

Глава 25. Организация запросов Power Query

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

[Предыдущая глава](#) [Содержание](#)

При объединении двух таблиц может возникнуть ошибка, если данные поступают из разных источников – *Formula.Firewall*. В примере, который использовался в нескольких последних главах, вы сохраняли Excel-файл со сценарием в таком месте, что исходные файлы хранились в подкаталоге. Что может пойти не так? Изменение области видимости...¹

Давайте поэкспериментируем, чтобы увидеть, как возникает ошибка. Откройте *Query Organization.xlsx*. Пройдите по меню *Данные* → *Получить данные* → *Из файла* → *Из книги*. Перейдите к файлу *Departments.xlsx*. Выберите таблицу *EmployeeDepts*.

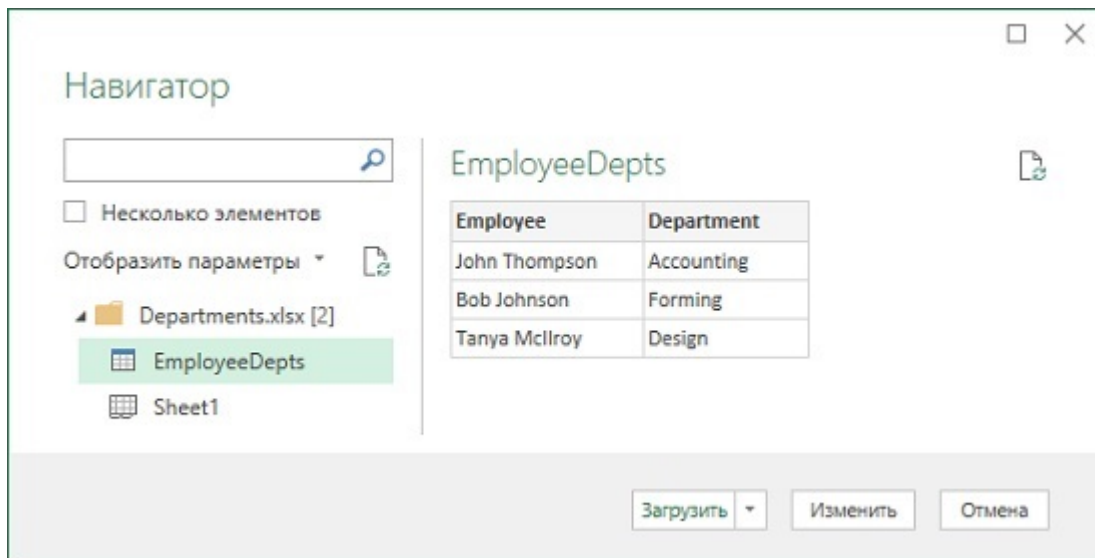


Рис. 25.1. Импорт таблицы *EmployeeDepts*

Поскольку данные выглядят аккуратно, можно без дополнительных изменений загрузить их. Щелкните стрелку на кнопке *Загрузить* → *Загрузить в...* → *Только создать подключение*. В области *Запросы и подключения* появится новый запрос – *EmployeeDepts*:

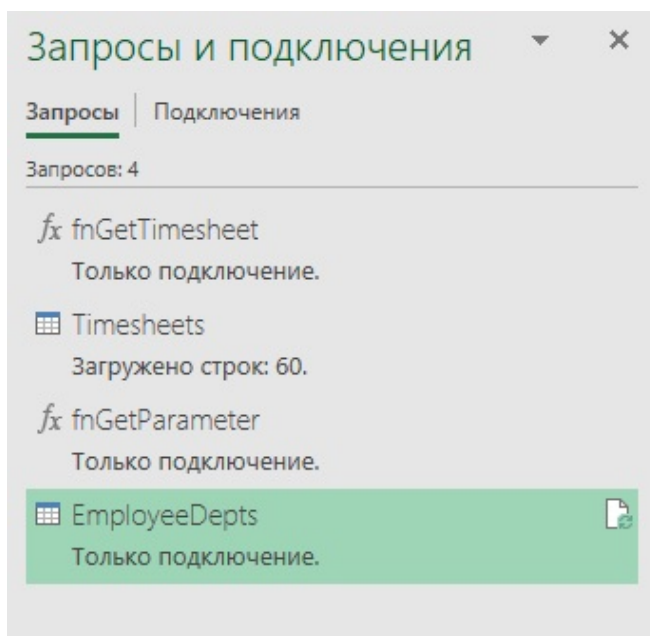


Рис. 25.2. Доступные запросы и функции

¹ Не уверен, что правильно понял авторов. По английский это звучит: A scope change. – Прим. Багузина

Попробуем объединить запросы *Timesheets* и *EmployeeDepts*. Кликните правой кнопкой мыши на запрос *Timesheets* → *Изменить*. В редакторе Power Query пройдите по меню *Главная* → *Объединить запросы*. Выберите в нижней части окна таблицу *EmployeeDepts*. Выберите столбец *Employee* в каждой таблице:

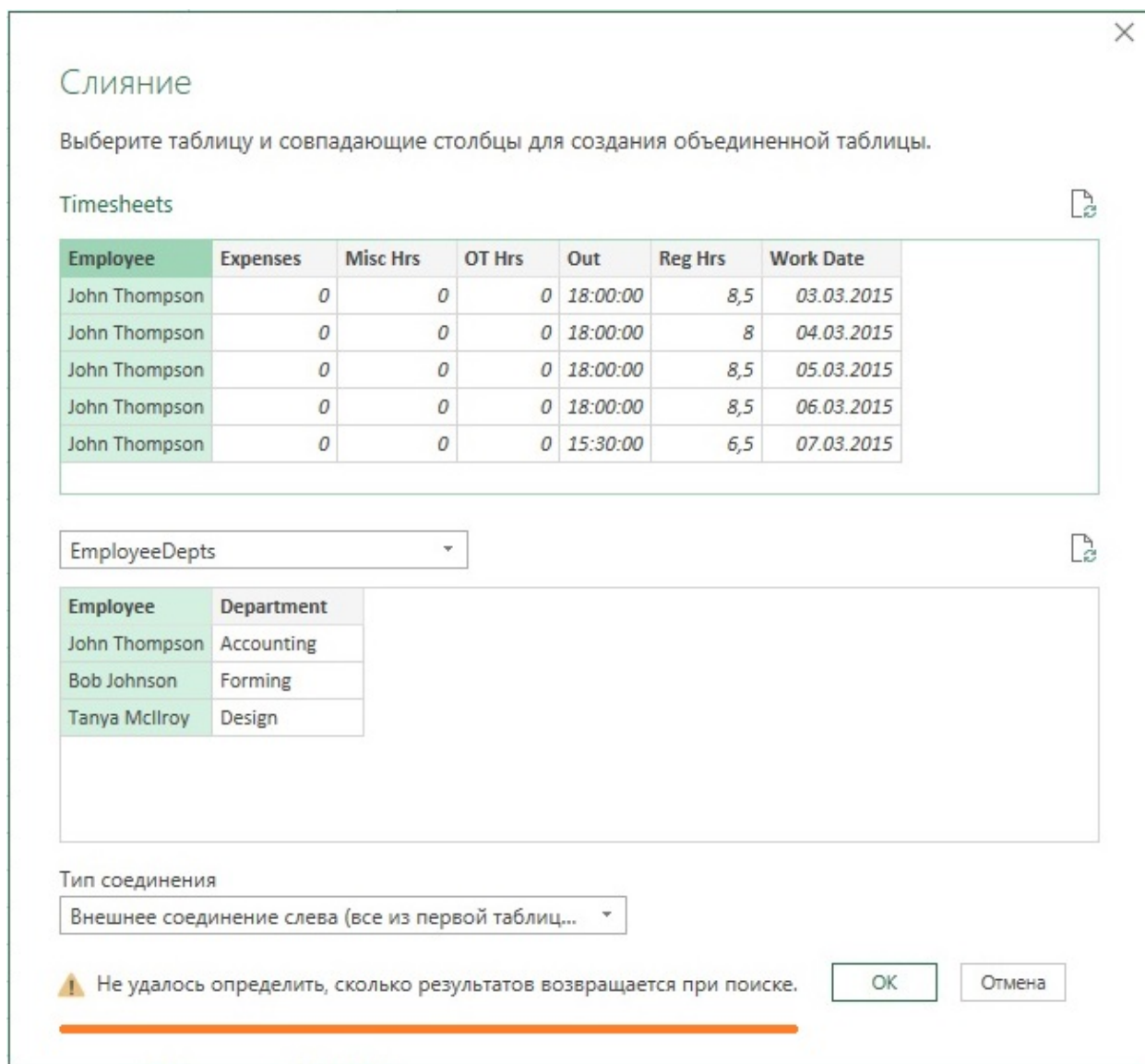


Рис. 25.3. Объединение запросов

Обратите внимание, что Power Query не удалось определить, сколько результатов вернуть. Тем не менее, попробуем нажать Ок.

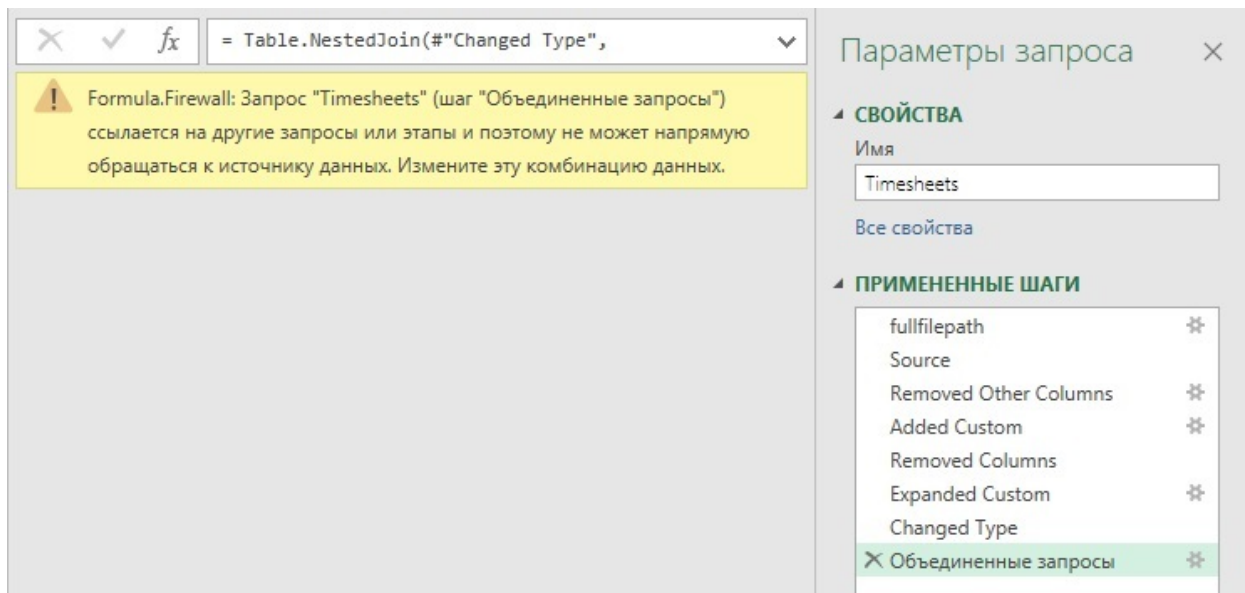


Рис. 25.4. Ошибка *Formula.Firewall*

Это сообщение немного пугает. На самом деле оно означает, что Power Query пытается подключиться к двум различным наборам данных в рамках одного запроса. Даже, если вы работаете с двумя таблицами из одной книги Excel, они рассматриваются как два набора данных, и Power Query это не нравится. В редакторе Power Query пройдите по меню *Файл* → *Отменить и закрыть*.

В Excel в области *Запросы и подключения* кликните правой кнопкой мыши на запрос *Timesheets* → *Загрузить в...* → *Только создать подключение*. Появится предупреждение *Возможна потеря данных*. Кликните *Ок*. Таблица в Excel исчезнет, а у запроса *Timesheets* изменится свойство на *Только подключение*. Повторите попытку соединить запросы *Timesheets* и *EmployeeDepts*. Теперь она увенчается успехом (в объединенный запрос из таблицы *EmployeeDepts* включите только столбец *Department*). Переименуйте запрос в *Consolidated*. Загрузите его на лист Excel.

	A	B	C	D	E	F	G	H
1	Employee	Expenses	Misc Hrs	OT Hrs	Out	Reg Hrs	Work Date	Department
2	John Thompson	0	0	0	18:00:00	8,5	03.03.2015	Accounting
3	John Thompson	0	0	0	18:00:00	8	04.03.2015	Accounting
4	John Thompson	0	0	0	18:00:00	8,5	05.03.2015	Accounting
5	John Thompson	0	0	0	18:00:00	8,5	06.03.2015	Accounting
6	John Thompson	0	0	0	15:30:00	6,5	07.03.2015	Accounting
7	John Thompson	0	0	0	18:00:00	8,5	10.03.2015	Accounting
8	John Thompson	0	0	0	18:00:00	8,5	11.03.2015	Accounting
9	John Thompson	0	0	0	18:00:00	6	12.03.2015	Accounting
10	John Thompson	0	0	0	18:00:00	8,5	13.03.2015	Accounting
11	John Thompson	0	0	0	18:00:00	8,5	14.03.2015	Accounting
12	Bob Johnson	0	0	0	18:00:00	9,5	03.03.2015	Forming
13	Bob Johnson	0	0	0	18:00:00	8	04.03.2015	Forming
14	Bob Johnson	0	0	0	18:00:00	8,5	05.03.2015	Forming

Рис. 25.5. Объединенный запрос

Создание промежуточных / загрузочных запросов

Чтобы не допустить появления ошибки *Formula.Firewall* лучше всего включить в разные запросы два этапа: (1) подключение к базе данных и (2) загрузку в объект назначения (лист Excel или модель Power Pivot). Промежуточные запросы используются для извлечения исходных данных из каждого источника данных. В них вы выполняете как можно больше преобразований, чтобы получить данные в форме, в которой они могут использоваться последующими запросами. Такие запросы имеют статус *Только подключение*. Загрузочные запросы предназначены для объединения данных из одного или нескольких промежуточных запросов и окончательной загрузки.

Объединение любых двух различных источников данных может привести к ошибке Formula.Firewall. Даже попытка ввести параметр из таблицы параметров Excel может вызвать эту ошибку. Если вы получите ошибку Formula.Firewall при попытке вызвать динамический параметр в середине запроса, попробуйте объявить переменную заранее. Это может решить проблему.

Предостережения

При обращении к базе данных необходимо как можно больше работы оставить на промежуточных запросах. Потому что только они используют преимущества сворачивания запросов для повышения производительности (только они имеют связь с базой данных). Использование сворачивания позволит Power Query передать задание обработки в базу данных. Как только вы переходите из промежуточного запроса во вторичный запрос, никакие команды не могут быть переданы обратно в базу данных для сворачивания. Это означает, что сам Power Query должен будет выполнять обработку.

И наоборот, если первичный запрос извлекает данные из источников, которые не поддерживают сворачивание (например, текстовые файлы или файлы Excel), рекомендуется создать один максимально простой промежуточный запрос, только чтобы извлечь данные. Затем вы можете создать множество вторичных запросов, чтобы изменять данные по мере необходимости. Поскольку нет возможности воспользоваться более эффективным механизмом обработки, вы хотите ввести данные один раз и далее стремиться к тому, чтобы каждый шаг работал с как можно меньшим количеством записей.

Не вводите много уровней промежуточных запросов. Сохранение понятного потока запросов поможет вам при отладке решения. Следует также учитывать, что Power BI не поддерживает решения с использованием промежуточных и загрузочных запросов. Если вы собираетесь перенести свое решение в Power BI, используйте один запрос.

Управление запросами

По мере создания все большего числа запросов вы обнаружите, что в области *Запросы и подключения* книги Excel нарастает хаос. Есть несколько возможностей навести порядок.

Новые запросы Power Query добавляет в список по мере их возникновения. Но можно изменить порядок следования запросов (одновременно, и в области *Запросы и подключения* книги Excel, и в навигаторе в редакторе Power Query). Щелкните запрос правой кнопкой мыши → выберите *Вверх* или *Вниз*.

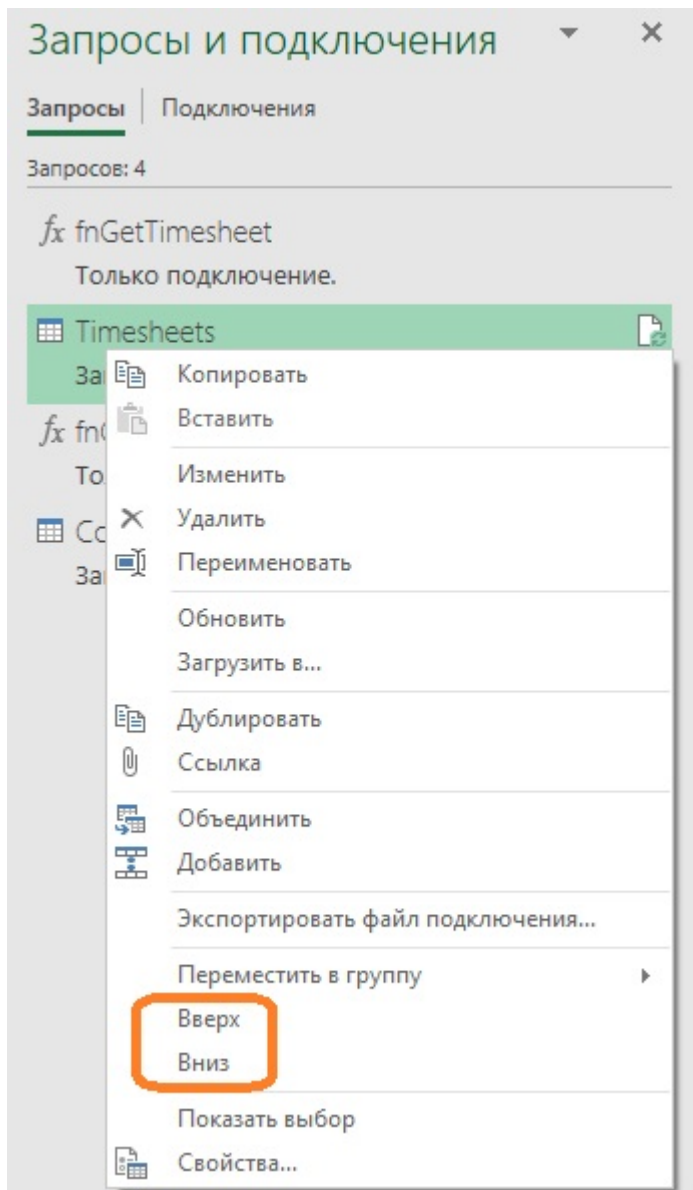


Рис. 25.6. Перемещение запроса

Группировка запросов

Особенно при использовании промежуточных и загрузочных запросов полезно создавать группы для хранения запросов. Щелкните правой кнопкой мыши запрос *fnGetTimesheet* → *Переместить в группу* → *Создать группу...* Назовите новую группу *Functions*:

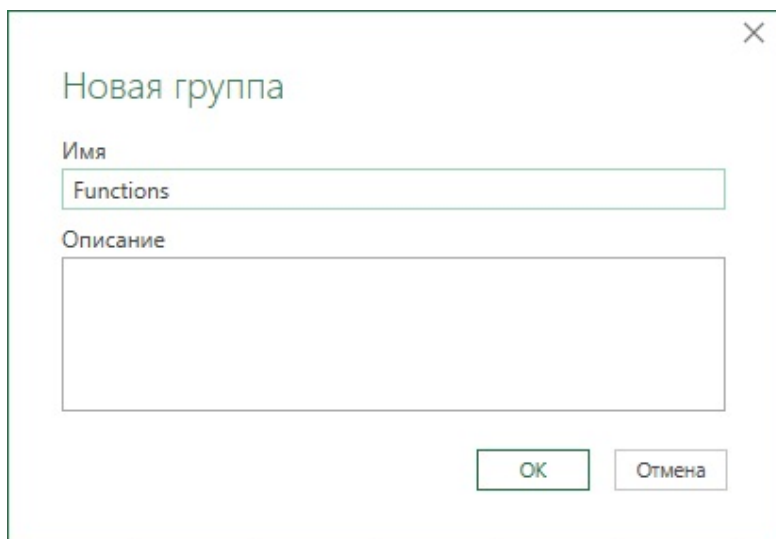


Рис. 25.7. Новая группа для функций

Переместите функции *fnGetParameter* и *fnGetParameter* в группу *Functions*. Создайте группу *Staging* (этапы), и переместите в нее промежуточные запросы *Timesheets* и *EmployeeDepts*. Создайте группу *Load* (загрузка), и переместите в нее запрос *Consolidated*.

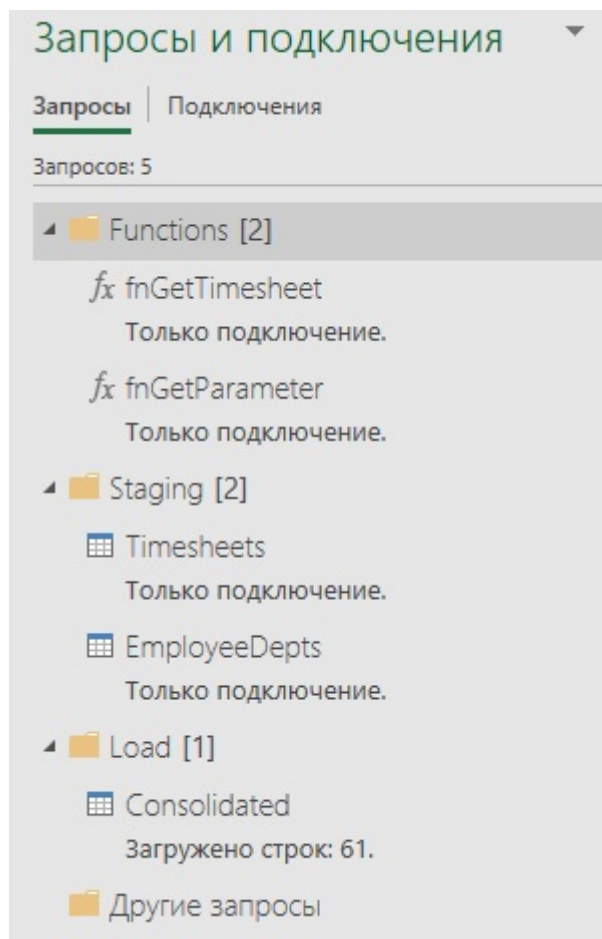


Рис. 25.8. Все запросы аккуратно организованы

Параллельно изменяется организация запросов в навигаторе редактора Power Query:

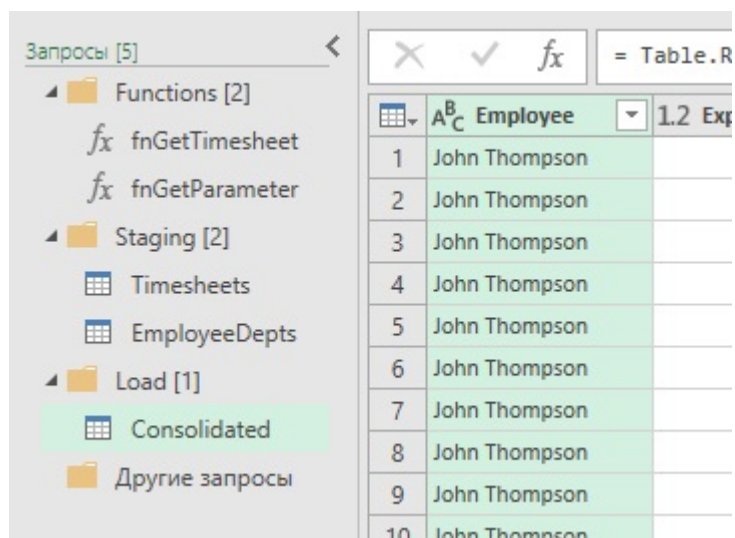


Рис. 25.9. Организация запросов в редакторе Power Query в области навигатора