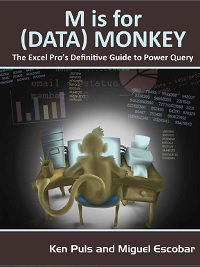
**Кен Пульс и Мигель Эскобар. Язык М для Power Query**

Power Query – относительно новый компонент Excel, используемый для извлечения данных из внешних источников. Перед помещением на лист Excel в редакторе Power Query данные можно обработать. M – имя языка программирования, используемого в Power Query. Книга является полным профессиональным руководством по языку М. Если вы ранее не использовали этот компонент Excel, рекомендую начать с [Марк Мур. Power Query](http://baguzin.ru/wp/?p=19403) и [Power Query в помощь](http://baguzin.ru/wp/?p=20373).

Ken Puls and Miguel Escobar. M is for (Data) Monkey.[[1]](#footnote-1) The Excel's Pro's Definitive Guide to Power Query. – Holy Macro! Books, 2015. – 308 p.



[Введение: революция в подготовке данных](#_Введение:_революция_в)

[Глава 1. Основы импорта данных](#_Глава_1._Основы)

Глава 2. Настройки по умолчанию

Глава 3. Основные операции добавления данных

Глава 4. Импорт всех файлов из папки

Глава 5. Агрегирование листов Excel

Глава 6. Обратный разбор сводной таблицы

Глава 7. Импорт больших текстовых файлов

Глава 8. Импорт из баз данных

Глава 9. Объединение таблиц и запросов

Глава 10. Куда загружаются запросы

Глава 11. Определение типов данных

Глава 12. Импорт веб-данных

Глава 13. Загрузка данных из Exchange

Глава 14. Группировка и суммирование

Глава 15. Транспонирование и обработка сложных сводных таблиц

Глава 16. Автоматизация обновления

Глава 17. Формулы Power Query

Глава 18. Условная логика

Глава 19. Объекты Power Query

Глава 20. Понимание языка М

Глава 21. Создание пользовательских функций

Глава 22. Продвинутая условная логика

Глава 23. Таблицы динамических параметров

Глава 24. Динамические календари

Глава 25. Организация запросов

### Введение: революция в подготовке данных

Анализируя данные, мы, как правило, выполняем следующие действия:

* преобразовываем данные в соответствии с потребностями,
* добавляем новые данные к существующему набору,
* объединяем несколько наборов данных,
* добавляем столбцы или проводим вычисления с набором данных.

Нас могут называть *обезьянами данных*, но чаще более мягко – информационными работниками. Независимо от того, что записано в должностных инструкциях, наша роль – очистить данные и превратить их в информацию. Наша работа вряд ли будет отмечена на ежегодном собрании, но она важна для организаций. К сожалению, даже если работа сделана хорошо, результаты анализа могут вызывать подозрения.

Хотя Excel имеет мощный набор инструментов, чтобы выполнить бизнес-анализ, преобразование необработанных данных всегда было проблемой. На самом деле, именно на подготовку данных мы зачастую тратим б***о***льшую часть времени. До сих пор основные усилия профессионалов Excel по извлечению и подготовке данных были направлены на изучение богатого арсенала формул Excel, языка VBA и SQL-запросов. Два последних метода используются немногими, так как требуют значительных инвестиций времени в их освоение.

Power Query решает проблему подготовки данных. Этот инструмент прост в освоении и имеет интуитивно понятный интерфейс. Каждое действие запоминается, а код при желании можно просмотреть и обновить. Когда дело доходит до импорта, очистки и преобразования данных, вы можете понять Power Query быстрее, чем освоить формулы Excel. С Power Query обрабатывать грязные данные проще, чем с VBA.

Power Query позволяет выполнять многие преобразования с помощью команд меню. Этим командам соответствует код на языке М, но вы не обязаны его знать. Код создается автоматически. Если же вы решите воспользоваться языком М, то сможете писать гораздо более эффективные запросы.

С Power Query вы можете загружать данные в одно из трех мест: листы Excel, модель данных Power Pivot, Power BI, или ограничиться подключением к источнику данных. Последний пункт может показаться странным, но он означает, что некоторые запросы могут быть использованы в качестве источника данных для других запросов.

Эта книга была написана с использованием Power Query версии 2.24, выпущенной в июле 2015 г.[[2]](#footnote-2) Чтобы установить Power Query в Excel 2010 и Excel 2013 загрузите надстройку с сайта [Microsoft](https://www.microsoft.com/ru-RU/download/details.aspx?id=39379). Начиная с Excel 2016 Power Query уже встроен.

Мы много лет изучали Power Query, а теперь и сами [обучаем](https://www.powerquery.training/) работе в этой программе.

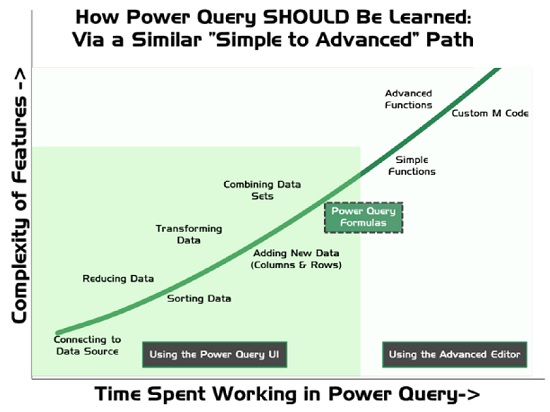


Рис. 1. Кривая обучения Power Query

В версиях Excel 2010 и Excel 2013 команды Power Query расположены на отдельной *вкладка Power Query*. Начиная с версии Excel 2016 Microsoft зачем-то сжал все команды в группу *Получить и преобразовать данные* на вкладке *Данные*.

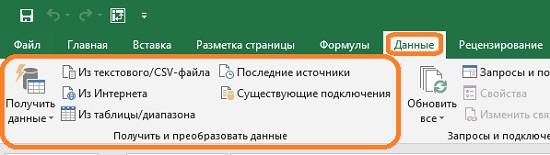


Рис. 2. Размещение команд Power Query в Excel 2016 и более поздних версиях

Оригинальные Excel-файлы, используемые в этой книге, доступны по [ссылке](https://www.powerquery.training/book-files/).

### Глава 1. Основы импорта данных

Бизнес-информация часто хранится в:

* TXT-файлах с каким-то символом в качестве разделителя
* CSV-файлах с разделителем запятая
* книгах Excel

Процесс импорта файла с разделителями следует базовому процессу ETL (extract, transform and load – извлечение, преобразование и загрузка). Чтобы импортировать файл создайте новую (пустую) книгу Excel, перейдите на вкладку *Данные*. В области *Получить и преобразовать данные* кликните *Из текстового/CSV-файла* и выберите Ch01-Delimited.txt. Excel запускает окно предварительного просмотра:

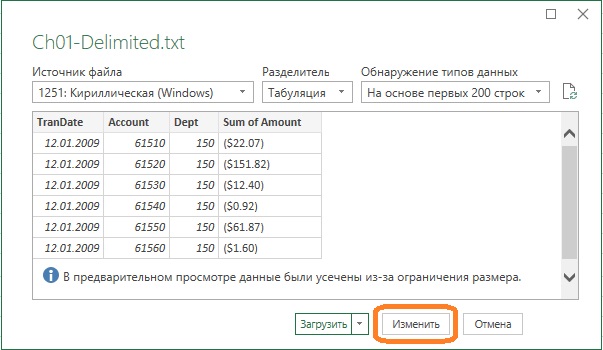


Рис. 3. Окно предварительного просмотра Power Query

Кликните на кнопке *Изменить*. Откроется окно редактора Power Query:

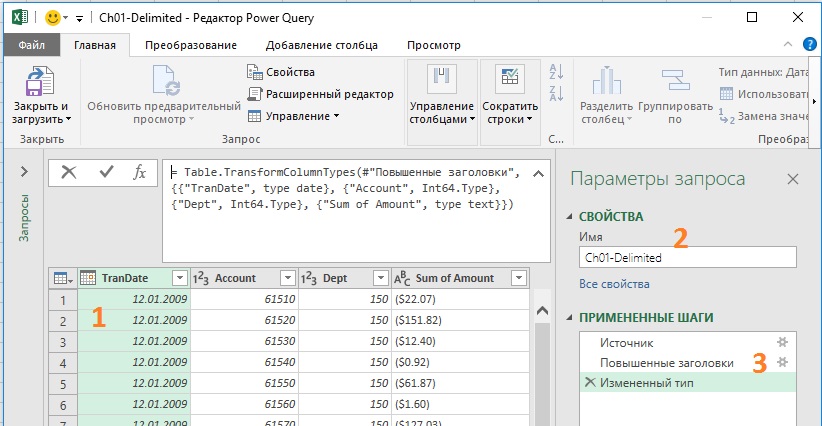


Рис. 4. Окно редактора Power Query

Данные отображаются в табличном формате с заголовками (1). Запрос автоматически получил имя (2), такое же, как и имя файла. В области *Примененные шаги* перечислены три действия (3). Если вы кликните на первом шаге – *Источник*, то увидите немного иное представление данных:

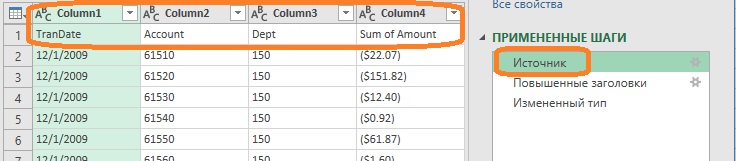


Рис. 5. Данные на первом шаге запроса

Power Query импортировал данные, проанализировал их и автоматически обработал:

* разбил на столбцы, используя разделитель табуляция;
* определил, что данные в каждом столбце во всех строках, кроме первой, однородны; а в первой строке отличаются;
* добавил второй шаг, «подняв» первую строку в заголовки;
* на третьем шаге попытался определить типы данных в каждом столбце; обратите внимание на значки возле заголовков на рис. 4; столбец *TranDate* определил, как даты, а столбцы *Account* и *Dept*, как целочисленные. И это правильно. А вот со столбцом *Sum of Amount* не справился, посчитав, что данные в нем текстовые. (Это связано с тем, что Power Query применил русские региональные настройки к данным в американском стиле с десятичным разделителем точка.)

Вы можете перемещаться по шагам, чтобы увидеть, что Power Query делал с данными. Чтобы сделать заголовки более читаемыми, их можно изменить. Выберите последний шаг в области *Примененные шаги*. Если этого не сделать, то ваши манипуляции попытаются записаться за тем шагом, на котором вы находитесь. А это не всегда возможно, и вносит некоторый хаос.

Кликните правой кнопкой мыши на заголовке столбца *TranDate*, выберите *Переименовать*, удалите часть заголовка *Tran*. Переименуйте последний столбец (рис. 6). Если вы внимательно следили за областью *Примененные шаги*, то заметили, что при переименовании столбца *TranDate* был создан новый шаг (1). Однако переименование столбца *Sum of Amount* не привело к созданию еще одного шага. Вместо этого два шага были объединены вместе, что хорошо видно в коде (2). Power Query пытается сохранить эффективность путем слияния подобных операций при построении запроса.

Изображение выглядит как снимок экрана

Описание создано с очень высокой степенью достоверности

Рис. 6. Переименование столбцов привело к созданию четвертого шага

Power Query также позволяет изменить тип данных, выбранных автоматически. Если, например, столбцы *Account* и *Dept* должны быть отформатированы как текст:

* Выберите столбец *Account*
* Удерживая клавишу Ctrl, выберите столбец *Dept*
* Щелкните правой кнопкой мыши заголовок любого выбранного столбца, пройдите по меню *Тип изменения* –> *Текст*.

Изображение выглядит как текст, снимок экрана

Описание создано с высокой степенью достоверности

Рис. 7. Изменение типа данных

Поля *Account* и *Dept* будут выровнены влево, указывая на то, что они отформатированы как текст, а не значения, в соответствии со стандартами Excel. В поле *Примененные шаги* появится пятый шаг – *Измененный тип 1*.

Изображение выглядит как снимок экрана

Описание создано с очень высокой степенью достоверности

Рис. 8. Значения в столбцах *Account* и *Dept* выровнены влево

Почему же изменение типа не слились с аналогичным на третьем шаге? Power Query не знает, был ли автоматический третий шаг важным, поэтому реагирует безопасным способом и создает новый шаг. Кстати, если теперь вы попробуете выделить столбец *Dept* и изменить тип данных назад на *Целое число*, то Power Query «сообразит», что вы (а не автомат) делаете что-то странное, и предложит отменить предыдущее изменение:

Изображение выглядит как снимок экрана

Описание создано с очень высокой степенью достоверности

Рис. 9. Предупреждение, что более раннее действие, возможно является лишним

Присвойте запросу говорящее имя: просто отредактируйте текст в поле *Имя*. Теперь можно завершить запрос загрузкой. Если кликнуть на кнопку *Закрыть и загрузить* на вкладке *Главная*, данные по умолчанию загружаются в Таблицу на лист Excel. Имя Таблицы совпадет с именем запроса. Справа появится панель *Запросы и подключения*. В ней предоставлена ключевая информация о запросах. Например, количество загруженных строк и обнаруженных ошибок. Из этой панели удобно находить запросы и управлять ими: редактировать, удалять, обновлять.

Изображение выглядит как снимок экрана

Описание создано с очень высокой степенью достоверности

Рис. 10. Загрузка запроса в Таблицу на лист Excel

Запрос может показать наличие ошибок (цифра 1 на рис. 10). Не переживайте: это означает, что региональные настройки не соответствуют формату данных в файле, подготовленных для США. В частности, дата представлена в формате ММ.ДД.ГГГГ. Поэтому Power Query используя российские региональные настройки, счел, что значение 12.13.2009 является ошибкой, и не загрузил данные в Excel. Позже мы покажем, как решить эту проблему.

В будущем вы можете обновить данные одним из следующих методов:

* Щелкните правой кнопкой мыши Таблицу на листе Excel, выберите *Обновить*
* Щелкните правой кнопкой мыши на имени запроса на панели *Запросы и подключения*, выберите *Обновить*
* Перейдите в Excel на вкладку *Данные*, кликните кнопку *Обновить все*

Каждый раз при выполнении одной из этих команд Excel запускает Power Query, чтобы открыть исходный файл, выполнить записанные шаги и поместить данные в таблицу Excel. Пока вы сохраняете новый файл поверх того, который вы использовали для создания запроса, и пока структура файла не меняется, обновление будет работать автоматически.

Если вы закрыли панель *Запросы и подключения*, ее можно снова отобразить пройдя по меню *Данные* –> *Запросы и подключения.*

Созданный Excel-файл можно сохранить в форматах .XLSX, .XLSB или .XLSM.

**Подключение к табличным данным в Excel.** Из открытой книги Excel можно импортировать данные, хранящиеся на листе, в именованном диапазоне, или в динамическом именованном диапазоне. В первом случае данные могут храниться не в *Таблице*, а в табличном формате (рис. 11). Данные содержатся в файле с примерами Excel Data.xlsx.

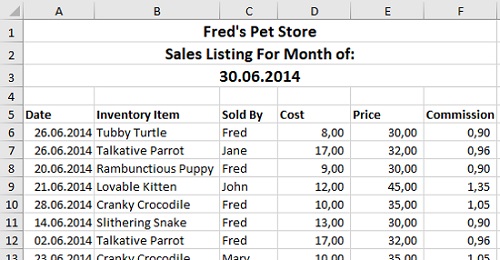


Рис. 11. Импортируемые данные находятся в табличном формате

Чтобы отправить эти данные в Power Query кликните на любую ячейку в диапазоне (1 на рис. 12), перейдите на вкладку *Данные* (2) и нажмите кнопку *Из таблицы/диапазона* (3).

Изображение выглядит как снимок экрана

Автоматически созданное описание

Рис. 12. Поместить данные из таблицы в запрос Power Query

Excel откроет окно *Создание таблицы* (рис. 13), и предложит подтвердить 1) границы таблицы и 2) включает ли набор данных заголовки. После внесения изменений (если таковые требуются) и подтверждения, Excel превратит выбранный диапазон в *Таблицу* и откроет окно редактора Power Query. Если на листе данные уже содержались в *Таблице,* Excel сразу запустит редактор Power Query.

Изображение выглядит как снимок экрана

Автоматически созданное описание

Рис. 13. Окно *Создание таблицы*

Обратите внимание (рис. 14):

* Заголовки были импортированы из *Таблицы*, поэтому нет шага повышения заголовков.
* В качестве имени используется имя вновь созданной *Таблицы*.

Изображение выглядит как снимок экрана

Автоматически созданное описание

Рис. 14. Небольшие отличия в первых шагах работы Power Query

При нажатии кнопки *Закрыть и загрузить* Excel создаст таблицу на новом листе, используя имя запроса. Однако это имя конфликтует с именем существующей таблицы в Excel. Тогда к имени новой таблицы будет добавлено подчеркивание и число (рис. 15). Чтобы управлять процессом наименования, прежде чем нажать кнопку *Закрыть и загрузить* измените имя таблицы в окне редактора Power Query с *Таблица1* на *FromTable*. И лишь затем нажмите кнопку *Закрыть и загрузить.*

Изображение выглядит как снимок экрана

Автоматически созданное описание

Рис. 15. Данные из запроса загружаются на новый лист

Поскольку новая таблица ничем не отличается от ранее имевшейся в Excel, то такое действие не имеет самостоятельного смысла. Мы показали эту процедуру просто для иллюстрации того, как подключиться и загрузить данные из таблицы Excel.

**Подключение к именованным диапазонам.** Если вы не хотите, чтобы стиль таблицы применялся к вашим данным, можете подключиться к *диапазону* Excel. В качестве примера используйте лист *NamedRange* приложенного файла Excel. Данные в нем идентичны данным в предыдущем примере (см. рис. 11). Чтобы подключиться к именованному диапазону сначала определите именованный диапазон, а затем создайте новый запрос.

Итак, выберите ячейки A5:F42 (рис. 16), перейдите в поле *Имя*, введите Date, нажмите Enter.

Изображение выглядит как снимок экрана

Автоматически созданное описание

Рис. 16. Определение именованного диапазона

После определения имени, диапазон можно выбрать с помощью стрелки раскрывающей список имен (рис. 17). Независимо от того, где вы находитесь в книге, вы перейдете на этот лист, и данные в названном диапазоне будут выбраны.

Изображение выглядит как снимок экрана

Автоматически созданное описание

Рис. 17. Доступ к данным через имя диапазона

Убедитесь, что выбран весь именованный диапазон и его имя отображается в поле *Имя*. Создайте новый запрос, пройдя по меню *Данные* –> *Из таблицы/диапазона*. Power Query не создаст *Таблицу* из исходных данных, а будет ссылается на именованный диапазон. Интерфейс Power Query выглядит так же, как при импорте файлов с разделителями, а не как при подключении к таблице Excel. Одной из особенностей *Таблиц* Excel является то, что они имеют строку заголовка. Для именованного диапазона заголовок необязателен. Power Query подключается к источнику данных и выполняет его анализ, чтобы выяснить, как обрабатывать данные. Подобно текстовым файлам с разделителями, он определяет строку, которая выглядит как заголовки, повышает ее и применяет типы данных к столбцам.

Имя запроса по умолчанию – это имя именованного диапазона. Опять же, хороший стиль –изменить имя, иначе Excel сам добавит символы, чтобы избежать дублирования имен.

**Динамические именованные диапазоны.** *Таблицы* Excel автоматически расширяются по вертикали и горизонтали по мере добавления новых данных. Это замечательная особенность *Таблиц*. Но, с другой стороны, *Таблицы* автоматически форматируются. Можно создать динамический именованный диапазон, который расширяется по мере роста данных.

На листе Dynamic Excel-файл вы найдете еще одну копию исходных данных. Настройте динамический именованный диапазон, который будет расширяться при добавлении новых записей:

* Перейдите на вкладку *Формулы*, нажмите кнопку *Диспетчер имен*, и в открывшемся окне нажмите *Создать*.
* Введите имя DynamicRange (рис. 18)
* Введите формулу: =Dynamic!$A$5:ИНДЕКС(Dynamic!$F:$F;ПОИСКПОЗ(99^99;Dynamic!$A:$A))[[3]](#footnote-3)
* Нажмите Ok

Изображение выглядит как снимок экрана

Автоматически созданное описание

Рис. 18. Создание динамического диапазона

Теперь вы можете ссылаться на этот именованный диапазон в Power Query, но поскольку это динамический диапазон, вы не можете выбрать его в поле *Имя* в Excel. А раз вы не можете выбрать его, как вы можете поместить его в Power Query? Трюк в том, чтобы создать пустой запрос, а затем указать Power Query, к какому диапазону вы хотите подключиться. Пройдите по меню *Данные* –> *Получить данные* –> *Из других источников* –> *Пустой запрос* (рис. 19). Откроется окно редактора Power Query. Если строка формул не отображается, перейдите на вкладку *Просмотр* и установите флажок *Строка формул.*

Изображение выглядит как снимок экрана, текст

Автоматически созданное описание

Рис. 19. Создание пустого запроса

В строке формул введите: =Excel.CurrentWorkbook(). Нажмите Enter. Отобразится таблица со списком всех объектов этой книги Excel, к которым можно подключиться:

Изображение выглядит как снимок экрана

Автоматически созданное описание

Рис. 20. Список всех объектов, которые Power Query видит в текущей книге Excel

Кликните на поле Table в строке DynamicRange. Power Query загрузит динамический диапазон:

Изображение выглядит как снимок экрана

Автоматически созданное описание

Рис. 21. Динамический диапазон в редакторе Power Query

Если посмотреть на перечень примененных шагов, то можно увидеть, что Power Query

* подключился к источнику данных – книге Excel
* перешел к таблице DynamicRange
* проанализировав содержимое продвинул заголовки столбцов
* преобразовал типы данных

Вам осталось присвоить запросу имя *FromDynamicRange*, кликнуть на кнопку *Закрыть и загрузить*, чтобы поместить данные из запроса на новый лист Excel.

**Общие замечания.** Если это возможно, предпочитайте решения на основе *Таблиц* Excel. Они проще в настройке и последующем обслуживании. К динамическим диапазонам обращайтесь если только вы точно понимаете, помечу не подходят *Таблицы*.

1. Язык назвали М, в частности, потому, что буквы с A по L уже были заняты)) А название книги отсылает к английскому букварю, в котором букве М соответствует [изображение](https://seaurchinstudio.com/products/monkey) обезьянки – M is for Monkey

    [↑](#footnote-ref-1)
2. Перевод и скриншоты выполнены в Excel 2019 на 64-разрядной версии 2.61. – *Прим. Багузина* [↑](#footnote-ref-2)
3. Указанная формула поддерживает добавление новых строк, но не новых столбцов. Еще больше возможностей доступно с использованием функции [СМЕЩ](http://baguzin.ru/wp/?p=7710). [↑](#footnote-ref-3)