

Глава 9. Объединение таблиц и запросов Power Query

Это продолжение перевода книги Кен Пульс и Мигель Эскобар. Язык M для Power Query. Главы не являются независимыми, поэтому рекомендую читать последовательно.

[Предыдущая глава](#) [Содержание](#) [Следующая глава](#)

Одна из классических проблем, которую решают профессионалам Excel – это объединение двух таблиц данных, и последующее создание сводной таблицы. Как правило, обработка основывается на функциях ВПР() или ИНДЕКС(ПОИСКПОЗ()). Power Query представил еще один относительно простой метод объединения двух таблиц. Предположим, вы хотите объединить две Таблицы, расположенные на листе Excel:

	A	B	C	D	E	F	G	H
5	Inventory Items							
6								
7	SKU Number	Brand	Type	Unit	Pack Quantit	Sales Price	Unit Cos	Margin
8	510007	Budweiser	Lager	Cans	15	29,5	24,4	5,1
9	510010	Canadian	Lager	Cans	6	12	9,95	2,05
10	510014	Canterbury	Ale	Cans	6	11,5	9,5	2
11	510019	Corona Extra	Lager	Bottles	6	13,5	11,25	2,25
12	510021	Corona Grande	Lager	Bottles	1	4,5	3,7	0,8
13	510032	Granville Island	Ale	Bottles	6	13	10,95	2,05
14	510037	Guinness	Stout	Cans	4	13	10,99	2,01
15	510038	Heineken	Lager	Bottles	6	14,5	11,95	2,55
16	510046	Kokanee	Lager	Tall Cans	6	15,5	13,05	2,45
17	510057	Miller	Lager	Bottles	12	24	20,2	3,8
18	510059	OK Springs	Lager	Bottles	6	13	10,99	2,01
19	510065	OK Springs	Ale	Bottles	6	13	10,99	2,01
20								
21	Sales Transactions							
22								
23	Date	SKU Number	Brand	Sales				
24	2014-03-12	510007	Budwei	64				
25	2014-03-12	510010	Canadia	45				
26	2014-03-12	510014	Canterb	62				
27	2014-03-12	510019	Corona	64				
28	2014-03-14	510019	Corona	24				
29	2014-03-12	510021	Corona	38				
30	2014-03-14	510021	Corona	31				
31	2014-03-12	510032	Granvill	24				

Рис. 9.1. Исходные Таблицы на листе Excel

Создание запроса-подключения

Чтобы объединить запросы, они должны существовать. Наличие *Таблицы* в Excel недостаточно; Power Query должен распознавать данные как запрос. Откройте файл Merge.xlsx. Чтобы загрузить таблицу *Inventory* в Power Query, щелкните любую ячейку в ней, пройдите по меню *Данные* → *Из таблицы/диапазона*. Таблица будет импортирована в Power Query и отобразится в окне редактора. На вкладке *Главная* щелкните раскрывающийся список в нижней части кнопки *Заккрыть и загрузить*. Выберите *Заккрыть и загрузить в...* → *Только создать подключение*.

Запрос отображается в области *Запросы и подключения* книги Excel, но Power Query не создал новую таблицу ни на текущем, ни на новом листе Excel:

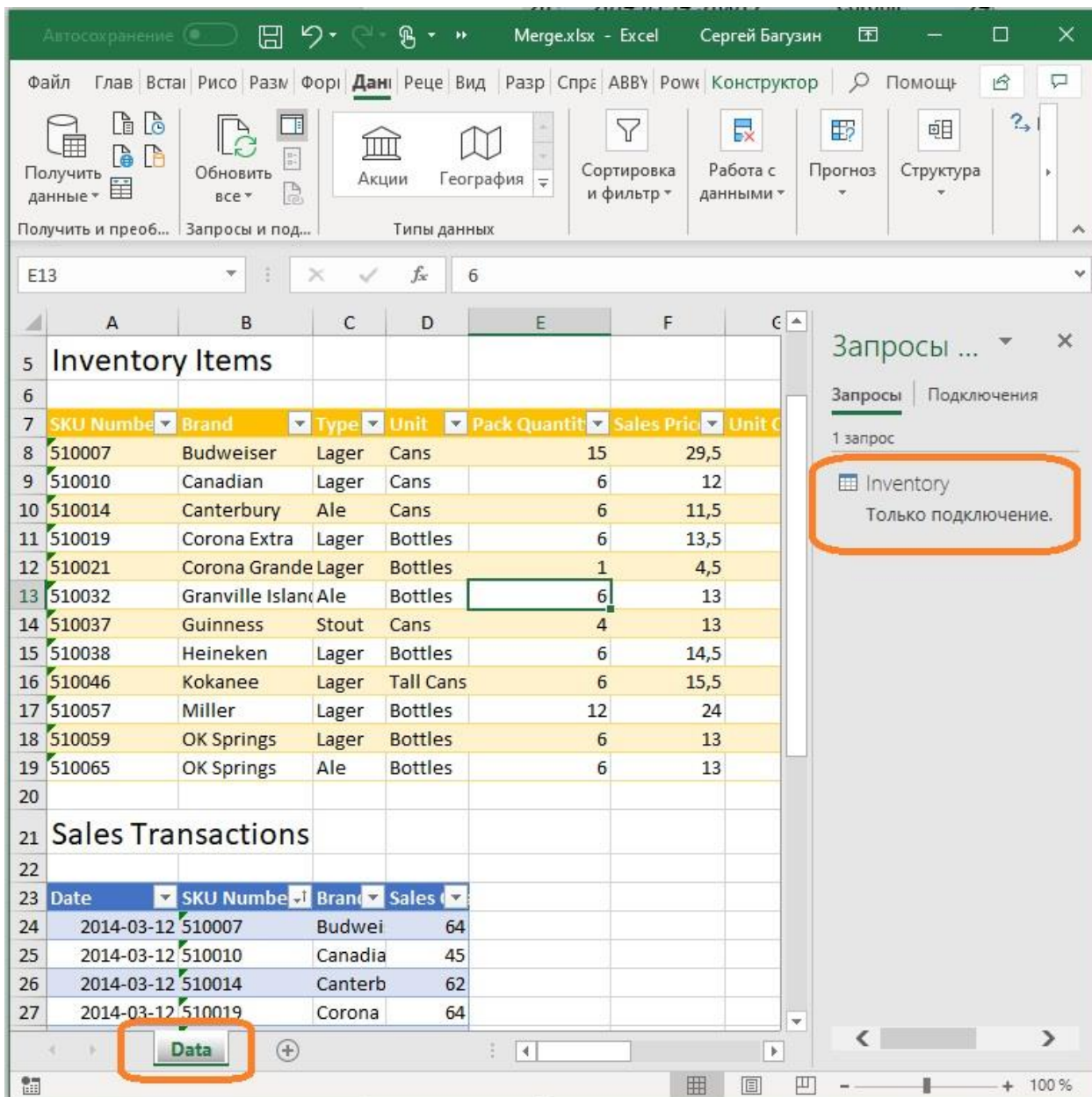


Рис. 9.2. Новый запрос, созданный только для подключения к Таблице

Если вы поспешили, и как обычно, кликнули *Заккрыть и загрузить*, то на новом листе увидите новую *Таблицу*. Чтобы исправить ситуацию, щелкните правой кнопкой мыши запрос на панели *Запросы и подключения*. В редакторе Power Query пройдите по меню *Главная* → стрелочка на кнопке *Заккрыть и загрузить* → *Заккрыть и загрузить в...* → *Только создать подключение*. Power Query предупредит, что вы собираетесь удалить данные. Подтвердите.

И наоборот, если вы создали запрос только для подключения, то в дальнейшем можете через этот интерфейс добавить *Таблицу* на лист Excel.

Теперь импортируйте в Power Query таблицы *Sales*. Аналогичным образом загрузите ее только создав подключение.

Объединение запросов

Данные → *Получить данные* → *Объединить запросы* → *Объединить*. Откроется диалоговое окно *Слияние*. Выберите таблицу *Sales* в верхнем раскрывающемся списке, и таблицу *Inventory* – в нижнем. Однако, кнопка *Ок* по-прежнему не подсвечена:

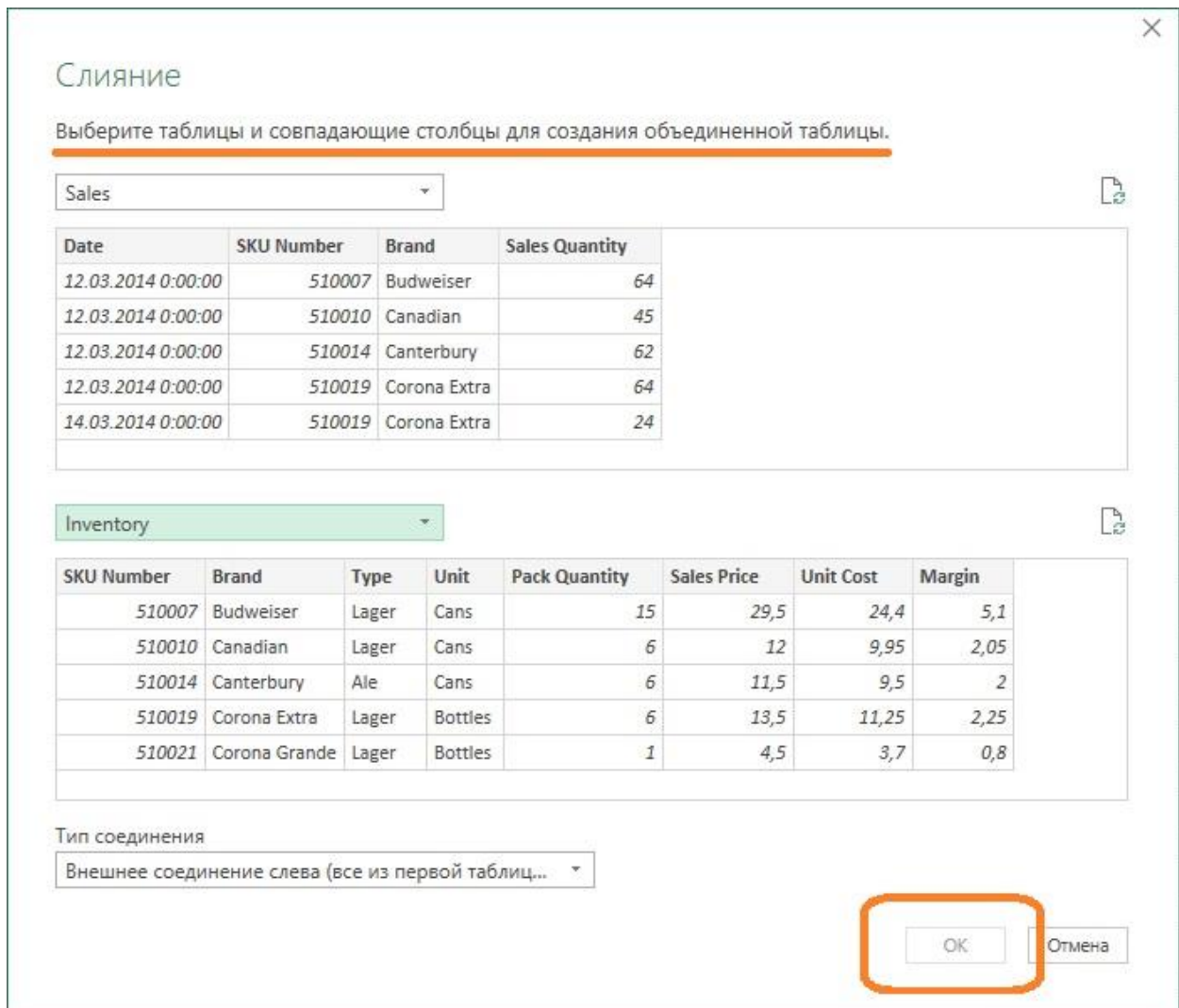


Рис. 9.3. Вы выбрали таблицы, но почему вы не можете продолжить?

Еще раз внимательно прочтите инструкцию, содержащуюся под заголовком окна. Power Query не знает, какие поля вы хотите использовать для выполнения слияния. Вам нужно в каждой таблице выделить столбцы с идентификатором. Причем так, чтобы в одной таблице этот столбец содержал уникальные значения, а в другой таблице значения могут повторяться. Такая связь называется «один ко многим», и ее использование – лучший способ гарантировать, что вы получите результаты, соответствующие вашим ожиданиям.

Power Query также поддерживает соединения один-к-одному и многие-ко-многим.

В нашем примере столбец *SKU Number* содержит уникальные продукты в таблице *Inventory*. Столбец *SKU Number* представлен и в таблице *Sales*. Здесь значения *SKU Number* могут повторяться много раз. Используйте этот столбец для связывания таблиц. Щелкните заголовок *SKU Number* в каждой таблице:

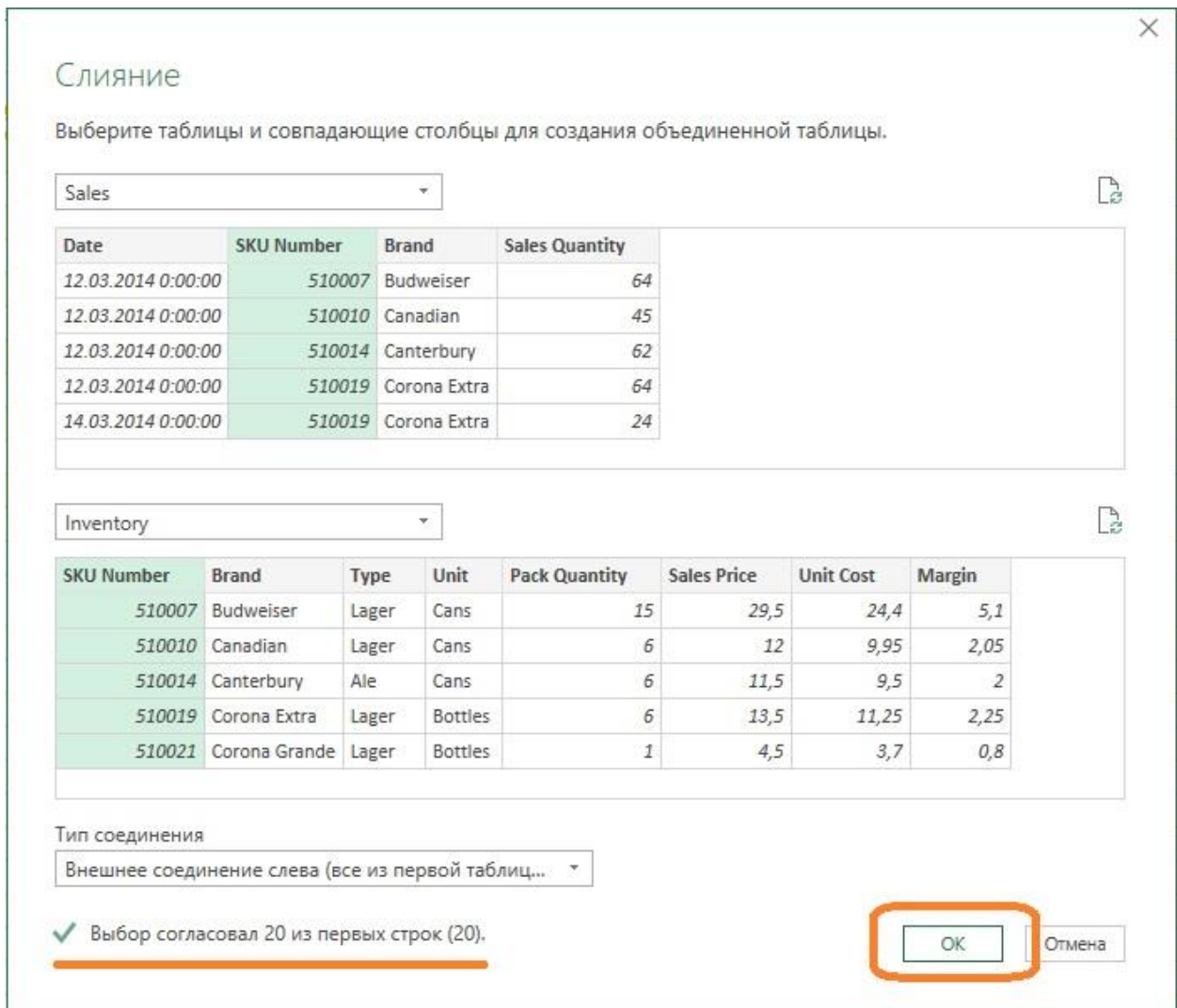


Рис. 9.4. Столбцы для связывания выбраны корректно

Обратите внимание, что предложенные вами столбцы для установления связи были проверены Power Query, и одобрены. Если, например, в таблице *Sales* вы выделите столбце *Brand* кнопка *Ок*, по-прежнему будет неактивна, а в нижней части окна *Слияние* появится сообщение об ошибке.

Нажмите кнопку *Ок*. Откроется редактор Power Query, в котором вторая таблица представлена одним столбцом:

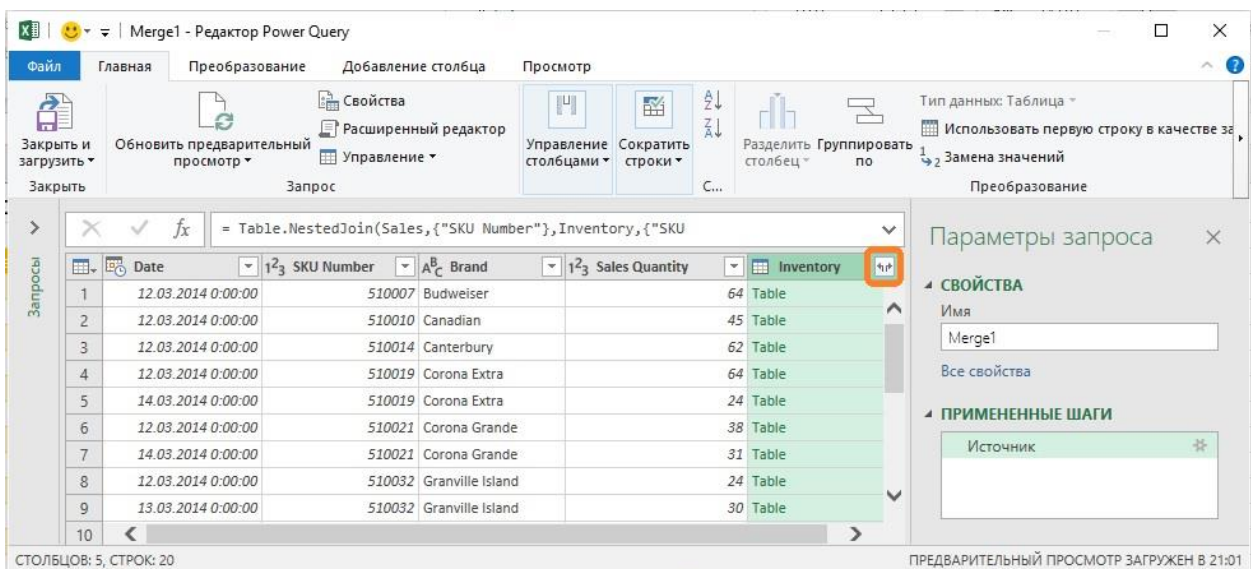


Рис. 9.5. Новый (последний) столбец таблицы *Sales* содержит соответствующие записи таблицы *Inventory*

Разверните столбец *Inventory*. Вопрос только в том, какие столбцы таблицы *Inventory* вам нужны. Итак, щелкните значок развернуть. Снимите галочку со столбцов, которые уже есть в таблице *Sales* (SKU Number and Brand), снимите флажок *Использовать исходное имя столбца как префикс*, нажмите *Ok*. Теперь сведения о продукте объединены с продажами:

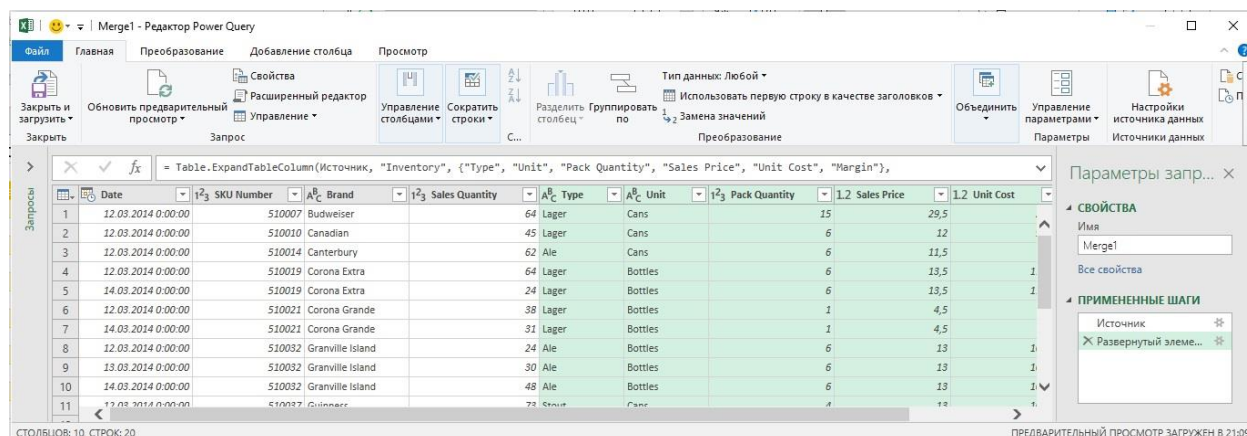


Рис. 9.6. Детали из таблицы *Inventory* объединены с данными таблицы *Sales*

Переименуйте запрос *OneToMany*. *Главная* → *Закреть и загрузить*. На листе Excel отобразится 20 строк *Таблицы*, как если бы свою работу выполнила функция ВПР:

	A	B	C	D	E	F	G	H	I	J
1	Date	SKU Number	Brand	Sales Quantity	Type	Unit	Pack Quantity	Sales Price	Unit Cost	Margin
2	12.03.2014 0:00:00	510007	Budweiser	64	Lager	Cans	15	29,5	24,4	5,1
3	12.03.2014 0:00:00	510010	Canadian	45	Lager	Cans	6	12	9,95	2,05
4	12.03.2014 0:00:00	510014	Canterbury	62	Ale	Cans	6	11,5	9,5	2
5	12.03.2014 0:00:00	510019	Corona Extra	64	Lager	Bottles	6	13,5	11,25	2,25
6	14.03.2014 0:00:00	510019	Corona Extra	24	Lager	Bottles	6	13,5	11,25	2,25
7	12.03.2014 0:00:00	510021	Corona Grande	38	Lager	Bottles	1	4,5	3,7	0,8
8	14.03.2014 0:00:00	510021	Corona Grande	31	Lager	Bottles	1	4,5	3,7	0,8
9	12.03.2014 0:00:00	510032	Granville Island	24	Ale	Bottles	6	13	10,95	2,05
10	13.03.2014 0:00:00	510032	Granville Island	30	Ale	Bottles	6	13	10,95	2,05
11	14.03.2014 0:00:00	510032	Granville Island	48	Ale	Bottles	6	13	10,95	2,05
12	12.03.2014 0:00:00	510037	Guinness	73	Stout	Cans	4	13	10,99	2,01
13	13.03.2014 0:00:00	510037	Guinness	76	Stout	Cans	4	13	10,99	2,01
14	14.03.2014 0:00:00	510037	Guinness	74	Stout	Cans	4	13	10,99	2,01
15	12.03.2014 0:00:00	510038	Heineken	30	Lager	Bottles	6	14,5	11,95	2,55
16	13.03.2014 0:00:00	510038	Heineken	30	Lager	Bottles	6	14,5	11,95	2,55
17	14.03.2014 0:00:00	510038	Heineken	44	Lager	Bottles	6	14,5	11,95	2,55
18	12.03.2014 0:00:00	510046	Kokanee	57	Lager	Tall Cans	6	15,5	13,05	2,45
19	12.03.2014 0:00:00	510057	Miller	43	Lager	Bottles	12	24	20,2	3,8
20	12.03.2014 0:00:00	510059	OK Springs	76	Lager	Bottles	6	13	10,99	2,01
21	12.03.2014 0:00:00	510065	OK Springs	78	Ale	Bottles	6	13	10,99	2,01
22										

Рис. 9.7. Слияние на основе связи один-ко-многим

Многие-ко-многим

При выполнении процедуры слияния нужно быть внимательным. Если вы попытаетесь выполнить слияние иным образом, вы обнаружите иной результат:

- Данные → Получить данные → Объединить запросы → Объединить.
- Выберите *Sales* в верхней части, а *Inventory* – в нижней
- Щелкните заголовок *Brand* в каждой таблице
- Нажимать *Ok*
- Щелкните значок развернуть
- Снимите галочку со столбцов *SKU Number and Brand*
- Снимите флажок *Использовать исходное имя столбца как префикс*
- Нажмите *Ok*
- Переименовать запрос *ManyToMany*
- *Главная* → *Закреть и загрузить*

Новая процедура слияния отличается двумя аспектами: (1) таблицы *Sales* и *Inventory* в окне *Слияние* переставлены местами, (2) для связи выбран столбец *Brand*. Тем не менее, в *Таблице* на листе Excel появилось 22 записи – на 2 больше, чем исходное количество транзакций. Чтобы понять, почему это произошло, вернитесь в редактор Power Query, перейдите к первому шагу запроса *ManyToMany* (цифра 1 на рис. 9.8). Если вы перейдете к строке 19 и щелкните пробел справа от слова *Table* (2), вы увидите предварительный просмотр данных в таблице, которые при слиянии будут объединены в таблице *Sales* (3).

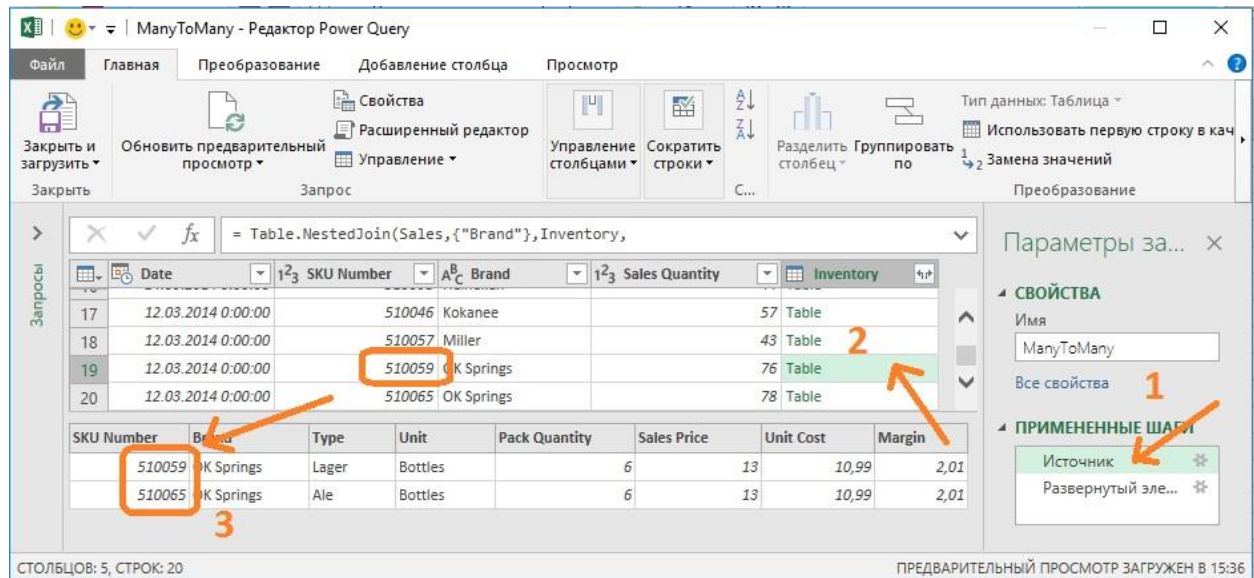


Рис. 9.8. Слияние многие-ко-многим в действии

В предыдущем слиянии вы связали данные на основе столбца *SKU Number*. В этом примере слияние осуществляется на основе столбца *Brand*. Однако бренду *OK Springs* соответствует два артикула в таблице *Inventory*. На основе этого примера вы можете увидеть, что нужно быть осторожным при создании слияний, чтобы не попасться в ловушку многие-ко-многим.