**Преобразование сведения в столбец в Power Query**

Это фрагмент книги [Гил Равив. Power Query в Excel и Power BI: сбор, объединение и преобразование данных](https://baguzin.ru/wp/?p=24451).

[Предыдущий раздел](https://baguzin.ru/wp/?p=24620) [К содержанию](https://baguzin.ru/wp/?p=24451#content)  Следующий раздел

В предыдущих разделах вы познакомились с различными сценариями отмены свертывания столбцов, узнали, каким образом данное преобразование помогает трансформировать обобщенные (сводные) таблицы в таблицы фактов. Но иногда исходная таблица структурирована наоборот, слишком «плоско». Будто кто-то злоупотребил отменой свертывания столбцов и сформировал исходную таблицу с несколькими строками, по идее принадлежащими одному и тому же объекту. Здесь показано, как применить преобразование, обратное к отмене свертывания столбцов – сведение в столбец.

Загрузите файл C07E04.xlsx. Откройте новую книгу Excel. Пройдите *Данные* –> *Получить данные* –> *Из файла* –> *Из книги*. Выберите файл C07E04.xlsx и кликните Импорт. В окне *Навигатор* выберите *Revenues*, а затем щелкните *Преобразовать данные*. В окне редактора Power Query изучите таблицу, содержащую транзакции компании AdventureWorks. Она показана в верхней части рисунка:



Рис. 1. Исходная таблица и искаженная таблица после отмены свертывания столбцов

Последние четыре столбца представляют четыре числовых атрибута для каждой транзакции продаж: столбец SubTotal представляет сумму продаж до вычета налогов и расходов на фрахт, TaxAmt — сумму налога с продаж, Freight — стоимость доставки, а TotalDue — общую сумму (TotalDue = SubTotal + TaxAmt + Freight). Таблица продаж в верхней части рисунка сформирована хорошо. Каждая транзакция указана в отдельной строке.

Выберите четыре последних столбца и пройдите *Преобразование* –> *Отменить свертывание столбцов*. Результат показан в нижней части рис. 1. Таблица с отменой сворачивания столбцов сложнее для анализа, поскольку контекст строки теперь содержит только один из четырех числовых атрибутов.

Допустим, что эта таблица и служит исходной, и нужно реверсировать отмену свертывания столбцов. Выберите столбец *Атрибут* и пройдите *Преобразование* –> *Столбец сведения*. Настройте окно *Столбец сведения*:



Рис. 2. Окно *Столбец сведения*

Щелкните Ok. Восстановлена прежняя структура таблицы.

#### Применение преобразования сведения в столбец к таблицам с многострочными записями

Преобразование отмены свертывания столбцов применяется к числовым столбцам. Столбцы привязки помогают поддерживать контекст каждой транзакции в процессе операции сведения в столбец. Ниже показано, что произойдет, если преобразование отмены свертывания столбцов было применено ко всей таблице, и как разрешить это затруднение.

Откройте новую книгу Excel. Пройдите *Данные* –> *Получить данные* –> *Из файла* –> *Из книги*. Выберите файл C07E04.xlsx и кликните Импорт. В окне *Навигатор* выберите *Revenues*, а затем щелкните *Преобразовать данные*. Выберите первый столбец и нажмите Ctrl+A (англ.) для выбора всех столбцов. Пройдите *Преобразование* –> *Отменить свертывание столбцов*.



Рис. 3. Отмена свертывания для всех столбцов

Применение преобразования сведения в столбец на этом этапе не приведет к цели, поскольку утеряны столбцы привязки. Сохраните запрос. Мы к нему вернемся. Пройдите *Главная* –> *Закрыть и загрузить*.

#### Как работать с парами атрибут/значение

Прежде чем научиться работать с парами атрибут/значение, обговорим важную роль формата атрибут/значение. Этот формат часто встречается, он особенно характерен для файлов журналов и неструктурированных данных, которые не отформатированы как JSON или XML.

Почему этот формат столь распространен? Представьте, что нужно импортировать данные из системы, где большой набор сущностей может включать сотни различных атрибутов, настраиваемых пользователями системы. Немногие из этих атрибутов являются обязательными, остальные — необязательны.

Экспорт подобного набора данных в текстовые файлы в обычном табличном формате неэффективен. В результате формируется разреженная таблица с сотнями столбцов и пустыми ячейками, поскольку большинство атрибутов необязательные или отсутствуют. Для того чтобы не экспортировать разреженные таблицы, которые раздувают файл десятками и даже сотнями символов-разделителей для пустых значений, разработчики соответствующей системы могут реализовать экспорт в виде последовательности многострочных записей пар атрибут/значение. Таким образом, можно найти этот формат довольно распространенным, особенно если нужно импортировать журналы или экспортированные текстовые файлы.

#### Работа с фиксированным числом атрибутов

Начнем с простого сценария, когда можно предположить, что каждая многострочная запись занимает фиксированное число строк. Затем применим общий подход, когда каждая запись всегда начинается с одного и того же атрибута.

Таблица на рис. 3 содержит группы из семи строк пар атрибут/значение. Для сведения столбцов нужно восстановить контекст каждой группы и найти способ применить для каждой группы уникальные номера перед выполнением операции сведения в столбец.

На панели *Запросы* щелкните правой кнопки мыши на запросе *Revenues* и выберите команду *Ссылка*. Переименуйте новый запрос в *Revenues - Fixed Number of Attributes*. Пройдите *Добавить* *столбец* –> *Столбец индекса*. Новый столбец содержит порядковый номер, начиная с нуля для первой строки. На следующем шаге применим целочисленное деление индекса на 7. Целочисленное деление — это деление, при котором отбрасывается остаток. Этот шаг позволит преобразовать текущий индекс из отдельного индекса строки в групповой индекс, поскольку каждая группа состоит из семи строк.

Выберите столбец *Индекс*. Пройдите *Преобразование* –> *Столбец "Количество"* –> *Стандартный* –> *Целочисленное деление*. В окне *Целочисленное деление* введите 7 щелкните Ok. Столбец индекса имеет нулевое значение для первой группы из семи строк, 1 – для второй группы из семи строк, 2 – для третьей группы строк и т.д.

Не добавляйте новый столбец с целочисленным делением. Примените целочисленное деление к имеющемуся столбцу *Индекс*.

Выберите столбец *Атрибут*, пройдите *Преобразование* –> *Столбец сведения*. Настройте окно также, как на рис. 2. Щелкните Ok. Удалите столбец *Индекс*. Теперь в таблице успешно сведены столбцы. Сохраните отчет для последующего использования. Просмотреть готовое решение можно в файле C07E05 - Solution.xlsx.

#### Обработка нефиксированного числа атрибутов

Иногда приходится работать с многострочными записями, которые включают необязательные атрибуты. В подобных ситуациях нельзя предположить, что имеется фиксированное число строк для каждой записи. Более общая методика основано на ином предположении: каждая запись начинаться с одного и того же имени атрибута. В то время как остальные атрибуты могут быть необязательными, первый атрибут является обязательным. Для преобразования многострочных записей в таблицу фактов воспользуйтесь следующим алгоритмом:



Рис. 4. Обобщенный алгоритм сведения в столбцы

Откройте файл, сохраненный в предыдущем разделе. Откройте редактор Power Query. На панели *Запросы* правой кнопки мыши щелкните на запросе *Revenues* и выберите *Ссылка*. Переименуйте новый запрос в *Revenues - Fixed First Attribute*. Этот новый запрос будет использоваться для сведения в столбцы данных с учетом предположения о том, что каждая новая многострочная запись начинается с атрибута *Country*.

Пройдите *Добавить* –> *Столбец индекса*. Необходимо идентифицировать индексы строк с помощью атрибута *Country* в столбце *Атрибут*. Пройдите *Добавить столбец* –> *Условный столбец*. Настройте окно *Добавить условный столбец*:



Рис. 5. Настройка окна *Добавить условный столбец*

Щелкните Ok.

Выберите столбец *Row ID* и пройдите *Преобразование* –> *Заполнить* –> *Вниз*. Удалите столбец *Индекс*. Идентификатор строк (Row ID) теперь содержит число 0 для первой группы, число 7 для второй группы, число 14 для третьей группы и т.д. Если применять эту методику для многострочных записей с необязательными атрибутами, числа в идентификаторе строк (Row ID) не будут кратны 7 или любому другому фиксированному числу. Неважно, какие значения представлены в идентификаторе строки. Важно лишь то, что для каждой записи или группы строк он состоит из уникальных значений.

Теперь, когда восстановлен контекст каждой записи, можно перейти к завершению этапа сведения в столбец. Выберите столбец *Атрибут*, пройдите *Преобразование* –> *Столбец сведения*. Настройте окно *Столбец сведения* как на рис. 2. Щелкните Ok. Удалите столбец *Row ID*. Можно просмотреть файл решения C07E05 - Solution.xlsx.