**Глава 17. Множественные таблицы данных Power Pivot**

Это продолжение перевода книги Роб Колли. Формулы DAX для Power Pivot. Главы не являются независимыми, поэтому рекомендую начать сначала.

[Предыдущая глава](http://baguzin.ru/wp/?p=20080) [Содержание](http://baguzin.ru/wp/?p=19780#content) Следующая глава

До сих пор мы работали только с одной таблицей данных – Продажи [Sales]. Таблицы данных обычно представляют бизнес-процессы, и в этом случае таблица Sales описывает трансакции продаж. Очевидно, что реальный бизнес включает более одного бизнес-процесса, и трансакции собираются в несколько таблиц.

Добавим еще один бизнес-процесс – Сервисные звонки [Service Calls]. Колл-центр принимает звонки от клиентов, и фиксирует: ID клиента, дату и продолжительность звонка, ID продукта, в связи с которым был звонок, а также субъективную оценку того, как была решена проблема. Эти данные доступны нам в конце каждого дня в формате CSV (рис. 17.1; чтобы получить картинку, как на рис. 17.1 я импортировал данные из CSV-файла в Excel; для этого создал новый файл Excel и прошел по меню Данные –> Получить и преобразовать данные –> Из текстового/CSV=файла).



Рис. 17.1. Записи звонков

Загрузим эти данные в нашу модель. Откройте файл ch17\_MultipleDataTables.xlsx, пройдите по меню Power Pivot –> Управление и в окне Power Pivot пройдите по меню Главная –> Получить внешние данные –> Из других источников –> Текстовый файл.

В нашей модели появилась таблица Service Calls (рис. 17.2).



Рис. 17.2. Таблица Service Calls, импортированная в модель данных

Свяжите таблице в верхней части модели с Service Calls (рис. 17.3).



Рис. 17.3. Связи, созданные для таблицы Service Calls

Обратите внимание, мы не создали связи между таблицами Sales и Service Calls. Вы никогда не связываете таблицы данных друг с другом! Это связано с тем, что в каждом столбце ключа таблицы данных есть дубликаты. Например, мы принимаем несколько вызовов в один день и продаем несколько продуктов в один день, поэтому таблицы Sales и Service Calls содержат несколько строк для данной даты. Всегда связывайте таблицы данных только с набором таблиц поиска, но не друг с другом (рис. 17.4).



Рис. 17.4. (а) Никогда не связывайте таблицы данных друг с другом; (б) таблицы данных связывайте только с набором таблиц поиска

Определим некоторые меры:

[Calls] = COUNTROWS(ServiceCalls)

[Total Call Minutes] = SUM(ServiceCalls[Minutes])

[Avg Call Length] = DIVIDE([Total Call Minutes];[Calls])

Директор колл-центра сообщил, что минута работы колл-центра стоит $1,75. Подсчитаем стоимость обслуживания звонков:

[Cost of Calls] = [Total Call Minutes] \* 1,75

С новыми мерами наш анализ сервисных звонков становится наглядным (рис. 17. 5).



Рис. 17.5. Сводная таблица сервисных звонков

### Объединение мер из разных таблиц данных в одной сводной

В традиционном Excel, анализ нескольких таблиц данных сводится к отображению двух сводных таблиц бок о бок (рис. 17.6).



Рис. 17.6. Традиционный анализ двух таблиц данных в Excel

Такое представление сталкивается с трудностями, когда строки относятся к разным продуктам/датам/покупателям (рис. 17.7).



Рис. 17.7. Некоторые элементы отображаются в одной сводной, но не в другой

Мы часто говорим, что для пользователей Excel изучение Power Pivot – это скорее отучение, чем изучение чего-то нового. И одна из важных заповедей – не объединяйте данные в одну таблицу (что является стандартным подходом при построении обычных сводных таблиц). Несколько таблиц данных обычно представляют различные бизнес-процессы и очень непохожи друг на друга.

В модели данных меры из разных таблиц могут отражаться в одной сводной! Определенные один раз, меры используются в любом месте. Перестройка сводной не потребует переделки мер. Последние автоматически отреагируют на новые срезы. В нашем примере вы можете разместить меры [Service Calls] и меры [Sales] в одной сводной (рис. 17.8).



Рис. 17.8. Две меры, принадлежащие разным таблицам данных в одной сводной

Просто измените фильтры, и данные автоматически подстроятся (рис. 17.9).



Рис. 17.9. Меры подстраиваются под изменение фильтров

### Гибридные меры

Иногда мы хотим использовать меру, основанную на двух наборах данных. Например, мы можем уточнить прибыль, уменьшив ее на стоимость обслуживания звонков (рис. 17.10):

[Net Profit] = [Profit] - [Cost of Calls]



Рис. 17.10. Чистая прибыль

Теперь, когда у нас есть новая мера – [Net Profit], вам не нужны бок о бок [Profit] и [Cost of Calls]. Вы можете убрать их и сосредоточиться только на чистой прибыли. Вы можете использовать в качестве инструмента визуализации Power View, чтобы собрать быстрый отчет о прибыли (рис. 17.11).



Рис. 17.11. Чистая прибыль на дашборде, выполненном в Power View (см. файл. ch17\_MultipleDataTables\_Excel2013\_Dashboard; к сожалению, у меня файл работал некорректно)

Обратите внимание: Power Pivot является сердцем всех продуктов Power BI, а связи и меры – ядро Power Pivot. Используя их, вы находитесь на пути к овладению миром данных. Последующая визуализация на основе Power Pivot становится простой и удобной. Используя меры, вы можете нарезать кубы, используя любой из доступных атрибутов, и быстро построить любой отчет.

### Проблемы нескольких таблиц данных

Если вы создадите сводную на основе таблицы [Sales], а затем добавите из таблицы [Service Calls] меру [Calls], то результат будет неожиданным (рис. 17.12).



Рис. 17.12. Мера [Calls] возвращает неожиданный результат

Мера [Calls] выводит во всех ячейках одно и то же число – 7374, что является общей суммой всех звонков. Что случилось с обещанием «определить один раз, использовать везде»? Проблема связана с тем, что мы в область сводной таблицы *Строки* поместили столбец [Order Date] из таблицы [Sales]. Но это значение из таблицы данных [Sales], а не из таблицы поиска [Calendar] (рис. 17.13). Поэтому для меры [Calls] никакие фильтры по дате не были установлены.



Рис. 17.13. Неверное поле в области *Строки* сводной таблицы

Запомните хорошее правило: используйте поля из таблиц поиска (а не таблиц данных) в строках, столбцах, фильтрах и срезах при анализе мер из разных таблиц данных. Это хорошая практика даже в моделях, которые имеют одну таблицу данных. Замените [Order Date] из таблицы [Sales] на [Date] из таблицы [Calendar]:



Рис. 17.14. Мера [Calls] правильно отображаются, когда в троках находится поле [Date] из таблицы [Calendar]

Лучший способ избежать попадания в эту ловушку – скрыть некоторые из столбцов таблицы данных. Для этого войдите в модель данных, перейдите в режим *Диаграмма*, щелкните правой кнопкой мыши на поле и выберите *Скрыть от клиентских средств* (рис. 17.15).



Рис. 17.15. Спрячьте поле [Order Date] из таблицы [Sales], чтобы оно стало невидимым в сводной таблице

Вот несколько рекомендаций:

* В идеале ваши **таблицы данных** должны содержать только числовые столбцы (например, количество и сумма), а также столбцы-ключи или ID (Date, ProductKey, CustomerKey), которые используются для подключения к справочникам.
* Все остальные столбцы должны «жить» в связанных **таблицах поиска**.
* Визуально проверьте таблицы данных, нет ли в них текстовых столбцов? Как правило, следует перенести такие столбцы в **таблицу поиска**.
* Таким образом, в идеальной **таблице данных** вы можете скрыть все столбцы! Ключевые столбцы (которые формируют связи) можно скрыть, так как у вас уже есть их «близнецы» в таблицах поиска. Числовые столбцы также можно скрыть, так как вы определили явные меры для них.

### Таблица данных, подключенная к некоторым, но не ко всем таблицам поиска

Если вы захотите посмотреть на звонки в разрезе территорий продаж, то не увидите связи (рис. 17.16, а). И это не удивительно, так как в модели данных таблицы [Sales Territory] и [Service Calls] не связаны (рис. 17.16, б). Если наша бизнес-модель изменится, и мы начнем управлять сервисными центрами на территориальной основе, мы добавим столбец [Territory] в таблицу [ServiceCalls]. Но сейчас у нас работает единый колл-центр, который обслуживает всех наших клиентов.



Рис. 17.16. Звонки не связаны с территориями продаж: а) ни в сводной таблице, б) ни в модели данных