

Power BI – How to Transform and Clean Data

(Power Query)

Introduction

In this document, I tried to use Power BI (Power Query) to transform and clean data.

Step 1) Download the EXCEL

Excel Sample Data:

You can find the below excel file in my github.



Financial
Sample.xlsx

Step 2) Load the Data

Excel file import

Step 2.1) Open the power BI desktop.

Step 2.2) Click “Import data from excel” and choose the file “Financial Sample.xlsx” to import.

Step 2.3) checked the “Financials” worksheet and you can see the Date column

Navigator

Display Options ▾

- 2.4 Financial Sample.xlsx [2]
 - ☒ financials
 - ☐ Sheet1

OGS	Profit	Date	Month Number	Month Name	Year
16185	16185	1/1/2014	1	January	2014
13210	13210	1/1/2014	1	January	2014
21780	10890	1/6/2014	6	June	2014
8880	4440	1/6/2014	6	June	2014
24700	12350	1/6/2014	6	June	2014
393380	136170	1/12/2014	12	December	2014
9210	4605	1/3/2014	3	March	2014
7554	22662	1/6/2014	6	June	2014
18990	18990	1/6/2014	6	June	2014
4635	13905	1/6/2014	6	June	2014
24700	12350	1/6/2014	6	June	2014
319860	13327.5	1/7/2014	7	July	2014
239500	47900	1/8/2014	8	August	2014
10730	4292	1/9/2014	9	September	2014
41400	1725	1/10/2013	10	October	2013
6150	3075	1/12/2014	12	December	2014
2920	2920	1/2/2014	2	February	2014
9740	4870	1/2/2014	2	February	2014
7554	22662	1/6/2014	6	June	2014
261560	90540	1/6/2014	6	June	2014
1101	3303	1/7/2014	7	July	2014
4415	1766	1/8/2014	8	August	2014
5490	2745	1/9/2013	9	September	2013

Load Transform Data Cancel

Step 3) Transform the Data

Step 3.1) Click “Transform Data” button, then you will see the Power Query screen.

The screenshot shows the Power Query Editor interface. The 'financials' table is loaded, displaying 22 rows of data. The columns are: Segment, Country, Product, Discount Band, Units Sold, Manufacturing Price, Sale Price, and Gross Sales. The data includes various segments like Government, Midmarket, and Channel Partners across different countries and products.

Step 3.2) Find the column, you want to transform.

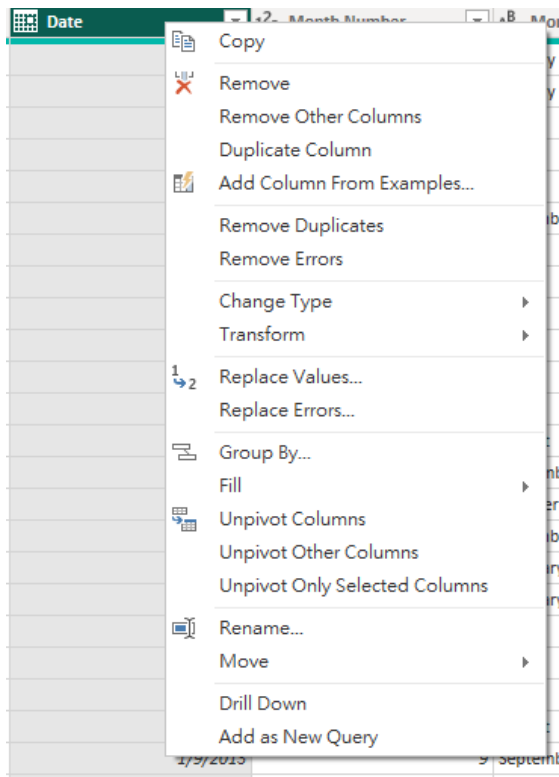
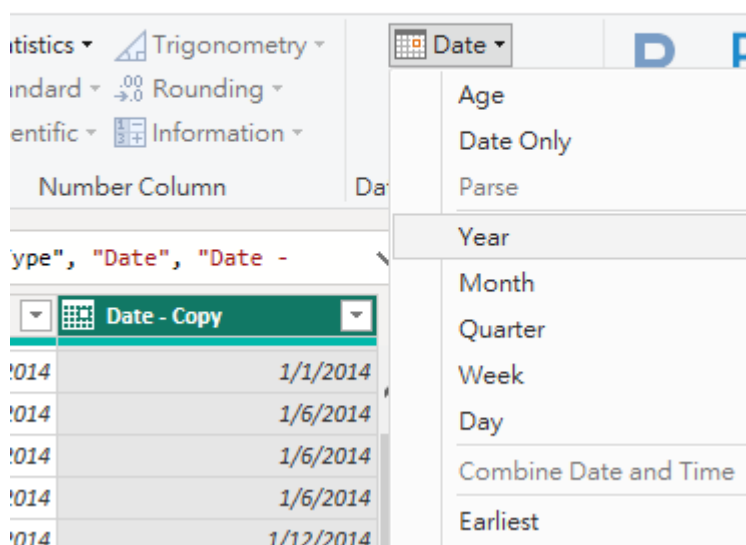
In this example, we will use the “Date” column to transform

Click the “Date” Column header

Click the “Transform” Tab

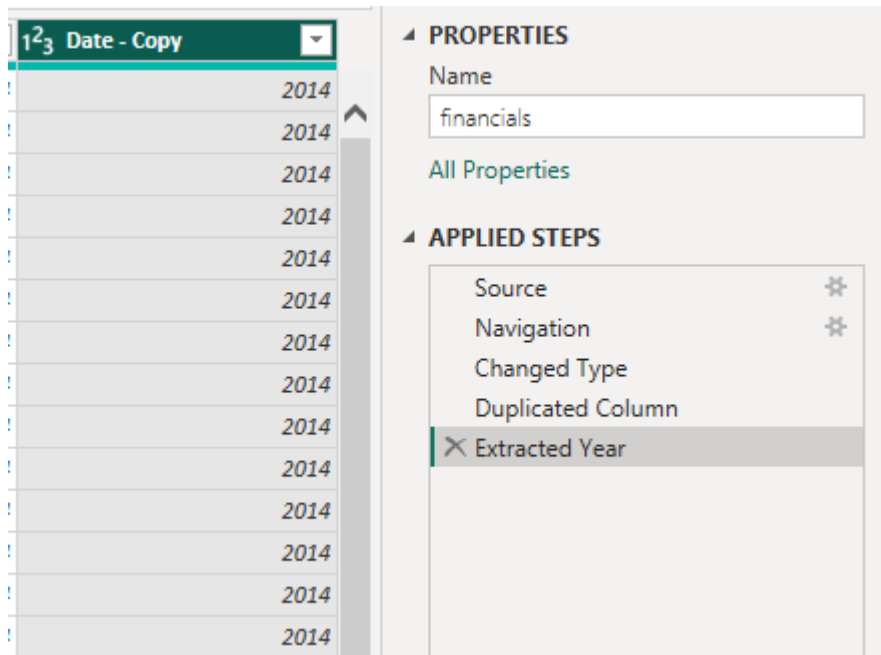
The screenshot shows the Power Query Editor interface with the 'financials' table. The 'Date' column is highlighted, and the 'Transform' tab is active. The formula bar shows the query definition: `= Table.TransformColumnTypes(financials_Table,{{"Segment", type text}, {"Country", type text}, {"Product", type text}, {"Discount Band", type text}, {"Units Sold", type text}, {"Profit", type number}, {"Date", type date}, {"Month Number", type number}})`. The table data is as follows:

	Profit	Date	Month Number
1	16185	1/1/2014	1
2	13210	1/1/2014	1
3	10890	1/6/2014	6
4	4440	1/6/2014	6
5	12350	1/6/2014	6
6	136170	1/12/2014	12
7	4605	1/3/2014	3
8	22662	1/6/2014	6
9	18990	1/6/2014	6
10	13905	1/6/2014	6
11	12350	1/6/2014	6
12	13327.5	1/7/2014	7
13	47900	1/8/2014	8
14	4292	1/9/2014	9
15	1725	1/10/2013	10
16	3075	1/12/2014	12

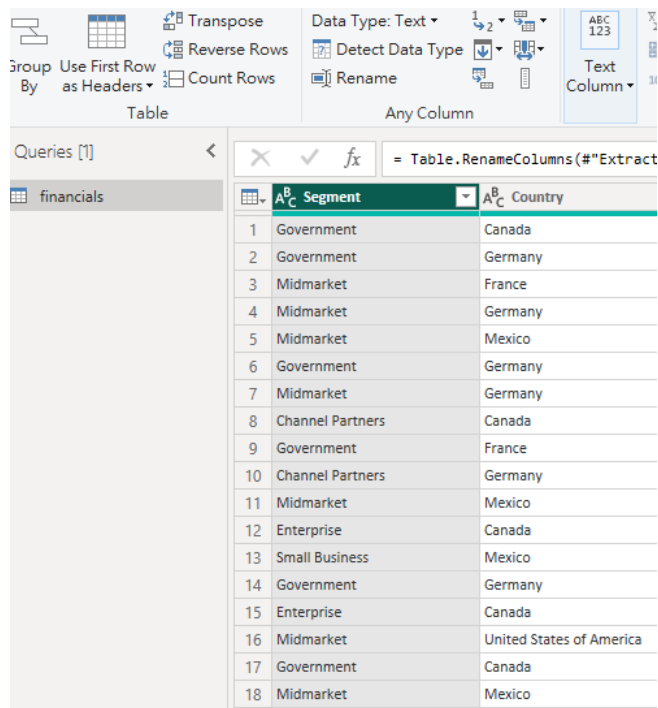
Step 3.3) Extract the “Date” column and form a new column (Year)**Right Click the “Date” column and Click “Duplicate Column”****Then you can see the new column “Date - copy”****Step 3.4) Use the new column “Date - copy” to extract the Year****Click “Date - Copy” column and Click the “Date” dropdown on top****Click “Year”**

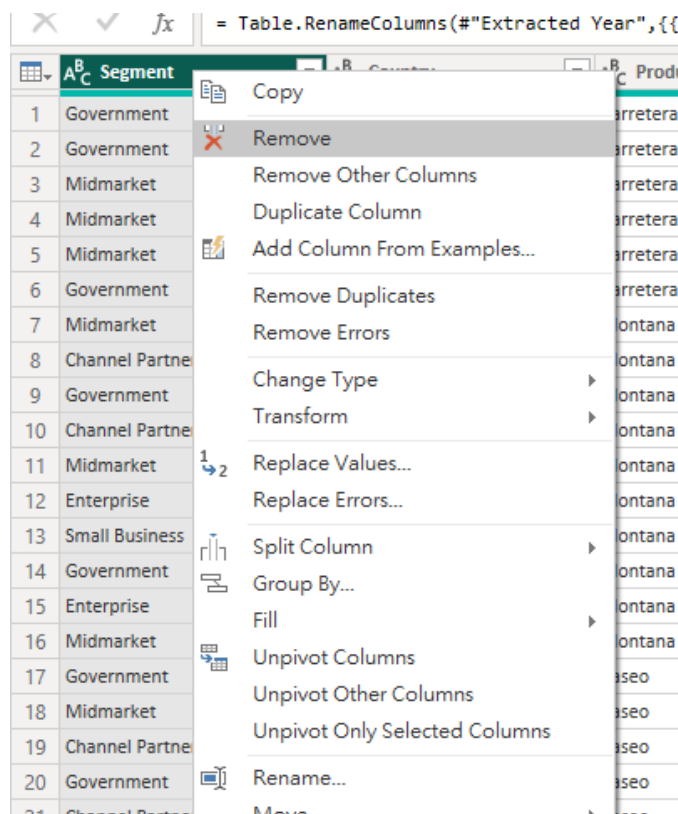
Step 3.5) The year will be extract, the Transform process is completed

Please rename the column name to “Year”,

**Step 4) Clean the Data****Step 4.1) In this example, you want to remove the column Segment**

Click “Segment” Column



Step 4.2) Right Click “Segment” Column and Click “Remove”**Step 4.3) The “Segment” column will be removed**

Formula Bar: `= Table.RemoveColumns(#"Renamed Columns",{ "Segment" })`

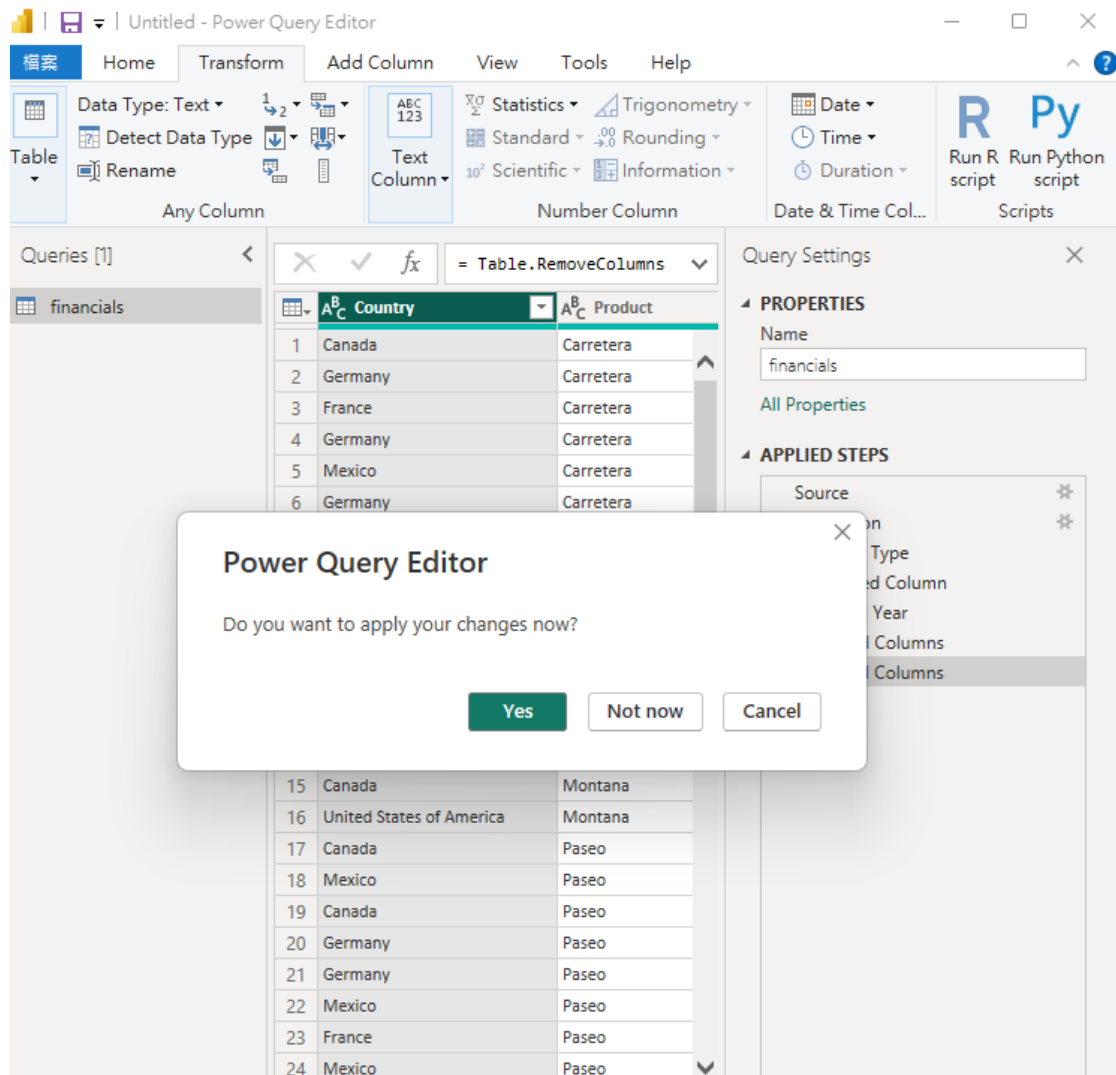
	Country	Product	Discount Band
1	Canada	Carretera	None
2	Germany	Carretera	None
3	France	Carretera	None
4	Germany	Carretera	None
5	Mexico	Carretera	None
6	Germany	Carretera	None
7	Germany	Montana	None
8	Canada	Montana	None
9	France	Montana	None
10	Germany	Montana	None
11	Mexico	Montana	None
12	Canada	Montana	None
13	Mexico	Montana	None
14	Germany	Montana	None
15	Canada	Montana	None
16	United States of America	Montana	None

Step 5) Close and Apply the Change

Step 5.1) Close the Power Query

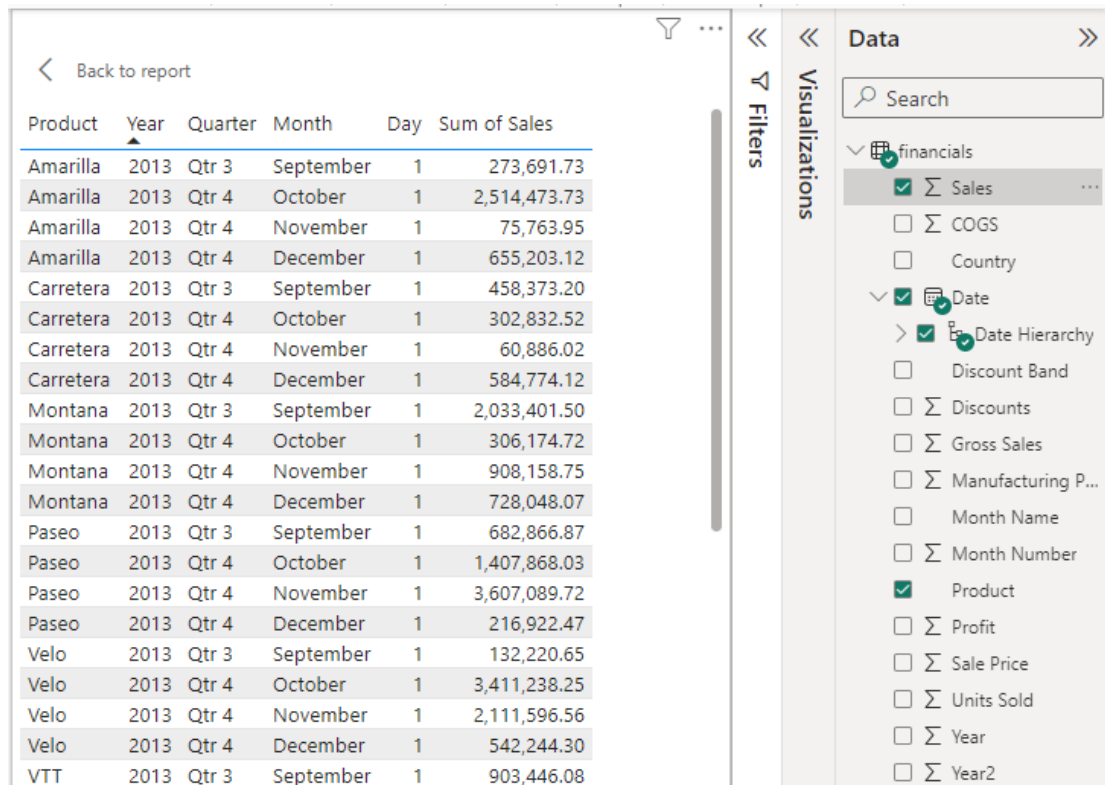
It will ask you “Do you want to apply your changes now?”

Click “Yes”



Step 6) Close and Apply the Change

Step 6.1) After close all and apply the change , You can select the fields you want and display



The screenshot displays the Power BI interface. On the left, a table shows sales data with columns: Product, Year, Quarter, Month, Day, and Sum of Sales. The table lists 20 rows of data for various products like Amarilla, Carretera, Montana, Paseo, Velo, and VTT across different quarters and months of 2013. On the right, the 'Data' pane is visible, showing a search bar and a list of fields. The 'financials' section is expanded, showing 'Sales' selected with a checkmark. Other fields like COGS, Country, Date, Date Hierarchy, Discount Band, Discounts, Gross Sales, Manufacturing P..., Month Name, Month Number, Product, Profit, Sale Price, Units Sold, Year, and Year2 are listed with checkboxes.

Product	Year	Quarter	Month	Day	Sum of Sales
Amarilla	2013	Qtr 3	September	1	273,691.73
Amarilla	2013	Qtr 4	October	1	2,514,473.73
Amarilla	2013	Qtr 4	November	1	75,763.95
Amarilla	2013	Qtr 4	December	1	655,203.12
Carretera	2013	Qtr 3	September	1	458,373.20
Carretera	2013	Qtr 4	October	1	302,832.52
Carretera	2013	Qtr 4	November	1	60,886.02
Carretera	2013	Qtr 4	December	1	584,774.12
Montana	2013	Qtr 3	September	1	2,033,401.50
Montana	2013	Qtr 4	October	1	306,174.72
Montana	2013	Qtr 4	November	1	908,158.75
Montana	2013	Qtr 4	December	1	728,048.07
Paseo	2013	Qtr 3	September	1	682,866.87
Paseo	2013	Qtr 4	October	1	1,407,868.03
Paseo	2013	Qtr 4	November	1	3,607,089.72
Paseo	2013	Qtr 4	December	1	216,922.47
Velo	2013	Qtr 3	September	1	132,220.65
Velo	2013	Qtr 4	October	1	3,411,238.25
Velo	2013	Qtr 4	November	1	2,111,596.56
Velo	2013	Qtr 4	December	1	542,244.30
VTT	2013	Qtr 3	September	1	903,446.08