**Глава 3. Простое объединение таблиц в Power Query**

Это продолжение перевода книги Кен Пульс и Мигель Эскобар. Язык М для Power Query. Главы не являются независимыми, поэтому рекомендую читать сначала.

[Предыдущая глава](http://baguzin.ru/wp/?p=20889) [Содержание](http://baguzin.ru/wp/?p=20852#content) Следующая глава

Перед профессионалами Excel часто встают задачи объединения данных из нескольких однотипных таблиц.[[1]](#footnote-1) Power Query может делать это автоматически.

В папке примеров есть три CSV-файла: Jan 2008.csv, Feb 2008.csv и Mar 2008.csv. Начните с импорта первого файла:

* Создайте новую книгу Excel
* Создайте запрос *Данные* –> *Из текстового/CSV-файла*
* Выберите файл Jan 2008.csv



Рис. 3.1. Импортированный CSV-файл Jan 2008.csv содержит одну ошибку

Power Query импортирует файл и автоматически выполнит следующие действия:

* Продвинет первую строку в заголовки.
* Задаст типы данных.

Поскольку исходный файл Jan 2008.csv содержит данные в стандарте США, следует удалить шаг *Измененный тип*, и повторно назначить типы данных, используя для столбцов *TranDate* и *Sum of Amount* локальные установки США (а не РФ, действующие по умолчанию; подробнее см. [предыдущую главу](http://baguzin.ru/wp/?p=20889)). Переименуйте столбцы *TranDate* –> *Date* и *Sum of Amount* –> *Amount*. Нажмите кнопку *Закрыть и загрузить*. Данные будут импортированы на лист Excel в виде, как на рис. 1.

Данные всё еще содержат одну ошибку – общие итоги. Вернитесь в редактор Power Query. Выделите столбец *Date*, кликните *Удалить строки* –> *Удалить ошибки*. Нажмите *Закрыть и загрузить*. Строка с итогами будет удалена.

Повторите операции для импорта Feb 2008.csv и Mar 2008.csv. Когда вы закончите, у вас будет три таблицы в книге Excel, каждая на своем листе. Чтобы объединить таблицы создайте новый запрос. Пройдите по меню *Получить данные* –> *Объединить запросы* –> *Добавить*:



Рис. 3.2. Меню объединения запросов

Откроется диалоговое окно *Добавление* (рис. 3.3). Доступ к окну *Добавление* можно получить и из редактора Power Query. Для этого в редакторе перейдите на вкладку *Главная* и пройдите по меню *Добавить в запросы* –> *Добавить запросы в новый*. (рис. 3.4).



Рис. 3.3. Окно *Добавление*



Рис. 3.4. Доступ к окну *Добавление* из редактора Power Query

Диалоговое окно *Добавление* объединяет запросы Power Query, а не таблицы Excel. Упорядочьте запросы в правом окне, чтобы данные располагались последовательно. Нажмите Ok. Power Query создаст новый запрос *Append1*, который включает один шаг:



Рис. 3.5. Новый объединенный запрос *Append1*

У вас может возникнуть соблазн прокрутить запрос вниз, чтобы увидеть, все ли ваши записи вошли в него. К сожалению, это займет много времени, так как бегунок работает не так как вы привыкли в Excel. При перемещении вниз новые строки будут подгружаться довольно медленно. Причина в том, что Power Query может использоваться для обработки больших наборов данных. Представьте, что вы подключаетесь к набору данных, из 5 миллионов строк, но хотите вытащить записи только для отдела №150. Power Query осуществляет как бы «предварительный просмотр», который должен дать достаточно информации для определения ключевой структуры данных. Вы выполните преобразования в данных предварительного просмотра и создаёте шаблон. Во время загрузки всех строк Power Query обрабатывает этот шаблон, извлекая только необходимые записи. Это намного эффективнее, чем загрузка всех данных в книгу и последующая обработка каждой строки и столбца.

Но если вы не видить все данные, как вы проверите, что объединенный запрос корректен? Переименуйте запрос *Append1* –> *Transactions*. Кликните *Закрыть и загрузить*.



Рис. 3.6. Новый запрос суммирует все строки трех запросов

Вы также можете создать сводную таблицу, и убедиться, что Excel корректно объединил запросы:



Рис. 3.7. Сводная таблица на основе данных из запроса *Transactions*

### Объединение запросов с разными заголовками

Ниже показана ситуация, когда пользователь забыл переименовать столбец *TranDate* в запросе Mar 2008. При объединении запросов Jan 2008 и Mar 2008 получится:



Рис. 3.8. Столбец *TranDate*, полный нулевых значений в январе, и столбец *Date*, полный нулевых значений в марте

Чтобы исправить это, откройте запрос Mar 2008, переименуйте столбец *TranDate* –> *Date*. Сохраните запрос Mar 2008. Откройте запрос *Transactions*. Как только вы откроете запрос, вы увидите, что он уже исправлен – столбец *TranDate* отсутствует. Чтобы поправить таблицу *Transactions* на листе Excel, просто обновите ее.

1. На самом деле, Power Query поддерживает два типа объединений:



В английском варианте, это Merge Queries и Append Queries. Первая опция позволяет объединять таблицы, исключая строки-дубли и проводя иные интеллектуальные операции с данными. Вторая опция просто добавляет каждый последующий набор в конец существующего. Пиктограммы довольно неплохо иллюстрируют это. Настоящая заметка посвящена второй опции. [↑](#footnote-ref-1)