

Power BI - Power Query Tips

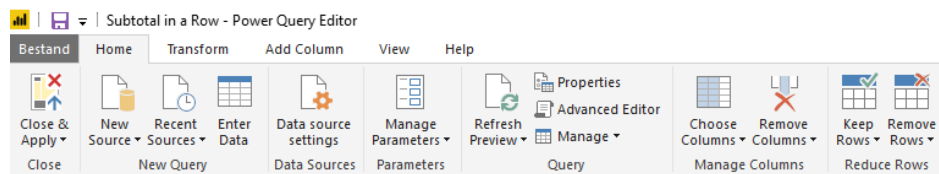
Read a folder with CSV-files

Connect to folder



1. **Home > External data > Get Data > More**
2. Select **Folder** and click **Connect**
3. Select the folder and click on **OK**
4. When the contents of the folder visible click on **Edit**
5. Apply filters to keep only the files you need

Remove unnecessary columns

1. Click on the column **Content**, hold down <CTRL> and click on **Name**
2. Choose **Home > Manage Columns > Remove Columns > Remove Other Columns**



Open the CSV-files

1. Create a new column: **Add Column > General > Custom Column**
2. Insert a name for this column "Data" and type the following function:
=Csv.Document([Content])
3. Click on **OK** to confirm
4. Open the tables by clicking on  in the column header 
5. Select the columns you would like to import
If the separator (; or ,) is not recognized correctly you will need to split the columns:
6. Click on the created column and choose **Home > Transform > Split Column > By Delimiter**
7. Select the separator and click on **OK**

Remove the column Content

Right-click in the "Content" and select **Delete**

Create column headers


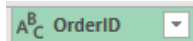
Mark the top row as column headers **Home > Transform > Use First Row as Headers**

Define the datatype for every column

1. Select all columns via <CTRL>+A
2. **Transform > All Columns > Detect Data Type**

Remove column header rows

The column headers in the various files are also imported and need to be deleted.

1. Click on the folder button  for example for the first column 
2. Choose **Text filters... > Does not Equal...** and enter the name of this column

Note: Power Query case sensitive

Remove empty rows

CSV files can contain empty rows.

1. Click on any of the columns that should never be empty
2. **Home > Reduce Rows > Remove Rows > Remove Blank Rows**

Remove errors

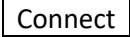


Upon the detection (or assigning) of datatypes errors could occur, for example there are rows that do not contain date whereas the others do.

1. Select all columns via <CTRL>+A
2. **Home > Reduce Rows > Remove Rows > Remove Errors**

Power BI - Power Query Tips

Read a folder with XLSX-files





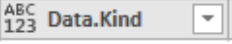
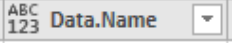
Connect a folder

1. **Home > External data > Get Data > More**
2. Select **Folder** and click 
3. Select the folder and click on 
4. When the contents of the folder visible click on 
5. Apply filters to keep only the files you need



Remove unnecessary columns

1. Click on the column **Content**, hold down <CTRL> and click on **Name**
2. Choose **Home > Manage Columns > Remove Columns > Remove Other Columns**

Open the XLSX files

1. Create a new column: **Add column > General > Custom Column**
2. Insert a name for this column: "Data" and type the following function:
=Excel.Workbook([Content])
3. Click on  to confirm
4. Open the workbooks by clicking on  in the column  click  to confirm
5. In the column  filter on one of the information types (Table or Sheet).
6. In the column  filter on the name of the worksheets (or tables).
7. Select column "Data.Data" and select **Home > Manage Columns > Remove Columns > Remove Other Columns**

Open the worksheets/tables in the XLSX files

1. Click on  in the only column 
2. Select the columns to import
3. Deselect the option **Use original column name as prefix**

Remove the column Content

Right-click the column "Content" and select **Delete**

Note: The options below are redundant if the source data was saved in tables.

Create column headers


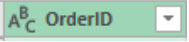
Mark the top row as column headers **Home > Transform > Use First Row as Headers**

Define the datatype for every column

1. Select all columns via <CTRL>+A
2. **Transform > All Columns > Detect Data Type**

Remove column header rows

The column headers in the various files are also imported and need to be deleted.

1. Click on the folder button  for example for the first column 
2. Choose **Text filters... > Does not Equal...** and enter the name of this column

Note: Power Query case sensitive

Remove empty rows

CSV files can contain empty rows.

1. Click on any of the columns that should never be empty
2. **Home > Reduce Rows > Remove Rows > Remove Blank Rows**

Remove errors


Upon the detection (or assigning) of datatypes errors could occur, for example there are rows that do not contain date whereas the others do.

1. Select all columns via <CTRL>+A
2. **Home > Reduce Rows > Remove Rows > Remove Errors**

Power BI - Power Query Tips

Use a Wizard to read a folder with XLSX-files

Connect to folder

1. **Home > External data > Get Data > More...**
2. Select  **Folder** and click on **Connect**


Start the Wizard

3. Click on **Combine** and choose **Combine & Edit / Combine & Transform data**

The wizard will use one of the files as an example.

4. Select the file to be used as an example

Select the object to be extracted from each file. [Learn more](#)

Example File:  First file

5. Select the Table, Sheet or Named range and click on **OK**

◀  Sample File Parameter1 [3]

 Facts_Orders

 FactsOrders

 FactsOrders1

Tip: Files with error scan be skipped ☐ Skip files with errors

Result

Queries [5]

- Transform File from PQ_YearlyOrders [3]
 - Sample Query [2]
 - Sample File Parameter1 (Sample File)
 - Sample File
 - Transform Sample File from PQ_YearlyOrders**
 - Transform File from PQ_YearlyOrders
- Other Queries [1]
 - PQ_YearlyOrders

Transform Sample ...: Transformations applied to this query will be copied to the Fx Transform File...

Fx Transform File ...: a function that applies steps to individual files.

PQ_YearlyOrders: The primary query

Power BI - Power Query Tips

Merge tables (VLOOKUP)

In Excel lots of people use VLOOKUP to combine information from 2 tables. With Power Query this is similarly possible, and since this only takes time whenever the Query is refreshed you will experience a substantial increase in performance.

Fact table vs Dimension table

You will be expanding a table with facts (transactions), with information saved in another table (dimension), for example Employees, Customers, Categories etc.

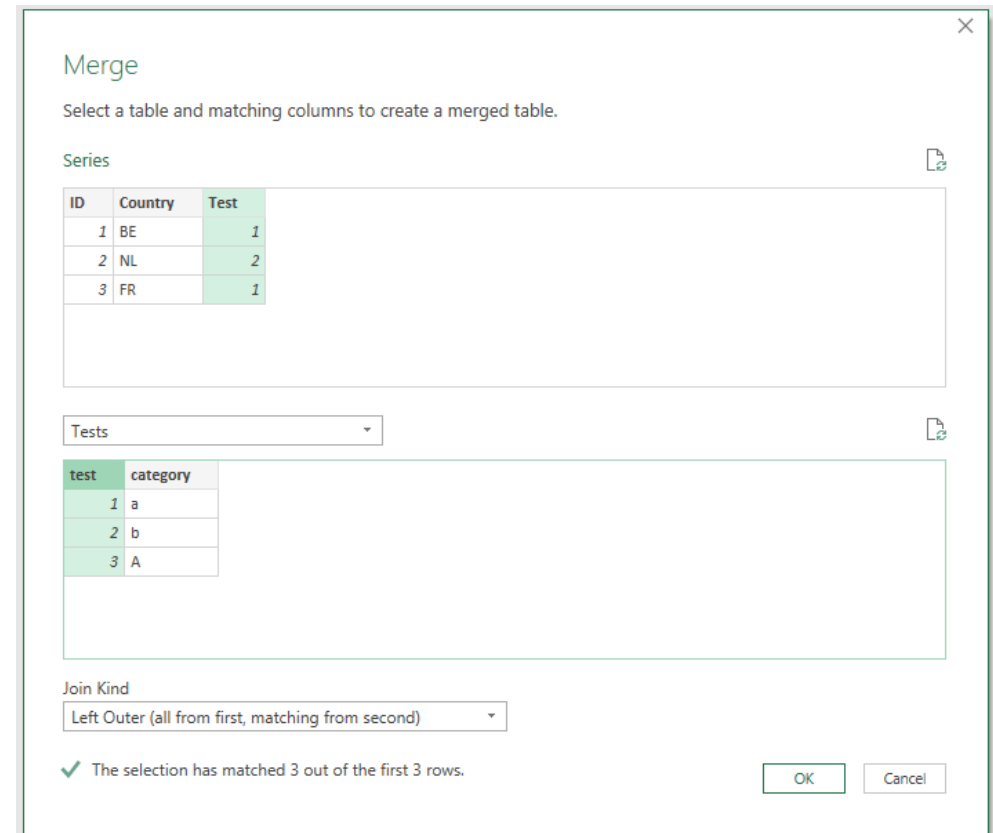
Create the query with the extra information from the dimension

1. Create a standard query containing the extra info, make sure that the primary key column is also included.
2. Choose **Home > Query > Properties...** and disable ☐ **Enable load to report**.
This will ensure that the data will not be part of the data model.

Create the query to the fact table

1. Create a basic query, make sure that you include the external key column, you need this column to create a link with the dimension table
2. Choose **Home > Combine > Merge Queries**
3. In the upper part of the window you see the current query, in lower part you select the dimension query
4. Mark in both tables the key column.
5. In normal situations select **Type join: Left Outer**, in other words the fact table will be enriched with the information from the Dimension table.

Important: Power Query is case sensitive. If necessary, use the transformation **Transform > Text Column > Format > UPPERCASE** to create keys that are equal.



Tip By means of a **Right Outer join**, you can filter a table.

All lines from the first table will be removed that do not have a corresponding key in the table below.

Power BI - Power Query Tips

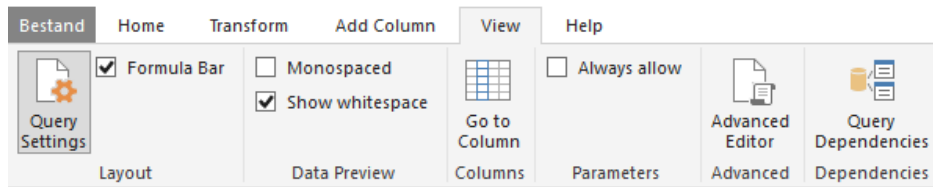
Create a Date table

Import the date table

1. Copy the code below:

```
// start of date code
let
    StartDate = #date(2017,1,1),
    #"Set length" = List.Dates(StartDate, Number.From(DateTime.LocalNow()) - Number.From(StartDate) + 365, #duration(1,0,0,0)),
    #"Converted to Table" = Table.FromList(#"Set length", Splitter.SplitByNothing(), null, null, ExtraValues.Error),
    #"Rename Column Date" = Table.RenameColumns(#"Converted to Table",{{"Column1", "Date"}}),
    #"Set as Date Column" = Table.TransformColumnTypes(#"Rename Column Date",{{"Date", type date}}),
    #"Inserted Year" = Table.AddColumn(#"Set as Date Column", "Year", each Date.Year([Date]), Int64.Type),
    #"Inserted Quarter" = Table.AddColumn(#"Inserted Year", "Quarter", each Date.QuarterOfYear([Date]), Int64.Type),
    #"Inserted Month" = Table.AddColumn(#"Inserted Quarter", "Month", each Date.Month([Date]), Int64.Type),
    #"Inserted Month Name" = Table.AddColumn(#"Inserted Month", "Month Name", each Date.MonthName([Date], "en-BE"), type text),
    #"Insert mmm YYYY" = Table.AddColumn(#"Inserted Month Name", "mmmYYYY", each Text.Upper ( Date.ToText([Date], "MMM yyyy", "en")
    )),
    #"Insert mm YYYY" = Table.AddColumn(#"Insert mmm YYYY", "mm YYYY", each Text.Upper ( Date.ToText([Date], "MM.yyyy", "en") )),
    #"Insert Monthkey" = Table.AddColumn(#"Insert mm YYYY", "MonthKey", each Date.ToText([Date], "yyyyMM", "en")),
    #"Changed Type" = Table.TransformColumnTypes(#"Insert Monthkey",{{"mmmYYYY", type text}, {"mm YYYY", Int64.Type}, {"MonthKey",
    Int64.Type}}
    in
    #"Changed Type"
// end of date code
```

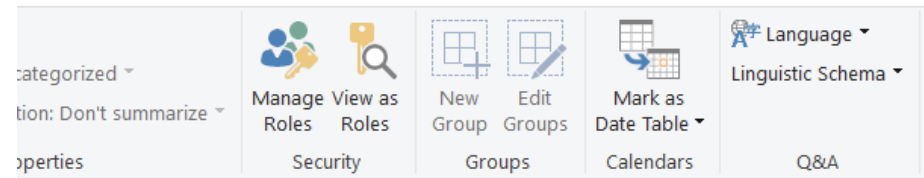
2. Choose **Power Query > Get Data > From other sources > Blank Query**
3. Start the advanced editor: **View > Advanced > Advanced editor**



4. Remove the original text and paste the copied code
5. Click on **Done**
6. Change the name of the query to "Dates"
7. Review the code and change the StartDate in the first step
8. **Close & Apply**

Mark the table as a Date Table

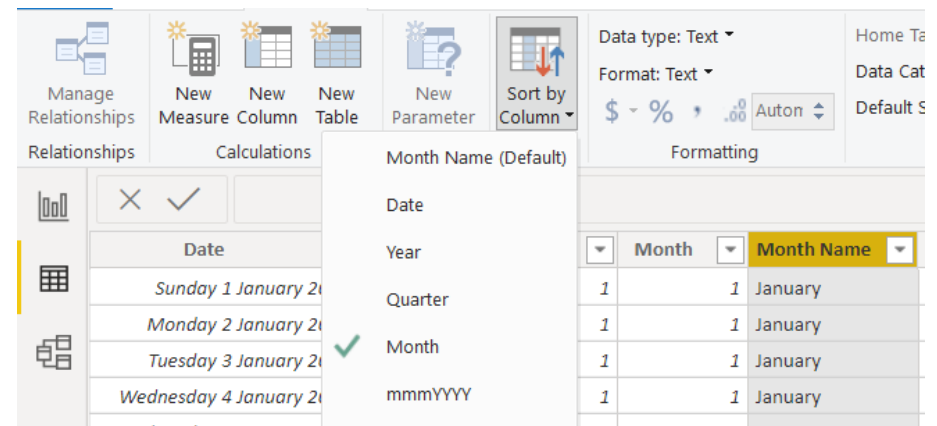
1. In Power BI go to the Data view
2. Navigate to the created table "Dates"
3. Click **Modeling > Calendars > Mark as Date Table > Mark as Date Table**
4. Select the "Date" column as unique identifier and click **OK**



Change the sorting for "Month name" column(s)

The "Month name" will be sorted alphabetically, use the following method to apply the correct sort-order.

5. Select the "Month name"-column
Click **Modeling > Sort > Sort by Column > ...**
6. Indicate which column should be used to sort the "Month name"-column (e.g. "Month" or "Month number")



Power BI - Power Query Tips

Turning Metadata into a column

Sometimes exports contain metadata in the top-rows. For multiple reasons we want this data to appear as a column, repeating the same values multiple times.

Before

	A	B	C	D	E	F
1	ReportDate	22/03/2019				
2	Source	DB-ext-Sales				
3						
4	ID	Customer	Status	Duration		
5	2019003	Goldendox	Closed	5		
6	2019004	Goldendox	Closed	4		
7	2019005	Indonlam	Closed	4		
8	2019006	An-fix	Closed	2		
9	2019007	Goldendox	Closed	6		

After

ID	Customer	Status	Duration	ReportDate
2019003	Goldendox	Closed	5	Friday 22 March 2019
2019004	Goldendox	Closed	4	Friday 22 March 2019
2019005	Indonlam	Closed	4	Friday 22 March 2019
2019006	An-fix	Closed	2	Friday 22 March 2019

Import the Source file

1. **Home > External data > Get Data > Excel**
2. Select **File** and click **Open**
3. Select the sheet and click on **Edit** / **Transform data**
4. Change the name of the query, press **<Enter>** to confirm

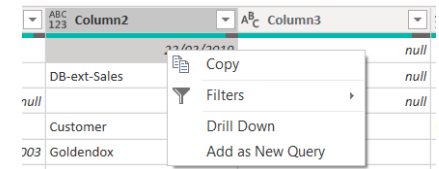
Activate the formula bar

If the formula bar is not active, activate it

1. **View > Layout > ☒ Formula Bar**

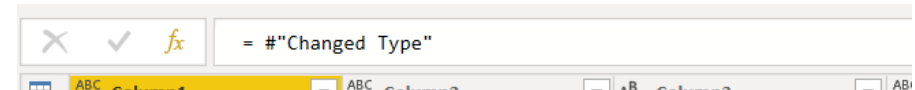
Capture the metadata from the first line

1. Go to the last query-step in the Query Settings-pane, most likely "Changed Type"
2. Right-click on the value to capture and choose **Drill Down**
3. Right-click on the step Column1 and change the name to "Capture reportDate"



Restore the table

1. Click on **fx** next to your formula bar
2. Change the **= #"Capture ReportDate"** in to **= #"Changed Type"** to refer to a previous state of the query



Tip: Power Query is Case-sensitive

3. Right-click on the step Custom1 and change the name to "Restore table"

Clean up

1. Use **Home > Reduce Rows > Remove Rows > Remove top rows** to clean the unnecessary rows
2. Use **Home > Transform > Use First Row as headers**

Insert the captured value

1. Click on **Add Column > Custom Column**
2. Insert a name for the custom column, e.g. "ReportDate"
3. Insert the formula **= #"Capture ReportDate"**

You refer to the name given to step where the metadata was captured.

Apply the correct data type

Click on **ABC 123** of **ABC 123 ReportDate** and select the correct data type

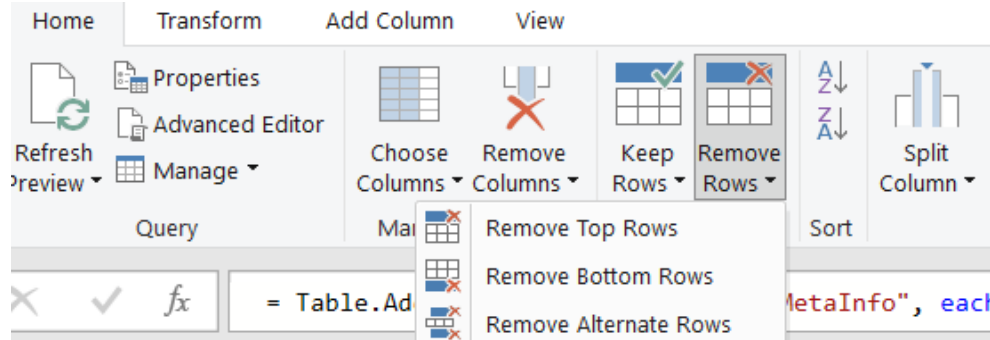
Finish the query

Home > Close & Apply the query

Power BI - Power Query Tips

Removing meta information

Often imported data starts or ends with meta information at the top or bottom of the table. This can obviously be removed via the option **Remove Rows > Remove Top Rows**.



But in some situations, you do not know how many rows at the top need to be removed. In those case the following method will give you a solid and flexible solution.

Before

	A	B	C	D	E	F	G	H	I
1	Titel	Orders and Deals							
2	ReportDate	09/01/2019							
3	DataStore	EU_Main:386301							
4									
5	OrderID	Productkey	Productcolour	Quantity	PricePerUnit	CurrencyCode	BtwTvaVat	Orderdate	Customerkey
6	O0000001	P0037	Green	2	15750,73	EUR	0,21	03/01/2012	C00025
7	O0000002	P0038	Green	2	16070,68	GBP	0,19	03/01/2012	C00009
8	O0000003	P0026	Green	3	9380,22	GBP	0,19	03/01/2012	C00074

After

	OrderID	Productkey	Productcolour	Quantity	PricePerUnit	CurrencyCode	BtwTvaVat	Orderdate
	O0000337	P0050	Green	1	20030,3	EUR	0,21	Friday 15 March 2013
	O0000480	P0047	Green	1	28250,97	EUR	0,21	Monday 1 July 2013
	O0000545	P0004	Orange	1	2486,75	EUR	0,21	Thursday 15 August 2013
	O0000715	P0012	Green	1	12020,6	EUR	0,21	Wednesday 1 January 2014
	O0000725	P0012	Violet	1	12200,83	EUR	0,21	Thursday 9 January 2014

Test your columns

The idea about this technique is a meta info will result in multiple columns in a single row which are empty. In the picture above starting from column C.

Insert a conditional column

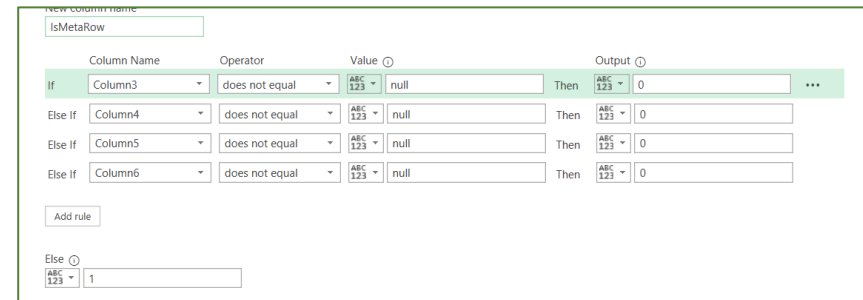
Assuming that the data is already loaded and your are in the Power Query editor. If needed, remove the steps like **Promoted headers** and **Changed type**.

1. Choose **Add Column > General > Conditional column**
2. Enter the name of the column, e.g. "IsMetaRow"

Test the columns that should be empty

Column3, Column4... are always empty in the meta info rows, so if one of them is not empty, it is not a meta info row.

1. Enter the following rule:
If Column3 does not equal null then 0
2. Click on **Add Rule** and enter the next rule
If Column4 does not equal null then 0
3. Repeat this for as many as necessary
4. Click in Else box and type 1



5. Click on **OK**

Filter the IsMetaRow column

Click on **ABC 123 IsMetaRow** and deselect 1

Remove the IsMetaRow

Right-click the IsMetaRow and choose **Remove**

Clean up

Use **Home > Transform > Use First Row as headers**