

Глава 11. Определение типов данных в 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](#)

Power Query автоматически устанавливает типы данных. Причем, от версии к версии делает это всё лучше и лучше. Тем не менее, встречаются ситуации, когда его выбор покажется вам странным. А иногда Power Query и вовсе не может распознать тип данных.

	A	B	C	D	E	F
1	Fred's Pet Store					
2	Sales Listing For Month of:					
3	31.05.2014					
4						
5	Date	Inventory Item	Sold By	Cost	Price	Commission
6	10.05.2014	Lovable Kitten	Fred	12,00	45,00	1,35
7	09.05.2014	Talkative Parrot	John	17,00	32,00	0,96
8	24.05.2014	Lovable Kitten	Jane	12,00	45,00	1,35
9	20.05.2014	Adorable Kitty Cat	John	25,00	35,00	1,05
10	08.05.2014	Lovable Kitten	Jane	12,00	45,00	1,35
11	23.05.2014	Sleepy Gerbil	Jane	25,00	39,00	1,17

Рис. 11.1. Исходная Таблица в Excel

Power Query распознает несколько типов данных (больше, чем Excel):

- Число: десятичное, целое, валюта
- Дата и время: дата/время, дата, время, дата/время/часовой пояс, продолжительность
- Текст: текст
- Логическое значение: ИСТИНА/ЛОЖЬ
- Объект: бинарный (двоичный)
- Неопределенный: любой

Наибольшие неудобства доставляет тип данных *Любой*. Его Power Query использует, чтобы указать, что не знает, какой тип данных выбрать. Данные, определенные как *Любой*, могут принимать различные формы при загрузке или ссылке в будущем.

Почему важны типы данных?

Чтобы продемонстрировать важность типов данных, вы создадите запрос, и удалите типы данных, определенные Power Query по умолчанию. Откройте файл Defining Data Types.xlsx (рис. 11.1). Выберите любую ячейку в Таблице на листе May → Данные → Из таблицы/диапазона. В редакторе Power Query в области ПРИМЕНЕННЫЕ ШАГИ удалите шаг *Измененный тип*. В этот момент типы данных во всех столбцах не определены:

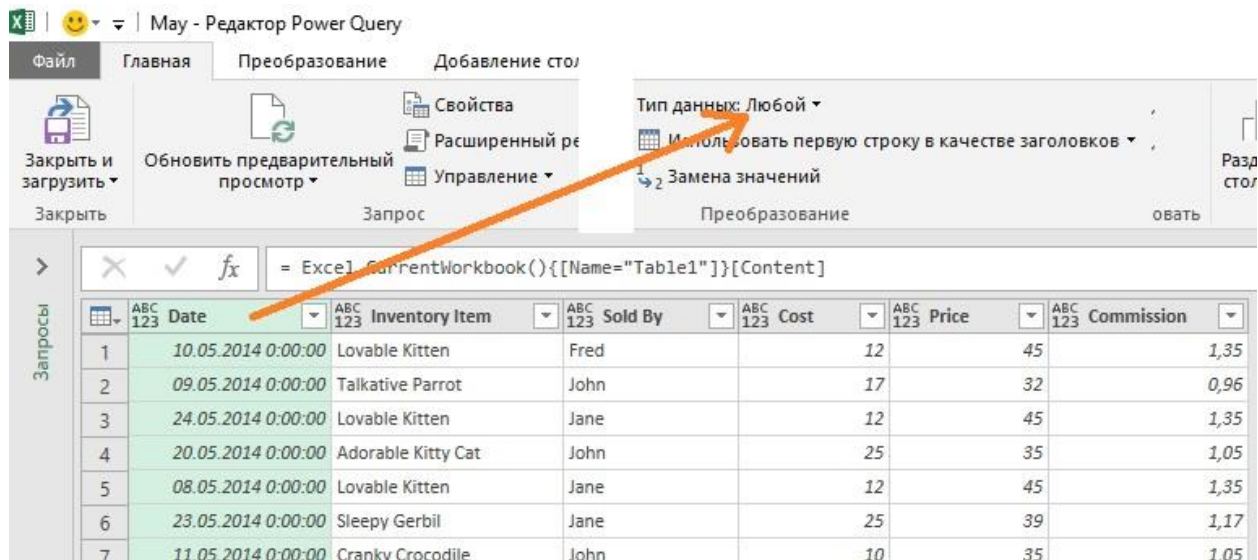


Рис. 11.2. В столбце *Data* тип данных *Любой*

Интересно, что каждый столбец в этом наборе данных имеет один и тот же тип данных – *Любой*, но все они выглядят по-разному. Даты выглядят так, как будто они находятся в формате даты/времени, а числа выглядят так, как будто они являются числовым типом данных. Так в чем же дело? Выполните следующие действия, чтобы узнать:

- Переименуйте запрос – *May*
- *Главная* → *Закреть и загрузить*

На листе Excel появится *Таблица*:

	A	B	C	D	E	F
1	Date	Inventory Item	Sold By	Cost	Price	Commission
2	41769	Lovable Kitten	Fred	12	45	1,35
3	41768	Talkative Parrot	John	17	32	0,96
4	41783	Lovable Kitten	Jane	12	45	1,35
5	41779	Adorable Kitty Cat	John	25	35	1,05
6	41767	Lovable Kitten	Jane	12	45	1,35
7	41782	Sleepy Gerbil	Jane	25	39	1,17
8	41770	Cranky Crocodile	John	10	35	1,05

Рис. 11.3. Что случилось с этим столбцом *Date*?

Excel не распознал столбец *Date* как даты, и вывел порядковые номера дат. Хотя это легко исправить (форматируя столбец как даты), это указывает на более серьезную проблему.

Power Pivot и любой тип данных

Измените запрос и «приземлите» его на лист, и в модель данных. Для этого в файле Excel в области *Запросы и подключения* щелкните правой кнопкой мыши запрос *May* → *Загрузить в...* В открывшемся окне *Импорт данных* установите флажок *Добавить в модель данных* (оставив выбранным параметр *Таблица*, рис. 11.4). Нажмите *Ок*. Появится окно предупреждения об изменениях. Подтвердите.

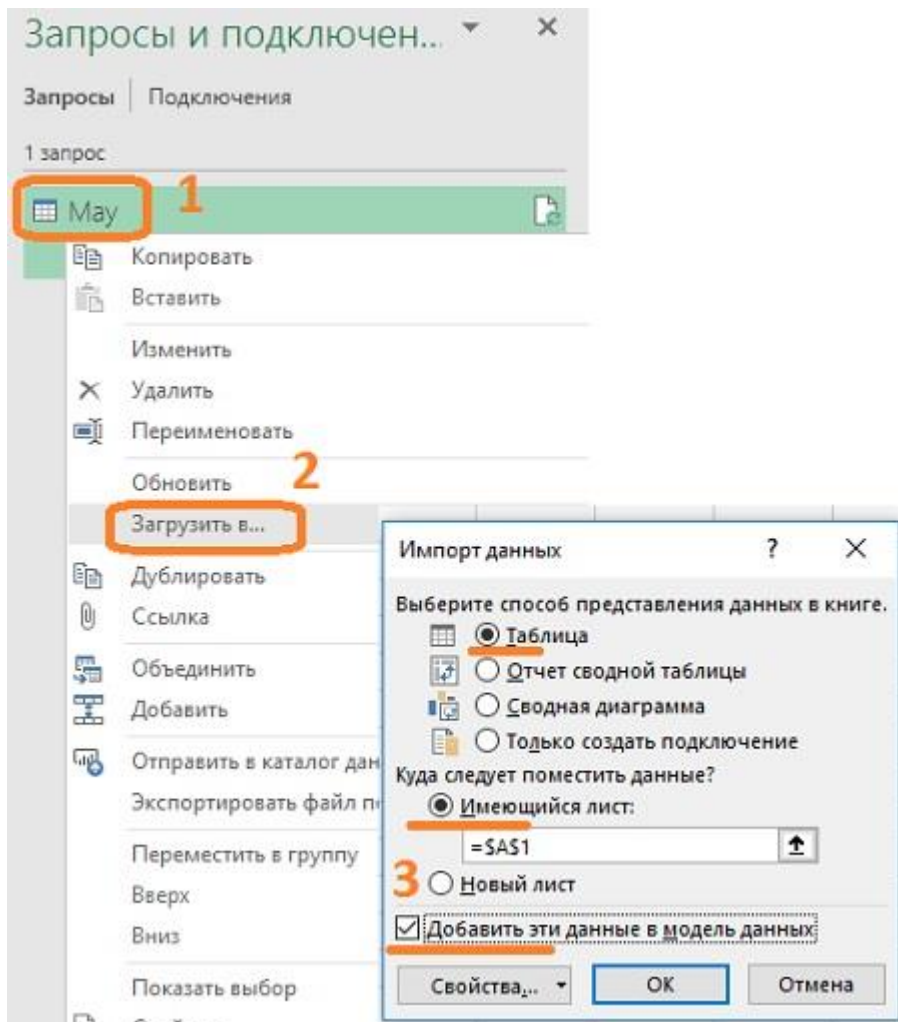


Рис. 11.4. Добавить данные в модель данных Power Pivot

Данные будут помещены в модель данных, а Таблица на листе Excel изменит свой вид:

	A	B	C	D	E	F
1	Date	Inventory Item	Sold By	Cost	Price	Commission
2	31.05.2014 0:00:00	Slithering Snake	Fred	13	30	0,9
3	31.05.2014 0:00:00	Rambunctious Puppy	John	9	30	0,9
4	21.05.2014 0:00:00	Tubby Turtle	John	8	30	0,9
5	04.05.2014 0:00:00	Tubby Turtle	Fred	8	30	0,9
6	31.05.2014 0:00:00	Rambunctious Puppy	Fred	9	30	0,9
7	21.05.2014 0:00:00	Tubby Turtle	Fred	8	30	0,9
8	29.05.2014 0:00:00	Slithering Snake	John	13	30	0,9

Рис. 11.5. Странные изменения Таблицы на листе Excel при добавлении данных в модель данных

Простое добавление соединения к модели данных Power Pivot изменяет поле обратно на дату. Или нет? Чтобы проверить, выберите любую ячейку в столбце A, перейдите на вкладку Главная, посмотрите на поле формат. Никаких изменений нет! Более того, вы можете применить любой

формат к данным в столбце A, и не увидите никаких изменений. Данные обрабатываются как текст.

Если вы откроете Power Pivot, вы увидите, что данные выглядят как дата/время, но тип данных – Текст. Это не позволит вам создать в Power Pivot таблицу календаря!

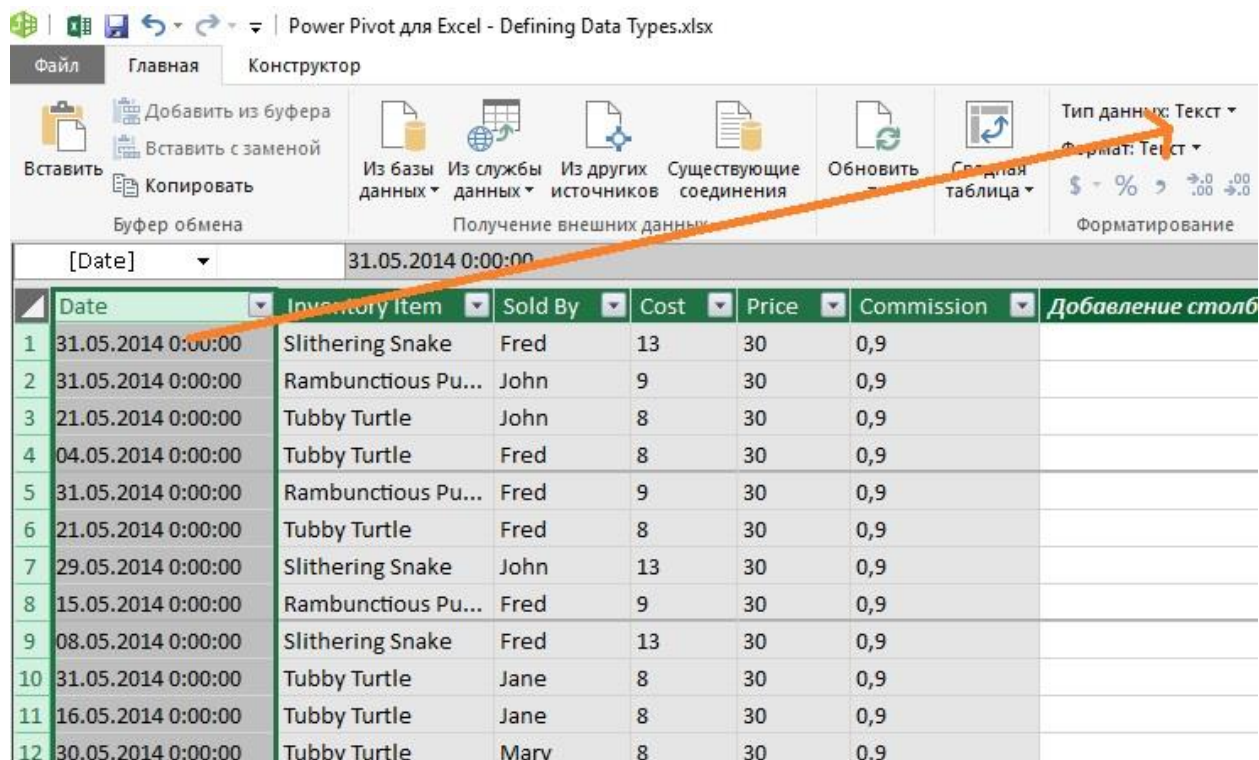


Рис. 11.6. Данные в Power Pivot, которые выглядят как даты, являются текстовыми

Если же вы попытаетесь изменить тип данных внутри Power Pivot, появится сообщение об ошибке:

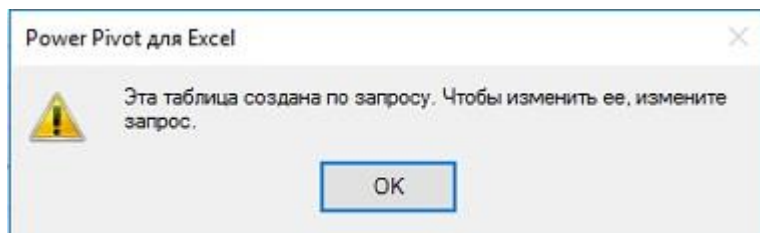


Рис. 11.7. Данные, импортированные по запросу, не могут быть изменены в Power Pivot

«Правильные» типы данных

Исправить эти проблемы довольно просто. Отредактируйте запрос *May* и установите релевантные типы данных:

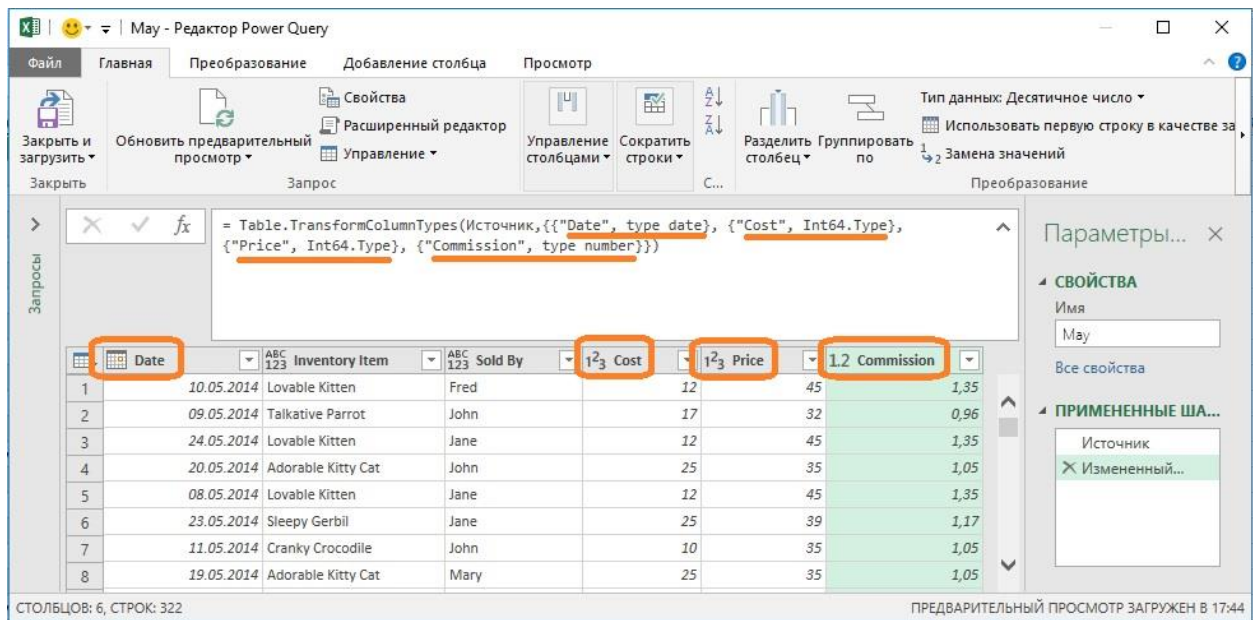


Рис. 11.8. Исправление типа данных у источника – в запросе

Проверьте: теперь, и на листе Excel, и в модели Power Pivot типы данных обрабатываются корректно. Помните, что Power Pivot форматирует даты в формате Дата/Время, причем 00:00:00 составляет 0 минут после полуночи указанной даты.

Объединение запросов с различными типами данных

Одна из типичных задач – объединение двух таблиц. Но что происходит, когда столбцы в этих таблицах имеют разные типы данных?

Поскольку запрос *May* уже исправлен, можно создать новый запрос для июньских данных без типов данных и посмотреть, что произойдет при их объединении. Выберите *Таблицу Jun*, пройдите по меню *Данные* → *Из таблицы/диапазона*. В редакторе Power Query удалите шаг *Измененный тип*. Переименуйте запрос *Jun*. *Главная* → *Заккрыть и загрузить...* → *Только создать подключение*.

У вас есть два запроса. В области *Запросы и подключения* щелкните правой кнопкой мыши на запросе *Jun* → *Добавить*. В окне *Добавление* выберите вторую таблицу – *May*. Нажмите *Ок*. Если вы проверите тип данных в столбце *Date*, то увидите, что он – *Любой*.

Закройте запрос без сохранения, и попробуйте в другом порядке. В области *Запросы и подключения* щелкните правой кнопкой мыши на запросе *May* → *Добавить*. В окне *Добавление* выберите вторую таблицу – *Jun*. Нажмите *Ок*. Не помогло – в столбце *Date*, по-прежнему тип *Любой*.

Не имеет значения, в каком порядке вы добавляете запросы: если типы данных в соединяемых таблицах не совпадают, они будут преобразованы в тип данных *Любой*. Исправьте типы данных в запросе *Jun* по аналогии с *May*, и лишь затем объедините запросы.

Рекомендации

Хотя Power Query довольно неплохо справляется с присвоением типов данных, мы рекомендуем после каждого импорта просмотреть шаг *Измененный тип* и проверить типы данных.