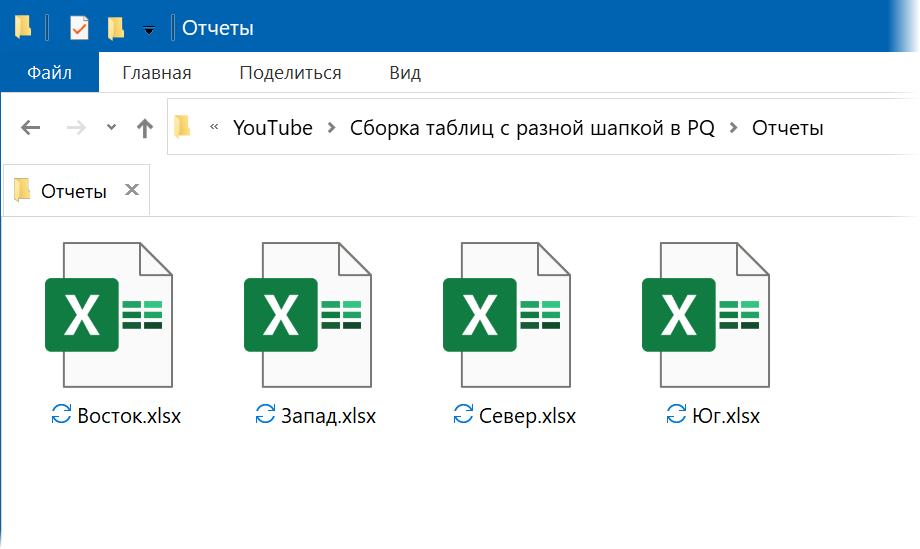
**Сборка таблиц с разными шапками из нескольких книг**

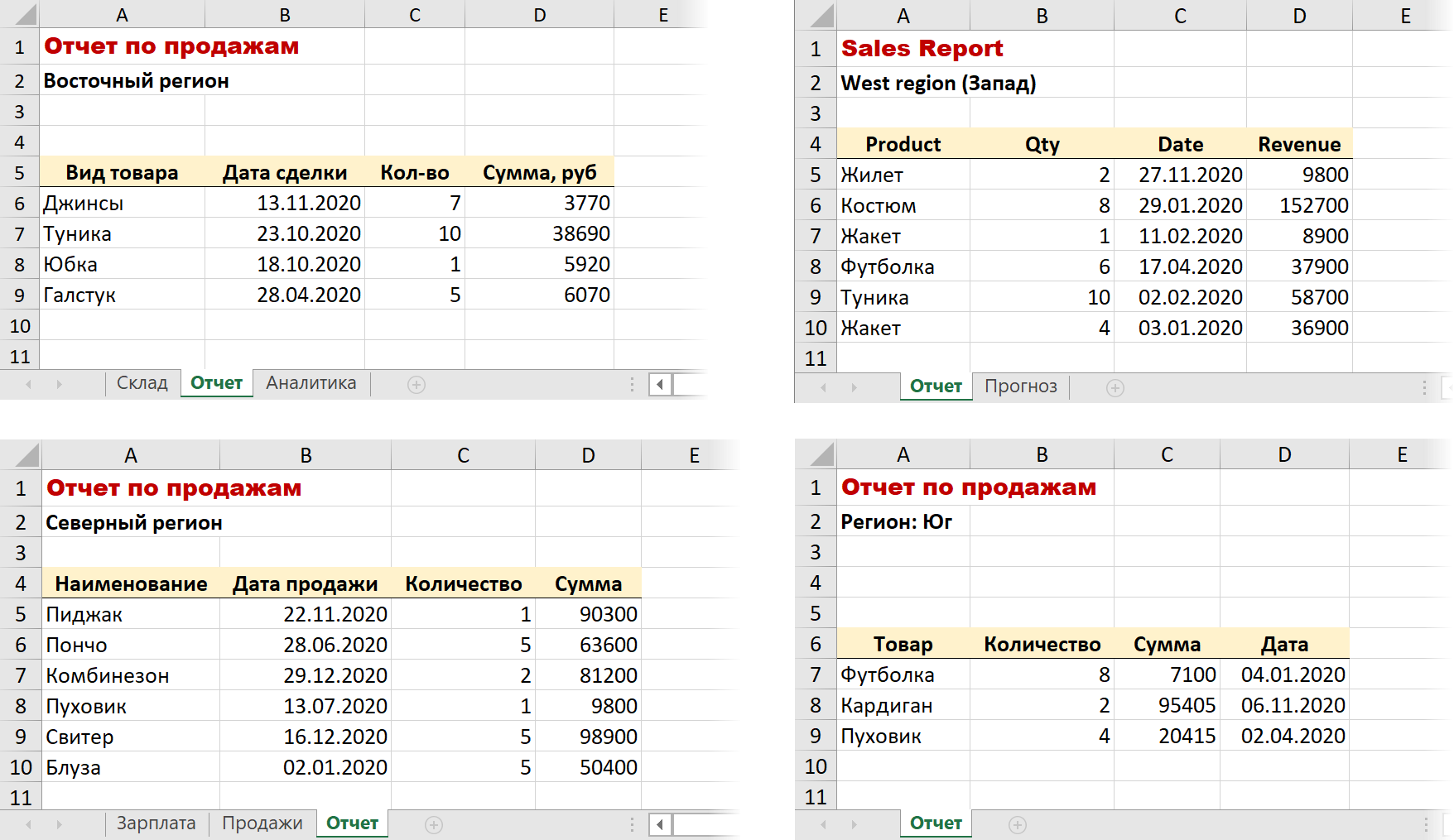
16692 20.05.2021 [Скачать пример](https://www.planetaexcel.ru/upload/iblock/1ec/2nmnjjmahzcoqq6xl4uad8bjbpgif0rj/collect-data-different-headers.zip)

**Постановка задачи**

Имеем несколько файлов (в нашем примере - 4 шт., в общем случае - сколько угодно) в одной папке **Отчеты**:



Внутри эти файлы выглядят вот так:



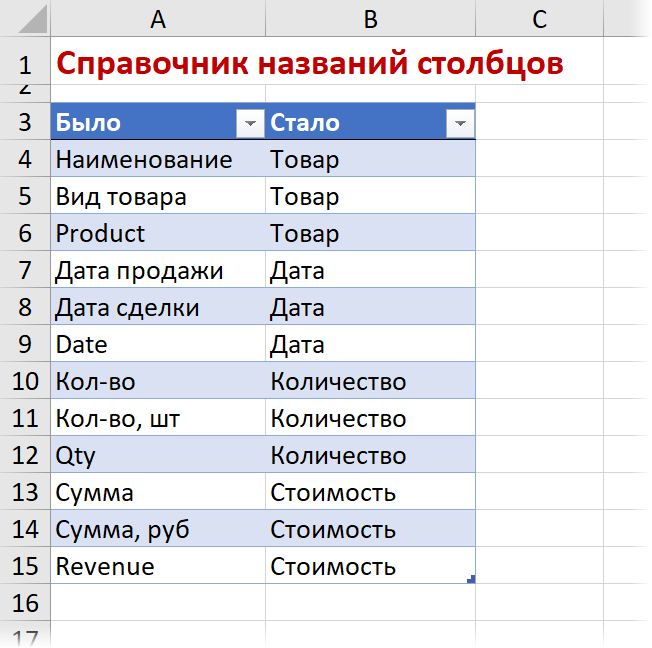
При этом:

* Нужный нам лист с данными всегда называется *Отчет*, номожет находиться в книге в любом месте.
* Помимо листа *Отчет*в каждой книге могут быть и другие листы.
* Таблицы с данными имеют разное количество строк и могут начинаться с разной строки на листе.
* Названия одних и тех же столбцов в разных таблицах может различаться (например, *Количество = Кол-во = Qty*).
* Столбцы в таблицах могут располагаться в разном порядке.

Задача: собрать данные по продажам из всех файлов с листа *Отчет*в одну общую таблицу, чтобы впоследствии построить по ней сводную или любую другую аналитику.

**Шаг 1. Готовим справочник названий столбцов**

Первое, что придётся сделать - подготовить справочник со всеми возможными вариантами названий столбцов и их правильной интерпретацией:

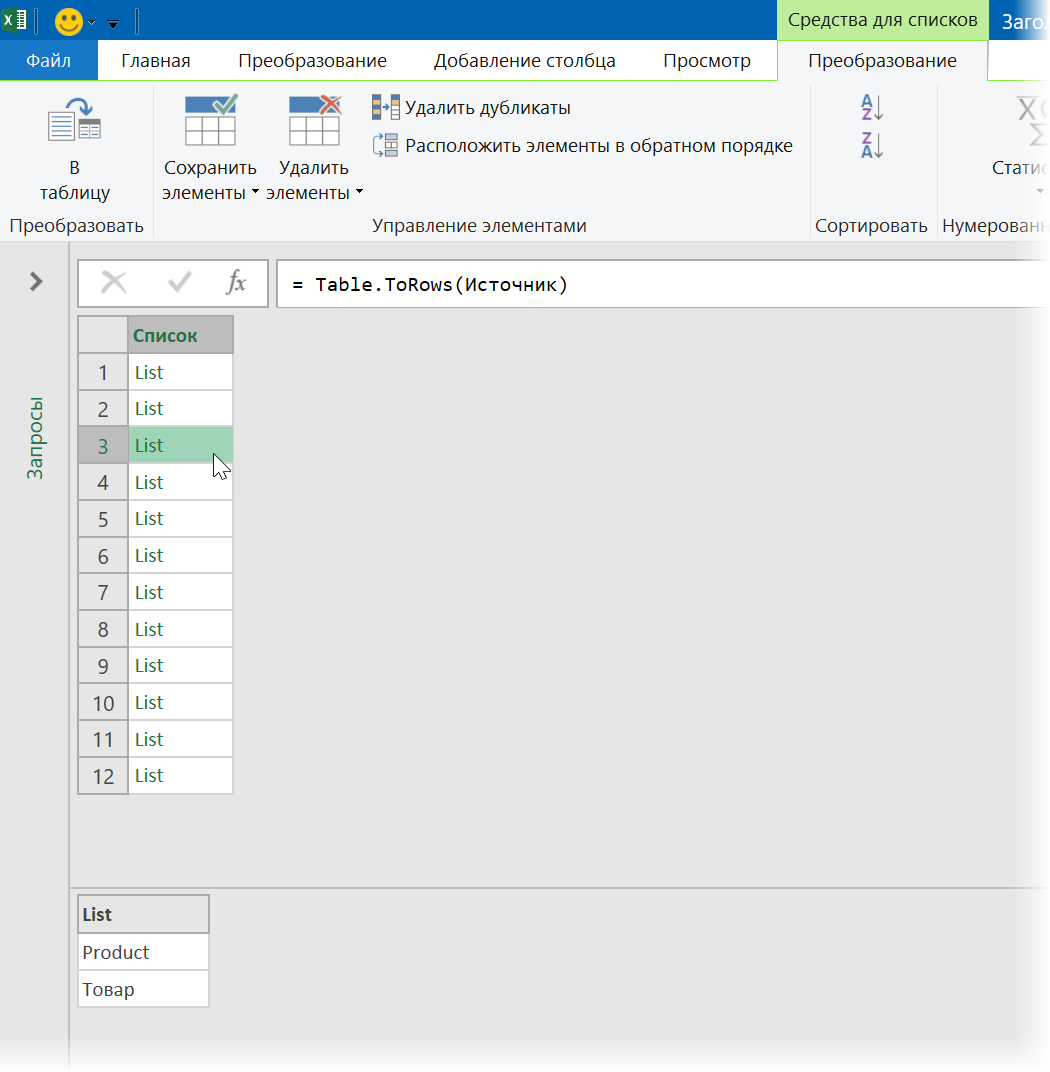


Конвертируем этот список в динамическую "умную" таблицу кнопкой Форматировать как таблицу на вкладке **Главная (Home - Format as Table)** или сочетанием клавиш Ctrl+T и загружаем в Power Query командой **Данные - Из таблицы/диапазона (Data - From Table/Range)**. В последних версиях Excel её переименовали в **С листа (From sheet)**.

В окне редактора запросов Power Query традиционно удаляем шаг **Измененный тип (Changed Type)** и добавляем вместо него новый шаг, нажав на кнопку ***fx***в строке формул (если её не видно, то включить её можно на вкладке **Просмотр**) и вводим туда формулу на встроенном в Power Query языке М:

=Table.ToRows(Источник)

Эта команда преобразует загруженную на предыдущем шаге **Источник** таблицу-справочник в список, состоящий из вложенных списков (List), каждый из которых, в свою очередь, представляет собой пару значений *Было-Стало* из одной строки:



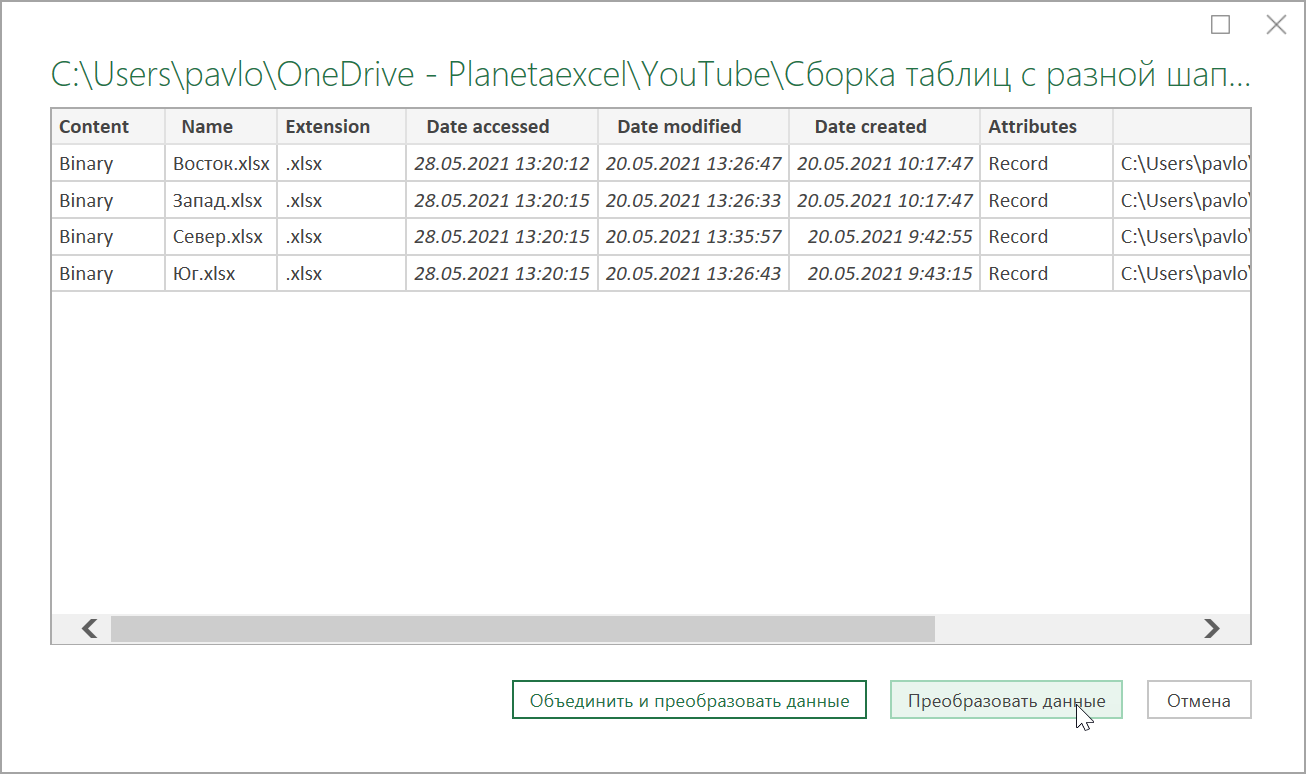
Такой вид данных потребуется нам чуть позже, при массовом переименовании шапок из всех загруженных таблиц.

Закончив преобразование, выбираем команды **Главная - Закрыть и загрузить - Закрыть и загрузить в...** и вид импорта **Только создать подключение (Home - Close&Load - Close&Load to... - Only create connection)** и возвращаемся обратно в Excel.

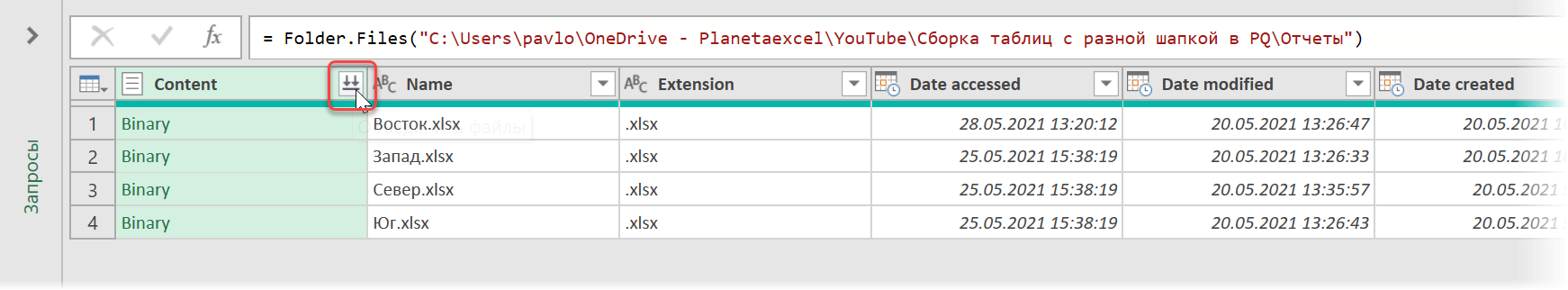
**Шаг 2. Грузим из всех файлов всё как есть**

Теперь давайте загрузим содержимое всех наших файлов из папки - пока в том виде, как есть. Выбираем команды **Данные - Получить данные - Из файла - Из папки (Data - Get Data - From file - From folder)** и затем папку, где лежат наши исходные книги.

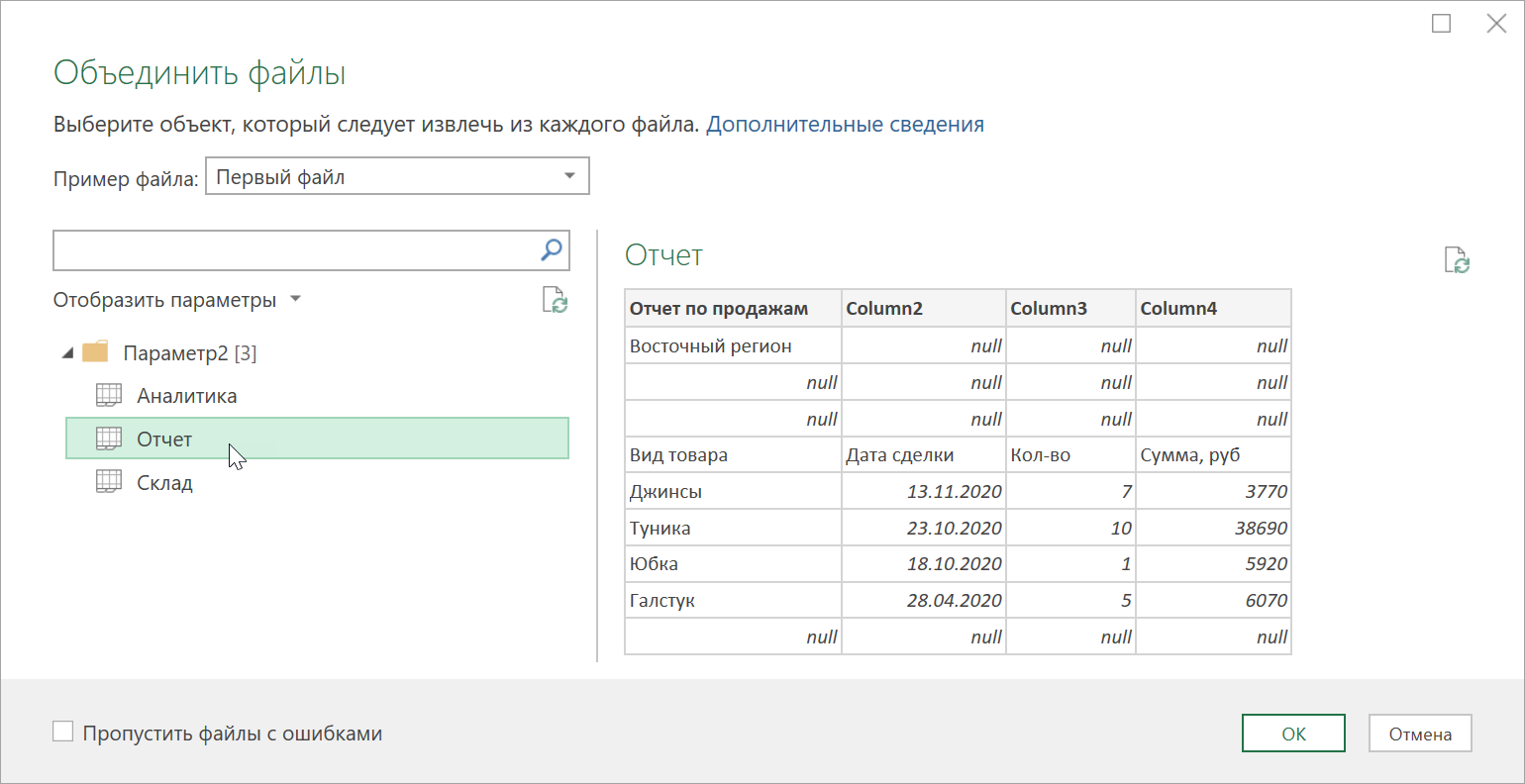
В окне предварительного просмотра жмем **Преобразовать (Transform)** или **Изменить (Edit)**:



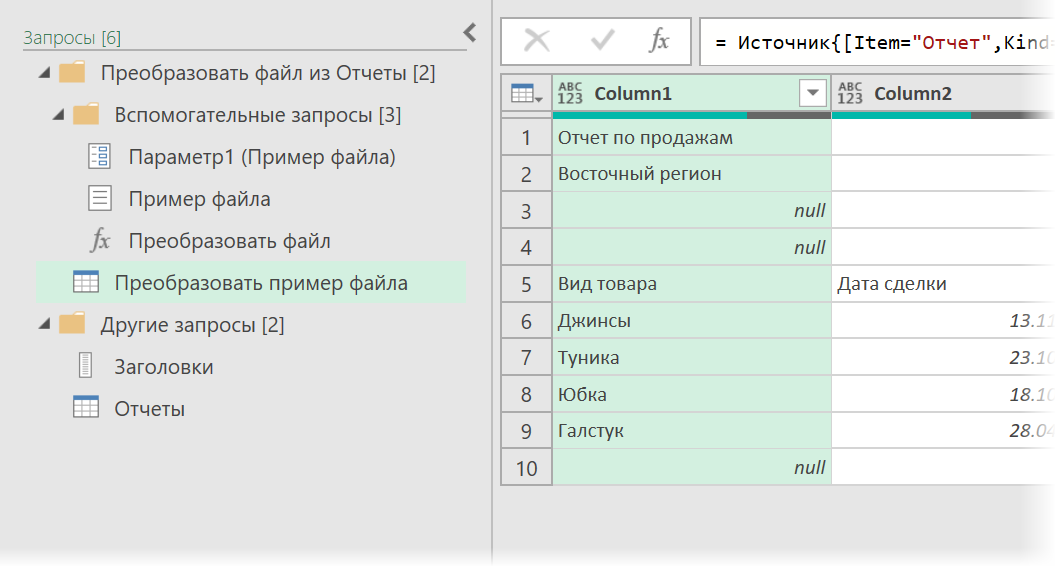
А затем разворачиваем содержимое всех загруженных файлов **(Binary)** кнопкой с двойными стрелками в шапке столбца **Content**:

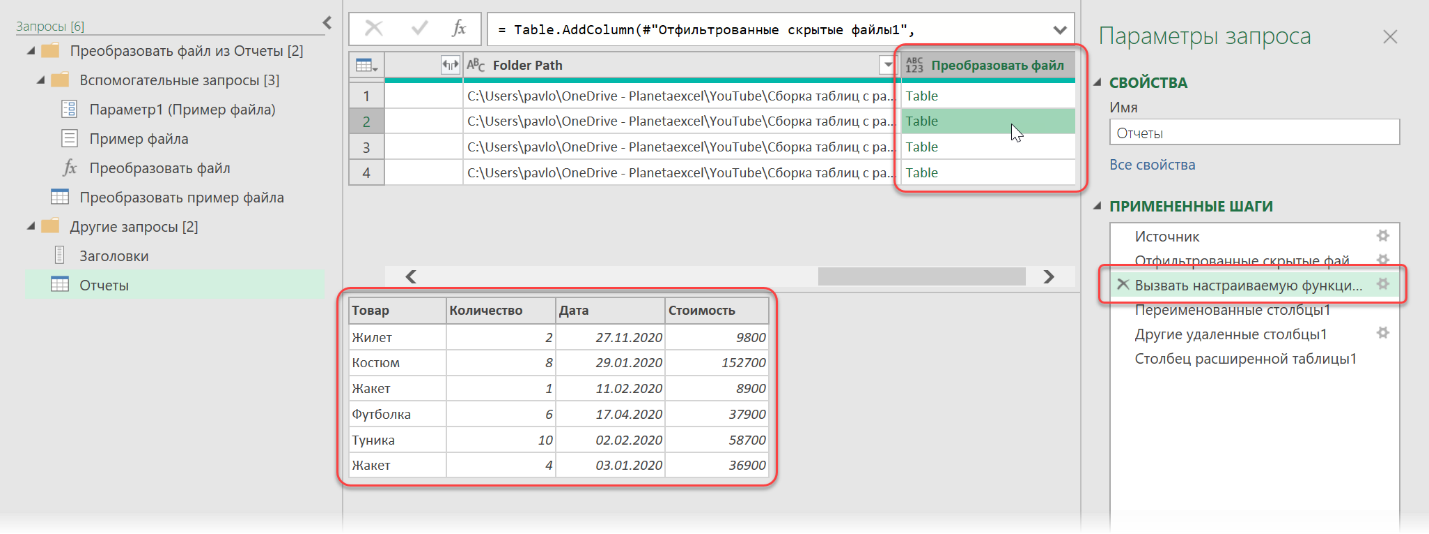
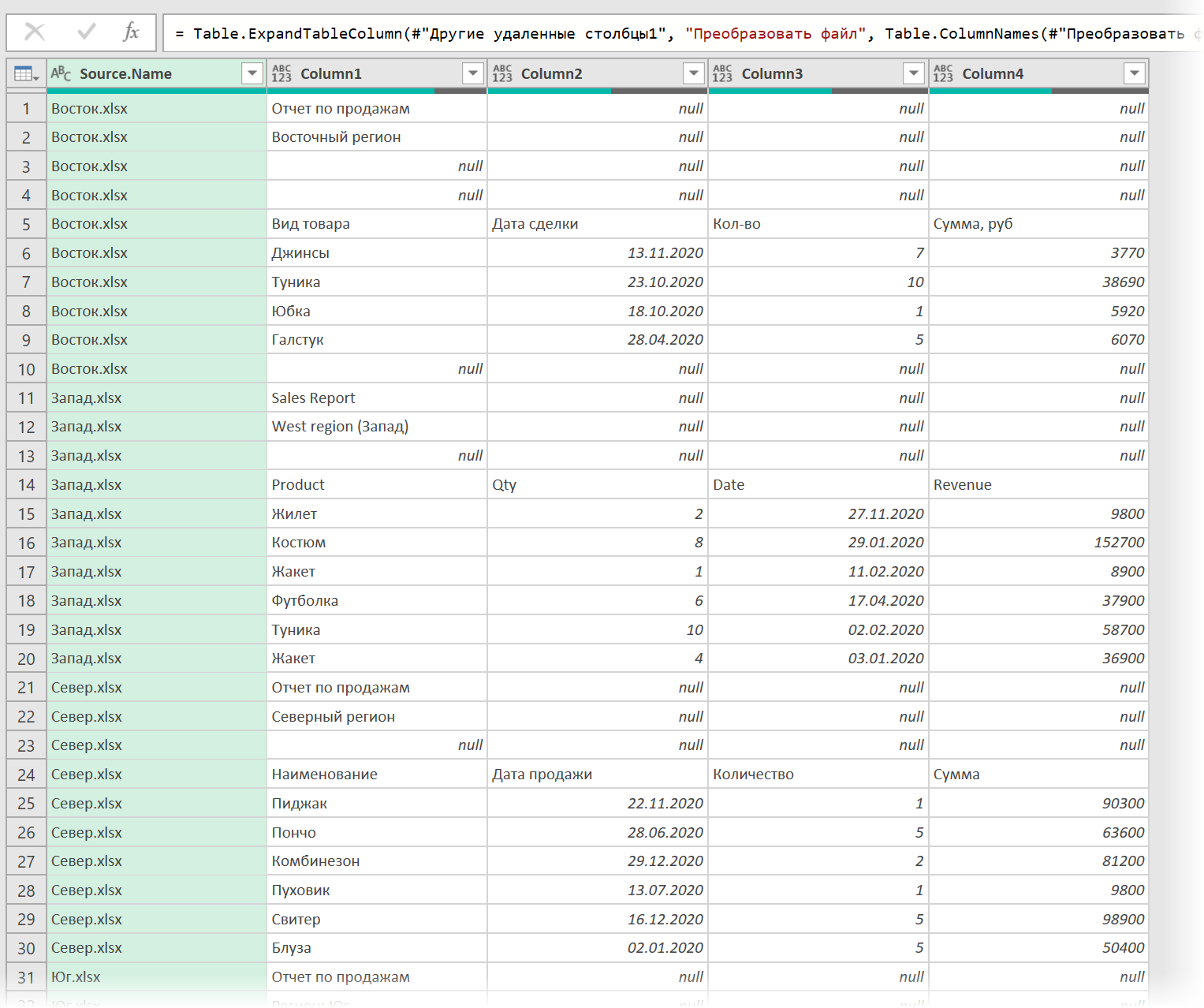


Power Query на примере первого файла (*Восток.xlsx*) спросит нас имя листа, который мы хотим взять из каждой книги - выберем *Отчет*и нажмём ОК:



После этого (на самом деле) произойдет несколько неочевидных для пользователя событий, последствия которых хорошо видно в левой панели:



1. Power Query возьмет первый файл из папки (у нас это будет *Восток.xlsx -*см. **Пример файла**) в качестве примера и импортирует его содержимое, создав запрос **Преобразовать пример файла**. В этом запросе будет несколько простых шагов типа **Источник**(обращение к файлу), **Навигация**(выбор листа) и, возможно, поднятие заголовков. Этот запрос умеет загружать данные только из одного конкретного файла *Восток.xlsx*.
2. На основе этого запроса будет создана связанная с ним функция **Преобразовать файл** (обозначена характерным значком **fx**), где файл-источник будет уже не константой, а переменной величиной - параметром. Таким образом, эта функция может извлекать данные из любой книги, которую мы ей подсунем в качестве аргумента.
3. Функция будет применена по очереди к каждому файлу (Binary) из столбца **Content** - за это отвечает шаг **Вызвать настраиваемую функцию** в нашем запросе, который добавляет к списку файлов столбец **Преобразовать файл** с результатами импорта из каждой книги:  
     
   
4. Лишние столбцы удаляются.
5. Содержимое вложенных таблиц разворачивается (шаг **Столбец расширенной таблицы**) - и мы видим итоговые результаты сборки данных из всех книг:  
     
   

**Шаг 3. Шлифуем**

На предыдущем скриншоте хорошо видно, что прямая сборка "как есть" получилась некачественной:

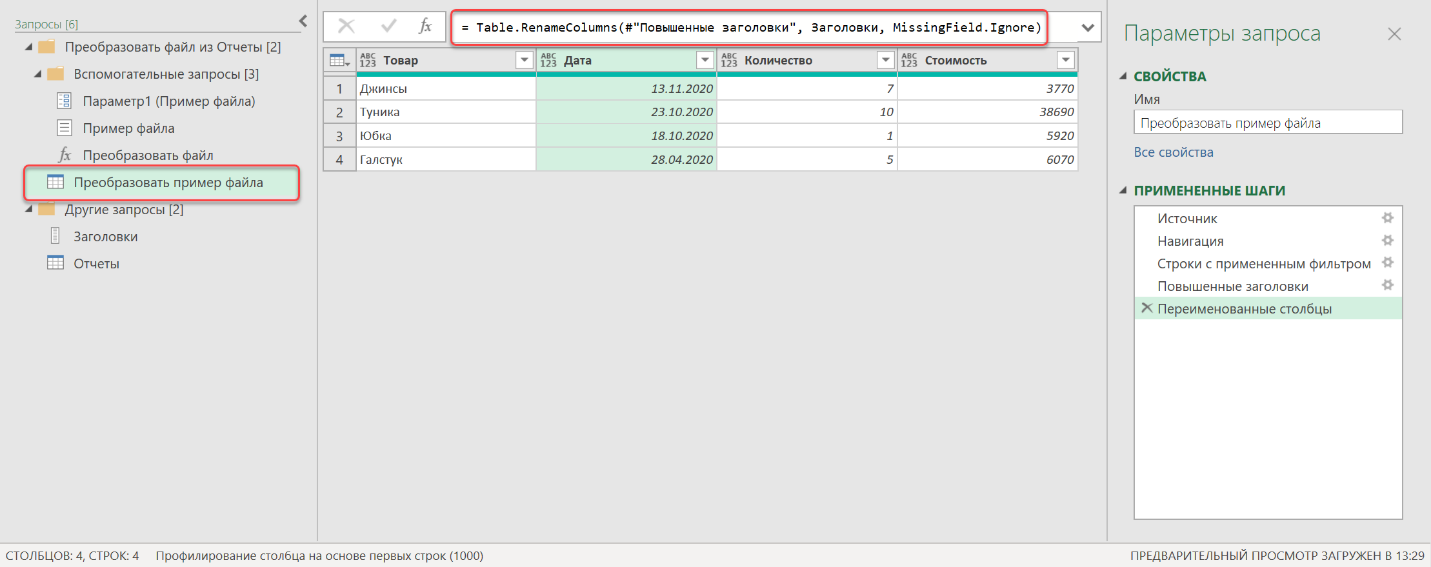
* Столбцы перепутаны.
* Много лишних строк (пустых и не только).
* Шапки таблиц не воспринимаются как шапки и перемешаны с данными.

Исправить все эти проблемы можно очень легко - достаточно подправить запрос Преобразовать пример файла. Все корректировки, которые мы в него внесём, автоматически попадут и в связанную с ним функцию Преобразовать файл, а значит будут использованы затем при импорте данных из каждого файла.

Открыв запрос **Преобразовать пример файла**, добавляем в него шаги по фильтрации ненужных строк (например, по столбцу **Column2**) и поднятие заголовков кнопкой **Использовать первую строку в качестве заголовков (Use first row as headers)**. Таблица станет выглядеть уже гораздо лучше.

Чтобы впоследствии столбцы из разных файлов автоматически встали друг под друга - они должны называться одинаково. Выполнить такое массовое переименование по созданному заранее справочнику, можно одной строкой М-кода. Снова нажмём на кнопку **fx** в строке формул и добавим функцию для переменования:

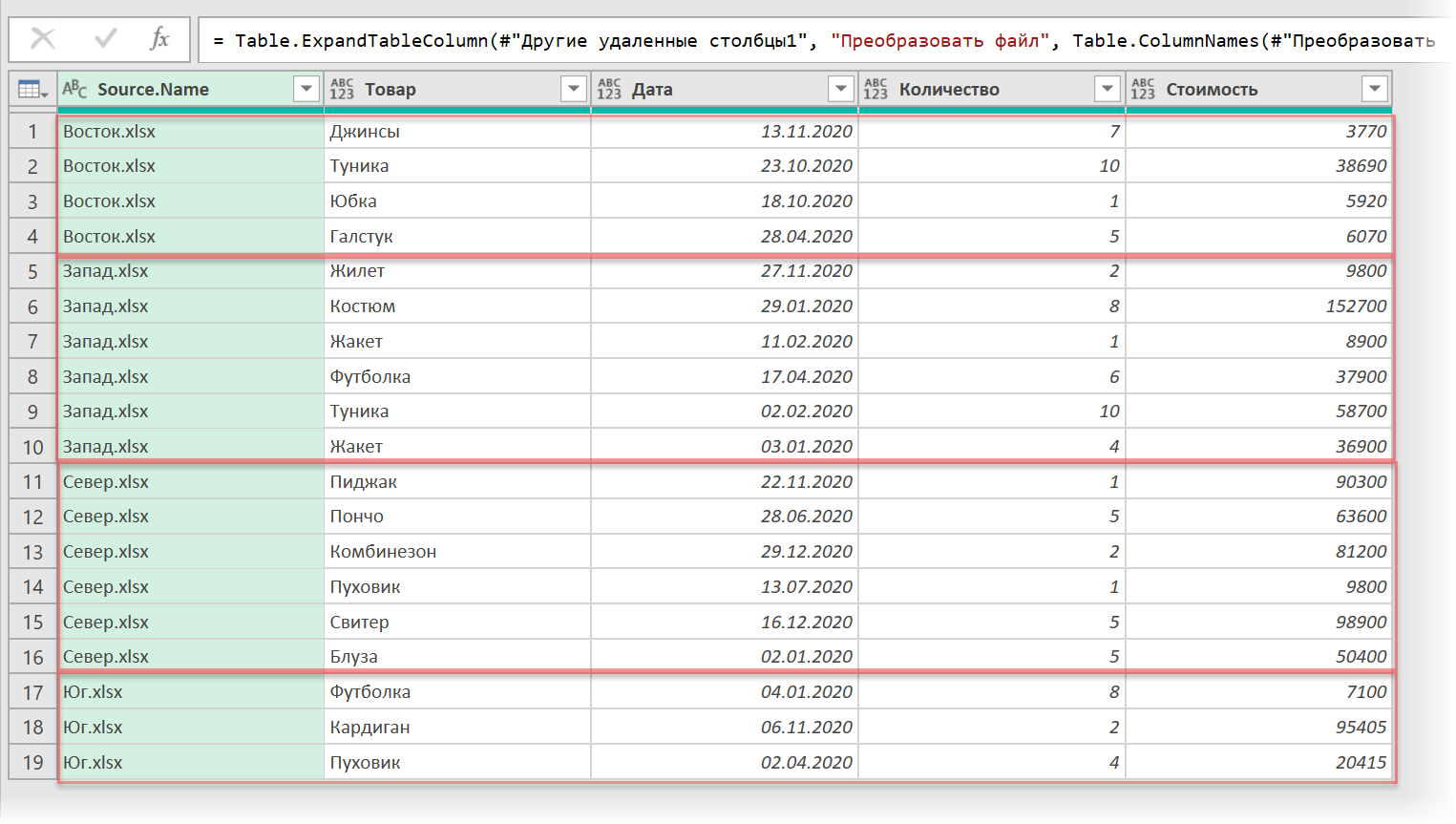
= Table.RenameColumns(#"Повышенные заголовки", Заголовки, MissingField.Ignore)



Эта функция берёт таблицу с предыдущего шага **Повышенные заголовки** и переименовывает в ней все столбцы по вложенному списку-справочнику **Заголовки**. Третий аргумент **MissingField.Ignore** нужен, чтобы на тех заголовках, которые есть в справочнике, но при этом отсутствуют в таблице, не возникала ошибка.

Собственно, всё.

Вернувшись в запрос **Отчеты**мы увидим уже совсем другую картинку - гораздо приятнее предыдущей:



**Ссылки по**