

Практическое применение Power Query в повседневной работе 2/39/22



Старший консультант отдела
налогового управления и
бухгалтерского учёта, PwC Латвия

Artis Vizbelis

Исполняя бухгалтерские или любые другие обязанности, когда обработка данных занимает много времени или требуется регулярно осуществлять повторяющиеся или схожие действия и процессы, важно иметь удобные и простые в использовании инструменты для выполнения задач. Планируя курсы Цифровой академии PwC, мы хотели бы рассказать об одном из инструментов, который существенно облегчит выполнение ваших повседневных задач. *Power Query* – инструмент для подготовки и трансформации данных, доступный в программе *Microsoft Power BI*, а начиная с версии 2016 года – также в *Microsoft Excel* и других программах. Этот инструмент позволяет пользователям извлекать данные как из различных внешних источников, так и из файлов, сохраненных на компьютере. С его помощью можно группировать, трансформировать данные и выполнять другие действия с ними. В данной статье мы приведем некоторые практические примеры, улучшающие представление о применимости и преимуществах такой функции.

Для чего нужно использовать *Power Query*

С помощью этого инструмента удобно обрабатывать и трансформировать данные, а также использовать сгруппированные данные для анализа. Оригинальные данные в источнике не изменяются. Потребуется всего один раз сформировать этапы обработки конкретного источника данных, и они будут сохранены. В следующий раз, выполняя работу, например, с данными следующего месяца, останется лишь нажать *Refresh* – и все ранее выполненные действия будут реплицированы, чтобы пользователю не пришлось повторно обрабатывать данные.

Инструмент имеет удобный интерфейс, который наглядно показывает, какие этапы обработки данных уже выполнены с их источником. Чтобы выполнить этапы трансформации данных, пользователю не требуются навыки программирования, поскольку действия по преобразованию данных можно выбрать из панели инструментов.

Несмотря на то что в *Excel* существует ограничение приблизительно на 1 миллион строк, *Power Query* позволяет использовать источники данных, содержащие 2 и больше миллионов строк. Однако это не означает, что можно наглядно отразить в таблице все строки в подробностях без предварительной обработки данных, но можно сгруппировать данные с помощью *Power Query* и отразить конечный результат в таблицах *Excel*.

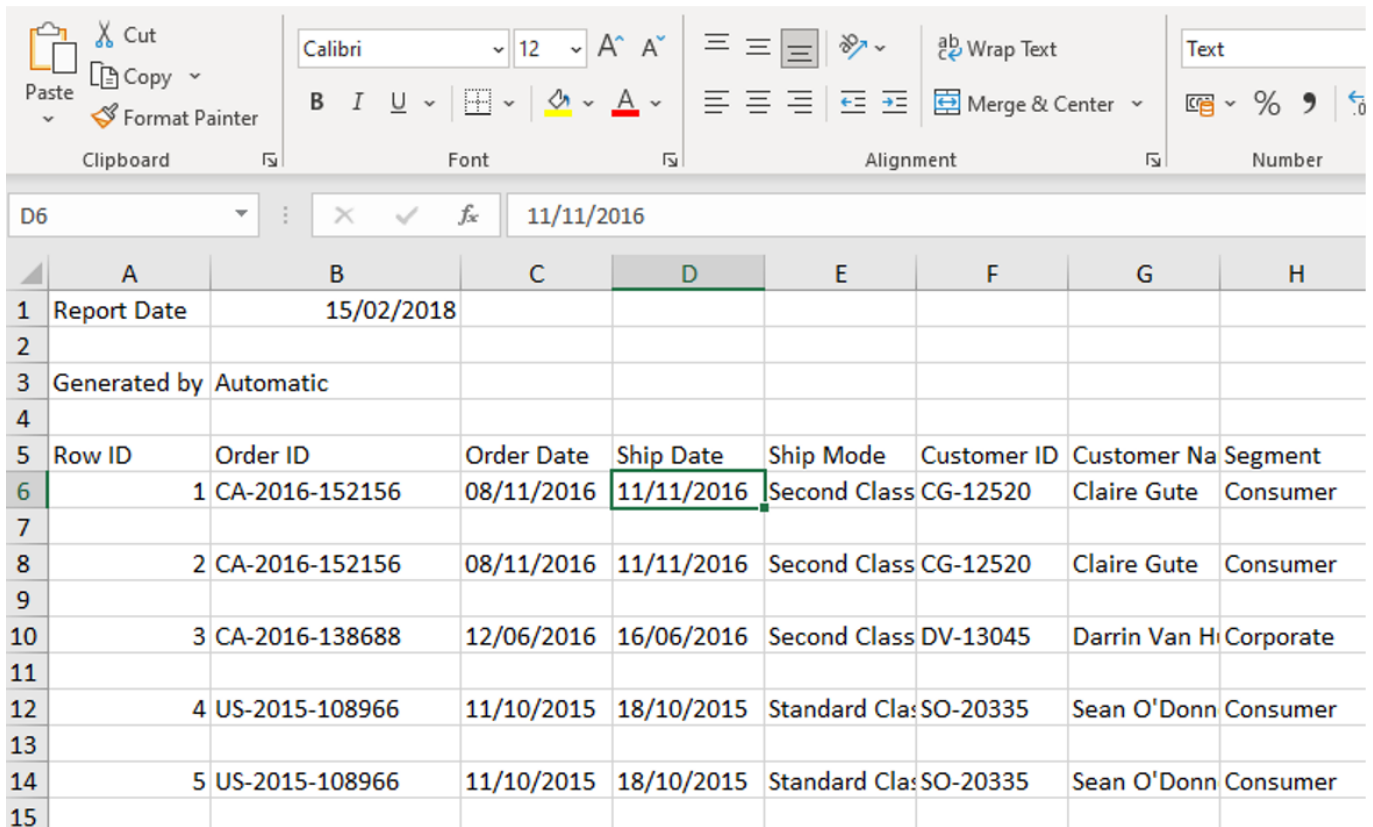
Как найти *Power Query* в программе *Excel*

В панели инструментов *Excel* инструмент *Power Query* находится в разделе *Get & Transform Data* группы *Data*. Выбрав *Get Data*, можно выбрать *Launch Power Query Editor*. Если в том же разделе выбрать загрузку данных, например, с помощью *From Text/CSV*, предлагается меню *Transform Data*, которое тоже откроет пользователю *Power Query*.

Практические примеры применения

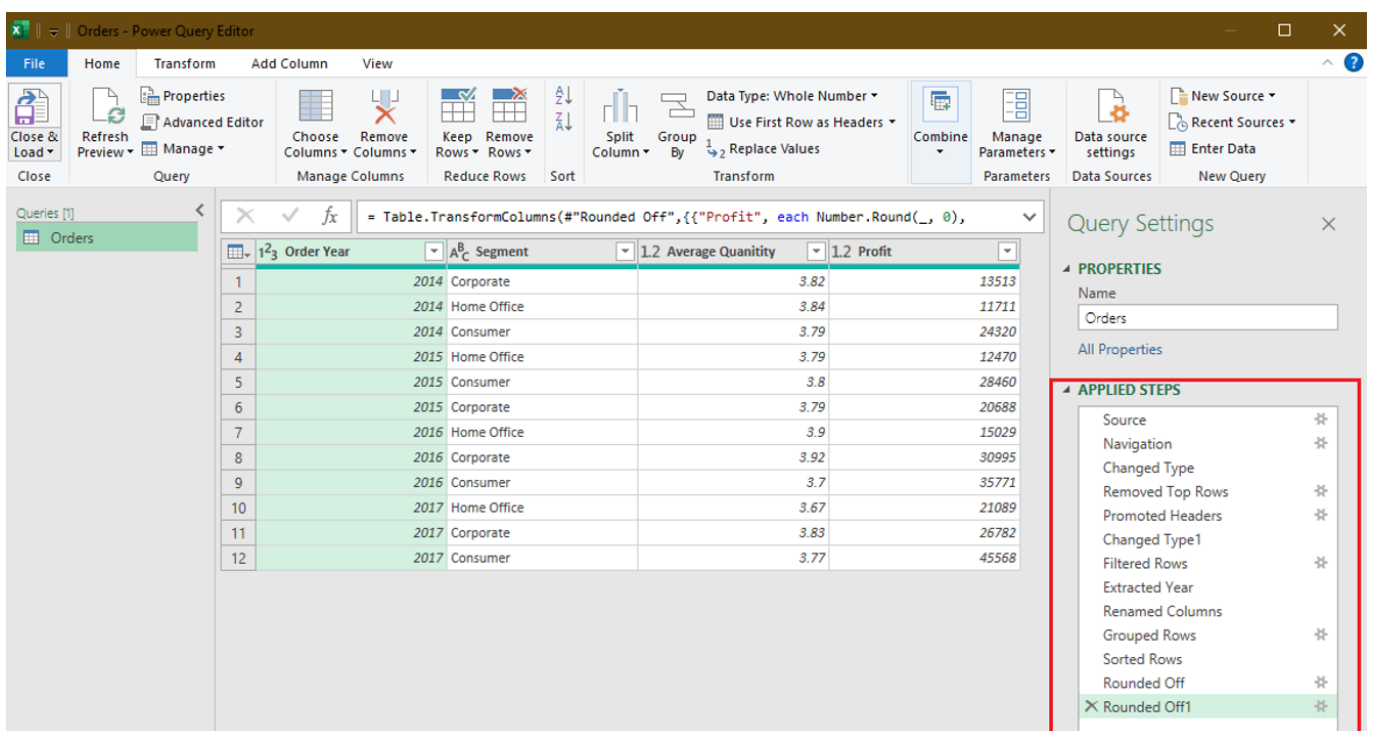
1. Обработка, группировка данных и отражение результата в таблице

В примере используются данные (20 000 строк), из которых первые четыре строки удаляются, а дата форматируется как *Text*, поэтому ее сначала нужно преобразовать, чтобы группировать по датам, и за каждой строкой данных следует одна пустая строка:



Report Date	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment
15/02/2018							
Generated by Automatic							
Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment
1	CA-2016-152156	08/11/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer
2	CA-2016-152156	08/11/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer
3	CA-2016-138688	12/06/2016	16/06/2016	Second Class	DV-13045	Darrin Van H	Corporate
4	US-2015-108966	11/10/2015	18/10/2015	Standard Class	SO-20335	Sean O'Donn	Consumer
5	US-2015-108966	11/10/2015	18/10/2015	Standard Class	SO-20335	Sean O'Donn	Consumer

Чтобы создать заготовку для такой обработки данных с помощью *Power Query*, в новом файле *Excel* выбираем этот файл в качестве источника данных, а затем выполняем необходимые действия для достижения желаемого результата:



Order Year	Segment	Average Quantity	Profit
2014	Corporate	3.82	13513
2014	Home Office	3.84	11711
2014	Consumer	3.79	24320
2015	Home Office	3.79	12470
2015	Consumer	3.8	28460
2015	Corporate	3.79	20688
2016	Home Office	3.9	15029
2016	Corporate	3.92	30995
2016	Consumer	3.7	35771
2017	Home Office	3.67	21089
2017	Corporate	3.83	26782
2017	Consumer	3.77	45568

APPLIED STEPS

- Source
- Navigation
- Changed Type
- Removed Top Rows
- Promoted Headers
- Changed Type1
- Filtered Rows
- Extracted Year
- Renamed Columns
- Grouped Rows
- Sorted Rows
- Rounded Off
- Rounded Off1**

С правой стороны видно, какие трансформации выполнены с исходными данными, например:

- *Removed Top Rows* – удалены первые строки из источника данных;
- *Changed Type* – формат текста изменен на дату;
- *Filtered Rows* – отобраны только заполненные строки;
- *Extracted Year* – из поля даты отражен только год;
- *Grouped Rows* – данные сгруппированы по году/сегменту/среднему количеству/прибыли.

В конечном итоге полученные данные отражаются в таблице *Excel* уже в сгруппированном виде:

	A	B	C	D
1	Order Year	Segment	Average Quantity	Profit
2	2014	Corporate	3.82	13513
3	2014	Consumer	3.79	24320
4	2014	Home Office	3.84	11711
5	2015	Home Office	3.79	12470
6	2015	Corporate	3.79	20688
7	2015	Consumer	3.8	28460
8	2016	Corporate	3.92	30995
9	2016	Home Office	3.9	15029
10	2016	Consumer	3.7	35771
11	2017	Consumer	3.77	45568
12	2017	Home Office	3.67	21089
13	2017	Corporate	3.83	26782
14				

В следующий раз при получении этих данных единственное, что потребуется сделать пользователю, – нажать *Refresh*, и все установленные в *Power Query* действия будут реплицированы автоматически.

2. Курсы валют в таблице *Excel* из внешнего источника данных

Второй пример использования *Power Query* – подключиться к внешнему источнику и загрузить из него в файл *Excel*, например, курсы валют за определенный период. Функциональность *Power Query* также позволяет создать заготовку, где указываем временной период, за который нужно обновить данные. В этом примере, изменив дату в ячейках B1 и B2 и обновив данные в таблице с помощью *Refresh*, мы обновляем курсы валют за выбранный период:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	From Date	01/09/2022												
2	To Date	20/09/2022												
3														
4	Date	AUD	BGN	BRL	CAD	CHF	CNY	CZK	DKK	GBP				
5	01/09/2022	1.4651	1.9558	5.2239	1.3169	0.9802	6.9017	24.488	7.4372	0.86472	7.8511	7.521	399.58	14878
6	02/09/2022	1.4671	1.9558	5.2153	1.3131	0.9839	6.9031	24.481	7.437	0.864				
7	05/09/2022	1.4616	1.9558	5.1407	1.3043	0.9747	6.8768	24.622	7.4364	0.863				
8	06/09/2022	1.4651	1.9558	5.13	1.3029	0.9745	6.9091	24.55	7.4365	0.857				
9	07/09/2022	1.4748	1.9558	5.1881	1.3037	0.975	6.8968	24.631	7.4365	0.86				
10	08/09/2022	1.4824	1.9558	5.2042	1.3134	0.9739	6.9564	24.543	7.4365	0.866				
11	09/09/2022	1.4704	1.9558	5.2087	1.307	0.9657	6.9543	24.536	7.4365	0.86				
12	12/09/2022	1.4749	1.9558	5.1933	1.3194	0.9667	7.0348	24.546	7.4365	0.867				
13	13/09/2022	1.4736	1.9558	5.1764	1.32	0.9669	7.0467	24.551	7.4366	0.867				
14	14/09/2022	1.4873	1.9558	5.1827	1.3177	0.9612	6.955	24.527	7.4366	0.864				
15	15/09/2022	1.4853	1.9558	5.1837	1.3172	0.9572	6.9852	24.518	7.4366	0.869				
16	16/09/2022	1.4894	1.9558	5.2279	1.3226	0.9579	6.9787	24.497	7.4366	0.8				
17	19/09/2022	1.495	1.9558	5.2886	1.3294	0.9658	7.0066	24.494	7.4373	0.877				
18	20/09/2022	1.4893	1.9558	5.2139	1.3268	0.9644	7.003	24.556	7.4368	0.873				
19														
20														
21														

Calibri 11 A A %

B I A

Cut

Copy

Paste Options:

Paste Special...

Smart Lookup

Refresh

Insert >

Delete >

Select >

Clear Contents

3. Извлечение данных из нескольких файлов с одинаковой разметкой

Посредством данного инструмента можно автоматически объединить файлы с одинаковой разметкой из одной папки и загрузить их в одну таблицу, предварительно в случае необходимости выполнив обработку конкретных данных. В этом случае остается лишь выбрать папку, где находятся все файлы, и *Power Query* автоматически создаст действия для объединения данных в одну таблицу и поле с названием файла, из которого помещены данные:

	A	B	C	D
1	Source.Name	Order Date	Profit	
2	Sample1.xls	2016	81,795.17	
3	Sample1.xls	2015	61,618.60	
4	Sample1.xls	2014	49,543.97	
5	Sample1.xls	2017	93,439.27	
6	Sample2.xls	2017	81,795.17	
7	Sample2.xls	2016	61,618.60	
8	Sample2.xls	2015	49,543.97	
9	Sample2.xls	2018	93,439.27	
10	Sample3.xls	2018	81,795.17	
11	Sample3.xls	2017	61,618.60	
12	Sample3.xls	2016	49,543.97	
13	Sample3.xls	2019	93,439.27	
14				

Выводы

В данной статье отражена лишь небольшая часть доступной функциональности. Следует добавить, что таким образом можно создавать заготовки для различных отчетов или для сравнения данных между системами, а также для выполнения других задач. Данный инструмент может заменить ряд типовых действий, осуществляемых на повседневной основе, – составление формул, копирование данных и др.

Если вас заинтересовала возможность применения данного или подобных ему инструментов в вашей конкретной ситуации, предлагаем присылать свои примеры и мы уже в ближайшее время пригласим всех желающих принять участие в *PwC's Digital Academy*, где предоставим решения присланных ситуаций.