

## Глава 15. Транспонирование и сложные случаи обратного разбора сводных таблиц 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](#)

Обратный разбор сводной таблицы в Power Query был рассмотрен в [главе 6](#). Однако, могут встречаться и более сложные ситуации, например, сводные таблицы с подкатегориями. Также иногда нужно преобразовать данные, расположенные горизонтально, в обычные вертикальные таблицы. Обе эти функции, к счастью, встроены в пользовательский интерфейс Power Query. Допустим, у вас есть красивая сводная таблица, которую нужно разобрать для последующего анализа:

	April				Account	Month	Measure	Value
	Actual	Budget	Variance					
Alcohol	15,747.28	25,012.00	(9,264.72)	?	Alcohol	April	Actual	15,747.28
Food	34,523.71	44,740.00	(10,216.29)		Food	April	Actual	34,523.71
Other	237.79	3,867.00	(3,629.21)		Other	April	Actual	237.79
					Alcohol	April	Budget	25,012.00
					Food	April	Budget	44,740.00
	50,508.78	73,619.00	(23,110.22)		Other	April	Budget	3,867.00

Рис. 15.1. Сводная таблица с подкатегориями требует разбора

Что делает эту проблему более сложной, чем стандартная процедура разбора сводной таблицы? Это – дополнительный уровень агрегирования: не только месяц, но и тип измерения (план, факт, расхождение). Когда вы разбираете сводную таблицу с одним уровнем агрегирования (например, по месяцам), вы используете имена столбцов в качестве данных, которые поместите в каждую соответствующую строку. А какие имена столбцов использовать здесь? Если вы используете вторую строку в качестве данных, вы потеряете важную информацию о месяце.

### Загрузка данных

Откройте файл UnPivot Sub Categories.xlsx:

	A	B	C	D	E	F	G	H
1			April			May		
2			Actual	Budget	Variance (\$)	Actual	Budget	Variance (\$)
3	Revenues							
4		Alcoholic Beverages	15 747,28	25 012,00	(9 264,72)	22 133,53	41 881,00	(19 747,47)
5		Food & Non-Alc Beverages	34 523,71	44 740,00	(10 216,29)	51 007,02	71 125,00	(20 117,98)
6		Incidental Revenue	237,79	3 867,00	(3 629,21)	186,59	4 334,00	(4 147,41)
7		Total Revenues	50 508,78	73 619,00	(23 110,22)	73 327,14	117 340,00	(44 012,86)
8								
9	Expenses							
10		Cost of Sales	21 977,81	24 296,00	2 318,19	23 442,63	40 284,00	16 841,37
11		Labour & Benefits	35 166,60	45 327,25	10 160,65	34 266,40	52 415,50	18 149,10
12		Other Operational	10 830,99	13 210,00	2 379,01	9 078,01	14 447,00	5 368,99
13		Total Expenses	67 975,40	82 833,25	14 857,85	66 787,04	107 146,50	40 359,46
14		Net Income (Loss)	(17 466,62)	(9 214,25)	(8 252,37)	6 540,10	10 193,50	(3 653,40)

Рис. 15.2. Исходный набор данных для разбора

Поскольку вы не хотите блокировать заголовки столбцов при импорте *Таблицы* в Power Query, вы можете использовать именованный диапазон. В Excel выберите диапазон A1:H14. Перейдите на вкладку *Формулы* → *Задать имя*. В окне *Создание имени* введите *Statement* в поле *Имя*, нажмите *Ok*.

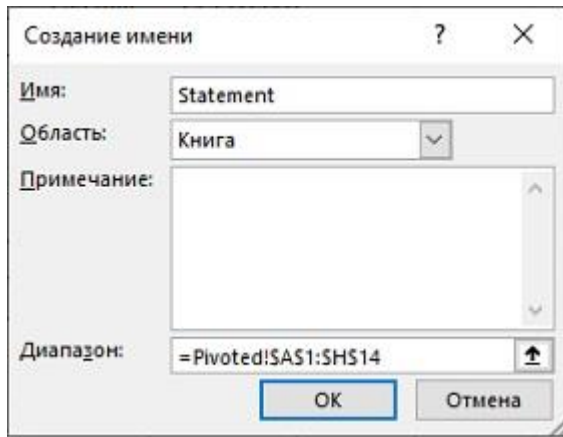


Рис. 15.3. Присвоение имени диапазону

Операцию присвоения имени можно выполнить проще, правда, не так явно. Выделите диапазон A1:H14, и введите *Statement* в область *Имя* в строке формул.

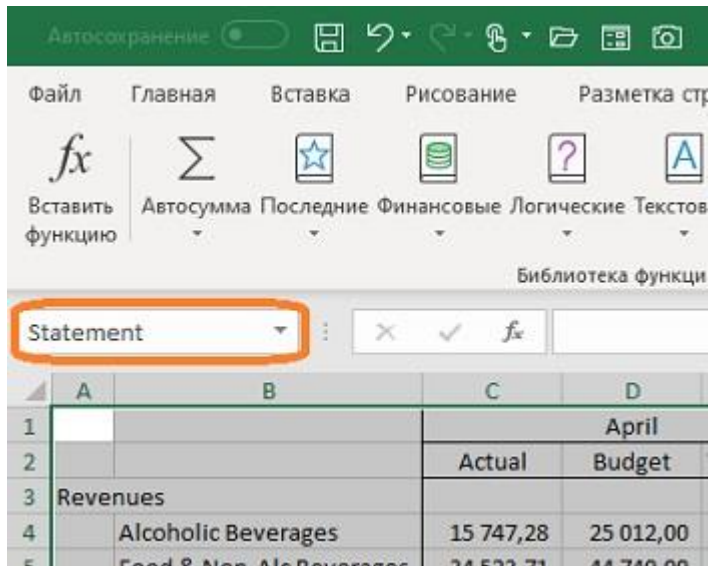


Рис. 15.4. Присвоение имени диапазону в строке формул

Создайте новый запрос, пройдя по меню *Данные* → *Из таблицы/диапазона*. В редакторе Power Query удалите все шаги, кроме первого – *Источник*.

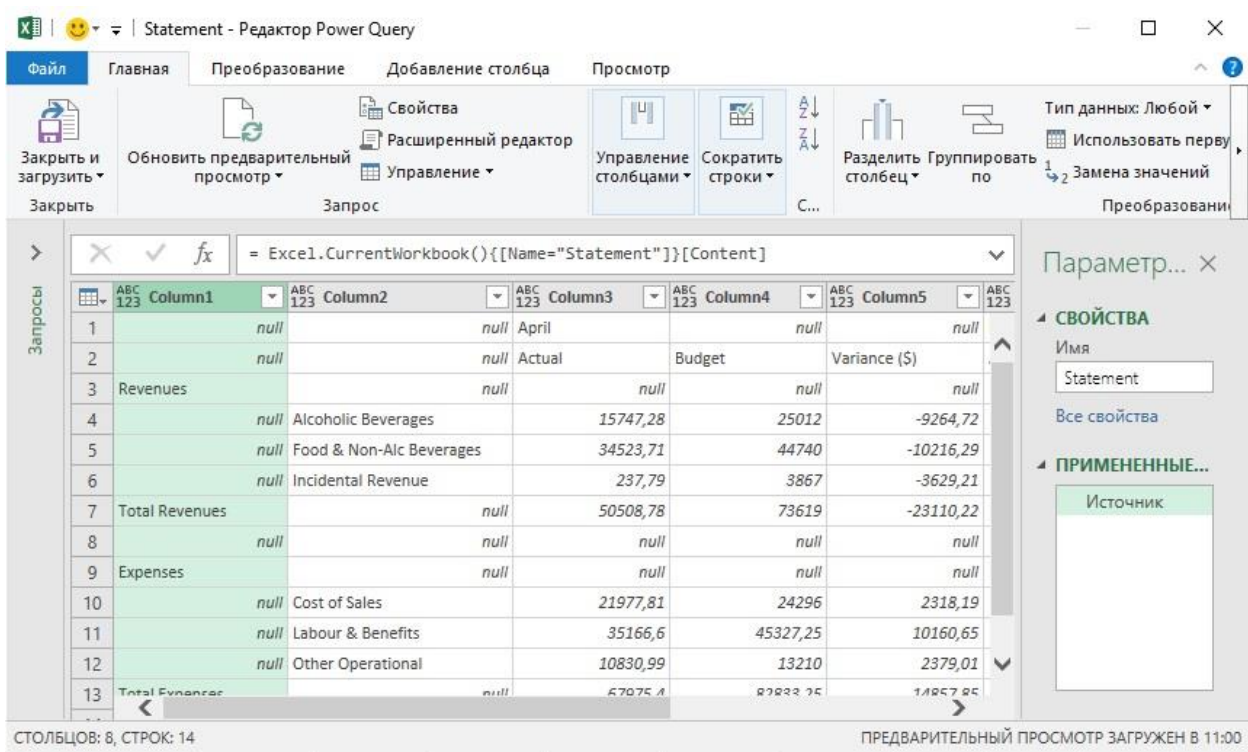


Рис. 15.5. Набор данных изобилует нулевыми значениями

Обратите внимание. Если в таблице Excel ячейка была пустой, значение в таблице Power Query равно *null*. Значение строки 1 столбца *Column4* также равно *null*.

### Заполнение по вертикали

Давайте начнем со столбца *Column1*. Несмотря на то, что столбец содержит много нулевых значений, вы не можете удалить его. Столбец *Column1* содержит информацию о классификации доходов и расходов. Проблема в том, что строки классификации не содержат всей важной информации. Ваша цель дополнить строки информацией о месяце и типе измерения (план, факт, расхождение). Кликните правой кнопкой мыши на столбце *Column1* → *Заполнить* → *Вниз*. Теперь у вас есть правильная классификация в каждой строке:

	ABC 123 Column1	ABC 123 Column2	ABC 123 Co
1	<i>null</i>	<i>null</i>	April
2	<i>null</i>	<i>null</i>	Actual
3	Revenues		<i>null</i>
4	Revenues	Alcoholic Beverages	
5	Revenues	Food & Non-Alc Beverages	
6	Revenues	Incidental Revenue	
7	Total Revenues		<i>null</i>
8	Total Revenues		<i>null</i>
9	Expenses		<i>null</i>
10	Expenses	Cost of Sales	
11	Expenses	Labour & Benefits	
12	Expenses	Other Operational	
13	Total Expenses		<i>null</i>
14	Net Income (Loss)		<i>null</i>

Рис. 15.6. Классификация счетов заполнена для нужных строк

Теперь вам нужно обработать заголовки столбцов. У них похожая проблема: апрель находится в *Column3* и отсутствует в *Column4* и *Column5*. К сожалению, в Power Query нет команды *Заполнить вправо*. Чтобы обойти это ограничение сначала примените...

## ...Транспонирование данных

Перейдите на вкладку к *Преобразование* → *Транспонировать*:

ABC 123	Column1	ABC 123	Column2	ABC 123	Column3	ABC 123	Column4	ABC 123	Column5
1		null		null	Revenues		Revenues		Revenues
2		null		null		null	Alcoholic Beverages		Food & Non-Alc B...
3	April		Actual			null	15747,28		34523,
4		null	Budget			null	25012		447
5		null	Variance (\$)			null	-9264,72		-10216,
6	May		Actual			null	22133,53		51007,
7		null	Budget			null	41881		711
8		null	Variance (\$)			null	-19747,47		-20117,

Рис. 15.7. Результат транспонирования таблицы

Кликните правой кнопкой мыши на столбце *Column1* → *Заполнить* → *Вниз*.

Поскольку таблица включает категории и подкатегории вы не можете применить разбор сводной таблицы на основе имен столбцов. Временно объедините *Column1* и *Column2*. Выберите *Column1*, удерживая нажатой клавишу *Ctrl*, выберите *Column2* → *Объединить столбцы*. Установите в качестве пользовательского разделителя вертикальную черту – |. (Символ | привлекателен тем, что он редко встречается в обычном наборе данных. Это означает, что последующее разделение на основе этого символа вряд ли нарушит данные.) Теперь у вас есть один столбец меток, которые можно использовать для заголовков:

ABC 123	Сведено	ABC 123	Column3	ABC 123	Column4	ABC 123	Column5	ABC 123	Column6
1			Revenues		Revenues		Revenues		Revenues
2			null		Alcoholic Beverages		Food & Non-Alc Be...		Incidental Revenue
3	April Actual		null		15747,28		34523,71		237,79
4	April Budget		null		25012		44740		3867
5	April Variance (\$)		null		-9264,72		-10216,29		-3629,21
6	May Actual		null		22133,53		51007,02		186,59
7	May Budget		null		41881		71125		4334
8	May Variance (\$)		null		-19747,47		-20117,98		-4147,41

Рис. 15.8. Объединение категорий и подкатегорий в один столбец

Переверните данные обратно и повысьте заголовки. Перейдите на вкладку к *Преобразование* → *Транспонировать*. *Главная* → *Использовать первую строку в качестве заголовков*. Переименуйте столбец | → *Class*, а столбец |\_1 → *Account*.

`= Table.RenameColumns("#Измененный тип",{{"|", "Class"}, {"|_1", "Account"}}`

ABC 123	Class	ABC 123	Account	1.2	April Actual	1.2	April Budget
1	Revenues		null		null		null
2	Revenues		Alcoholic Beverages		15747,28		25012
3	Revenues		Food & Non-Alc Beverages		34523,71		44740
4	Revenues		Incidental Revenue		237,79		3867
5	Total Revenues		null		50508,78		73619
6	Total Revenues		null		null		null
7	Expenses		null		null		null
8	Expenses		Cost of Sales		21977,81		24296
9	Expenses		Labour & Benefits		35166,6		45327,25
10	Expenses		Other Operational		10830,99		13210
11	Total Expenses		null		67975,4		82833,25
12	Net Income (Loss)		null		-17466,62		-9214,25

Рис. 15.9. Цель уже близка!

Дальнейшие процедуры стандартны для сценария разбора сводной таблицы. Отфильтруйте столбец *Account* и удалите нулевые значения. Выберите столбцы *Class* и *Account*. Щелкните правой кнопкой мыши один из выбранных заголовков → Отменить свертывание других столбцов:

	A <sup>B</sup> C Class	A <sup>B</sup> C Account	A <sup>B</sup> C Атрибут	1.2 Значение
1	Revenues	Alcoholic Beverages	April   Actual	15747,28
2	Revenues	Alcoholic Beverages	April   Budget	25012
3	Revenues	Alcoholic Beverages	April   Variance (\$)	-9264,72
4	Revenues	Alcoholic Beverages	May   Actual	22133,53
5	Revenues	Alcoholic Beverages	May   Budget	41881
6	Revenues	Alcoholic Beverages	May   Variance (\$)	-19747,47
7	Revenues	Food & Non-Alc Beverages	April   Actual	34523,71
8	Revenues	Food & Non-Alc Beverages	April   Budget	44740

Рис. 15.10. Почти правильная таблица

Осталось разделить столбец *Атрибут* и переименовать столбцы. Щелкните правой кнопкой мыши столбец *Атрибут* → *Разделить столбец* → *По разделителю* → *Пользовательский* → По символу |. Щелкните правой кнопкой мыши столбец *Атрибут.1* → *Переименовать* → *Месяц*. Щелкните правой кнопкой мыши столбец *Атрибут.2* → *Переименовать* → *Мера*. Щелкните правой кнопкой мыши столбец *Значение* → *Переименовать* → *Сумма*. Запрос обработан и готов к загрузке:

	A <sup>B</sup> C Class	A <sup>B</sup> C Account	A <sup>B</sup> C Месяц	A <sup>B</sup> C Мера	1.2 Сумма
1	Revenues	Alcoholic Beverages	April	Actual	15747,28
2	Revenues	Alcoholic Beverages	April	Budget	25012
3	Revenues	Alcoholic Beverages	April	Variance (\$)	-9264,72
4	Revenues	Alcoholic Beverages	May	Actual	22133,53
5	Revenues	Alcoholic Beverages	May	Budget	41881
6	Revenues	Alcoholic Beverages	May	Variance (\$)	-19747,47
7	Revenues	Food & Non-Alc Beverages	April	Actual	34523,71
8	Revenues	Food & Non-Alc Beverages	April	Budget	44740

Рис. 15.11. Сводная таблица с подкатегориями полностью разобрана

Эта техника не ограничивается только двумя уровнями заголовков. Если их больше, объедините все заголовки и подзаголовки в одну текстовую строку, и далее действуйте по описанному выше сценарию.

Если данные полностью разобраны, в дальнейшем легко построить, как исходную сводную таблицу, так и иные срезы:

	April			May		
	Actual	Budget	Variance (\$)	Actual	Budget	Variance (\$)
<b>Expenses</b>						
Cost of Sales	21 978	24 296	2 318	23 443	40 284	16 841
Labour & Benefits	35 167	45 327	10 161	34 266	52 416	18 149
Other Operational	10 831	13 210	2 379	9 078	14 447	5 369
<b>Expenses Итого</b>	<b>67 975</b>	<b>82 833</b>	<b>14 858</b>	<b>66 787</b>	<b>107 147</b>	<b>40 359</b>
<b>Revenues</b>						
Alcoholic Beverages	15 747	25 012	-9 265	22 134	41 881	-19 747
Food & Non-Alc Beverages	34 524	44 740	-10 216	51 007	71 125	-20 118
Incidental Revenue	238	3 867	-3 629	187	4 334	-4 147
<b>Revenues Итого</b>	<b>50 509</b>	<b>73 619</b>	<b>-23 110</b>	<b>73 327</b>	<b>117 340</b>	<b>-44 013</b>
<b>Общий итог</b>	<b>118 484</b>	<b>156 452</b>	<b>-8 252</b>	<b>140 114</b>	<b>224 487</b>	<b>-3 653</b>

Рис. 15.12. Исходные данные, восстановленные с помощью сводной таблицы

## Транспонирование вертикальных наборов данных в строки

Другой сложный сценарий преобразования связан с данными, которые располагаются группами вертикально, а разделяются пустыми строками:

Transactions				
10/14/2014				
Outback Steakhouse				
34.98				
10/16/2014				
Earl's Restaurant				
89.98				

Date	Vendor	Amount
10/14/2014	Outback Steakhouse	34.58
10/16/2014	Earl's Restaurant	89.58

Рис. 15.13. Задача – преобразовать вертикальные наборы в строки

Power Query не имеет стандартной команды для такой обработки, так что потребуется творческий подход и несколько трюков.

Откройте файл Transpose Stacked Tables.xlsx. Выберите область A1:A17. Пройдите по меню *Данные* → *Из таблицы/диапазона*. Добавьте индексный столбец. Перейдите на вкладку *Добавление столбца* → *Столбец индекса* → от 0.

ABC 123 Transactions	1.2 Индекс
26.10.2015 0:00:00	0
City of Vancouver	1
123456	2
34,98	3
null	4
null	5
27.10.2015 0:00:00	6
Earl's Restaurant	7
323511	8
89,98	9
null	10
null	11
28.10.2015 0:00:00	12
Best Buy	13
9272345	14
89,98	15

Рис. 15.14. Каждой строке таблицы присвоен индекс

Данные отдельных транзакций упорядочены: первая строка – дата, вторая – поставщик, третья – ID транзакции, четвертая и пятая – нули. Шестая строка соответствует первой строке второй транзакции. Добавим еще один столбец, возвращающий остаток от деления на шесть.

Выберите столбец *Индекс* → Перейдите на вкладку *Добавление столбца* → *Из числа* → *Стандартный* → *Остаток от деления* → 6. Теперь у вас есть список номеров строк по каждой транзакции:

	ABC 123 Transactions	1.2 Индекс	1.2 Остаток от деления
1	26.10.2015 0:00:00	0	0
2	City of Vancouver	1	1
3	123456	2	2
4	34,98	3	3
5	null	4	4
6	null	5	5
7	27.10.2015 0:00:00	6	0
8	Earl's Restauarant	7	1
9	323511	8	2
10	89,98	9	3
11	null	10	4
12	null	11	5
13	28.10.2015 0:00:00	12	0
14	Best Buy	13	1
15	9272345	14	2
16	89,98	15	3

Рис. 15.15. Данные одного типа пронумерованы одинаковыми цифрами

### Поворот данных

Если у вас есть список повторяющихся строк, вы можете их свернуть. Сначала это покажется странным, но вы увидите, насколько это красиво. Выберите столбец *Остаток от деления* → *Преобразование* → *Столбец сведения*. Когда вы даете команду *Столбец сведения*, значения в выбранном столбце будут использоваться в качестве заголовков новых столбцов (1 на рис. 15.16). Затем Power Query спрашивает, какой столбец вы хотите использовать для извлечения данных. В нашем примере – столбец *Transactions* (2). Обычно для данных выполняется суммирование значений, но в нашем примере мы не заинтересованы в этом. Нам нужны исходные значения, поэтому кликните на *Расширенные параметры* (3). Установите функцию агрегатного значения = *Не агрегировать* (4). Нажмите *Ок*.

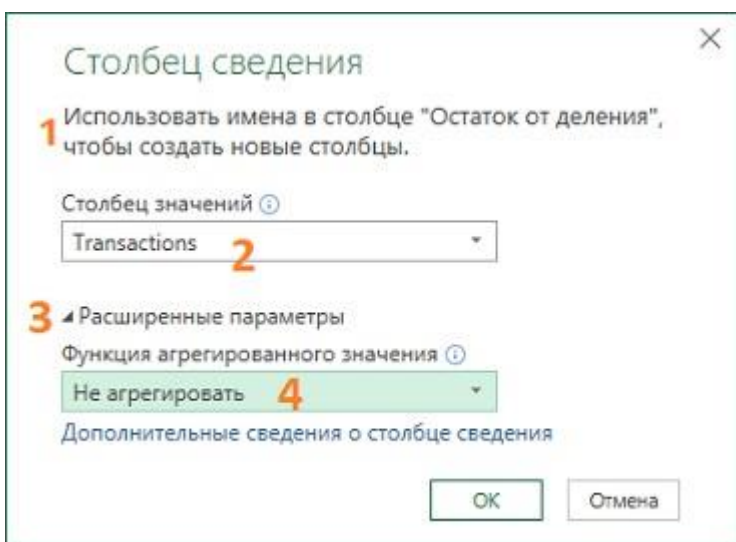


Рис. 15.16. Настройка параметров команды *Столбец сведения*

Таблица меняет форму и пустым ячейкам присваивается значение null

1.2 Индекс	ABC 123 0	ABC 123 1	ABC 123 2	ABC 123 3	ABC 123 4	ABC 123 5
1	0	26.10.2015 ...	null	null	null	null
2	1	null	City of Vanc...	null	null	null
3	2	null	null	123456	null	null
4	3	null	null	null	34,98	null
5	4	null	null	null	null	null
6	5	null	null	null	null	null
7	6	27.10.2015 ...	null	null	null	null
8	7	null	Earl's Restau...	null	null	null
9	8	null	null	323511	null	null
10	9	null	null	null	89,98	null
11	10	null	null	null	null	null
12	11	null	null	null	null	null
13	12	28.10.2015 ...	null	null	null	null
14	13	null	Best Buy	null	null	null
15	14	null	null	9272345	null	null
16	15	null	null	null	89,98	null

Рис. 15.17. Таблица сведения, полная нулевых значений

### Заполнение пробелов

Следующим шагом является заполнение недостающих данных по строкам. Выберите столбец 1 → удерживайте нажатой клавишу Shift → выберите столбец 5. Перейти на вкладку *Преобразование* → *Заполнить* → *Вверх*.

1.2 Индекс	ABC 123 0	ABC 123 1	ABC 123 2	ABC 123 3	ABC 123 4	ABC 123 5
1	0	26.10.2015 ...	City of Vanc...	123456	34,98	null
2	1	null	City of Vanc...	123456	34,98	null
3	2	null	Earl's Restau...	123456	34,98	null
4	3	null	Earl's Restau...	323511	34,98	null
5	4	null	Earl's Restau...	323511	89,98	null
6	5	null	Earl's Restau...	323511	89,98	null
7	6	27.10.2015 ...	Earl's Restau...	323511	89,98	null
8	7	null	Earl's Restau...	323511	89,98	null
9	8	null	Best Buy	323511	89,98	null
10	9	null	Best Buy	9272345	89,98	null
11	10	null	Best Buy	9272345	89,98	null
12	11	null	Best Buy	9272345	89,98	null
13	12	28.10.2015 ...	Best Buy	9272345	89,98	null
14	13	null	Best Buy	9272345	89,98	null
15	14	null	null	9272345	89,98	null
16	15	null	null	null	89,98	null

Рис. 15.18. Данные, заполненные по строкам

Нажмите фильтр столбца 0 → снимите null. Выберите столбцы от 0 до 3 → щелкните правой кнопкой мыши *Удалить другие столбцы*. Переименуйте столбцы 0 → *Дата* и отформатируйте его как дату, 1 → *Поставщик*, 2 → *ID транзакции* и отформатировать его как целое число, 3 → *Сумма* и отформатировать его как десятичное число. Переименуйте запрос *Transactions*.

Дата	ABC 123 Поставщик	123 ID транзакции	1.2 Сумма
26.10.2015	City of Vancouver	123456	34,98
27.10.2015	Earl's Restaurant	323511	89,98
28.10.2015	Best Buy	9272345	89,98

Рис. 15.19. Преобразованный набор данных готов к загрузке на лист Excel