Power Query Lab Reference

# Table of Contents

[What is Power Query? 2](#_Toc72449068)

[Getting data 3](#_Toc72449069)

[Transformations 4](#_Toc72449070)

[FactResellerSales 5](#_Toc72449071)

[DimReseller 7](#_Toc72449072)

[DimGeography 8](#_Toc72449073)

[DimEmployee 9](#_Toc72449074)

[DimSalesTerritory 10](#_Toc72449075)

[DimProductCategory 11](#_Toc72449076)

[DimProductSubcategory 12](#_Toc72449077)

[DimProduct 13](#_Toc72449078)

[DimProductSubcategory 15](#_Toc72449079)

[FactResellerSales (2) 16](#_Toc72449080)

[FactInternetSales 17](#_Toc72449081)

[Append1 18](#_Toc72449082)

[UN Population Data – Raw 19](#_Toc72449083)

# What is Power Query?

Power Query is a data transformation and data preparation engine. Power Query comes with a graphical interface for getting data from sources and a Power Query Editor for applying transformations. Because the engine is available in many products and services, the destination where the data will be stored depends on where Power Query was used. Using Power Query, you can perform the extract, transform, and load (ETL) processing of data.

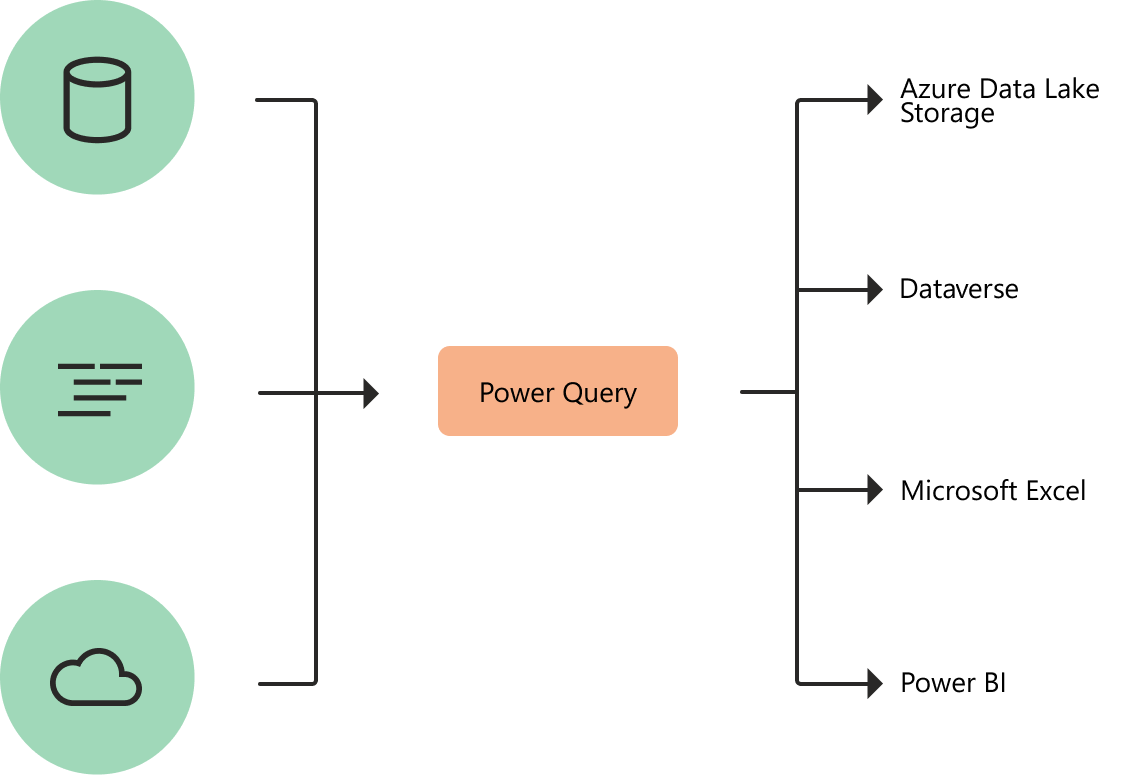


Diagram with symbolized data sources on the left, passing though Power query for transformation, and then going to various destinations, such as Azure Data Lake Storage, Dataverse, Microsoft Excel, or Power BI.

# Getting data

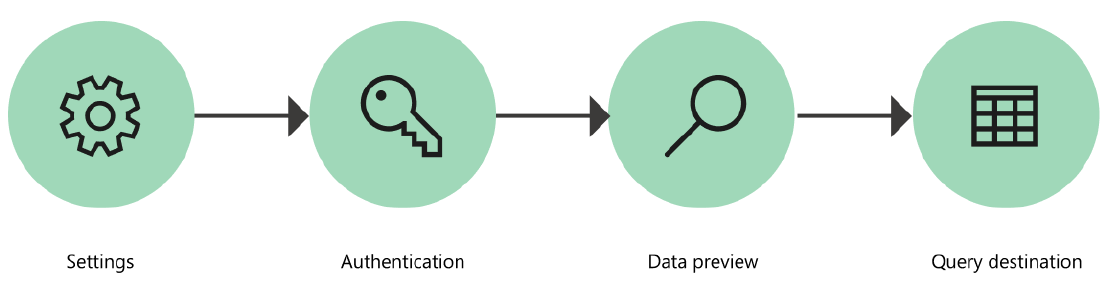
Power Query can connect to many different data sources so you can work with the data you need. This article walks you through the steps for bringing in data from Excel files to Power Query. Connecting to a data source with Power Query follows a standard set of stages before landing the data at a destination. The stages are:

1. Connection settings

2. Authentication (for select sources)

3. Data preview

4. Query destination



To connect to an Excel workbook from Power BI / Power Query:

1. Select the Excel option in the connector selection.
2. Browse for and select the Excel workbook you want to load. Then select Open.

(If the Excel workbook is online, use the Web connector to connect to the workbook.)

1. In Navigator, select the workbook information you want, then either select Load to load the data or Transform Data to continue transforming the data in Power Query Editor.

In the demo report from the training, we have connected with the following Excel files:

1. FactResellerSales
2. DimEmployee
3. DimGeography
4. DimProduct
5. DimProductSubcategory
6. DimProductCategory
7. DimReseller
8. DimSalesTerritory

To connect to an Excel workbook, authentication is not required by Power Query. For details, refer <https://docs.microsoft.com/en-us/power-query/get-data-experience>

# Transformations

Before loading data from a source into the data model, we need to see whether the data is clean and in the right shape so as to lend itself to the intended analysis. If not, we need to carry out the transformation of data in Power Query.

The transformation engine in Power Query includes many prebuilt transformation functions that can be used through the graphical interface of the Power Query Editor. These transformations can be as simple as removing a column or filtering rows, or as common as using the first row as a table header. There are also advanced transformation options such as merge, append, group by, pivot, and unpivot.

All these transformations are made possible by choosing the transformation option in the menu, and then applying the options required for that transformation.

In the dataset used in the training, the transformation steps applied are listed below.

# FactResellerSales

## Source

This step is automatically applied in the background when we identify the source of data.

## Navigation

For sources like Excel workbooks or databases like SQL Server, MySQL, Db2 etc. which can contain multiple tables, the Navigation step is also applied automatically in the background and we do not have to apply it manually as a transformation.

## Promoted Headers

If the first row of the sheet in Excel contains table headings, Promoted Headers transformation is applied automatically. This has happened for all the files mentioned hereinbefore. In cases where there are unwanted rows at the beginning on the sheet, we need to select Remove Top Rows from Home > Remove Rows.

## Changed Type

This step gets applied automatically after Promoted Headers when Power Query scans the values in all the columns and tries to detect and assign data types for each column. In case we delete this step, or need to apply it elsewhere later, we can first select all columns and go to Transform > Detect Data Type.

While changing data types, it is important from an optimization point of view to ensure that columns like all Key columns containing numeric values are assigned whole / decimal number type rather than text.

## Removed Other Columns

To remove unwanted columns, we can select the unwanted columns and go to Home > Remove Columns. However, this step can be tedious to modify later. So, we have selected the columns to be retained from Home > Choose Columns.

We can also get the same result by selecting Home > Remove Columns > Remove Other Columns after selecting the required columns.

## Split Columns

To eliminate the “SO” part from the SalesOrderNumber column we can simply select the column, go to Home > Replace Values, and replace “SO” with blank. However, to explore the Split Columns functionality, we selected the column, and from Transform > Split Column > By Number of Characters (2), split the text and numeric values in separate columns, so as to delete the column containing “SO” later.

## Removed Columns

The column containing the text “SO”, generated after splitting columns above, is to be deleted as it does not serve any purpose. The actual SO number is stored in the next column as a whole number.

## Changed Type

The data type of the numeric component of the SO number split into the next column needs to be changed to whole number.

## Added Custom Column

Go to Add Column > Custom Column and create a new column named “Profit (M)” which is equal to SalesAmount less TotalProductCost.

## Renamed Column

Rename column "SalesOrderNumber.2" to "SalesOrderNumber"

## Changed Type

Change data type of "OrderDate" to type date.

## Added Custom Column

Go to Add Column > Custom Column and create a new column named “FullDateKey” such that,

FullDateKey = Date.ToText([OrderDate], "yyyyMMdd")

## Changed Type

Change data type of " FullDateKey" to type text.

## Added Custom Column

Go to Add Column > Custom Column and create a new column named “Date\_Emp” such that,

Date\_Emp = [FullDateKey]&"\_"&Number.ToText([EmployeeKey],"0")

## Added Custom Column

Go to Add Column > Custom Column and create a new column named “Profit/Loss” such that,

Profit/Loss = if [#"Profit (M)"] >= 100 then "High Profit" else if [#"Profit (M)"] >= 0 then "Profit" else "Loss"

# DimReseller

## Source

This step is automatically applied in the background when we identify the source of data.

## Navigation

For sources like Excel workbooks or databases like SQL Server, MySQL, Db2 etc. which can contain multiple tables, the Navigation step is also applied automatically in the background and we do not have to apply it manually as a transformation.

## Promoted Headers

If the first row of the sheet in Excel contains table headings, Promoted Headers transformation is applied automatically. This has happened for all the files mentioned hereinbefore. In cases where there are unwanted rows at the beginning on the sheet, we need to select Remove Top Rows from Home > Remove Rows.

## Changed Type

This step gets applied automatically after Promoted Headers when Power Query scans the values in all the columns and tries to detect and assign data types for each column. In case we delete this step, or need to apply it elsewhere later, we can first select all columns and go to Transform > Detect Data Type.

While changing data types, it is important from an optimization point of view to ensure that columns like all Key columns containing numeric values are assigned whole / decimal number type rather than text.

## Removed Other Columns

To remove unwanted columns, we can select the unwanted columns and go to Home > Remove Columns. However, this step can be tedious to modify later. So, we have selected the columns to be retained from Home > Choose Columns.

We can also get the same result by selecting Home > Remove Columns > Remove Other Columns after selecting the required columns.

# DimGeography

## Source

This step is automatically applied in the background when we identify the source of data.

## Navigation

For sources like Excel workbooks or databases like SQL Server, MySQL, Db2 etc. which can contain multiple tables, the Navigation step is also applied automatically in the background and we do not have to apply it manually as a transformation.

## Promoted Headers

If the first row of the sheet in Excel contains table headings, Promoted Headers transformation is applied automatically. This has happened for all the files mentioned hereinbefore. In cases where there are unwanted rows at the beginning on the sheet, we need to select Remove Top Rows from Home > Remove Rows.

## Changed Type

This step gets applied automatically after Promoted Headers when Power Query scans the values in all the columns and tries to detect and assign data types for each column. In case we delete this step, or need to apply it elsewhere later, we can first select all columns and go to Transform > Detect Data Type.

While changing data types, it is important from an optimization point of view to ensure that columns like all Key columns containing numeric values are assigned whole / decimal number type rather than text.

## Removed Other Columns

To remove unwanted columns, we can select the unwanted columns and go to Home > Remove Columns. However, this step can be tedious to modify later. So, we have selected the columns to be retained from Home > Choose Columns.

We can also get the same result by selecting Home > Remove Columns > Remove Other Columns after selecting the required columns.

## Renamed Columns

Rename columns with user-friendly names, viz., "StateProvinceName" to "State", "EnglishCountryRegionName" to "Country", "PostalCode" to "PIN".

## Removed Duplicates

Select the "GeographyKey" column and go to Home > Remove Rows > Remove Duplicates. This will ensure that the values in this column remain unique after the data is updated and the report refreshes successfully. This step can be applied to all primary key columns in the data model.

# DimEmployee

## Source

This step is automatically applied in the background when we identify the source of data.

## Navigation

For sources like Excel workbooks or databases like SQL Server, MySQL, Db2 etc. which can contain multiple tables, the Navigation step is also applied automatically in the background and we do not have to apply it manually as a transformation.

## Promoted Headers

If the first row of the sheet in Excel contains table headings, Promoted Headers transformation is applied automatically. This has happened for all the files mentioned hereinbefore. In cases where there are unwanted rows at the beginning on the sheet, we need to select Remove Top Rows from Home > Remove Rows.

## Added Custom Column

Go to Add Column > Custom Column and create a new column named “Full Name” such that,

Full Name = [FirstName] & " " &[LastName]

## Removed Other Columns

To remove unwanted columns, we can select the unwanted columns and go to Home > Remove Columns. However, this step can be tedious to modify later. So, we have selected the columns to be retained from Home > Choose Columns.

We can also get the same result by selecting Home > Remove Columns > Remove Other Columns after selecting the required columns.

## Reordered Columns

Drag and drop column heading so that the columns are in the following sequence: "EmployeeKey", "ParentEmployeeKey", "SalesTerritoryKey", "Full Name", "Title", "BirthDate"

## Changed Type

This step gets applied automatically after Promoted Headers when Power Query scans the values in all the columns and tries to detect and assign data types for each column. In case we delete this step, or need to apply it elsewhere later, we can first select all columns and go to Transform > Detect Data Type.

While changing data types, it is important from an optimization point of view to ensure that columns like all Key columns containing numeric values are assigned whole / decimal number type rather than text.

# DimSalesTerritory

## Source

This step is automatically applied in the background when we identify the source of data.

## Navigation

For sources like Excel workbooks or databases like SQL Server, MySQL, Db2 etc. which can contain multiple tables, the Navigation step is also applied automatically in the background and we do not have to apply it manually as a transformation.

## Promoted Headers

If the first row of the sheet in Excel contains table headings, Promoted Headers transformation is applied automatically. This has happened for all the files mentioned hereinbefore. In cases where there are unwanted rows at the beginning on the sheet, we need to select Remove Top Rows from Home > Remove Rows.

## Changed Type

This step gets applied automatically after Promoted Headers when Power Query scans the values in all the columns and tries to detect and assign data types for each column. In case we delete this step, or need to apply it elsewhere later, we can first select all columns and go to Transform > Detect Data Type.

While changing data types, it is important from an optimization point of view to ensure that columns like all Key columns containing numeric values are assigned whole / decimal number type rather than text.

## Removed Columns

Select the "SalesTerritoryAlternateKey" column and remove it from Home > Remove Columns.

## Renamed Columns

Rename columns with user-friendly names, viz., "SalesTerritoryRegion" to "Region", "SalesTerritoryCountry" to "Country", "SalesTerritoryGroup" to "TerritoryGroup"

# DimProductCategory

No transformations applied manually.

# DimProductSubcategory

## Source

This step is automatically applied in the background when we identify the source of data.

## Navigation

For sources like Excel workbooks or databases like SQL Server, MySQL, Db2 etc. which can contain multiple tables, the Navigation step is also applied automatically in the background and we do not have to apply it manually as a transformation.

## Promoted Headers

If the first row of the sheet in Excel contains table headings, Promoted Headers transformation is applied automatically. This has happened for all the files mentioned hereinbefore. In cases where there are unwanted rows at the beginning on the sheet, we need to select Remove Top Rows from Home > Remove Rows.

## Changed Type

This step gets applied automatically after Promoted Headers when Power Query scans the values in all the columns and tries to detect and assign data types for each column. In case we delete this step, or need to apply it elsewhere later, we can first select all columns and go to Transform > Detect Data Type.

While changing data types, it is important from an optimization point of view to ensure that columns like all Key columns containing numeric values are assigned whole / decimal number type rather than text.

## Merged Queries

Select the DimProductSubcategory query from the queries list on the left of the preview pane in Query Editor, and go to Home > Merge Queries > Merge Queries. In the resulting Merge dialog box, select the DimProductCategory query as the second query. Select “ProductCategoryKey” columns in the preview pane of each of the queries in the dialog box and click OK to merge the queries.

## Expanded DimProductCategory

In the merged query, i.e., DimProductSubcategory, click on the diverging arrows button on the DimProductCategory column heading, to expand the table and from the dropdown select the “EnglishProductCategoryName” column only.

# DimProduct

## Source

This step is automatically applied in the background when we identify the source of data.

## Navigation

For sources like Excel workbooks or databases like SQL Server, MySQL, Db2 etc. which can contain multiple tables, the Navigation step is also applied automatically in the background and we do not have to apply it manually as a transformation.

## Promoted Headers

If the first row of the sheet in Excel contains table headings, Promoted Headers transformation is applied automatically. This has happened for all the files mentioned hereinbefore. In cases where there are unwanted rows at the beginning on the sheet, we need to select Remove Top Rows from Home > Remove Rows.

## Changed Type

This step gets applied automatically after Promoted Headers when Power Query scans the values in all the columns and tries to detect and assign data types for each column. In case we delete this step, or need to apply it elsewhere later, we can first select all columns and go to Transform > Detect Data Type.

While changing data types, it is important from an optimization point of view to ensure that columns like all Key columns containing numeric values are assigned whole / decimal number type rather than text.

## Removed Other Columns

To remove unwanted columns, we can select the unwanted columns and go to Home > Remove Columns. However, this step can be tedious to modify later. So, we have selected the columns to be retained from Home > Choose Columns.

We can also get the same result by selecting Home > Remove Columns > Remove Other Columns after selecting the required columns.

## Renamed Columns

Rename columns with user-friendly names, viz., "EnglishProductName" to "Product Name".

## Merged Queries

Select the DimProduct query from the queries list on the left of the preview pane in Query Editor, and go to Home > Merge Queries > Merge Queries. In the resulting Merge dialog box, select the DimProductSubcategory query as the second query. Select “ProductSubcategoryKey” columns in the preview pane of each of the queries in the dialog box and click OK to merge the queries.

## Expanded DimProductSubcategory

In the merged query, i.e., DimProduct, click on the diverging arrows button on the DimProductSubcategory column heading, to expand the table and from the dropdown select the “EnglishProductSubcategoryName” and “EnglishProductCategoryName” columns only.

# DimProductSubcategory

## Source

This step is automatically applied in the background when we identify the source of data.

## Navigation

For sources like Excel workbooks or databases like SQL Server, MySQL, Db2 etc. which can contain multiple tables, the Navigation step is also applied automatically in the background and we do not have to apply it manually as a transformation.

## Promoted Headers

If the first row of the sheet in Excel contains table headings, Promoted Headers transformation is applied automatically. This has happened for all the files mentioned hereinbefore. In cases where there are unwanted rows at the beginning on the sheet, we need to select Remove Top Rows from Home > Remove Rows.

## Changed Type

This step gets applied automatically after Promoted Headers when Power Query scans the values in all the columns and tries to detect and assign data types for each column. In case we delete this step, or need to apply it elsewhere later, we can first select all columns and go to Transform > Detect Data Type.

While changing data types, it is important from an optimization point of view to ensure that columns like all Key columns containing numeric values are assigned whole / decimal number type rather than text.

## Merged Queries

Select the DimProductSubcategory query from the queries list on the left of the preview pane in Query Editor, and go to Home > Merge Queries > Merge Queries. In the resulting Merge dialog box, select the DimProductCategory query as the second query. Select “ProductCategoryKey” columns in the preview pane of each of the queries in the dialog box and click OK to merge the queries.

## Expanded DimProductCategory

In the merged query, i.e., DimProductSubcategory, click on the diverging arrows button on the DimProductCategory column heading, to expand the table and from the dropdown select the “EnglishProductCategoryName” column only.

# FactResellerSales (2)

[Append Query Report]

## Source

This step is automatically applied in the background when we identify the source of data.

## Navigation

For sources like Excel workbooks or databases like SQL Server, MySQL, Db2 etc. which can contain multiple tables, the Navigation step is also applied automatically in the background and we do not have to apply it manually as a transformation.

## Promoted Headers

If the first row of the sheet in Excel contains table headings, Promoted Headers transformation is applied automatically. This has happened for all the files mentioned hereinbefore. In cases where there are unwanted rows at the beginning on the sheet, we need to select Remove Top Rows from Home > Remove Rows.

## Changed Type

This step gets applied automatically after Promoted Headers when Power Query scans the values in all the columns and tries to detect and assign data types for each column. In case we delete this step, or need to apply it elsewhere later, we can first select all columns and go to Transform > Detect Data Type.

While changing data types, it is important from an optimization point of view to ensure that columns like all Key columns containing numeric values are assigned whole / decimal number type rather than text.

## Removed Other Columns

To remove unwanted columns, we can select the unwanted columns and go to Home > Remove Columns. However, this step can be tedious to modify later. So, we have selected the columns to be retained, viz., "ProductKey", "EmployeeKey", "SalesTerritoryKey", "SalesAmount", "OrderDate", from Home > Choose Columns.

We can also get the same result by selecting Home > Remove Columns > Remove Other Columns after selecting the required columns.

## Added Custom Column

Go to Add Column > Custom Column and create a new column named “Channel” such that,

Channel = “Reseller”

# FactInternetSales

[Append Query Report]

## Source

This step is automatically applied in the background when we identify the source of data.

## Navigation

For sources like Excel workbooks or databases like SQL Server, MySQL, Db2 etc. which can contain multiple tables, the Navigation step is also applied automatically in the background and we do not have to apply it manually as a transformation.

## Promoted Headers

If the first row of the sheet in Excel contains table headings, Promoted Headers transformation is applied automatically. This has happened for all the files mentioned hereinbefore. In cases where there are unwanted rows at the beginning on the sheet, we need to select Remove Top Rows from Home > Remove Rows.

## Changed Type

This step gets applied automatically after Promoted Headers when Power Query scans the values in all the columns and tries to detect and assign data types for each column. In case we delete this step, or need to apply it elsewhere later, we can first select all columns and go to Transform > Detect Data Type.

While changing data types, it is important from an optimization point of view to ensure that columns like all Key columns containing numeric values are assigned whole / decimal number type rather than text.

## Removed Other Columns

To remove unwanted columns, we can select the unwanted columns and go to Home > Remove Columns. However, this step can be tedious to modify later. So, we have selected the columns to be retained, viz., "ProductKey", "CustomerKey", "SalesTerritoryKey", "SalesAmount", "OrderDate", from Home > Choose Columns.

We can also get the same result by selecting Home > Remove Columns > Remove Other Columns after selecting the required columns.

## Added Custom Column

Go to Add Column > Custom Column and create a new column named “Channel” such that,

Channel = “Internet”

# Append1

[Append Query Report]

## Source

To combine data from the FactResellerSales and FactInternetSales queries, go to Home > Append Queries > Append Queries as New. Select the Two Tables option. In the first dropdown select FactResellerSales, and in the second one select FactInternetSales. Click OK to create the new appended query.

# UN Population Data – Raw

[UN Population Report]

## Source

This step is automatically applied in the background when we identify the source of data.

## Navigation

For sources like Excel workbooks or databases like SQL Server, MySQL, Db2 etc. which can contain multiple tables, the Navigation step is also applied automatically in the background and we do not have to apply it manually as a transformation.

## Removed Top Rows

Since there are unwanted rows at the top of the data, go to Home > Remove Rows > Remove Top Rows, and enter the number of rows to be deleted from the top, viz., 15 in this case.

## Promoted Headers

Go to Home > Use First Row as Headers.

## Changed Type

This step gets applied automatically after Promoted Headers when Power Query scans the values in all the columns and tries to detect and assign data types for each column. In case we delete this step, or need to apply it elsewhere later, we can first select all columns and go to Transform > Detect Data Type.

While changing data types, it is important from an optimization point of view to ensure that columns like all Key columns containing numeric values are assigned whole / decimal number type rather than text.

## Unpivoted Other Columns

Select the columns "Index", "Country", "Continent", "Area", "Category", "Notes", "Country code". To unpivot the remaining columns, go to Transform > Unpivot Columns > Unpivot Other Columns.

## Renamed Columns

Rename columns with user-friendly names, viz., "Attribute" to "Year” and "Value" to "Population".