**Глава 23. Прояснение понятий «контекст фильтра» и «контекст строки»**

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=20468) [Содержание](http://baguzin.ru/wp/?p=19780#content) Следующая глава

В [главе 3](http://baguzin.ru/wp/?p=19780#ch3) мы описали кривую обучения DAX (по аналогии с кривой обучения Excel). И мы шли в гору по этому пути на протяжении двадцати уроков. В настоящей главе мы берем небольшой антракт, чтобы погрузиться в то, какие основы заложены в механизмы DAX. Вы сможете писать формулы DAX и без этого отступления, но, набив шишки, мы сочли полезным поделиться с вами нашим знанием.

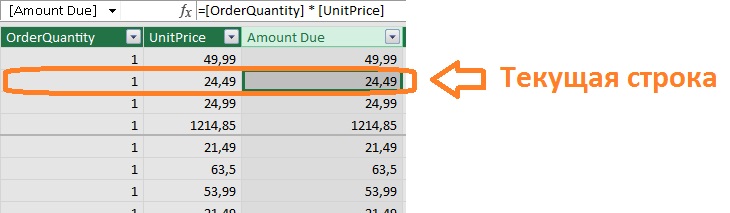


Рис. 23.1. Для вычисляемых столбцов контекст строки определяется автоматически

Итак, для начала мы можем дать два простых определения:

*Контекст строки = текущая строка*

*Контекст фильтра = набор фильтров в сводной таблице*

Контекст строки является ключевым компонентом вычисляемых столбцов. В Power Pivot, в отличие от Excel, нет ссылок в стиле A1 (строки не нумеруются). Контекст строки (текущая строка) автоматически определяется для вычисляемых столбцов. Вот почему вы можете определить вычисляемый столбец *Причитающаяся сумма* (Amount Due) формулой:

[OderQuantity] \* [Unit Price]

Когда мы ссылаемся на [Unit Price] или [OderQuantity], нет никакой двусмысленности, мы ссылаемся на эти значения в текущей строке (рис. 23.1). Вы не можете ссылаться на следующую строку или предыдущую строку (не имея ссылок в стиле A1; правда, вы можете использовать небольшой [трюк](https://powerpivotpro.com/2015/03/how-to-compare-the-current-row-to-the-previous-row-in-power-pivot/)). Но у вас есть текущая строка, и это позволяет вам писать формулы вычисляемых столбцов. Формулы, обрабатывающие по одной строке за раз.

### В мерах нет контекста строк

Попытка написать меру, используя ту же формулу, что и вычисляемый столбец…

[Total Amount Due] = [OrderQuantity] \* [Unit Price]

…приводит к ошибке. Мы уже знакомились с этим, когда формулировали [правило](http://baguzin.ru/wp/?p=19833#naked): если вы ссылаетесь на столбец в формуле меры, он всегда должен быть внутри какой-то функции. «Голые» столбцы в мерах недопустимы.

Меры не содержат контекста строки (понятие текущей строки не имеет смысла). Механизм DAX предполагает, что несколько строк останутся после того, как будут применены все фильтры. И как, например, столбец [Unit Price] может быть сокращен до одного значения для этих нескольких строк!?

Вот почему меры всегда требуют использовать какую-то функцию. Это нужно для агрегирования, чтобы свернуть несколько строк в одно значение. В тоже время в вычисляемом столбце ссылка на «голый» столбец является законной, потому что у нас всегда есть контекст строки, и нет необходимости «сворачивать» несколько значений в одно.

### Ключевой компонент мер – контекст фильтра

Рассмотрим таблицу Excel (рис. 23.2).

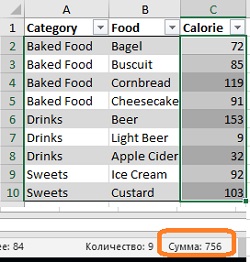


Рис. 23.2. Сумма калорий составляет 756

Если теперь взять фильтр по столбцу Category = "Drinks", сумма калорий уменьшится:

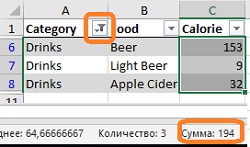


Рис. 23.3. Сумма калорий по напиткам

Можно сказать:

* сумма калорий для напитков составляет 194, или…
* сумма калорий для текущего контекста фильтра Calorie[Category] = "Drinks" составляет 194

Первый вариант ближе к человеческому языку, второй – языку DAX. Контекст фильтра – не что иное, как набор строк, которые были отфильтрованы. Контекст фильтра берется из координат, поступающих из сводной таблицы (областей Строки, Столбцы, Фильтры). Далее контекст фильтра перетекает по связям (под гору) из таблиц поиска в таблицы данных. И, наконец, контекст фильтра может быть изменен с помощью DAX-функции CALCULATE. Но фактически это похоже на простой фильтр, который мы выбрали в таблице Excel на рис. 23.3.

Еще раз, контекст фильтра – это набор координат, поступающих из сводной таблицы для текущей меры в конкретной ячейке сводной. Эти координаты, в свою очередь, приводят к отфильтрованному набору строк исходной таблицы, на которых считается арифметика меры.

### В вычисляемых столбцах нет контекста фильтра

Для иллюстрации давайте перейдем в модель данных на закладку (таблицу) Sales и создадим вычисляемый столбец с использованием формулы =SUM(Sales[SalesAmt])

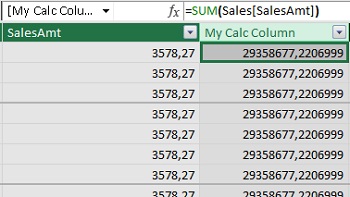


Рис. 23.4. Мы получили одно и то же значение для всех строк вычисляемого столбца при использовании функции агрегации, такой как SUM

Рассмотри результат подробнее:

1. Функции агрегации, такие как SUM, всегда игнорируют контекст строки и вместо этого работают с контекстом фильтра.
2. Но в вычисляемом столбце нет контекста фильтра. Однако это не означает, что мы не получаем никаких данных. На самом деле совсем наоборот...
3. Фразу «Нет контекста фильтра» можно заменить на «отсутствие фильтров». Таким образом, функция агрегации (SUM) берет в работу весь столбец [SalesAmt], т.е., складываются значения во всех строках столбца.

Контекст строки позволяет вычислить значение в одном столбце по формуле на основе значений в других столбцах той же строки. Контекст фильтра – это набор координат, поступающих из сводной таблицы, который отбирает нескольких строк. Контекст строки – основа вычисляемых столбцов. Контекст фильтра – основа формул меры. Эти понятия являются почти совершенными противоположностями.



Рис. 23.5. Контекст строки и контекст фильтра

### Связи и контекст фильтра

Что мы имеем в виду, когда говорим: контекст фильтра течет вниз по связям таблиц?

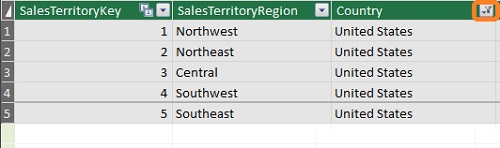


Рис. 23.6. Таблица продаж Sales фильтруется по стране [Country] = "United States"

Фильтр установлен в таблице SalesTerritory, которая связана с таблицей продаж Sales. Фильтр в таблице SalesTerritory «перетекает» в таблицу Sales по связи, установленной между таблицами (по ключевым столбцам SalesTerritoryKey, соединяющие две таблицы).

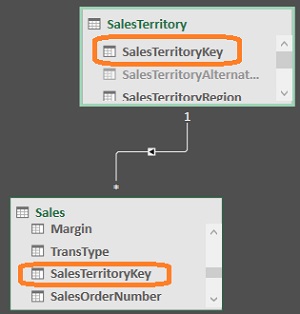


Рис. 23.7. SalesTerritory и Sales связаны

Если в таблице SalesTerritory, в столбце [Country] установить фильтр "United States", в столбце [SalesTerritoryKey] установится фильтр { 1, 2, 3, 4, 5}. Благодаря связи и в таблице Sales столбец Sales[SalesTerritoryKey] отфильтрует значения { 1, 2, 3, 4, 5}.

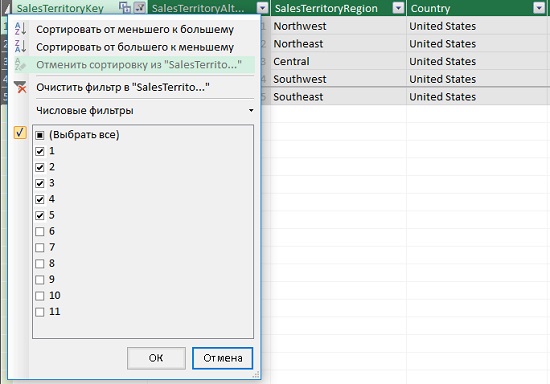


Рис. 23.8. Контекст фильтра в таблице SalesTerritory

**Связи и контекст строки.** Нет никакого взаимодействия между связями и контекстом строки (за исключением случаев, когда в формулах вычисляемых столбцов вы используете функции, основанные на связях, например, RELATED и RELATEDTABLE). Контекст строки не влияет на связи и не зависит от них.

### Исключения и новые определения

Исключения всё же есть. Х-функции (итераторы) создают контекст строки во время вычисления меры. Напомним, что [Х-функции](http://baguzin.ru/wp/?p=20080) (например, SUMX, RANKX и т.д.) и функция [FILTER](http://baguzin.ru/wp/?p=19926) проходят через <таблицу> (аргумент функции) по одной строке за раз. Таким образом, в этих функциях у вас есть контекст строки и функции агрегации вокруг ссылок на столбцы не являются обязательными. Например, внутри функции SUMX можно ссылаться на столбцы как на значения без функций агрегирования:

[Total Amount Due] = SUMX(Sales; Sales[Unit Price]\*Sales[Quantity])

Аналогично, внутри функции FILTER вы можете выполнять любую математику, используя голые ссылки на столбцы. Это очень удобно, когда, например, вы хотите отфильтровать строки, в которых значение в столбце A в два раза больше, чем в столбце B. При этом Х-функции не создают контекст строки для всей формулы. Контекст строки существует только внутри самой Х-функции.

### Функция CALCULATE создает контекст фильтра в вычисляемых столбцах

[Ранее](http://baguzin.ru/wp/?p=19833#step2) мы указывали на это в шаге 2 золотых правил DAX. Функция CALCULATE не только использует контекст фильтра, там, где он задан координатами сводной таблицы, но и создает (переопределяет) свой собственный контекст фильтра.

Функция CALCULATE преобразует текущий контекст строки в контекст фильтра. Вернемся к предыдущему примеру, где мы использовали функцию SUM() внутри вычисляемого столбца \*см. рис. 23.4). Создадим еще один вычисляемый столбец, вписав в него формулу:

=CALCULATE(SUM(Sales[SalesAmt]))

Если бы вы написали *меру*, как указано выше, используя CALCULATE без аргументов фильтра, это было бы странно. Но для *вычисляемого столбца*, это может быть интересно. CALCULATE принимает текущий контекст строки и «притворяется», что это на самом деле контекст фильтра:

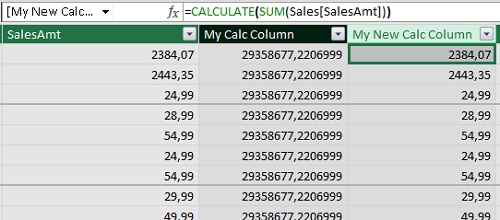


Рис. 23.9. Вычисляемый столбец дублирует столбец [SalesAmt]

Выше мы сказали, что в общем случае связи и контекст строки не взаимодействуют. Однако, перейдя к исключениям, мы можем использовать CALCULATE для отслеживания связей в вычисляемых столбцах. Давайте создадим вычисляемый столбец в таблице поиска Products (рис. 23.10). Вычисляемый столбец таблицы поиска дает сумму всех соответствующих (номеру продукта) строк из таблицы данных Sales.

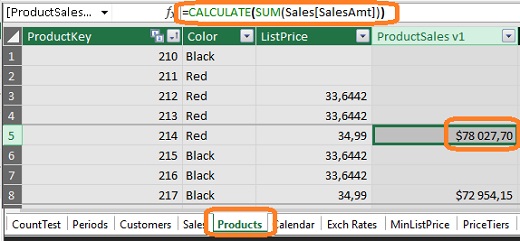


Рис. 23.10. Вычисляемый столбец таблицы поиска «видит» связь с таблицей данных

Как правило, такие столбцы подытогов в таблицах поиска не нужны, так как мы динамически вычисляем подытоги в сводных таблицах с помощью мер. Веская причина добавить такой вычисляемый столбец – подготовиться к классификации продуктов по корзинам, например, «топовые продукты (по продажам)», «аутсайдеры продаж» и т.п. Другими словами, нас обычно не интересует вычисляемые столбцы подытогов сами по себе, но мы используем их в качестве входных данных для другого вычисления. Того, которое разбивает строки на группы. А группы мы впоследствии используем для строк, столбцов, фильтров или срезов в сводных таблицах. Также группы могут использоваться в качестве фильтра в функции CALCULATE.

### Использование мер в контексте строки: настоящая круговерть

Ранее мы определили меру [Total Sales] = SUM(Sales[SalesAmt]). Что произойдет, если мы создадим вычисляемый столбец в таблице Products с формулой =[Total Sales]:

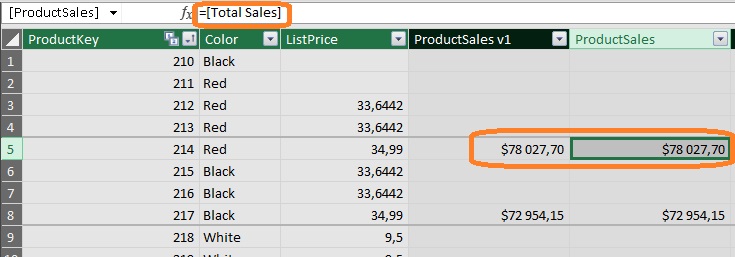


Рис. 23.11. Вычисляемый с помощью меры столбец дает удивительно правильные результаты

Так что же случилось с контекстом строки, который вроде бы не взаимодействует со связями? Всякий раз, когда вы ссылаетесь на меру, двигатель DAX делает вид, что вы «завернули» меру в функцию CALCULATE. Таким образом, контекст строки в вычисляемом столбце или внутри Х-функции, если вы ссылаетесь на меру, она будет действовать так, как если бы она была внутри CALCULATE. Т.е. контекст строки будет переведен в контекст фильтра. На рис. 23.11 формулу =[Total Sales] следует воспринимать, как = CALCULATE([Total Sales]).

Рассмотрим более сложный кейс. У нас есть мера для подсчета транзакций – Transactions = COUNTROWS(Sales). Определим новую меру для подсчета транзакций только самых дорогих товаров:

Transaction for Highest Price = CALCULATE(

COUNTROWS(Sales);

FILTER(

Products;

Products[ListPrice] = MAX(Products[ListPrice])

)

)

Из таблицы Products видно, что это товары с ценой по столбцу ListPrice = $3578,27:

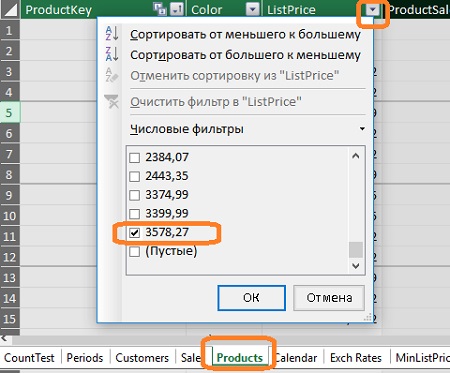


Рис. 23.12. Цена ListPrice самого дорогого товара

Вот как работает эта мера:

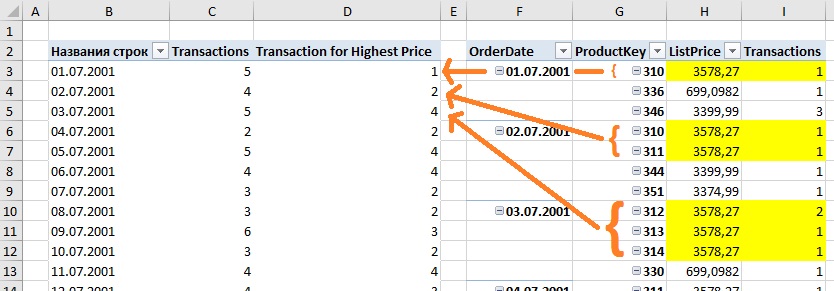


Рис. 23.13. Мера Transaction for Highest Price возвращает желаемый результат

В качестве лучшей практики мы часто рекомендуем:

* Строить меры шаг за шагом
* Повторно использовать меры везде, где это возможно

Следуя духу этих рекомендаций, мы могли бы определить промежуточную меру и повторно использовать ее:

Highest Price = MAX(Products[ListPrice])

Transaction for Highest Price BROKEN = CALCULATE(

COUNTROWS(Sales);

FILTER(

Products;

Products[ListPrice] = **[Highest Price]**

)

)

У-о-п-с… Эта мера, не работает, хотя она представляется очень похожей на предыдущую:

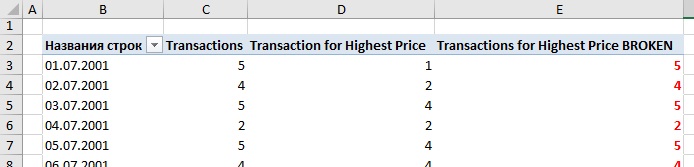


Рис. 23.14. Новая мера не работает

На наш взгляд, это наименее разумная вещь во всем языке DAX – единственное место, где мера возвращает другой результат, чем базовая формула той же меры. Давайте проанализируем работу меры строка за строкой (чтобы было удобно ссылаться на формулу, пронумеруем строки):

1. Transaction for Highest Price BROKEN = CALCULATE(
2. COUNTROWS(Sales);
3. FILTER(
4. Products;
5. Products[ListPrice] = **[Highest Price]**
6. )
7. )

Предварительно заметим, что функция FILTER в строке 3 является итератором, т.е. обрабатывает таблицу строка за строкой. Другими словами, FILTER создает контекст строки. В строке 5 мы использовали меру [Highest Price] в том месте, где у нас есть контекст строки.

А теперь давайте проследим работу меры для ячейки Е3 сводной таблицы (см. рис. 23.14). Для этой ячейки Sales[OrderDate] = 01.07.2001.

FILTER (строка 3) выполняет итерацию строка за строкой по таблице Products (4). При этом контекст фильтра из сводной таблицы Sales[OrderDate] = 01.07.2001 никак не фильтрует таблицу Products.

Поскольку COUNTROWS(Sales) будет возвращать значение только по тем ProductKey, которые были проданы 01.07.2001, мы сосредоточимся на ProductKey = { 336; 310; 346 }.

Для продукта ProductKey = 336 цена Products[ListPrice] = $699,09. Функция FILTER оценивает условие Products[ListPrice] = [Highest Price]. [Highest Price] – мера, т.е., работает так, будто она находится внутри CALCULATE. Таким образом, контекст строки Product[ProductKey] = 336 переходит в контекст фильтра. Что вернет мера [Highest Price] в этом контексте фильтра?

Filter Context: Product[ProductKey] = 336

[Highest Price] = MAX(Products[ListPrice])

Как вы уже догадались, $699,09. Т.е., возвращается не абсолютно максимальное значение [ListPrice] (во всей таблице), а максимальное значение в контексте фильтра.

Итак, строка 5 оценивает условие фильтра для строки ProductKey = 336 Products[ListPrice] =[Highest Price]:

$699,09 = $699,09

Без проблем: условие выполнено!

И так для всех продуктов: ProductKey = { 336; 310; 346 }. Поскольку обе стороны сравнения (5) всегда будут оценивать одно и то же значение.

Т.е., функция FILTER (3) в данном случае не предоставляет никаких дополнительных фильтров. Единственный фильтр поступает из исходной сводной по дате. Таким образом, наша сломанная (BROKEN) мера работает так же, как первоначальная Transactions = COUNTROWS(Sales), и всегда возвращает то же самое число.

### Почему же мера Transaction for Highest Price работает, а Transaction for Highest Price BROKEN – нет?

Приведем формулу меры еще раз:

Transaction for Highest Price = CALCULATE(

COUNTROWS(Sales);

FILTER(

Products;

Products[ListPrice] = MAX(Products[ListPrice])

)

)

Вы помните, какой результат мы получили, когда определили вычисляемый столбец в таблице Products как = SUM(Sales[SalesAmt])? Мы получили одно и то же число в каждой строке (см. рис. 23.4) – сумму всех строк таблицы Sales, так как не было контекста фильтра. Фрагмент формулы MAX(Products[ListPrice]) ведет себя аналогично. Поскольку нет контекста фильтра, наложенного на таблицу Products, он всегда возвращает максимальную цену для всех продуктов, которая составляет $3578,27.

Чтобы глубже понять, почему возникают такие различия, вспомним, как контекст строки и фильтра применяются к аргументам функции FILTER().

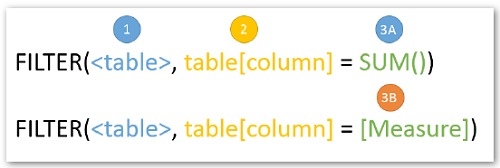


Рис. 23.15. Анатомия функции FILTER(): (1) применяется исходный контекст фильтра из сводной таблицы, (2) применяется контекст строки, генерируемый внутри функции-итератора FILTER, (3) применяется исходный контекст фильтра из сводной таблицы, (4) сначала генерируется контекст строки внутри функции FILTER, затем он переходит в контекст фильтра, и уже он применяется в формуле

Первый аргумент: <таблица>. Может быть именем таблицы (например, Products) или табличным выражением на основе функции, возвращающей таблицу (например, ALL(Products)). К аргументу <таблица> применяется исходный контекст фильтра, поступающий из сводной таблицы (если, конечно, вы не переопределите его с помощью функции ALL()).

Второй аргумент: <условие>. В нем оцениваться одно из следующих условий:

* table[column] – ссылка на столбец использует контекст строки, сгенерированный функцией FILTER (итератором). Обратите внимание, что обычно голая ссылка типа table[column] (без каких-либо функций агрегации вокруг столбца) не допускается, поскольку меры не имеют контекста строки. Но функция FILTER является итератором, она проходит строка за строкой (по таблице из первого аргумента) и генерирует контекст строки на каждом шаге итерации.
* функция SUM() или иная агрегация – всегда ссылается на контекст фильтра и игнорирует контекст строки, даже если контекст строки создан функцией FILTER. Единственный доступный контекст фильтра – тот, что исходит из сводной таблицы.
* CALCULATE(SUM()) – контекст строки, созданный каждой итерацией функции FILTER, будет преобразован в контекст фильтра, и сумма будет обрабатывать только отобранные строки, наследуя связи и т.п.
* [Мера] – будет вести себя так же, как CALCULATE(SUM()) или CALCULATE с любой другой агрегационной функцией, потому что ссылки на меры подразумевают использование функции CALCULATE.

Если от этой главы у вас закружилась голова, не волнуйтесь. Когда что-то таинственное начинает происходить, попробуйте пошаманить:

* Перенести функцию агрегирования внутрь CALCULATE().
* Заменить ссылку на меру ее базовой формулой.

### Ключевые моменты этой главы

Контекст строки – это знание текущей строки

Контекст фильтра – это набор координат фильтра, поступающих из сводной таблицы

Меры не имеют контекста строки, но… мы можем создать контекст строки внутри мер, используя функции-итераторы (Х-функции или FILTER).

Вычисляемые столбцы не имеют контекста фильтра, но… мы можем создать его, используя CALCULATE. CALCULATE трансформирует текущий контекст строки в контекст фильтра, который добавляется к существующему контексту фильтра. Важным следствием является то, что связи и перетекание фильтра, которую они обеспечивают, будут соблюдаться в рамках этого расчета.

Ссылки на меры, используемые в контексте строки, ведут себя так, как если бы они были обернуты в CALCULATE(),трансформируя существующий контекст строки в контекст фильтра.