**Глава 15. Транспонирование и сложные случаи обратного разбора сводных 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 последовательно.

[Предыдущая глава](http://baguzin.ru/wp/?p=21164) [Содержание](http://baguzin.ru/wp/?p=20852#content) Следующая глава

Обратный разбор сводной таблицы в Power Query был рассмотрен в [главе 6](http://baguzin.ru/wp/?p=21013). Однако, могут встречаться и более сложные ситуации, например, сводные таблицы с подкатегориями. Также иногда нужно преобразовать данные, расположенные горизонтально, в обычные вертикальные таблицы. Обе эти функции, к счастью, встроены в пользовательский интерфейс Power Query. Допустим, у вас есть красивая сводная таблица, которую нужно разобрать для последующего анализа:



Рис. 15.1. Сводная таблица с подкатегориями требует разбора

Что делает эту проблему более сложной, чем стандартная процедура разбора сводной таблицы? Это – дополнительный уровень агрегирования: не только месяц, но и тип измерения (план, факт, расхождение). Когда вы разбираете сводную таблицу с одним уровнем агрегирования (например, по месяцам), вы используете имена столбцов в качестве данных, которые поместите в каждую соответствующую строку. А какие имена столбцов использовать здесь? Если вы используете вторую строку в качестве данных, вы потеряете важную информацию о месяце.

### Загрузка данных

Откройте файл UnPivot Sub Categories.xlsx:



Рис. 15.2. Исходный набор данных для разбора

Поскольку вы не хотите блокировать заголовки столбцов при импорте *Таблицы* в Power Query, вы можете использовать именованный диапазон. В Excel выберите диапазон A1:H14. Перейдите на вкладку *Формулы* –> *Задать имя*. В окне *Создание имени* введите *Statement* в поле *Имя*, нажмите Ok.



Рис. 15.3. Присвоение имени диапазону

Операцию присвоения имени можно выполнить проще, правда, не так явно. Выделите диапазон A1:H14, и введите *Statement* в область *Имя* в строке формул*.*



Рис. 15.4. Присвоение имени диапазону в строке формул

Создайте новый запрос, пройдя по меню *Данные* –> *Из таблицы/диапазона*. В редакторе Power Query удалите все шаги, кроме первого – *Источник*.



Рис. 15.5. Набор данных изобилует нулевыми значениями

Обратите внимание. Если в таблице Excel ячейка была пустой, значение в таблице Power Query равно *null*. Значение строки 1 столбца *Column4* также равно *null*.

### Заполнение по вертикали

Давайте начнем со столбца *Column1*. Несмотря на то, что столбец содержит много нулевых значений, вы не можете удалить его. Столбец *Column1* содержит информацию о классификации доходов и расходов. Проблема в том, что строки классификации не содержат всей важной информации. Ваша цель дополнить строки информацией о месяце и типе измерения (план, факт, расхождение). Кликните правой кнопкой мыши на столбце *Column1* –> *Заполнить* –> *Вниз*. Теперь у вас есть правильная классификация в каждой строке:



Рис. 15.6. Классификация счетов заполнена для нужных строк

Теперь вам нужно обработать заголовки столбцов. У них похожая проблема: апрель находится в *Column3* и отсутствует в *Column4* и *Column5*. К сожалению, в Power Query нет команды *Заполнить вправо*. Чтобы обойти это ограничение сначала примените…

### …Транспонирование данных

Перейдите на вкладку к *Преобразование* –> *Транспонировать*:



Рис. 15.7. Результат транспонирования таблицы

Кликните правой кнопкой мыши на столбце *Column1* –> *Заполнить* –> *Вниз.*

Поскольку таблица включает категории и подкатегории вы не можете применить разбор сводной таблицы на основе имен столбцов. Временно объедините *Column1* и *Column2*. Выберите *Column1,* удерживая нажатой клавишу Ctrl, выберите *Column2* –> Объединить столбцы. Установите в качестве пользовательского разделителя вертикальную черту – |. (Символ | привлекателен тем, что он редко встречается в обычном наборе данных. Это означает, что последующее разделение на основе этого символа вряд ли нарушит данные.) Теперь у вас есть один столбец меток, которые можно использовать для заголовков:



Рис. 15.8. Объединение категорий и подкатегорий в один столбец

Переверните данные обратно и повысьте заголовки. Перейдите на вкладку к *Преобразование* –> *Транспонировать*. *Главная* –> *Использовать первую строку в качестве заголовков*. Переименуйте столбец | –> *Class*, а столбец |\_1 –> *Account*.



Рис. 15.9. Цель уже близка!

Дальнейшие процедуры стандартны для сценария разбора сводной таблицы. Отфильтруйте столбец *Account* и удалите нулевые значения. Выберите столбцы *Class* и *Account*. Щелкните правой кнопкой мыши один из выбранных заголовков –> Отменить свертывание других столбцов:



Рис. 15.10. Почти правильная таблица

Осталось разделить столбец *Атрибут* и переименовать столбцы. Щелкните правой кнопкой мыши столбец *Атрибут* –> *Разделить столбец* –> *По разделителю* –> *Пользовательский* –> По символу |. Щелкните правой кнопкой мыши столбец *Атрибут.1* –> *Переименовать* –> *Месяц*. Щелкните правой кнопкой мыши столбец *Атрибут.2* –> *Переименовать* –> *Мера*. Щелкните правой кнопкой мыши столбец *Значение* –> *Переименовать* –> *Сумма*. Запрос обработан и готов к загрузке:



Рис. 15.11. Сводная таблица с подкатегориями полностью разобрана

Эта техника не ограничивается только двумя уровнями заголовков. Если их больше, объедините все заголовки и подзаголовки в одну текстовую строку, и далее действуйте по описанному выше сценарию.

Если данные полностью разобраны, в дальнейшем легко построить, как исходную сводну таблицу, так и иные срезы:



Рис. 15.12. Исходные данные, восстановленные с помощью сводной таблицы

### Транспонирование вертикальных наборов данных в строки

Другой сложный сценарий преобразования связан с данными, которые располагаются группами вертикально, а разделяются пустыми строками:



Рис. 15.13. Задача – преобразовать вертикальные наборы в строки

Power Query не имеет стандартной команды для такой обработки, так что потребуется творческий подход и несколько трюков.

Откройте файл Transpose Stacked Tables.xlsx. Выберите область А1:А17. Пройдите по меню *Данные* –> *Из таблицы/диапазона*. Добавьте индексный столбец. Перейдите на вкладку *Добавление столбца* –> *Столбец индекса* –> от 0.



Рис. 15.14. Каждой строке таблицы присвоен индекс

Данные отдельных транзакций упорядочены: первая строка – дата, вторая – поставщик, третья – ID транзакции, четвертая и пятая – нули. Шестая строка соответствует первой строке второй транзакции. Добавим еще один столбец, возвращающий остаток от деления на шесть.

Выберите столбец *Индекс* –> Перейдите на вкладку *Добавление столбца* –> *Из числа* –> *Стандартный* –> *Остаток от деления* –> 6. Теперь у вас есть список номеров строк по каждой транзакции:



Рис. 15.15. Данные одного типа пронумерованы одинаковыми цифрами

### Поворот данных

Если у вас есть список повторяющихся строк, вы можете их свернуть. Сначала это покажется странным, но вы увидите, насколько это красиво. Выберите столбец *Остаток от деления* –> *Преобразование* –> *Столбец сведения*. Когда вы даете команду *Столбец сведения*, значения в выбранном столбце будут использоваться в качестве заголовков новых столбцов (1 на рис. 15.16). Затем Power Query спрашивает, какой столбец вы хотите использовать для извлечения данных. В нашем примере – столбец *Transactions* (2). Обычно для данных выполняется суммирования значений, но в нашем примере мы не заинтересованы в этом. Нам нужны исходные значения, поэтому кликните на *Расширенные параметры* (3). Установите функцию агрегатного значения = *Не агрегировать* (4). Нажмите Ok.



Рис. 15.16. Настройка параметров команды *Столбец сведения*

Таблица меняет форму и пустым ячейкам присваивается значение null



Рис. 15.17. Таблица сведения, полная нулевых значений

### Заполнение пробелов

Следующим шагом является заполнение недостающих данных по строкам. Выберите столбец 1 –> удерживайте нажатой клавишу Shift –> выберите столбец 5. Перейти на вкладку *Преобразование* –> *Заполнить* –> *Вверх*.



Рис. 15.18. Данные, заполненные по строкам

Нажмите фильтр столбца 0 –> снимите null. Выберите столбцы от 0 до 3 –> щелкните правой кнопкой мыши *Удалить другие столбцы*. Переименуйте столбцы 0 –> *Дата* и отформатируйте его как дату, 1 –> *Поставщик*, 2 –> *ID транзакции* и отформатировать его как целое число, 3 –> *Сумма* и отформатировать его как десятичное число. Переименуйте запрос *Transactions*.



Рис. 15.19. Преобразованный набор данных готов к загрузке на лист Excel