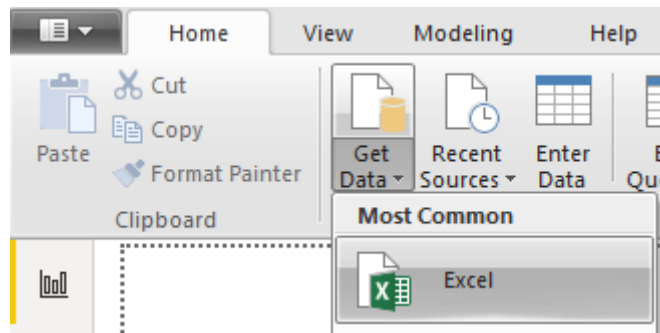


ETL & POWERQUERY

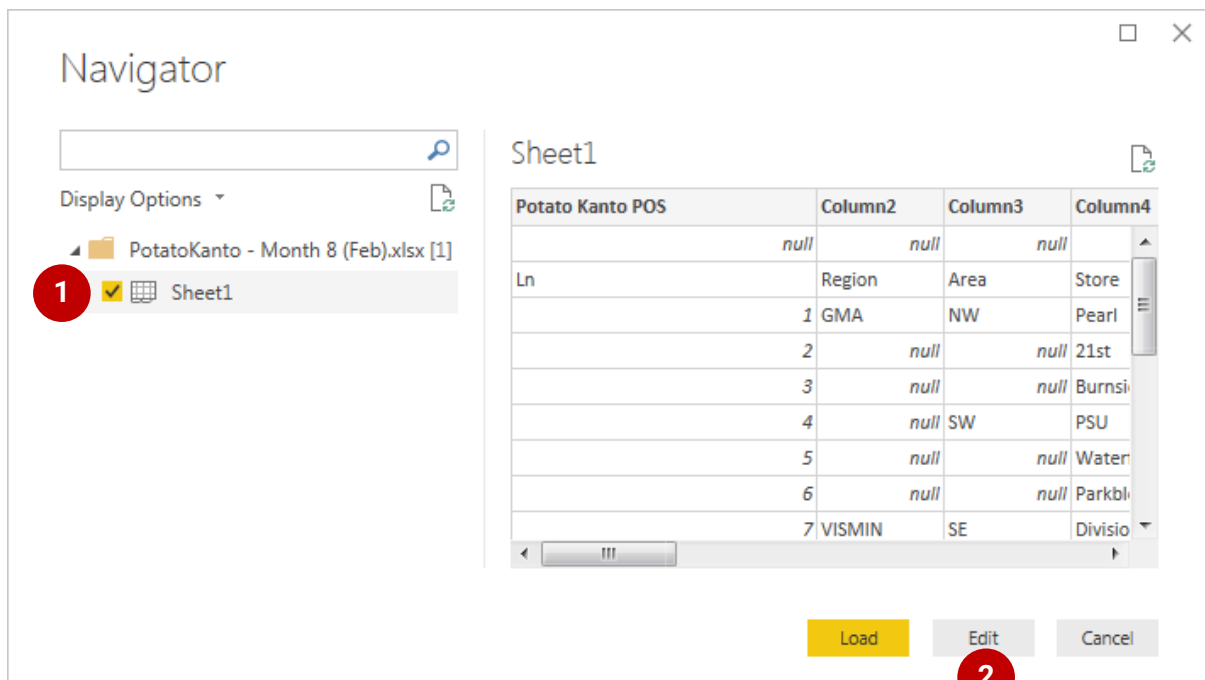
EXERCISE 1

This exercise aims to get the user familiar with the basic functionalities of PowerQuery, such as how to load data sets and apply transformations.

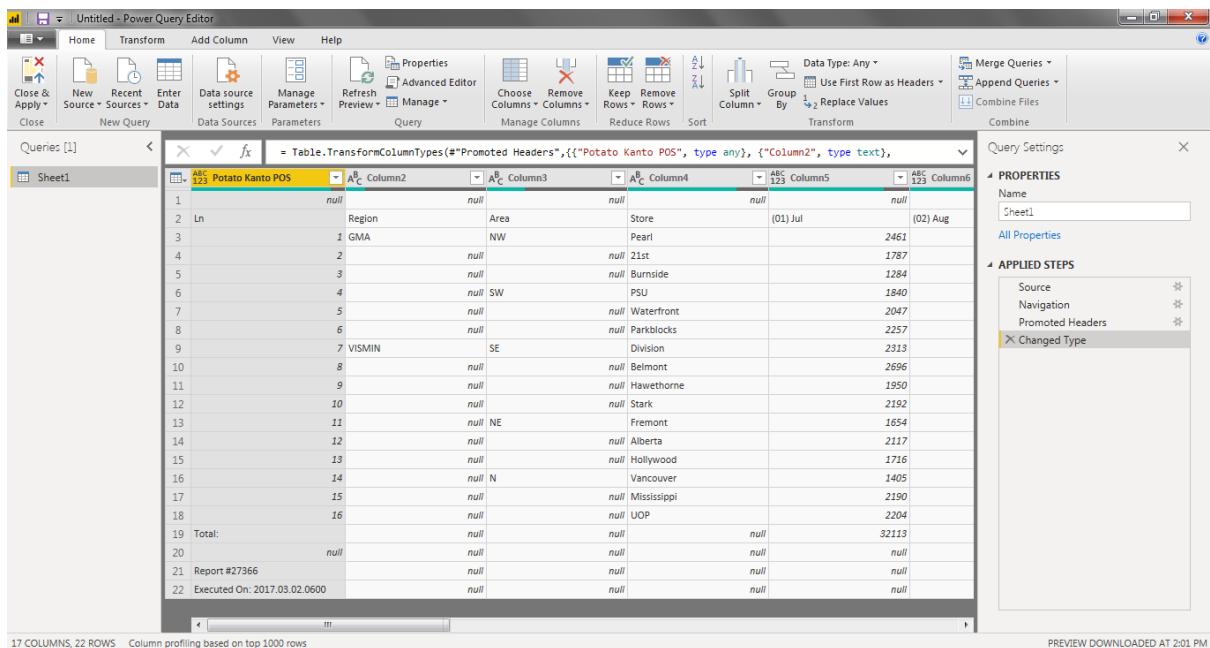
1. Launch a PowerBI instance.
2. Under the **Home** tab on the ribbon, click **Get Data > Excel**



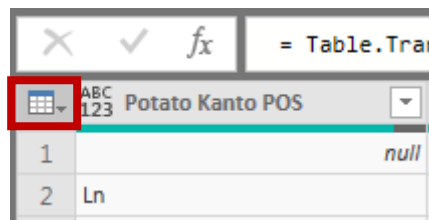
3. Browse to the folder containing the file "PotatoKanto – Month 8 (Feb).xlsx" and click **Open**.
4. On the **Navigator** window, select "Sheet1" and click **Edit**.



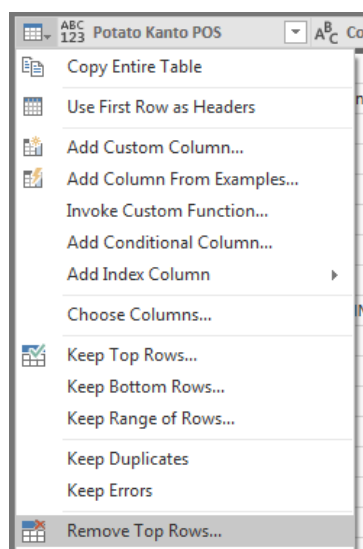
5. This will open the PowerQuery window.



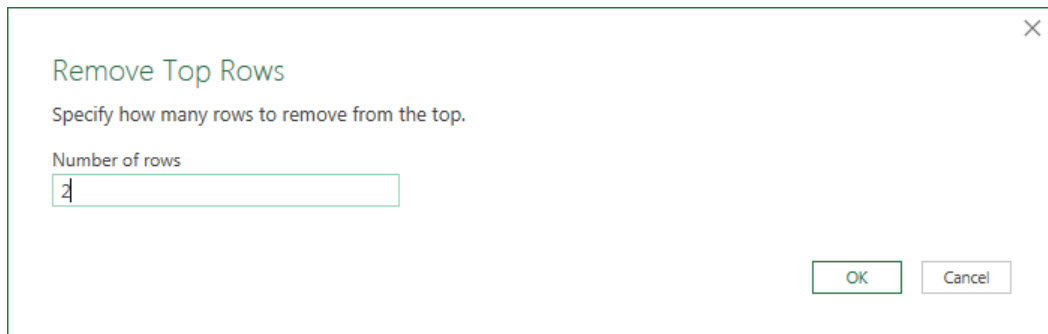
6. To access the menu of transformations available, select the icon to the left of the first column of the table.



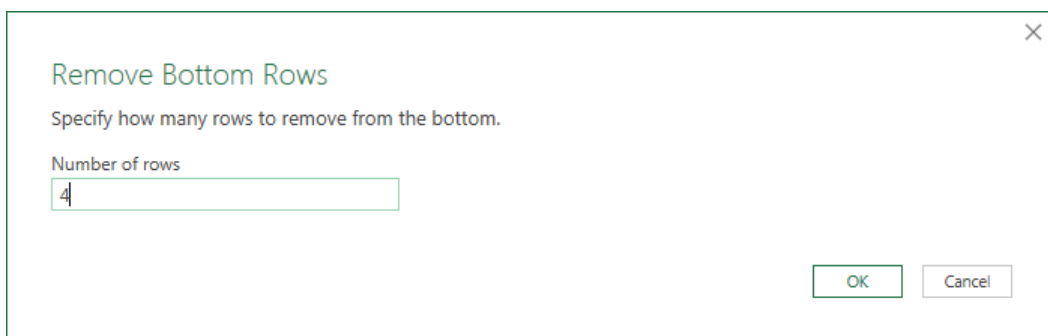
7. Select **Remove Top Rows**.



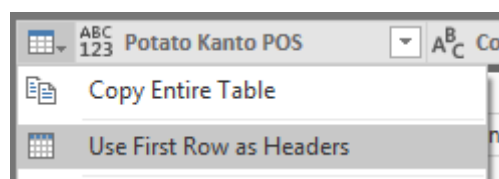
8. PowerQuery will prompt, asking for how many rows are to be removed. Input **2**, and click **OK**. Notice the top 2 rows have been removed from the table.



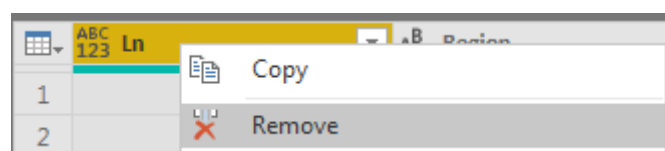
9. Access the same menu as in step 6, and click on **Remove Bottom Rows**.
10. Specify to remove **4** bottom rows and click **OK**. Notice the bottom 4 rows have been removed from the table.



11. Access the same menu in step 6, and select **Use First Row as Headers**.



12. Right-click on the **Ln** column, and select **Remove**. Do the same for column **YTD**.

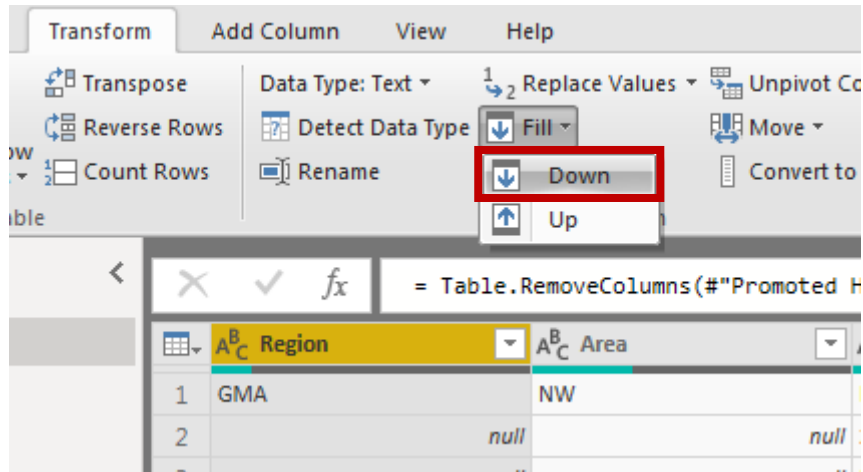


13. Click on the column header **Region** (a column has been selected when the column header is shaded yellow and the column values are shaded gray).

Under the **Transform** tab on the ribbon, select **Fill** then **Down**.

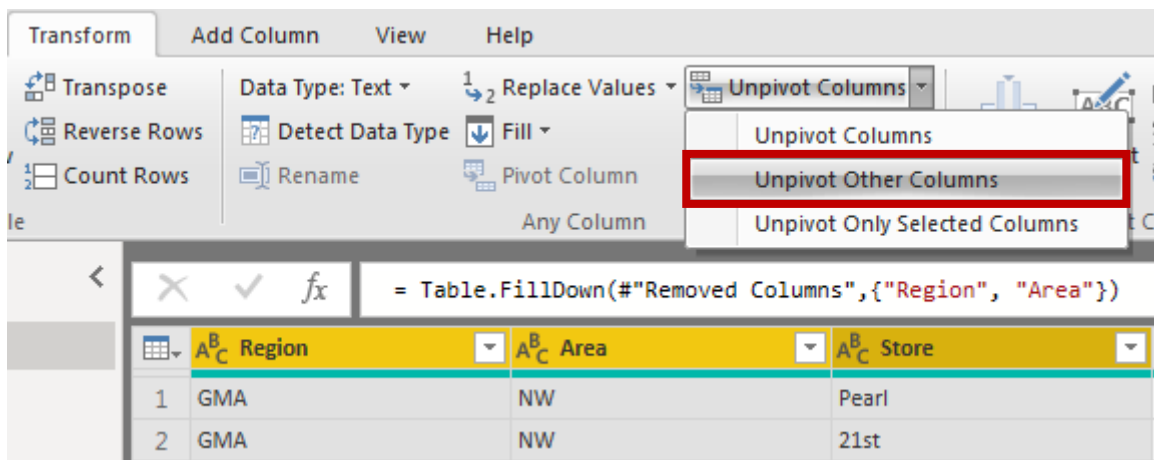
Do the same steps for column **Area**.

Note the *null* values are now filled with the values of the cells from above them.

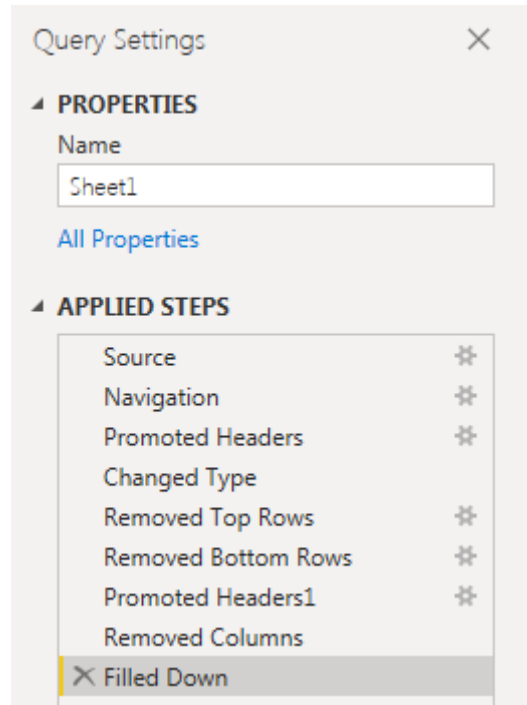


14. Select the columns **Region**, **Area**, and **Store** by clicking on **Region**, then while holding down [SHIFT] on the keyboard click on **Store**.

15. Under the **Transform** ribbon click on the dropdown arrow next to **Unpivot Columns** and select "**Unpivot Other Columns**".



16. The transform part is complete. Notice the right pane labelled **Query Settings**. It has recorded the applied steps the table went through to transform the data of the initial table.



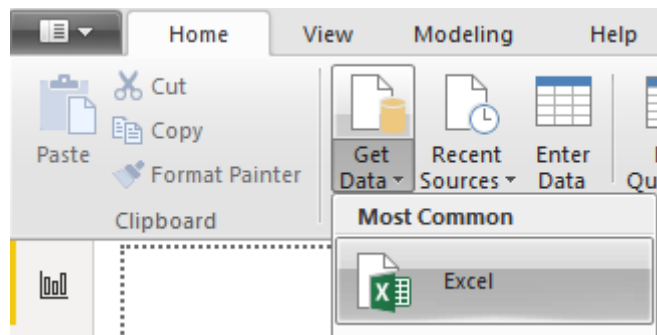
EXERCISE 2

This exercise aims to have the user create a User-Defined Function (UDF) from a series of steps done on a spreadsheet. The UDF then can be invoked to apply to several spreadsheets within a single folder.

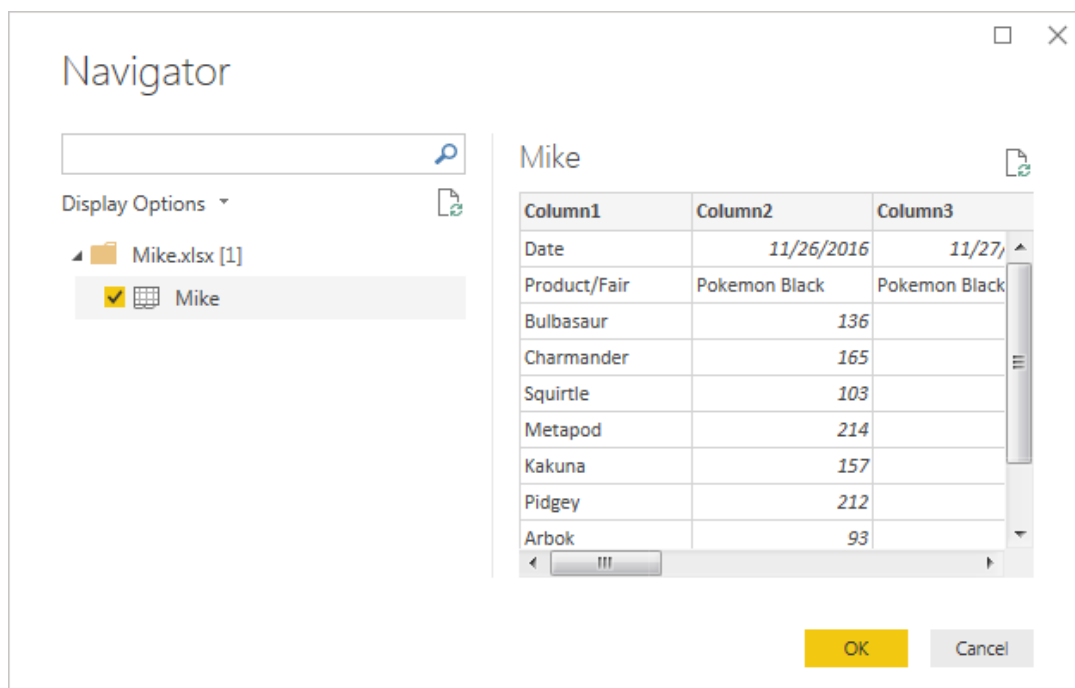
This is a way to make PowerQuery highly re-usable and portable, not to mention save a lot of time coding.

Performing the Transformations

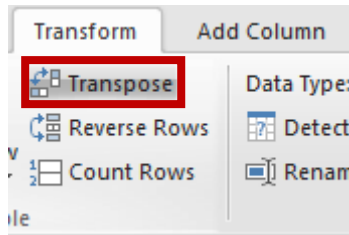
1. Launch an instance of Power BI.
2. Under the **Home** tab on the ribbon, click **Get Data > Excel**



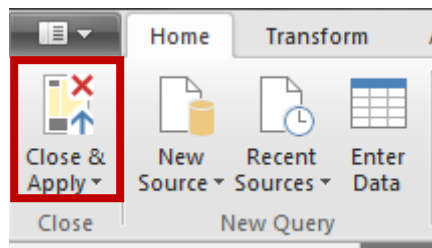
3. Browse to the folder containing the file "Mike.xlsx" in the "Exercise 2" folder and click **Open**.
4. On the **Navigator** window, select "Mike" and click **Edit**.



5. On the PowerQuery window, open the “**Transform**” tab on the ribbon and click “**Transpose**”.

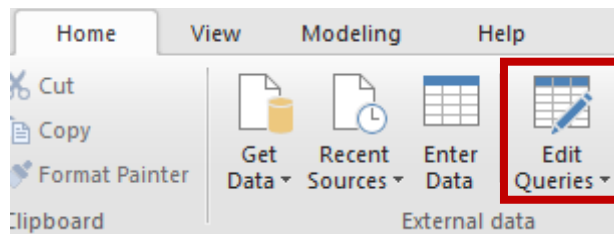


6. Execute **Use First Row as Headers**.
7. Select columns **Date** and **Product/Fair**. Then execute **Unpivot Other Columns**.
8. Select **Close & Apply** to load the data into the Power BI data model.

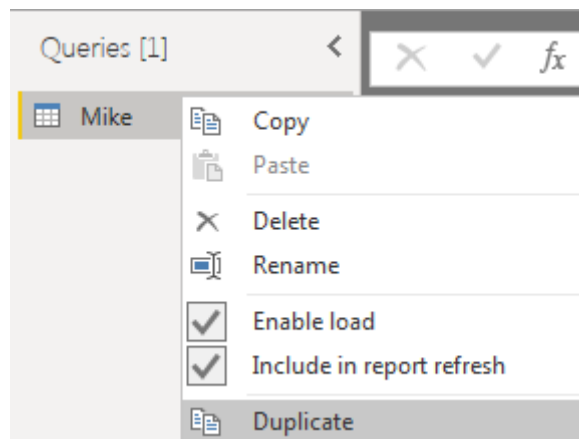


Creating the UDF

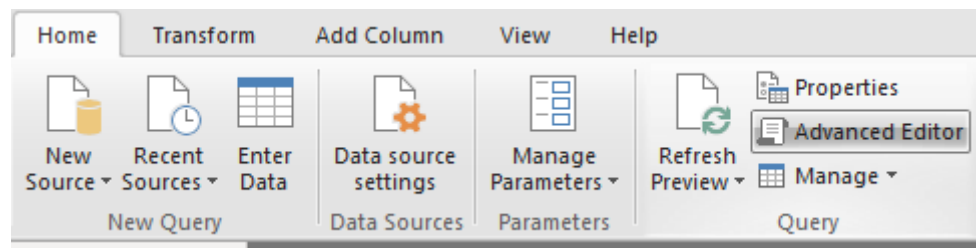
9. Return to PowerQuery by clicking on **Edit Queries**.



10. Duplicate the query named “Mike” by right-clicking on the “Mike” query under the list of Queries on the right pane, and selecting **Duplicate**.



11. Select the **Home** tab on the ribbon and click on **Advanced Editor**.

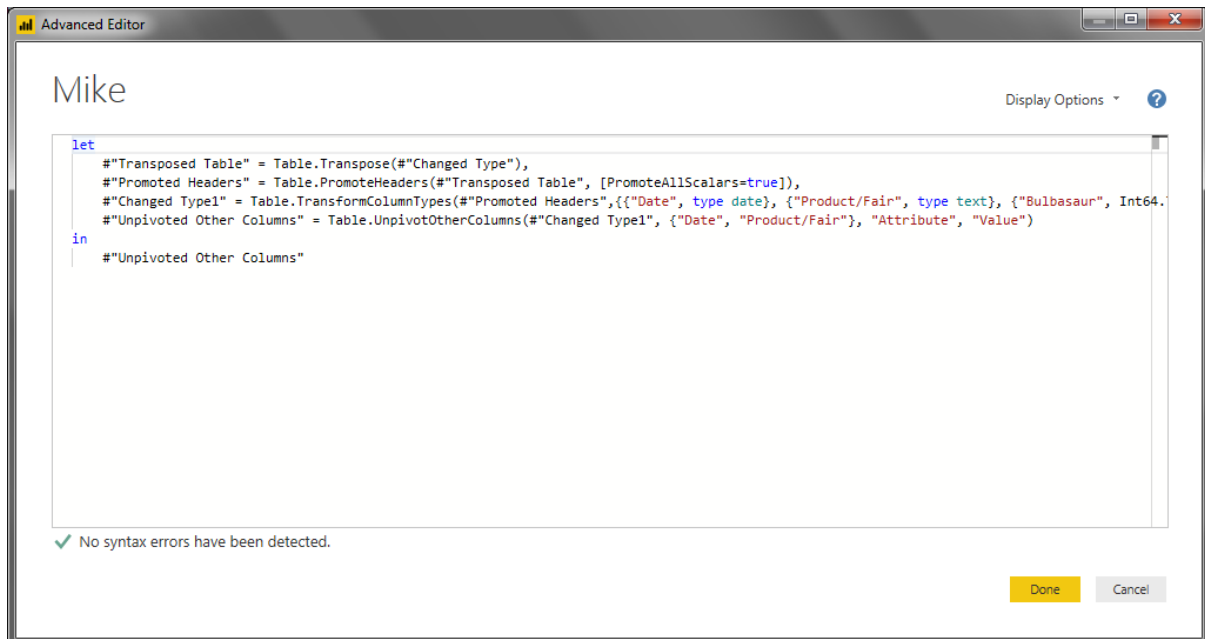


You should see another window as shown below:



This window contains the **M Code** for the applied steps.

12. Remove the line that contains with "Source = ...", "Mike_Sheet", and "#Changed Type" – three (3) lines removed in total. What remains should look like below screenshot.



13. On the very first line, before `let`, input the following:

`(InputTable as table)=>`

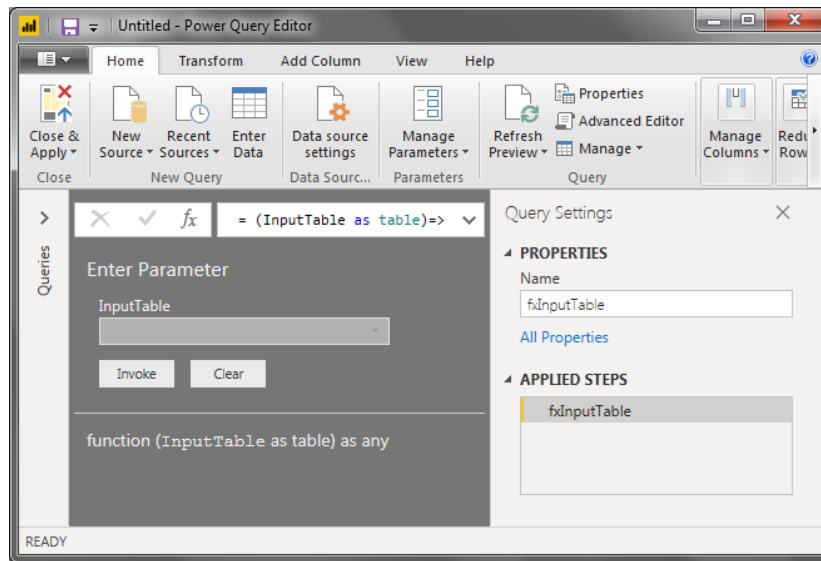
Replace `Table.Transpose("#Changed Type")`, with the following:

`Table.Transpose(InputTable)`



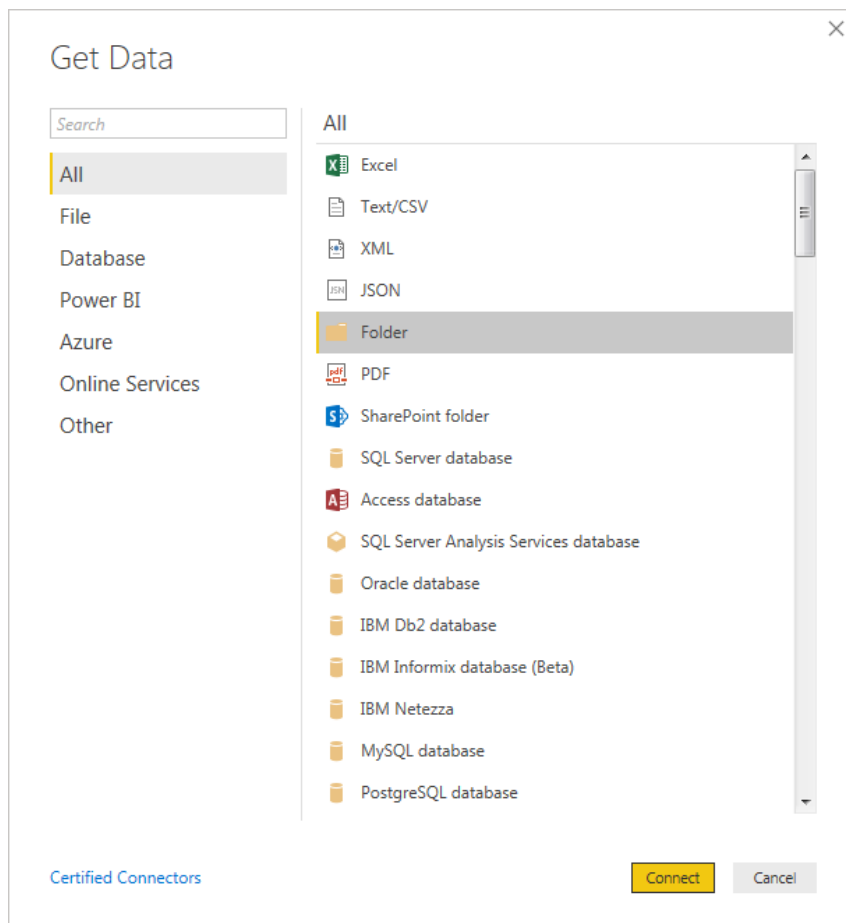
After making the edits, click **Done**

14. Rename this query to **fxInputTable** on the Workbook Queries pane.

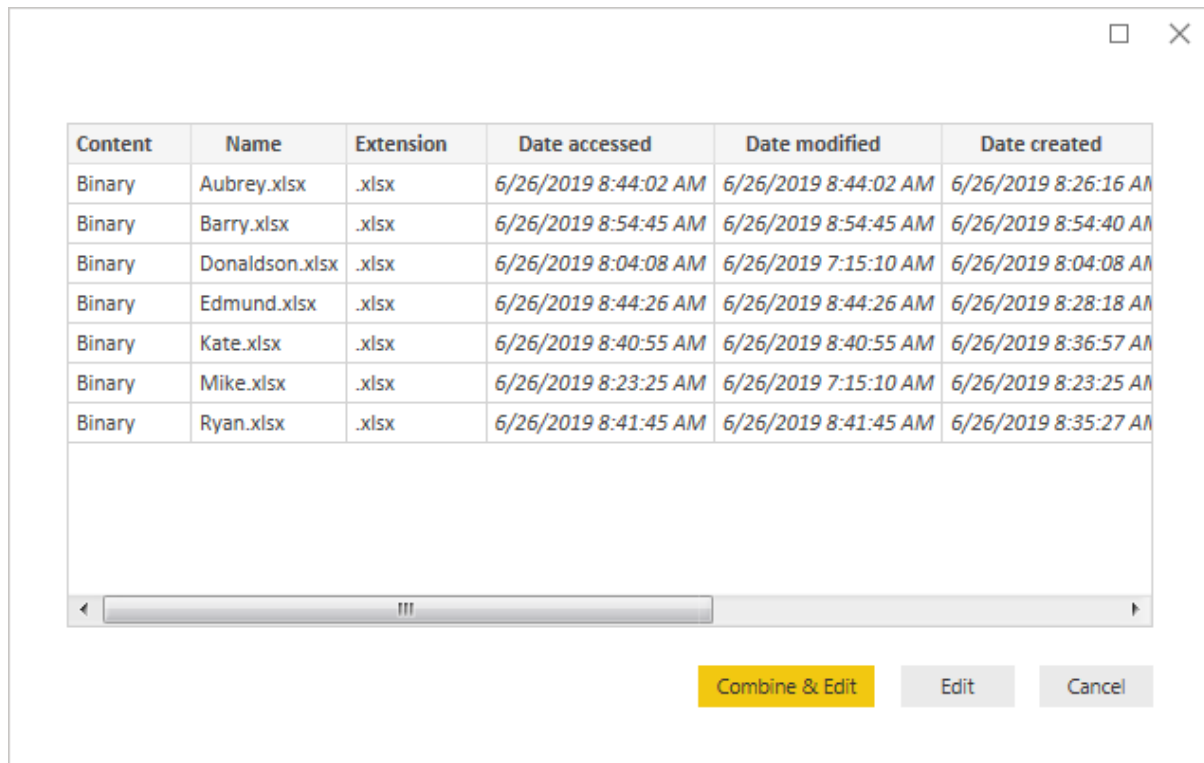


Invoking the UDF

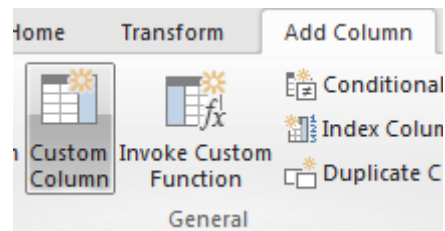
15. On the Home tab of the ribbon in the PowerQuery window, select **New Source > Folder > Connect**. Then browse to the folder **Data** within the **Exercise 2** folder.



16. The above step will load all metadata from the folder, and present another window similar to the screenshot below. Click **Edit** to continue.



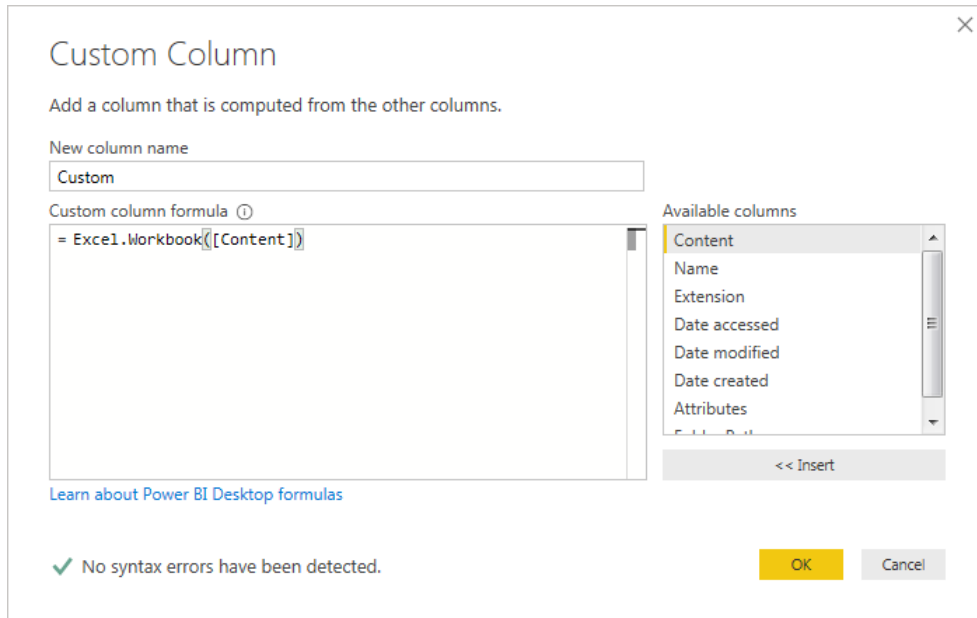
17. On the **Add Column** tab of the ribbon, click **Custom Column**.



18. Input the following on the Custom column formula:

Excel.Workbook([Content])

And hit **OK**.

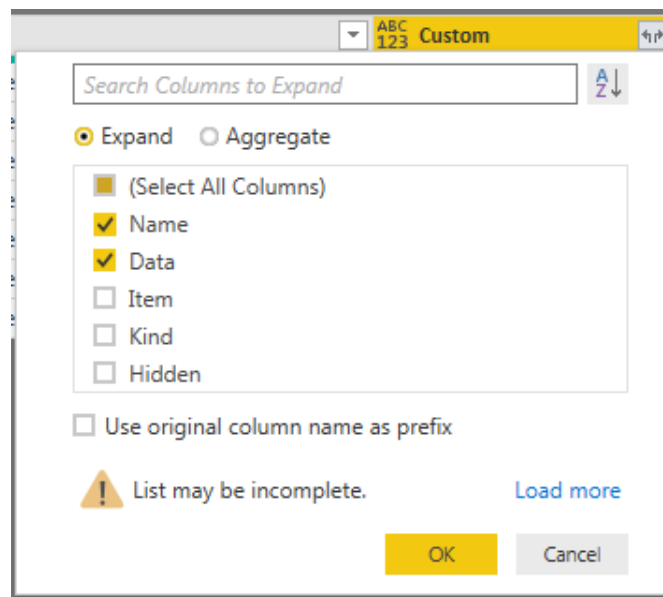


Notice another column, named **Custom**, added to the end of the table.

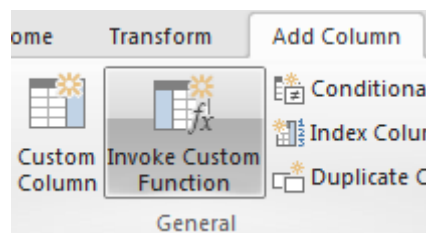
19. Click on the diverging arrows icon next to the newly added column name **Custom**.

ABC 123	Custom	
Table		
Table		
Table		
Table		
Table		
Table		
Table		

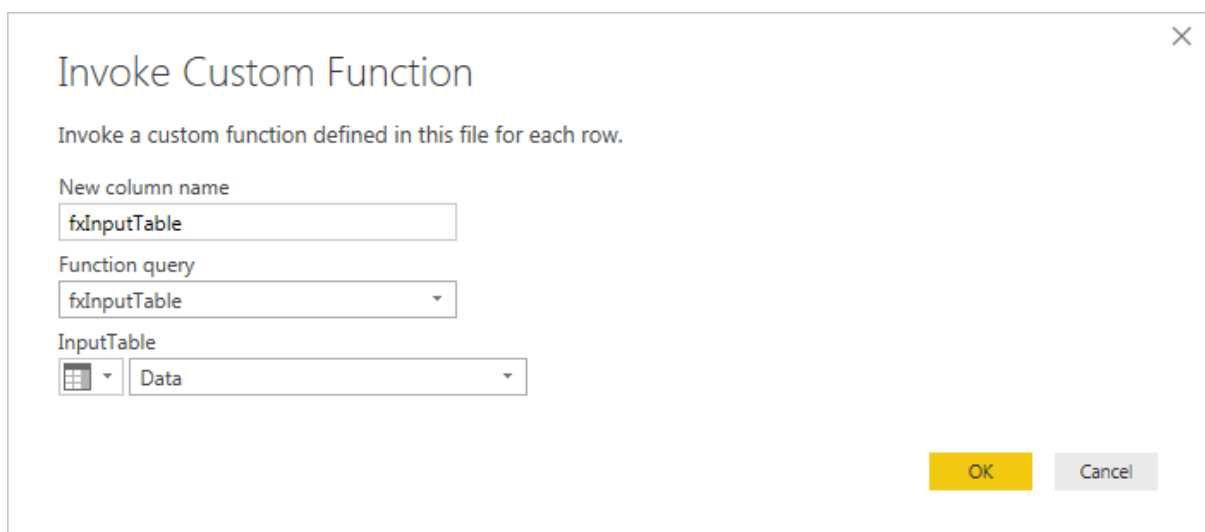
20. Select only **Name** and **Data**. Click **OK**.



21. On the **Add Column** tab of the ribbon, click **Invoke Custom Function**.



22. Under **Function query**, click on the dropdown arrow and select **fxInputTable**.
Another dropdown will appear. Select **Data** and click **OK** to accept.



23. Select the **Name.1** and newly created **fxInputTable** columns, right click on of the selected columns, and select **Remove Other Columns**.

ABC 123 Name.1	ABC 123 Data	ABC 123 fxInputTable
Aubrey	Table	Table
Barry	Table	Table
Donaldson	Table	Table

24. Expand the tables by clicking on the diverging arrows icon. Leave all columns selected, and click **OK**.

ABC 123 Name.1

ABC 123 fxInputTable

A


Z

↓

☒ Expand
 ☐ Aggregate

☒ (Select All Columns)
 ☒ Date
 ☒ Product/Fair
 ☒ Attribute
 ☒ Value

☐ Use original column name as prefix

 List may be incomplete.
 [Load more](#)

OK

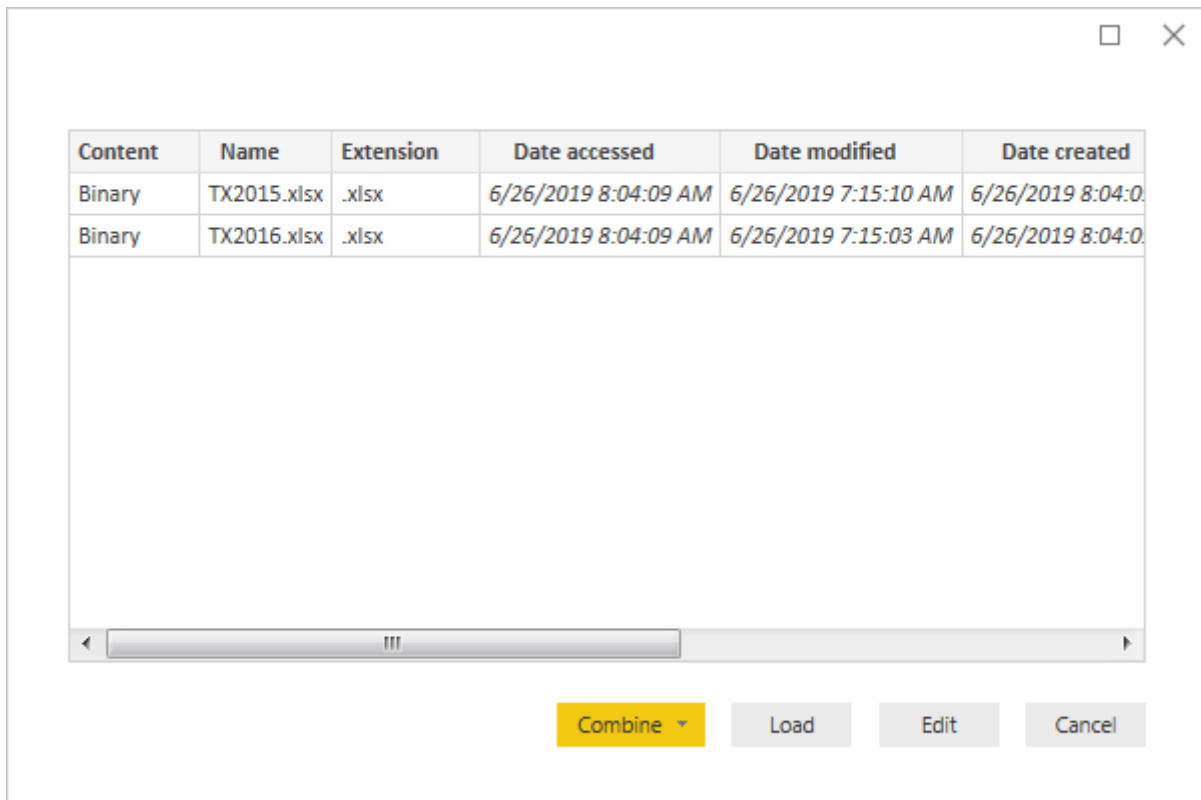
Cancel

25. Click on **Close & Apply**, to load the transformed dataset to PowerBI.

EXERCISE 3

This exercise aims to have the user apply their learnings from previous exercises and transform a set of transaction data that can be used in generating visualizations.

1. Launch an instance of Microsoft Power BI.
2. Load the files from folder **Exercise 3** on the USB provided for you. Hit **Edit** to continue.



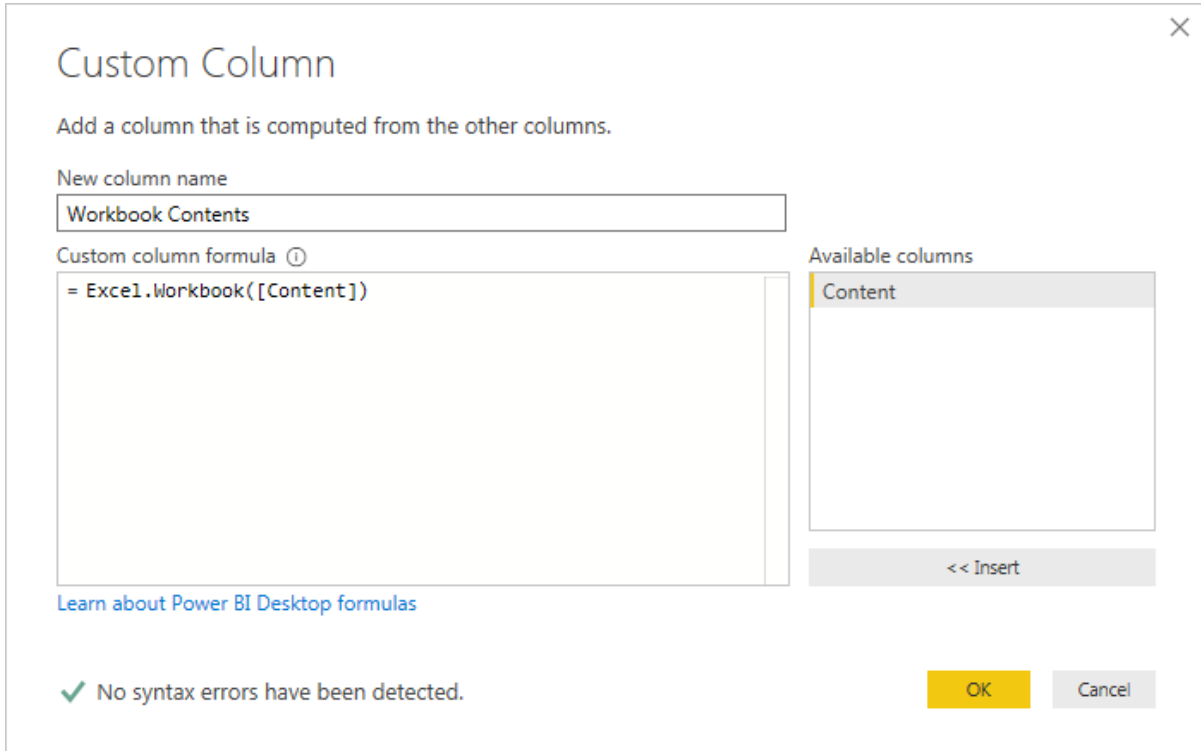
BEST PRACTICE OPTIONAL STEP:

Filter the Extension column to only include ".xlsx". This will ensure that only Excel files within the folder are processed.

3. **Remove Other Columns** such that only the column **Content** remains.

4. Add a **Custom Column**. Name this new column **Workbook Contents**. The Custom column formula is

Excel.Workbook([Content])



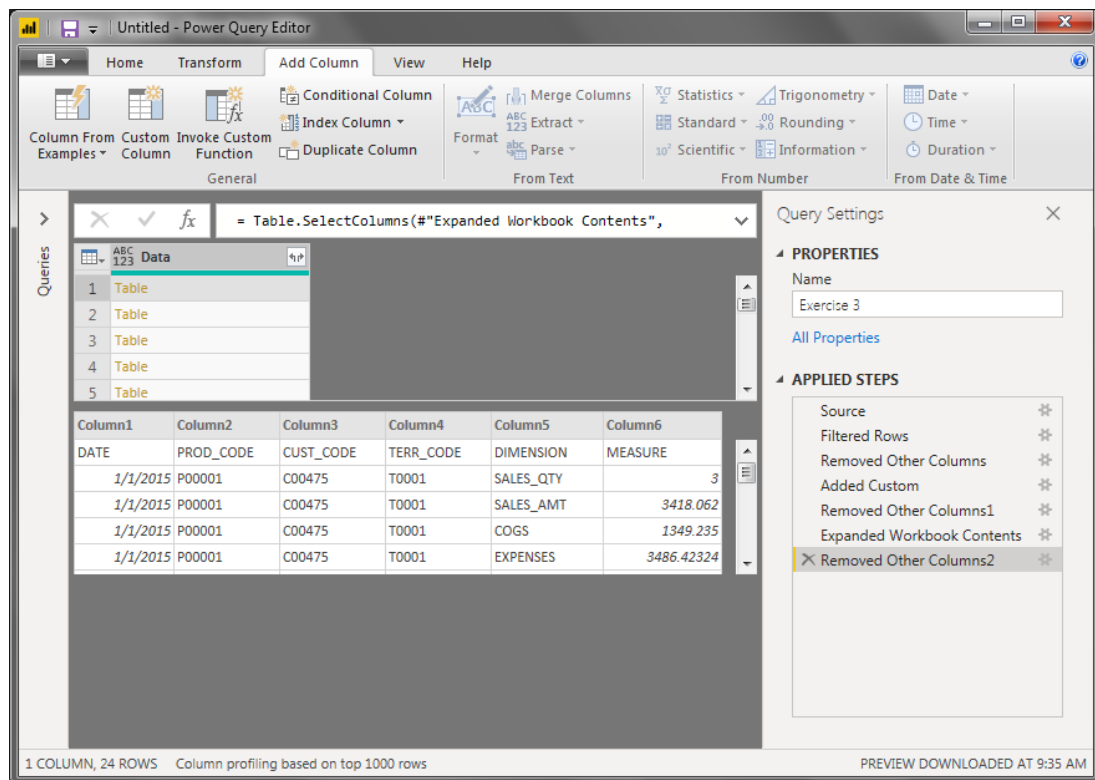
5. Remove the column **Content**.
6. Click on the diverging arrows icon next to **Workbook Contents** to expand. Select all but do not tick "Use original column name as prefix".

BEST PRACTICE OPTIONAL STEP:

Filter the Kind column to only include "Sheet". This will ensure that only Sheets inside each workbook are processed.

7. Remove **Other Columns** so that only the column **Data** remains.

8. Click on the one of the cells under column **Data**. You will get a preview of its contents.

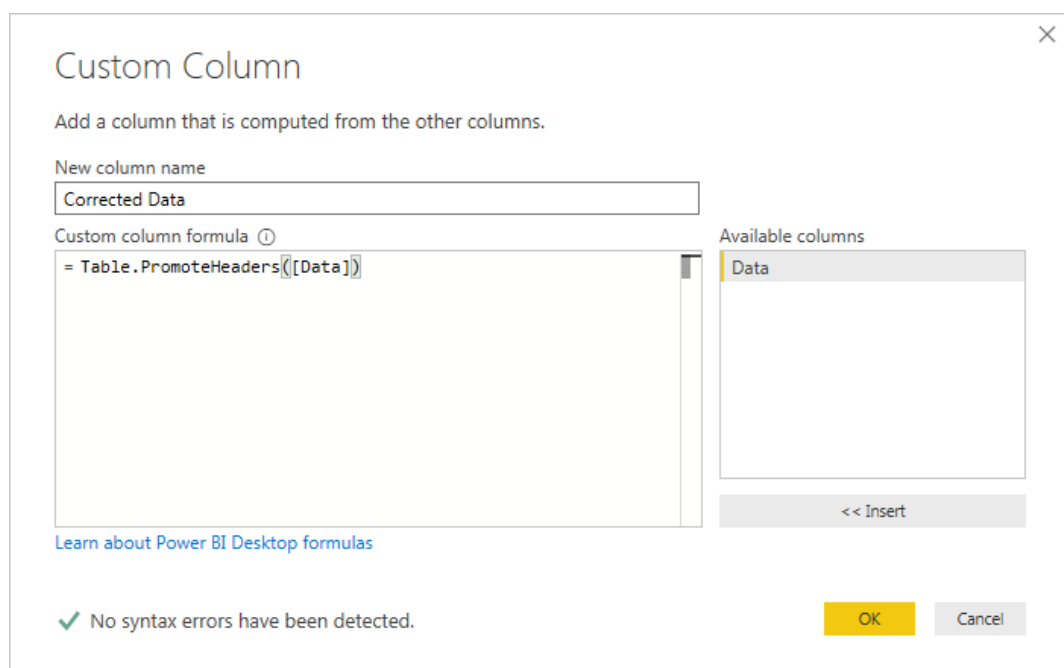


Notice that the column headers of each individual “Table” need to be corrected.

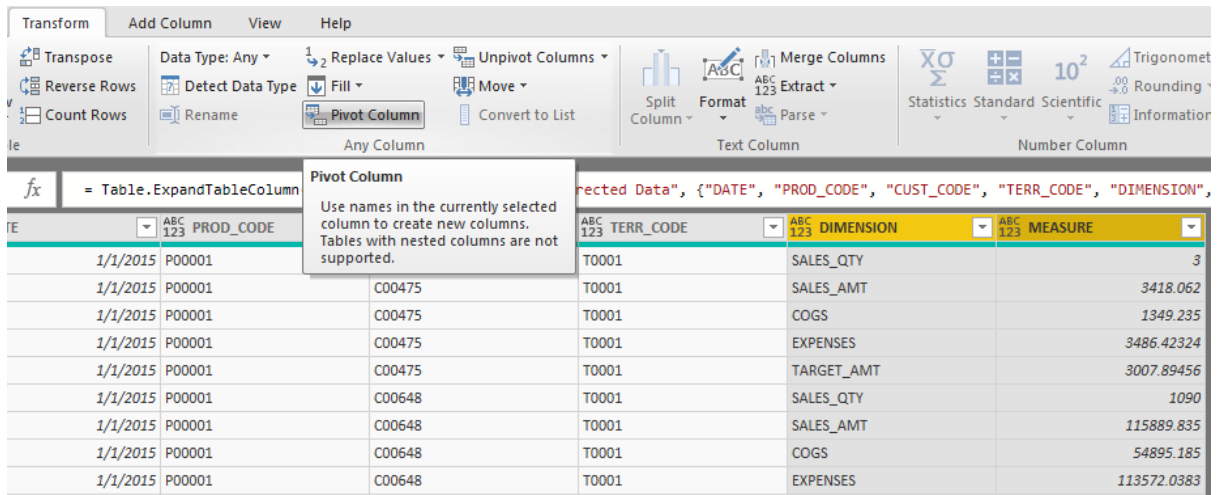
9. Add a custom column named **Corrected Data** with the following formula:

Table.PromoteHeaders([Data])

Then click OK.



10. Remove the column **Data** (only **Corrected Data** remains). Then, expand **Corrected Data** by clicking on the diverging arrows icon. Keep all fields ticked, and leave "Use original column name as prefix" unticked.
11. Select columns **DIMENSION** and **MEASURE**. Then under **Transform** tab on the ribbon, select **Pivot Column**.



DATE	PROD_CODE	TERR_CODE	DIMENSION	MEASURE
1/1/2015	P00001	C00475	SALES_QTY	3
1/1/2015	P00001	C00475	SALES_AMT	3418.062
1/1/2015	P00001	C00475	COGS	1349.235
1/1/2015	P00001	C00475	EXPENSES	3486.42324
1/1/2015	P00001	C00475	TARGET_AMT	3007.89456
1/1/2015	P00001	C00648	SALES_QTY	1090
1/1/2015	P00001	C00648	SALES_AMT	115889.835
1/1/2015	P00001	C00648	COGS	54895.185
1/1/2015	P00001	C00648	EXPENSES	113572.0383

12. Use "MEASURE" for Values Column.

Pivot Column

Use the names in column "DIMENSION" to create new columns.

Values Column ①

MEASURE

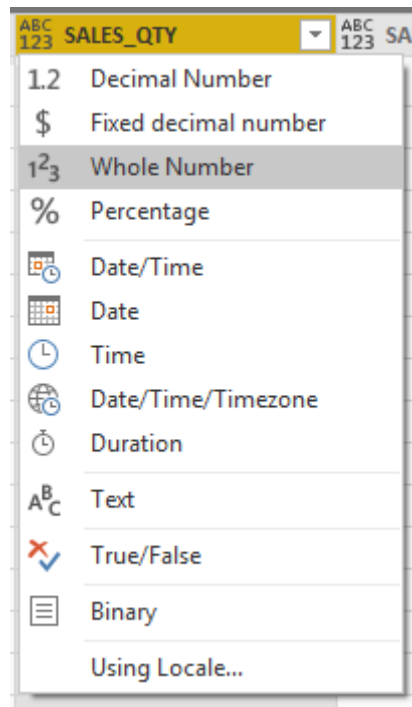
Advanced options

Learn more about Pivot Column

OK

Cancel

13. Set each column to the correct data type. This can be done by clicking on the ABC123 icon next to the column name, and selecting the appropriate data type.



The correct data types are as follows:

Column(s)	Type
DATE	Date
SALES_QTY	Whole Number
SALES_AMT, COGS, EXPENSES, TARGET_AMT	Decimal Number

14. **Close & Apply** to load the data to Power BI.