**Глава 22. Сложные связи модели 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=20444) [Содержание](http://baguzin.ru/wp/?p=19780#content) Следующая глава

В DAX связи очень важны. Но некоторые связи, как бы это сказать, сложнее, чем другие))

### Несколько связей между одними и теми же таблицами

Рассмотрим таблицы Sales и Calendar. Они связаны Calendar[Date] –> Sales[OrderDate]. Файл примера – ch22A\_ComplicatedRelationships.xlsx.

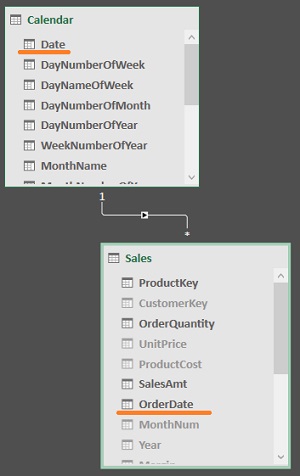


Рис. 22.1. Могут ли эти две таблицы иметь более одной связи?

Дата заказа [OrderDate] – не единственное поле даты в таблице продаж Sales. Есть еще дата доставки. Что делать, если мы иногда хотим проанализировать наши данные о продажах по дате доставки? Как мы с этим справимся? Создадим ли мы еще одну связь между этими двумя таблицами? Это вообще сработает?

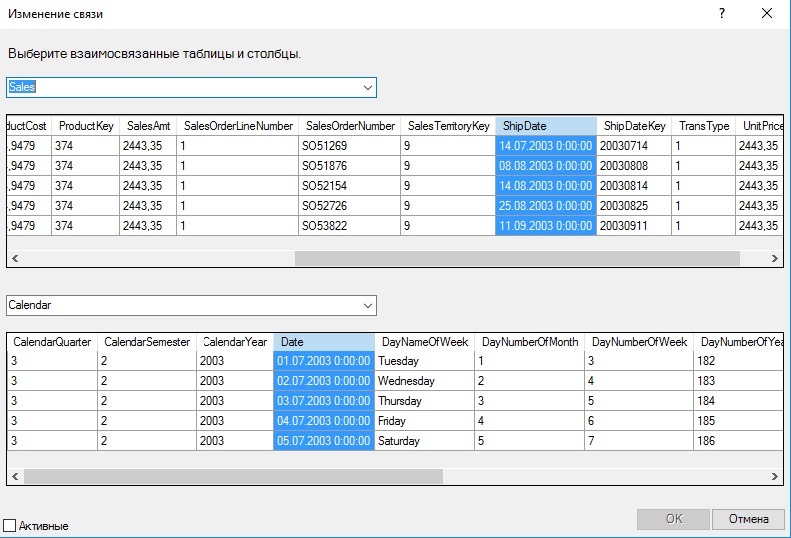


Рис. 22.2. Попытка создать связь Calendar[Date] –> Sales[ShipDate]

Возможно, вы ожидали ошибки при нажатии кнопки Ok в диалоговом окне выше. Но на самом деле это работает. На приведенной ниже диаграмме можно заметить, что вновь созданная связь представлена пунктирной линией, а не сплошной.

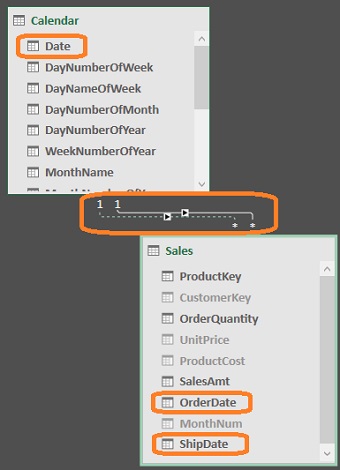


Рис. 22.3. Двойные свидания не всегда проходят гладко))

Легче понять, что происходит, если в окне Power Pivot перейти на вкладку *Конструктор* и кликнуть на копке *Управление связями* (рис. 22.4) Обратите внимание на столбец *Активно* для связей между таблицами продаж Sales и календарем Calendar. В один и тот же момент может быть активна только одна связь между двумя таблицами.

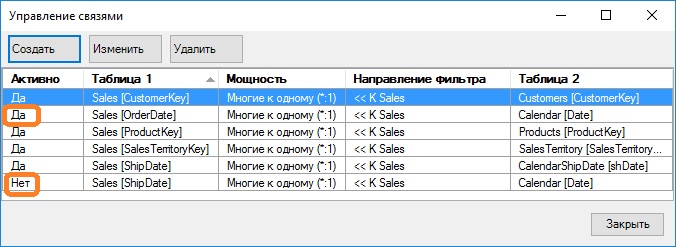


Рис. 22.4. В любой момент времени может быть активна только одна связь между двумя таблицами

Если нам нужно проанализировать данные о продажах по дате доставки, то можем перевернуть активные связи, нажав на кнопку *Изменить* (см. верхнюю часть рис. 22.4). Если у вас есть сводная таблица, показывающая меру [Total Sales], вот как она изменится при изменении активной связи:



Рис. 22.5. Мера [Total Sales] показывает объем продаж на основе даты заказа или даты доставки, в зависимости от того, какая связь активна

Это неудобно, поэтому давайте рассмотрим несколько более элегантных способов сделать то же самое.

### USERELATIONSHIP()

CALCULATE – замечательная функция, и она приходит нам на помощь еще раз. Создадим новую меру:

[Total Sales by Ship Date] =

CALCULATE (

[Total Sales],

USERELATIONSHIP (Sales[ShipDate], Calendar[Date])

)

Теперь мы можем поместить обе меры [Total Sales by Ship Date] и [Total Sales] (по дате заказа) в одну сводную.

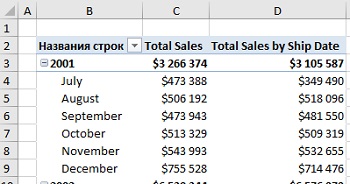


Рис. 22.5. Общий объем продаж по дате заказа и по дате отгрузки в одной сводной

### Связи Многие ко многим

Начнем с неудачного примера. Только что вы создали связь Calendar[Date] –> Sales[ShipDate]. Это связь *Один ко многим*, поскольку в таблице Calendar каждая дата уникальна, а в таблице Sales каждая дата может встречаться более одного раза. И это наиболее распространенный тип связи (и единственный, с которым мы имели дело до сих пор). А теперь давай сделаем что-нибудь... оригинальное. В таблице клиентов Customers у нас есть данные о дне рождении. Попробуем создать связь Customers[BirthDate] –> Sales[ShipDate].

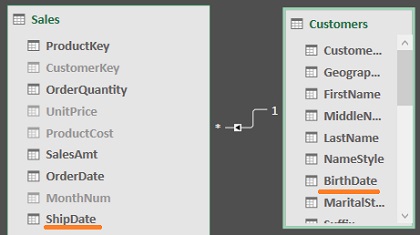


Рис. 22.6. Позволит ли PowerPivot связать Customers[BirthDate] –> Sales[ShipDate]?

Обратите внимание, что оба столбца содержат не уникальные значения:

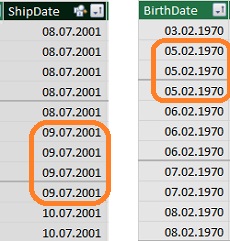


Рис. 22.7. У нас есть много повторяющихся значений для обоих столбцов

Давайте продолжим и попытаемся создать эту связь Customers[BirthDate] –> Sales[ShipDate]:

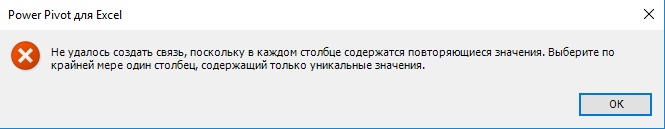


Рис. 22.8. Попытка создать связь приводит к ошибке

Power Pivot не позволит нам создать связь, потому что с каждой стороны повторяются (дублируются) значения. Повторим, что пример нереалистичен. Мы выбрали его, чтобы проиллюстрировать, что мы получим ошибку.

А вот еще один плохой пример. [Ранее](http://baguzin.ru/wp/?p=20299) мы говорили, что не следует связывать две таблицы данных друг с другом. В том примере у нас были таблицы звонков ServiceCalls и продаж Sales. И каждую из них мы связывали с общими для них таблицами поиска, но не друг с другом! Теперь попытаемся создать связь между ними Sales[OrderDate] –> ServiceCalls[CallDate]:

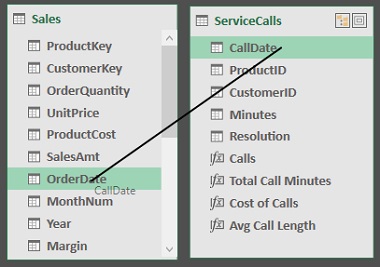


Рис. 22.9. Попытка связать две таблицы данных приводит к аналогичной ошибке

### Реальный мир – источник законных связей *Многие ко многим*

Теперь мы перейдем к описанию ситуации, когда связь *Многие ко многим* легитимна. Многие ко многим, или M2M сводится к тому, как организован ваш бизнес (или реальный мир), и в частности к концепции членства, когда некий объект (продукт, место, человек, …) одновременно принадлежит к двум родительским группам.

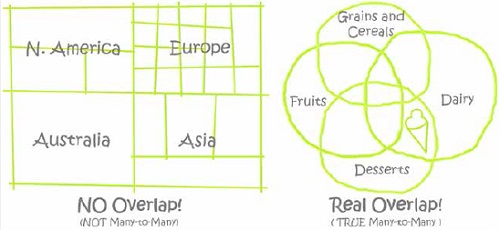


Рис. 22.10. Каждая страна и/или регион относится к одному континенту, но мороженое одновременно может быть и молочным и десертным

Файл с примерами – ch22B\_ComplicatedRelationships\_M2M.xlsx.

Это приводит к таблице поиска (FoodMultiCategory), которая для некоторых продуктом содержит две строки (рис. 22.11). Если вы попытаетесь создать связь между таблицей данных Sales и этой «сломанной» таблицей поиска, она завершится неудачей.

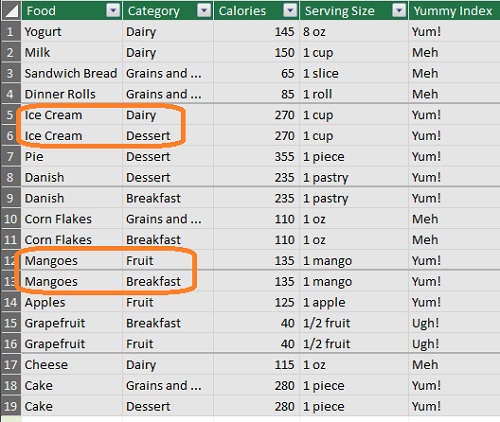


Рис. 22.11. Теперь у нас есть дубликаты в таблице поиска, и это проблема

Первое, что нам нужно сделать, это выделить столбец Category в отдельную таблицу поиска. Таким образом, таблица Food не будет содержать столбец [Category] и появится отдельная таблица Category.

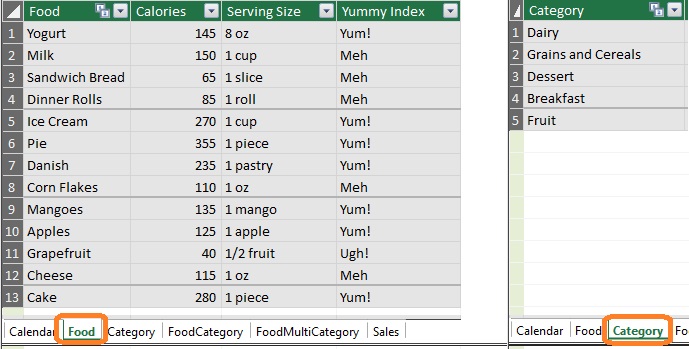


Рис. 22.12. Выделение столбца Category в отдельную таблицу

Теперь можно связать таблицы Sales и Food:

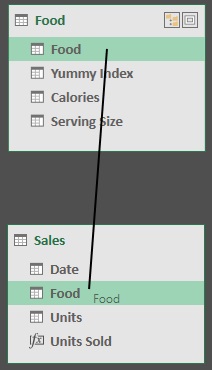


Рис. 22.13. Таблицы Sales и Food можно связать по полю [Food]

Но таблица Category оказалась неприкаянной, и вторая наша задача соединить ее с таблицей Food с помощью «моста» – переходной таблицы.

### Переходная таблица

Связь *Многие ко многим* между двумя объектами можно обрабатывать с помощью переходной таблицы:

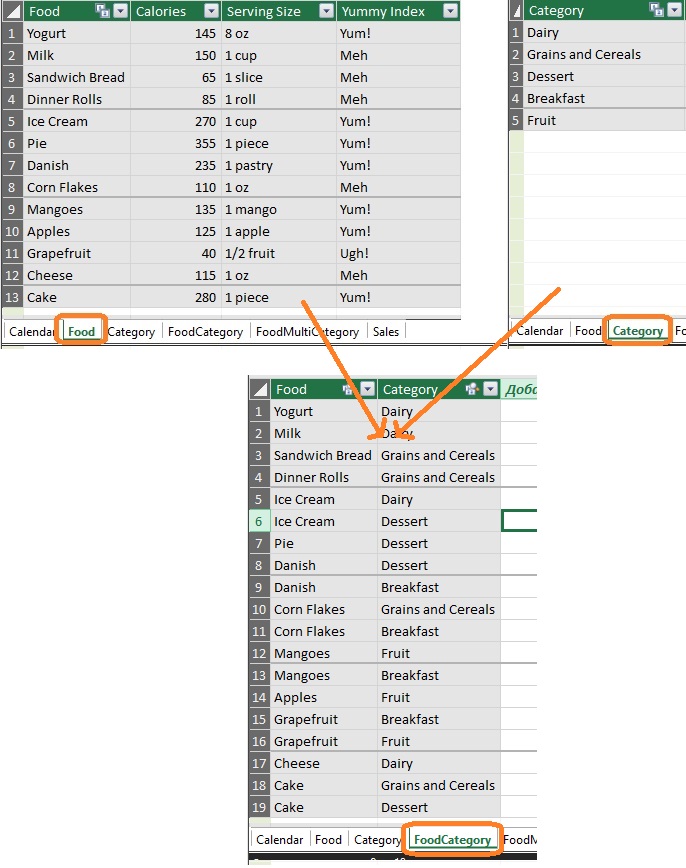


Рис. 22.14. Переходная таблица

Переходная таблица FoodCategory перечисляет для каждого продукта все категории, к которым относится этот продукт – по одной строке на родительскую категорию. Теперь мы можем создавать связи: FoodCategory[Food] –> Food[Food] и FoodCategory[Category] –> Category[Category].

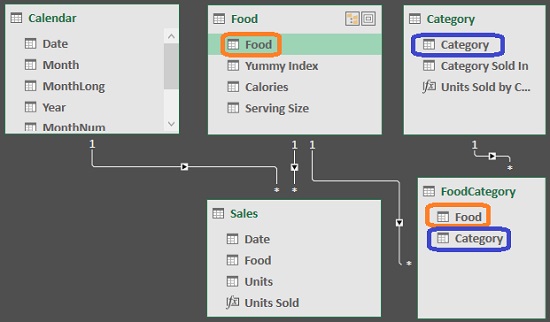


Рис. 22.15. FoodCategory можно связать с Food и Category

Но не всё так просто. Давайте определим базовую меру Units Sold = SUM(Sales[Units]). Сможем ли мы проанализировать продажи по категориям продуктов? Сводная работает не вполне корректно:

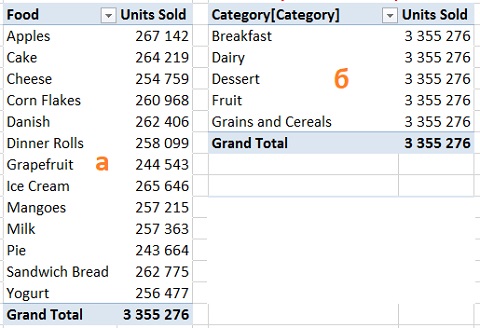


Рис. 22.16. Продажи (а) по продуктам – всё Ok; (б) продажи по категориям – мера не работает

Давайте разберем, что здесь происходит по шагам [золотых правил](http://baguzin.ru/wp/?p=19833) (например, для категории завтрак – Breakfest).

Шаг 1. Определите набор фильтров ячейки сводной таблицы: Category[Category] = "Breakfast"

Шаг 2. Измените набор фильтров, если используете функцию CALCULATE(): не применимо.

Шаг 3. Примените фильтры к таблице с исходными данными: используем фильтр "Breakfast" в таблице Category.

Шаг 4. Примените фильтры к таблицам поиска; фильтры передаются по направлению связей между таблицами; это приводит к тому, что в таблице данных будут отобраны только строки, соответствующие набору всех фильтров. Поскольку Category фильтруется на шаге 3, а таблица Category – это таблица поиска для таблицы FoodCategory, то фильтр течет вниз и применяется, как показано на рисунке:

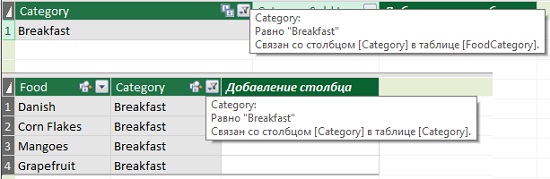


Рис. 22.17. Таблица FoodCategory (внизу) фильтруется установками таблицы поиска Category (вверху)

Шаг 5 и 6. Рассчитайте формулы и верните результат в ячейку: Units Sold = SUM(Sales[Units])

Подожди минутку... таблица Sales никогда не фильтровалась ни на одном из шагов выше! Таким образом, SUM(Sales[Units]) вернет сумму всех строк в таблице Sales – 3 355 276. И одно и то же число повторяется для каждой категории.

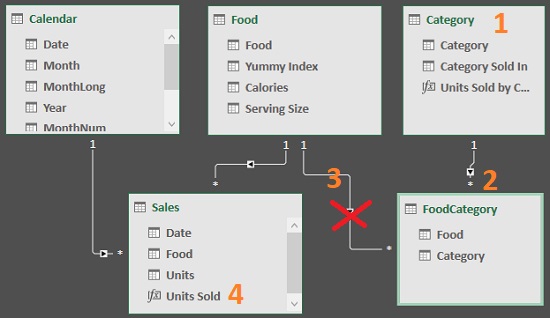


Рис. 22.18. Фильтры текут через связи вниз, но не вверх: 1) фильтры стартуют из таблицы Category; 2) фильтр течет вниз; 3) фильтр не будет течь вверх (по крайней мере, без специальной помощи); 4) мера Units Sold никогда не фильтруется

Оказывается, есть способ заставить фильтры течь вверх. Напишем новую меру:

Units Sold by Category = CALCULATE([Units Sold], FoodCategory)

Использование переходной таблицы в качестве аргумента фильтра в CALCULATE заставляет фильтры течь «вверх» по связи:

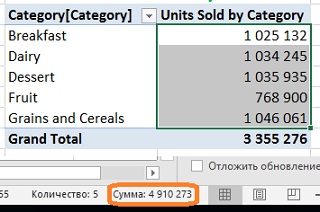


Рис. 22.19. Наша новой мера работает, как ожидалось

Обратите внимание, если вы суммируете пять строки, они не дадут общую сумму. Это легко объяснить, когда мы добавляем Food в сводную:

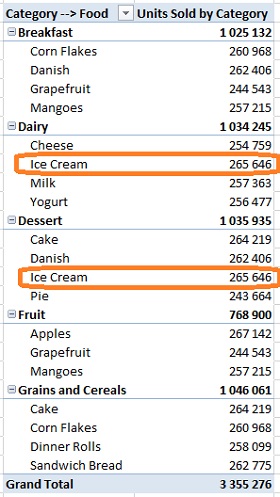


Рис. 22.20. Некоторые продукты (Food) встречаются в разных категориях (Category)

Помните, что мороженое попало в две категории; мы хотели, чтобы оно относилось, как молочным продуктам, так и к десертам. Таким образом, мороженое проявляется в двух местах. То же самое верно для других продуктов, относящихся к нескольким категориям. Следовательно, значения, показанные для любой одной категории, правильные (см. отдельные строки на рис. 20.19), но, если суммировать все пять строк, мы кое-что учтем дважды и итоговая сумма будет неверной. А вот итоги сводной таблицы суммируют «правильно», исключая двойной счет.

Мы рекомендуем вам просто запомнить этот шаблон работы со связями *Многие ко многим*:

М2М Measure = CALCULATE([Measure], BridgeTable)

Если вы хотите больше узнать по этой теме, рекомендую статьи [Джеффри](http://mdxdax.blogspot.com/2011/03/logic-behind-magic-of-dax-cross-table.html) и [Герхарда](http://blog.gbrueckl.at/2012/05/resolving-many-to-many-relationships-leveraging-dax-cross-table-filtering).

### Power BI Desktop

[Power BI Desktop](http://baguzin.ru/wp/?p=20444) имеет функцию, которая устраняет необходимость в приведенном выше подходе, но не необходимость в переходных таблицах. Удалите меру Units Sold by Category. Сохраните копию Excel-файла ch22B\_ComplicatedRelationships\_M2M.xlsx. Перейдите в Power BI Desktop и загрузите модель из Excel-файла (пройдите по меню *Файл* –> *Импортировать* –> *Из Excel*). Перейдите в режим *Модель* (рис. 22.21, цифра 1), кликните правой кнопкой мыши на связи таблиц Food и FoodCategory (2), выберите *Свойства*, и в открывшемся окне *Изменение* *связи* измените направление кросс-фильтрации на *Двунаправленное* (3).

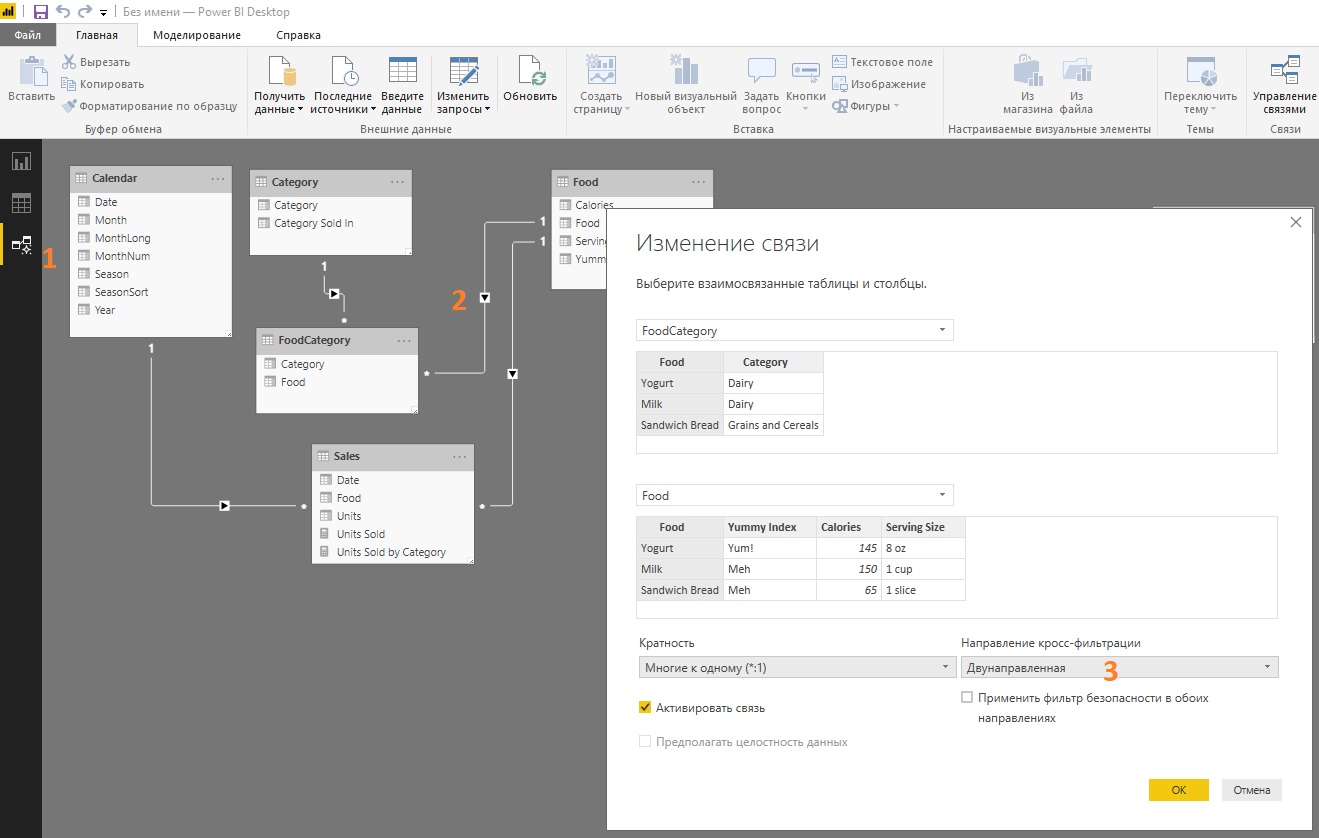


Рис. 22.21. Изменение свойств связи

Мера…

Units Sold = SUM(Sales[Units])

… прекрасно работает, даже когда вы помещаете Category в область строк сводной таблицы. Никаких дополнительных работ не требуется!

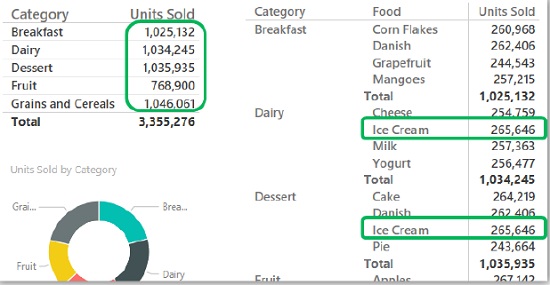


Рис. 22.22. В Power BI Desktop обычные формулы могут работать и для связей *Многие ко многим*