

Объединение таблиц: забудьте про ВПР

*Любая достаточно развитая технология неотличима от волшебства.
(Артур Кларк)*

Теперь давайте на примерах рассмотрим второй тип слияния запросов в Power Query – объединение (merge). Предположим, что у нас имеются три таблицы: с данными по продажам, прайс-лист на товары и справочник по регионам и менеджерам:

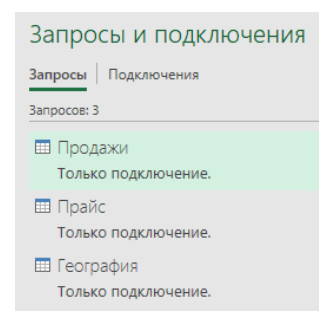
Продажи				Прайс			География			
Дата	Товар	Количество	Продавец	Продукт	Цена	Скидка	Город	Менеджер	Статус	Регион
17.10.2017	Джемпер	1	Дмитрий	Джемпер	9055	0%	Новосибирск	Дмитрий	менеджер	Сибирь
02.06.2017	Брюки	3	Дмитрий	Брюки	4625	10%	Екатеринбург	Елена	менеджер	Сибирь
15.11.2017	Кардиган	4	Елена	Кардиган	5810	0%	Санкт-Петербург	Андрей	менеджер	Запад
01.10.2017	Свитер	3	Андрей	Свитер	7925	0%	Москва	Анастасия	стажер	Центр
25.01.2017	джинсы	3	Анастасия	Джинсы	6435	0%	Москва	Анна	менеджер	Центр
22.04.2017	Кофта	4	Дмитрий	Кофта	735	0%	Екатеринбург	Сергей	менеджер	Сибирь
17.11.2017	Пуловер	4	Анна	Пуловер	5225	0%	Новосибирск	Максим	стажер	Сибирь
27.11.2017	Свитер	2	Елена	Юбка	8980	20%	Екатеринбург	Александр	менеджер	Сибирь
16.10.2017	Юбка	4	Сергей	Блуза	7900	0%	Екатеринбург	Мария	стажер	Сибирь
12.12.2017	Блуза	4	Анастасия	Комбинезон	4940	0%	Санкт-Петербург	Дарья	стажер	Запад
24.07.2017	Кардиган	1	Максим	Рубашка	9285	5%				
20.05.2017	Комбинезон	3	Елена	Водолазка	9255	0%				
23.10.2017	Джемпер	2	Дмитрий	Бриджи	2220	5%				
20.11.2017	Рубашка	2	Максим	Лосины	9070	0%				
04.11.2017	Джинсы	4	Анна	Шорты	8150	15%				
09.05.2017	Кофта	1	Андрей							
23.09.2017	Брюки	2	Александр							
23.07.2017	Кардиган	2	Дмитрий							
22.08.2017	Джинсы	1	Анастасия							
23.02.2017	Водолазка	1	Максим							

На выходе необходимо посчитать суммарную выручку по каждому товару (с учетом скидок) для каждого региона. При этом количество проданного у нас в первой таблице, цены надо подтягивать из второй, а информация о регионах для каждого менеджера есть только в третьей. Если решать эту задачу классическим путем, то пришлось бы как минимум к первой таблице добавлять пару вычисляемых столбцов с функциями **ВПР (VLOOKUP)**, **ПОИСКПОЗ (MATCH)** и **ИНДЕКС (INDEX)**, а потом вычислять итоги с помощью сводной таблицы. Давайте посмотрим, как всё вышеперечисленное можно изящно сделать не формулами, а через Power Query.

Загружаем все таблицы как подключения

Сначала превратим таблицы в «умные», дадим им понятные имена (я назвал их *Продажи*, *Прайс* и *География* соответственно) и загрузим их в Power Query по очереди, используя кнопку **Из таблицы / диапазона** на вкладке **Данные (Data → From Table / Range)** и вернемся потом обратно в режиме **Только создать подключение (Only Create Connection)**.

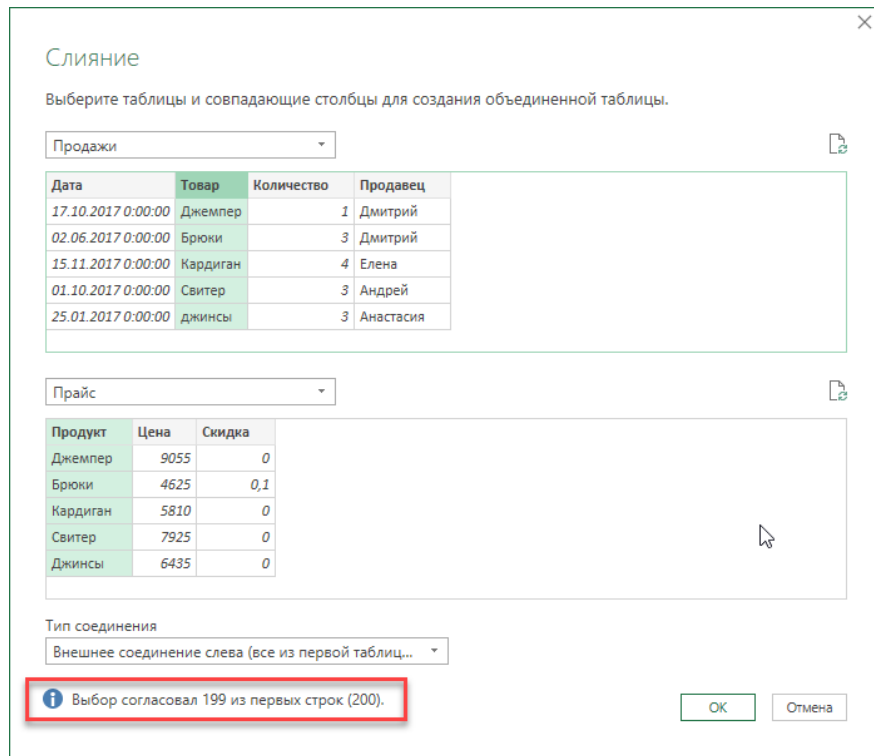
Чуть позже в главе **Загрузка «умных» таблиц в Power Query макросом** мы научимся быстро делать массовую загрузку сразу всех «умных» таблиц из книги в Power Query, но пока этот относительно скучный шаг надо проделать вручную.



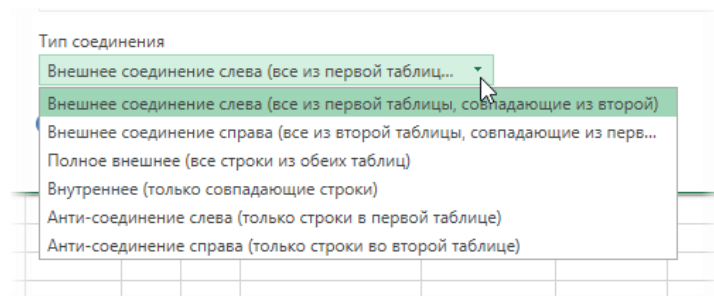
Выполняем слияние

Чтобы добавить цены из прайса к таблице продаж, на вкладке **Данные** выберем команду **Получить данные → Объединение запросов → Объединить (Data → Get Data → Combine queries → Merge)**.

В открывшемся окне сверху выберем таблицу, к которой мы хотим добавить данные (*Продажи*), а снизу таблицу, откуда мы хотим их подставить (*Прайс*). Затем выделим мышкой столбцы *Товар* и *Продукт* в обеих таблицах, по которым должен идти поиск и подстановка:



Обратите внимание, что в нижней части окна есть выпадающий список, где можно выбрать тип соединения:



Чуть позже мы подробно разберем каждый вариант, а пока что достаточно знать, что первый пункт из этого списка – **Внешнее соединение слева (Left Outer Join)** – это как раз то, что делает функция **ВПР**, и то, что нам нужно в этом примере.

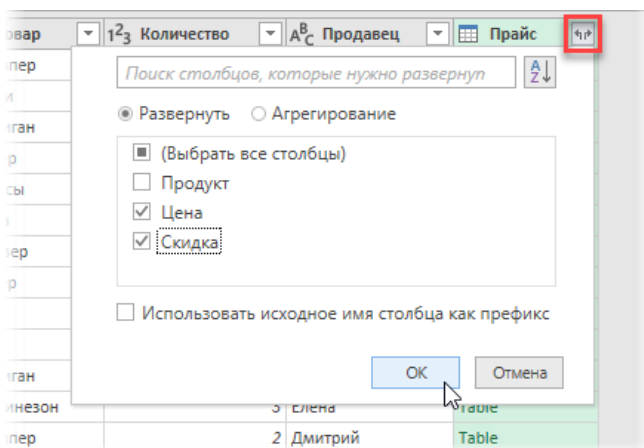
Ещё один важный нюанс в том, что внизу окна Power Query напишет фразу «**Выбор согласовал 199 из первых строк (200)**». В переводе на человеческий язык это означает, что он смог найти только 199 цен к 200 наших исходных сделок, т. е. один из товаров в таблице *Продажи* был не найден в прайс-листе. Чуть позже мы разберемся и с этим, а пока нажмем **OK**.

Должно открыться окно редактора запросов, где мы увидим новый запрос со стандартным именем *Merge1*, в котором будет наша исходная таблица продаж с добавленным к ней столбцом *Прайс*. В каждой ячейке этого столбца будет таблица с фрагментом прайс-листа, соответствующим данному товару. Увидеть её содержимое можно, щелкнув мышью в белый фон ячейки со словом **Table** (но не в слово **Table!**):

	Дата	Товар	Количество	Продавец	Прайс
1	17.10.2017 0:00:00	Джемпер		1 Дмитрий	Table
2	02.06.2017 0:00:00	Брюки		3 Дмитрий	Table
3	15.11.2017 0:00:00	Кардиган		4 Елена	Table
4	01.10.2017 0:00:00	Свитер		3 Андрей	Table
5	25.01.2017 0:00:00	джинсы		3 Анастасия	Table
6	22.04.2017 0:00:00	Кофта		4 Дмитрий	Table
7	17.11.2017 0:00:00	Пуловер		4 Анна	Table
8	27.11.2017 0:00:00	Свитер		2 Елена	Table
9	16.10.2017 0:00:00	Юбка		4 Сергей	Table
10	12.12.2017 0:00:00	Блуза		4 Анастасия	Table
11	24.07.2017 0:00:00	Кардиган		1 Максим	Table

Продукт	Цена	Скидка
Брюки	4625	0,1

Теперь развернем вложенные таблицы, используя кнопку с двойными стрелками в заголовке столбца *Прайс*. Из выпадающего списка можно выбрать те столбцы прайс-листа, которые мы хотим подставить, и снять флажок **Использовать исходное имя столбца как префикс** (Use original column name as prefix), чтобы новые столбцы назывались просто *Цена* и *Скидка*, а не *Прайс.Цена* и *Прайс.Скидка*:



После нажатия на **OK** мы достигнем желанной цели: к нашей таблице продаж добавятся колонки с ценами и скидками из прайс-листа:

	Дата	Товар	Количество	Продавец	Цена	Скидка
1	17.10.2017 0:00:00	Джемпер		1 Дмитрий	9055	0
2	23.10.2017 0:00:00	Джемпер		2 Дмитрий	9055	0
3	02.06.2017 0:00:00	Брюки		3 Дмитрий	4625	0,1
4	15.11.2017 0:00:00	Кардиган		4 Елена	5810	0
5	24.07.2017 0:00:00	Кардиган		1 Максим	5810	0
6	01.10.2017 0:00:00	Свитер		3 Андрей	7925	0
7	27.11.2017 0:00:00	Свитер		2 Елена	7925	0
8	25.01.2017 0:00:00	джинсы		3 Анастасия	null	null
9	04.11.2017 0:00:00	Джинсы		4 Анна	6435	0
10	22.04.2017 0:00:00	Кофта		4 Дмитрий	735	0
11	09.05.2017 0:00:00	Кофта		1 Андрей	735	0

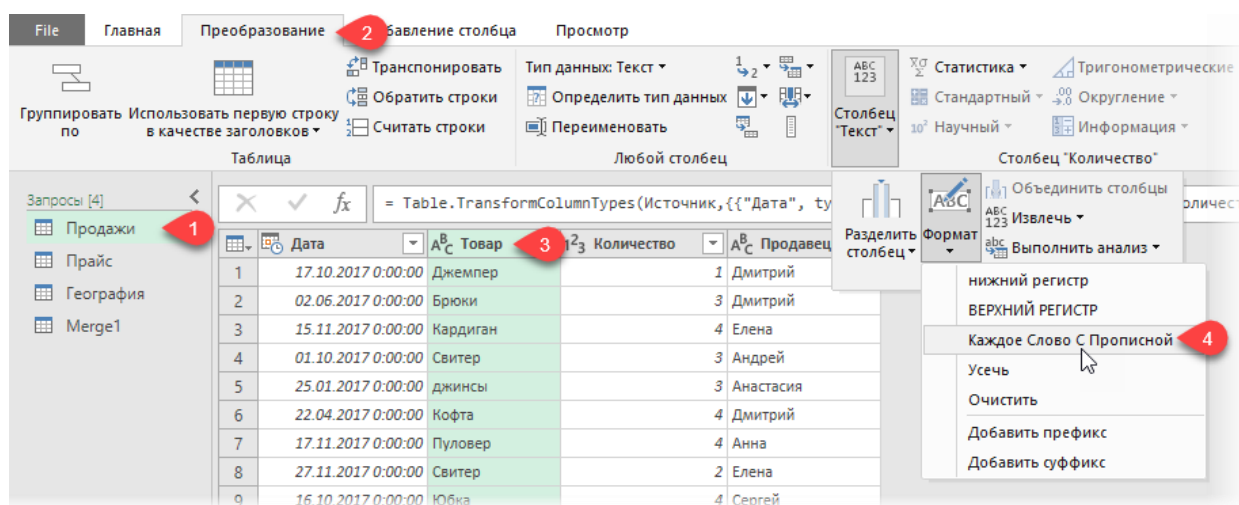
Исправляем ошибки

Думаю, вам известно, что когда функция **ВПР** (VLOOKUP) не находит искомого значения, то она выдает ошибку **#Н/Д** (#N/A). Power Query же реагирует на подобные ситуации по-другому, что хорошо видно в 8-й строке нашей итоговой таблицы. Поскольку (в отличие от **ВПР**) объединение происходит с учётом регистра, «джинсы» с маленькой буквы не нашлись в прайс-листе (где они есть, но с заглавной), и напротив них в таблице появилось значение **null**.

null – это специальное слово (и тип данных) в Power Query, обозначающее пустоту¹. В отличие от Excel, где пустая ячейка может быть чем угодно – от формулы, выводящей пустую строку "", до форматирования «белым на белом», в Power Query этот вопрос имеет однозначный ответ: **null** – значит пусто.

Вопрос, что с этими **null** теперь делать. Тут есть три варианта.

1. **Ничего не делать.** Ячейки с **null** в Power Query потом на листе Excel превратятся просто в пустые ячейки.
2. **Заменить **null** на что-то полезное.** Щёлкнув правой кнопкой мыши по заголовку столбца с ошибками, можно выбрать в контекстном меню команду **Заменить ошибки (Replace Errors)** и ввести значение, на которое вы хотите их заменить (например, на 0 или на «здесь ошибка»).
3. Если ошибку можно **исправить на лету**, как в нашем случае, то проще всего будет сделать это тут же. Для этого переключимся в запрос **Продажи** через левую панель в редакторе, выделим столбец **Товар** и исправим в нем регистр с помощью команды **Формат → Каждое Слово С Прописной** на вкладке **Преобразование (Transform → Format → Capitalize Each Word)**:

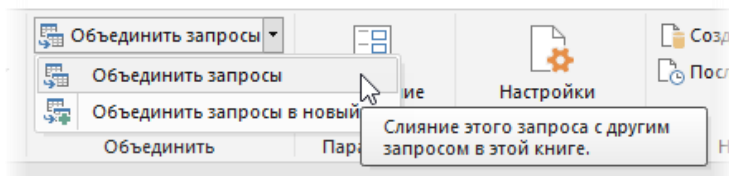


Вернувшись после этого в запрос **Merge1**, мы увидим, что ошибка исчезла.

Объединение в этом же запросе

Чтобы объединить две таблицы, не обязательно создавать отдельный третий запрос: во многих случаях бывает удобнее выполнить это здесь же, на месте, в текущем запросе, с которым мы имеем дело.

Давайте подобным образом добавим к нашей сборке данные из третьей таблицы **География**. Для этого, находясь в запросе **Merge1**, выберем на вкладке **Главная** команду **Объединить запросы → Объединить запросы (Home → Merge Queries → Merge Queries)**:



Мы увидим уже знакомое окно, где нужно будет выбрать запрос **География** в качестве второй таблицы для объединения и выделить общие столбцы (**Продавец** и **Менеджер**) в этих таблицах, как мы делали ранее:

¹ Подробнее про **null** см. главу **Тип null** ближе к концу этой книги.

Слияние

Выберите таблицу и совпадающие столбцы для создания объединенной таблицы.

Merge1

Дата	Товар	Количество	Продавец	Цена	Скидка
17.10.2017 0:00:00	Джемпер	1	Дмитрий	9055	0
23.10.2017 0:00:00	Джемпер	2	Дмитрий	9055	0
02.06.2017 0:00:00	Брюки	3	Дмитрий	4625	0,1
15.11.2017 0:00:00	Кардиган	4	Елена	5810	0
24.07.2017 0:00:00	Кардиган	1	Максим	5810	0

География

Город	Менеджер	Статус	Регион
Новосибирск	Дмитрий	менеджер	Сибирь
Екатеринбург	Елена	менеджер	Сибирь
Санкт-Петербург	Андрей	менеджер	Запад
Москва	Анастасия	стажер	Центр
Москва	Анна	менеджер	Центр

Тип соединения
Внешнее соединение слева (все из первой таблиц...

✓ Выбор согласовал 200 из первых строк (200).

OK Отмена

После нажатия на **OK** и разворачивания вложенных таблиц, как в предыдущем случае, мы увидим подставленные данные из третьей таблицы – столбцы **Город**, **Статус** и **Регион**:

Дата	Товар	Количество	Продавец	Цена	Скидка	Город	Статус	Регион
17.10.2017 0:00:00	Джемпер	1	Дмитрий	9055	0	Новосибирск	менеджер	Сибирь
23.10.2017 0:00:00	Джемпер	2	Дмитрий	9055	0	Новосибирск	менеджер	Сибирь
02.06.2017 0:00:00	Брюки	3	Дмитрий	4625	0,1	Новосибирск	менеджер	Сибирь
22.04.2017 0:00:00	Кофта	4	Дмитрий	735	0	Новосибирск	менеджер	Сибирь
15.11.2017 0:00:00	Кардиган	4	Елена	5810	0	Екатеринбург	менеджер	Сибирь
27.11.2017 0:00:00	Свитер	2	Елена	7925	0	Екатеринбург	менеджер	Сибирь
01.10.2017 0:00:00	Свитер	3	Андрей	7925	0	Санкт-Петербург	менеджер	Запад
09.05.2017 0:00:00	Кофта	1	Андрей	735	0	Санкт-Петербург	менеджер	Запад
25.01.2017 0:00:00	Джинсы	3	Анастасия	6435	0	Москва	стажер	Центр
24.07.2017 0:00:00	Кардиган	1	Максим	5810	0	Новосибирск	стажер	Сибирь

Теперь можно добавить к нашей таблице вычисляемый столбец для выручки. Для этого на вкладке **Добавление столбца** (Add Column) нажмем на кнопку **Настраиваемый столбец** (Custom Column) и введем в открывшееся окно имя нового столбца и формулу для его расчёта:

Настраиваемый столбец

Имя нового столбца
Выручка

Пользовательская формула столбца:
=[Количество]*[Цена]*(1-[Скидка])

Доступные столбцы:
Дата
Товар
Количество
Продавец
Цена
Скидка
Город

<< Вставить

Сведения о формулах Power Query

✓ Синтаксические ошибки не обнаружены.

OK Отмена

Что интересно, после нажатия на **ОК** и создания вычисляемого столбца для выручки можно совершенно безнаказанно удалить колонки, которые участвуют в её вычислении, например **Количество**, **Цену** и **Скидку**. Это никак не повлияет на столбец **Выручка**:

	Дата	Товар	Продавец	Город	Статус	Регион	Выручка
1	17.10.2017 0:00:00	Джемпер	Дмитрий	Новосибирск	менеджер	Сибирь	9055
2	23.10.2017 0:00:00	Джемпер	Дмитрий	Новосибирск	менеджер	Сибирь	18110
3	02.06.2017 0:00:00	Брюки	Дмитрий	Новосибирск	менеджер	Сибирь	12487,5
4	22.04.2017 0:00:00	Кофта	Дмитрий	Новосибирск	менеджер	Сибирь	2940
5	15.11.2017 0:00:00	Кардиган	Елена	Екатеринбург	менеджер	Сибирь	23240
6	27.11.2017 0:00:00	Свитер	Елена	Екатеринбург	менеджер	Сибирь	15850
7	01.10.2017 0:00:00	Свитер	Андрей	Санкт-Петербург	менеджер	Запад	23775
8	09.05.2017 0:00:00	Кофта	Андрей	Санкт-Петербург	менеджер	Запад	735
9	25.01.2017 0:00:00	Джинсы	Анастасия	Москва	стажер	Центр	19305

Для большинства пользователей, привыкших к формулам на листе Excel, этот момент будет слегка необычным. Однако, если подумать, это даёт нам возможность избавляться в запросах от ненужных более столбцов в любой момент, не задумываясь о том, в каких вычислениях они раньше участвовали. А чем раньше вы избавитесь от лишних данных, тем быстрее будет работать ваш запрос.