой не сможет пройти мой путь, поскольку в процессе обучения я по большей части черпал свои знания у троих своих друзей (впоследствии – авторов этой книги), которых можно было отнести к числу наиболее продвинутых экспертов по Power Query/M в мире. Я помню, мы частенько шутили с Мелиссой и Риком о том, что они *избраны* для ответа своим шедевром на шедевр Марко Руссо и Альберто Феррари «Подробное руководство по DAX» (<https://dmkpress.com/catalog/computer/data/978-5-97060-859-3>), но в то же время я понимал, что для каждого из них по отдельности это было бы слишком монументальной задачей.

В то время я тестировал приложения, которые в большом количестве писал Грегори Деклер. Мы часто затрагивали тему Power Query и языка M при обсуждении возможного будущего этих приложений. Однажды Грег сказал мне, что работает над черновиком полного руководства по языку M – той самой книги,

о которой мы частенько шутили с Мелиссой и Риком (и о появлении которой я грезил в своих мечтах). Он спросил меня, кого бы я порекомендовал ему в со-авторы, и угадайте, что я ему ответил (сюрприз, сюрприз: спойлер на обложке).

Овладение языком М – процесс небыстрый и нелегкий. Но если вы работаете в аналитике, разрабатываете отчеты в Power BI, трудитесь на ниве исследования данных или бизнес-анализа или взаимодействуете с Excel, желая развить свои профессиональные навыки, то вам, как и мне, изучение языка М с такими блестящими и терпеливыми наставниками, какие были у меня, тоже покажется легкой увеселительной прогулкой!

Брайан Джулиус,
эксперт и инструктор по Power BI

Об авторах

Грегори Деклер (Gregory Deckler) – семикратный обладатель Microsoft MVP в области Data Platform, активный блогер и член сообщества Power BI, в активе которого более 6000 ответов на вопросы посетителей. Грег является автором множества книг по Power BI, включая первое и второе издания книги «Learn Power BI», второе издание «Power BI Cookbook» и «Mastering Power BI», а также «DAX Cookbook». Кроме того, Грег создал несколько полезных инструментов для Power BI и регулярно размещает видеоконтент на своих каналах в YouTube: «Microsoft Hates Greg» и «DAX For Humans».

Я хотел бы сердечно поблагодарить моих драгоценных соавторов по этой книге, у которых я многому научился. Также я выражаю признательность моему сыну Рокету за поддержку всех моих начинаний. Особая благодарность Брайану Джулиусу, вдохновившему меня на написание этой книги, – без него она бы не увидела свет.

Рик де Гроот (Rick de Groot) – консультант по Power BI, блогер и автор каналов в YouTube из Нидерландов. Заручившись богатым опытом работы в сфере финансов, он посвятил более 14 лет развитию навыков в области анализа данных, работая с Power Query и языком M. Рик является независимым консультантом и преподает Power BI, Power Query и DAX. Он регулярно публикует полезные статьи в своих блогах «Power Query How» и «BI Gorilla», которые стали очень важными и популярными ресурсами для тех, кто интересуется языком M. Стремление Рика делиться своими знаниями с окружающими вылилось в две номинации Microsoft MVP в области Data Platform.

Я бы хотел искренне поблагодарить Брайана Джулиуса. Его вера в мои силы подтолкнула меня к тому, чтобы больше делиться с окружающими своими знаниями в области Power Query и развить навыки владения языком M. Без его поддержки мое имя вряд ли могло появиться на обложке этой книги. Кроме того, Брайан принимал активное участие в обсуждении содержания книги и очень помог с выстраиванием логических цепочек. Я очень признателен тебе, Брайан.

Также я бы хотел поблагодарить свою семью – ваши терпение и поддержка при написании этой книги значат для меня очень много.

Кроме того, я выражаю благодарность каждому из будущих читателей. Все наши дискуссии оказались очень ценными.

Мелисса де Корте (Melissa de Kort) – вдохновленный оптимизатор, славящийся своим умением значительно упрощать задачи при помощи Power Query. Богатый опыт работы в структуре крупных компаний помог Мелиссе внести существенный вклад в развитие сообщества. Она является настоящим адептом рас-

пространения знаний, а в ее актив входят блоги, учебники, курсы и вебинары, способствующие росту популярности Power Query и языка M.

Моя искренняя благодарность Марселю и Сэму за наполнение радостью тех моментов, которые мы провели вместе.

Также я выражаю глубокую признательность моим соавторам по книге и Брайану Джулиусу. Благодарю мою семью за бесконечное терпение и поддержку. Кроме того, хотелось бы отметить всех, кто щедро делится своими знаниями, наполняя сообщество новыми идеями и их воплощениями. Ваш вклад неоценим!

О рецензентах

Эрин Островски (Erin Ostrowsky) – цифровой кочевник, стремящаяся создавать здоровые и позитивные отношения везде, где появляется. Она обожает STEAM – не забывайте об искусстве! – и всегда старается решать проблемы позитивно и весело – этого так не хватает нашему миру! Эрин любит учиться, творить, заводить друзей, путешествовать и писать о своей жизни. В настоящее время она руководит компанией Fit 4 Duty Data, занимающейся консалтингом и обучением.

Я хотела бы поблагодарить Бога, моих друзей и мою семью – вы делаете этот мир лучше, особенно в трудные времена!

Вахид Доустимадж (Vahid Doustimaj) – признанный Microsoft MVP и сертифицированный преподаватель с более чем 15-летним опытом. Вахид – настоящий технофил, интересующийся управлением проектами и аналитикой данных. В настоящее время работает в Австралии руководителем аналитического проекта. Является активным участником сообществ Microsoft Fabric и Power BI и организует мероприятия для персидского подразделения Power BI.

Я бы хотел поблагодарить свою жену за бесконечное терпение и помощь. Ее поддержка стала моей путеводной звездой.

Ахмед Ойелово (Ahmed Oyelowo) – управляющий партнер, инструктор и консультант компании Foresight BI & Analytics Global Solutions, сертифицированный аналитик данных Microsoft Power BI и аналитик данных Azure Enterprise. Также обладает сертификатом МСТ и является автором самого популярного курса по Power BI на Udemy (90 000+ студентов). Четырехкратный обладатель Microsoft MVP в области Data Platform.

О переводчике



Александр Гинько, обладающий богатым опытом работы в сфере ИТ и более десяти лет посвятивший переводам книг и статей на самые разные темы, в последние годы специализируется на переводе книг в области бизнес-аналитики и программирования для издательства «ДМК Пресс» по направлениям Python, R, SQL, Power BI, DAX, Excel, Power Query, Tableau... На данный момент в активе Александра уже более 20 книг, включая одну авторскую, и он продолжает плодотворно работать над переводом новых книг. Читателей

этой книги могут заинтересовать следующие переводы Александра на тему Power BI, DAX и Power Query, ставшие бестселлерами:

- «Подробное руководство по DAX» (<https://dmkpress.com/catalog/computer/data/978-5-97060-859-3>);
- «Анализ данных при помощи Microsoft Power BI и Power Pivot для Excel» (<https://dmkpress.com/catalog/computer/data/978-5-97060-858-6>);
- «Приручи данные с помощью Power Query в Excel и Power BI» (<https://dmkpress.com/catalog/computer/data/978-5-93700-105-4>) и др.

Помимо перевода книг, Александр ведет свой канал в Telegram (https://t.me/alexanderginko_books), на котором вы можете из первых уст получить ответы на все интересующие вас вопросы об уже переведенных книгах, находящихся в работе и запланированных на будущее. Также на канале можно найти промокоды на все книги Александра для покупки книг на сайте издательства «ДМК Пресс» с большими скидками.

Конец ознакомительного фрагмента.

Приобрести книгу можно

в интернет-магазине

«Электронный универс»

e-Univers.ru