Using Power BI

ETL software is used to automate the redundant manipulation of data. In this exercise you will use Power BI to:

Extract the Data file(s) from Excel

Transform the data

Load the transformed data into Excel

The sales data for *Cuddly Stuffed Animals* are stored in an Excel sheet titled January Sales Data. The column headings are in the ninth row, some rows have partial duplicate data and some rows are blank.

	A	B	C	D	E	F	G	H	I
7					Cuddly Stuffed Animals				
8					January Sales Data				
9	Sales Order #	Salesperson #	Salesperson	Title	Profit Margin	Region #	Region	State	Product
10	35005	1304	Martin Van Bure	Sales Associate I	1,908.00		3	South	TN	Stuffed Horse
11									Stuffed Horse
12									
13	35006	1302	Grover Clevelan	Senior Sales Associate	936.00		1	Midwest	NE	Stuffed Panda
14									Stuffed Panda
15									
16	35009	1301	Thomas Jeffers	Sales Associate I	2,028.00		2	Northeast	PA	Stuffed Panda
17					2,028.00				Stuffed Panda
18									
19	35011	1305	George Washin	Sales Associate I	1,560.00		1	Midwest	MN	Stuffed Panda
20					1,560.00				Stuffed Panda
21									
22	35014	1303	Ulysses Grant	Senior Sales Associate	1,215.00		1	Midwest	MN	Stuffed Giraffe
23					1,215.00				Stuffed Giraffe
24									
25	35017	1305	George Washin	Sales Associate I	708.00		2	Northeast	RI	Stuffed Elephant
26					708.00				Stuffed Elephant
27									
28	35025	1303	Ulysses Grant	Senior Sales Associate	1,272.00		2	Northeast	ME	Stuffed Horse
29									Stuffed Horse
30									
31	35026	1301	Thomas Jeffers	Sales Associate I	1,716.0					uin
32									uin
33									
34	35027	1301	Thomas Jeffers	Sales Associate I	1,215.0					fe
35									fe
36									
37	35032	1303	Ulysses Grant	Senior Sales Associate	1,584.0					uin
38									uin

Introduction | **January Sales Data** | Products | Salesperson | Regi

Sample Report

Product	Units	Sales	Profit Margin
Stuffed Elephant	4,392	$109,800.00	$ 64,782.00
Stuffed Giraffe	8,052	$169,092.00	$ 90,585.00
Stuffed Horse	12,120	$266,640.00	$160,590.00
Stuffed Panda	7,488	$172,224.00	$ 97,344.00
Stuffed Penguin	5,016	$100,320.00	$ 55,176.00
Stuffed Unicorn	3,060	$ 73,440.00	$ 40,545.00
Total	40,128	$891,516.00	$509,022.00

Note: You cannot extract data from an open Excel file so make sure that the Excel file is closed before starting to input the data.
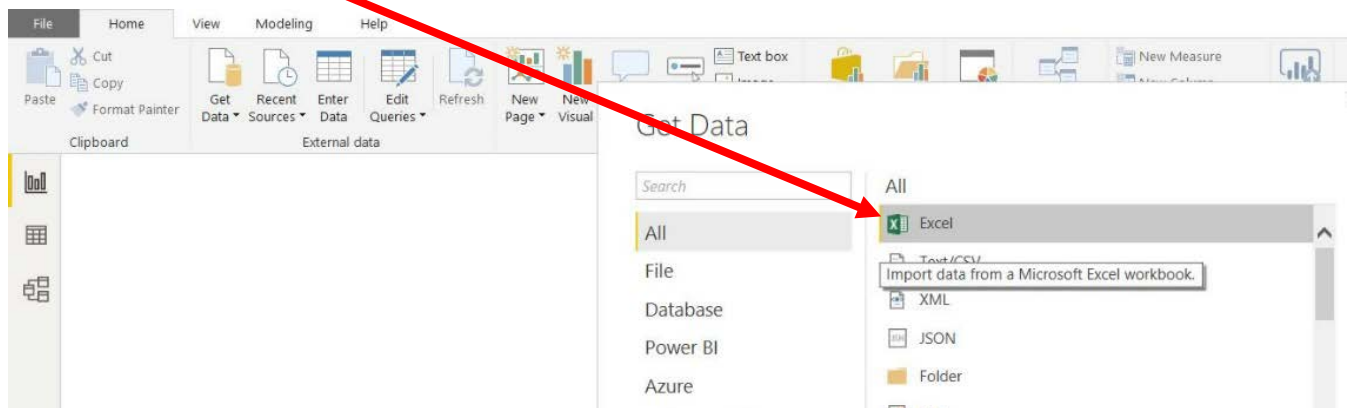
Start Power BI Desktop

Operation 1: <mark>Get Data</mark> – Input Excel file

Select "<mark>Get Data</mark>" from the "External data" group under the "Home" tab.
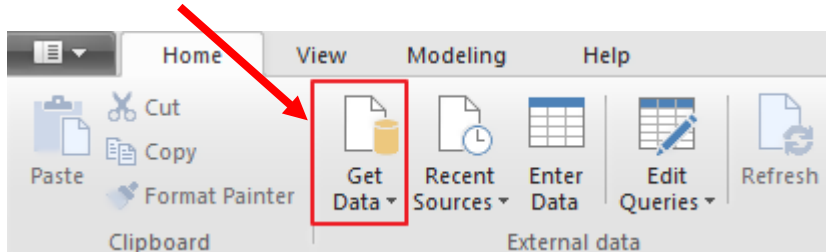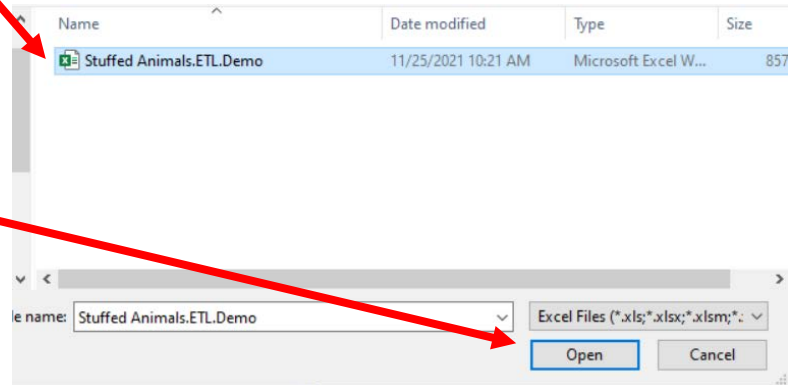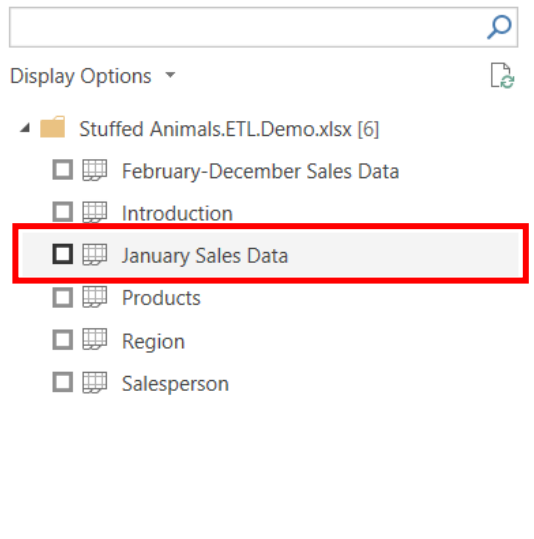

Select "Excel".

Select the file.

Select "Open".

Highlighting the desired sheet does **not** activate the Load and Edit buttons.

Double click on "January Sales Data" or you can select the check box on the left of "January Sales Data".

The data preview will be displayed.

Select "Transform Data" or "Edit" and a new window called Power Query Editor opens.

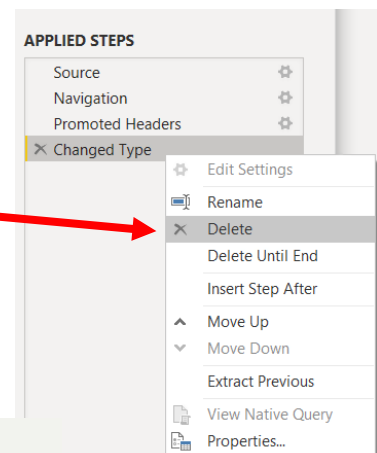In the Power Query window, Power BI Desktop records each data modifications in the "Applied Steps" section.
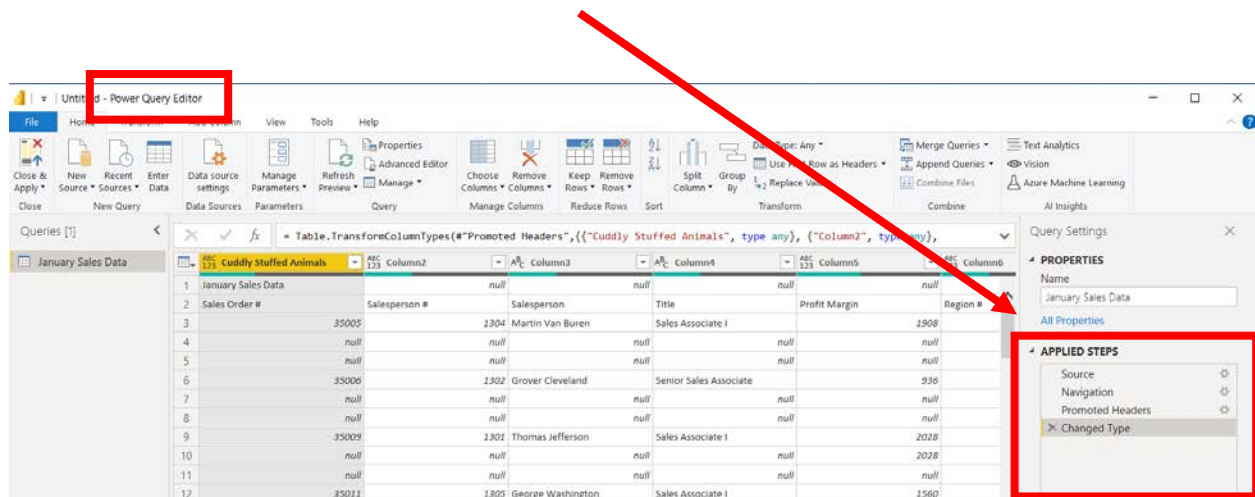
The program added the steps that it performed as Power BI inputted the data. Note that it did not correctly find the row that contains the "Headers". Therefore, some steps must be deleted.
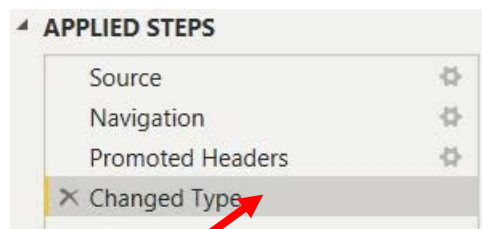
Select the step, "Changed Type".

Right click and select "Delete".

Delete the step "Promoted Headers", leaving us with two steps.

Operation 2: <mark>Fix the Headers</mark> – The field headers are in the third row of the source document. Remove the top two rows to move the headers to the first row and then promote them to the Query Editors' headers.

	ABC 123 Column1	ABC 123 Column2	ABC 123 Column3	ABC 123 Column4	ABC 123 Column5	ABC 123 Column6
1	Cuddly Stuffed Animals	null	null	null	null	
2	January Sales Data	null	null	null	null	
3	Sales Order #	Salesperson #	Salesperson	Title	Profit Margin	Region #
4	35005	1304	Martin Van Buren	Sales Associate I	1908	
5	null	null	null	null	null	
6	null	null	null	null	null	
7	35006	1302	Grover Cleveland	Senior Sales Associate	936	
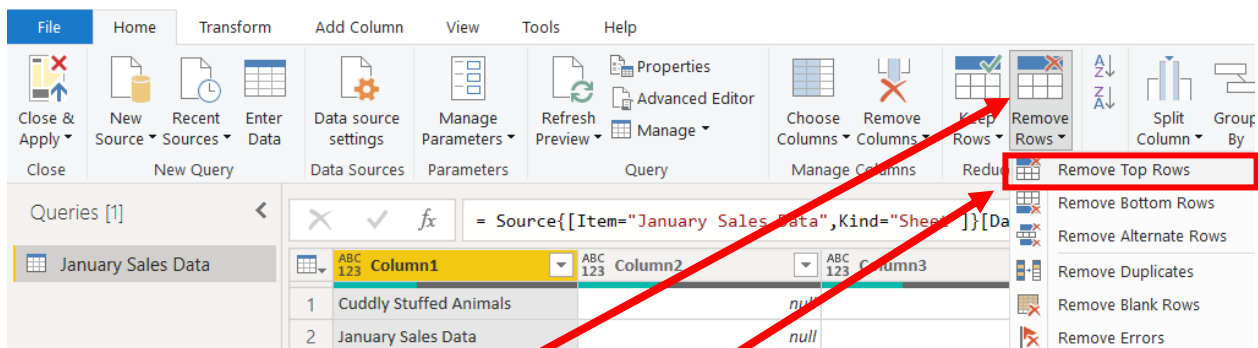
Select "Remove Rows".

Select "Remove Top Rows".

Enter a "2" to remove top three rows.

Remove Top Rows

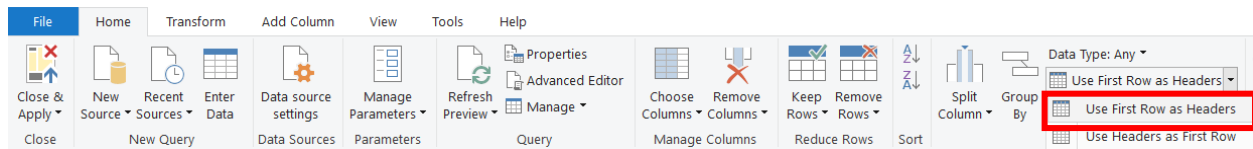Specify how many rows to remove from the top.

Number of rows

2

Then click "OK".

OK   Cancel

	ABC 123 Column1	ABC 123 Column2	ABC 123 Column3	ABC 123 Column4	ABC 123 Column5	ABC 123 Column6
1	Sales Order #	Salesperson #	Salesperson	Title	Profit Margin	Region #
2	35005	1304	Martin Van Buren	Sales Associate I	1908	

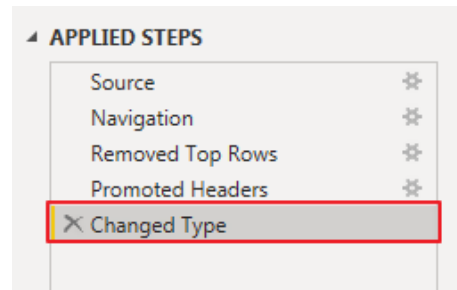To promote the first row in the current data to Query Editors' headers, select "Use First Row as Headers".

Note: After you perform a transformation, the changes are recoded in the "Applied Steps" section. If you incorrectly did a step you can easily delete the step and do it again.

Operation 3: <mark>Filter Rows</mark> – Exclude rows that do not contain Sales Order Number

You can observe that there are several rows that do not contain a sales order number. We can add a filter step which will remove such rows. Note that we are not deleting anything from the source data file. This will just filter such data out and exclude these records from the final table.

Click on the "drop-down" next to "Sales Order #", Uncheck "null" and Click "OK".
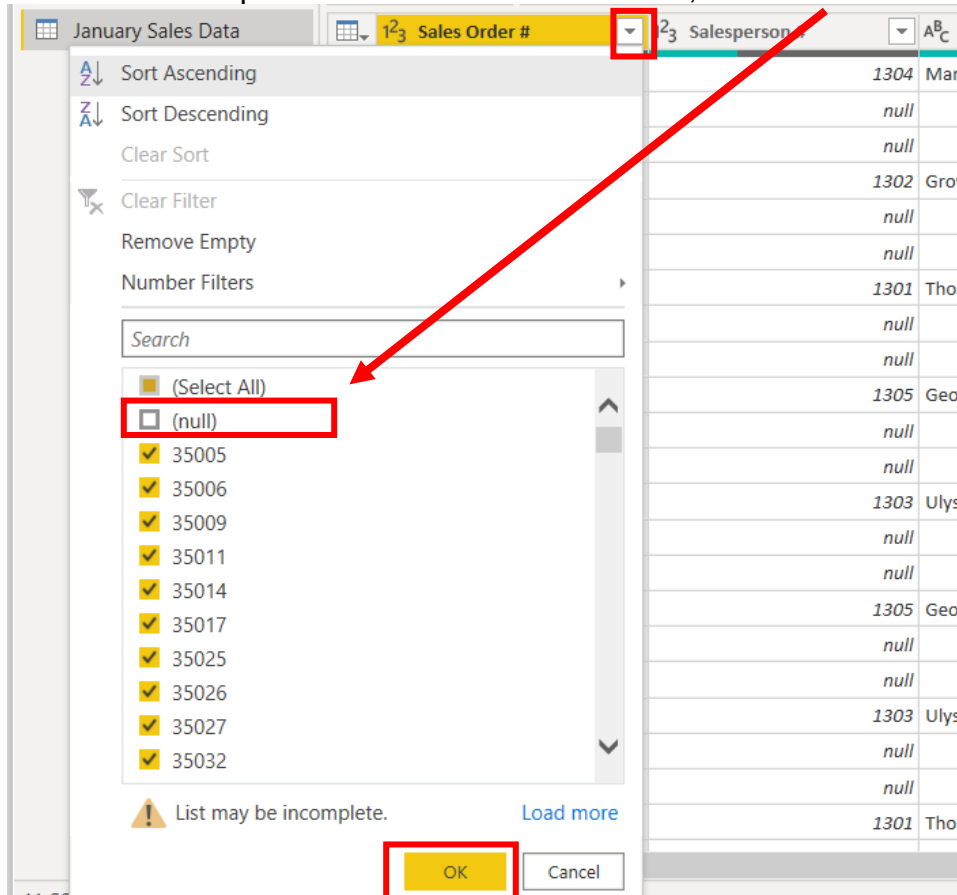
Notice that "Filtered Rows" step is automatically added on the right side in the "Applied Steps" section.
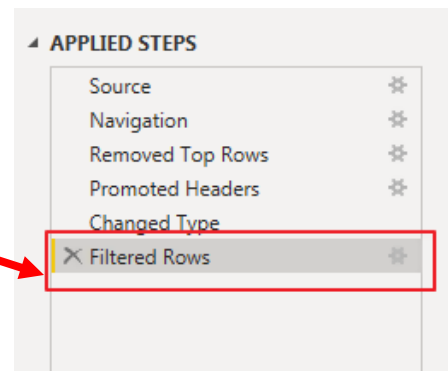
The Data in the Fields Must Conform to Predefined Rules.

In the original Excel file, The Cuddly Stuffed Animals company provided the only acceptable values for "Product", "Salesperson" and "Region". Sometime data is entered incorrectly into the data.
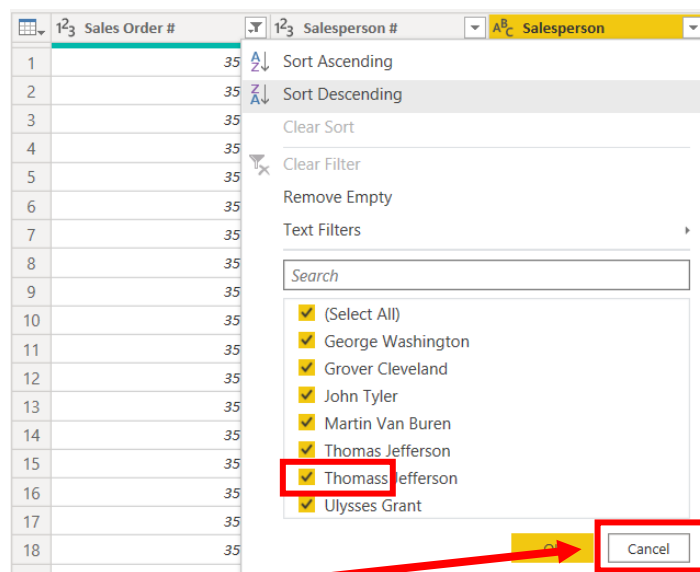
Code	Salesperson
1301	Thomas Jefferson
1302	Grover Cleveland
1303	Ulysses Grant
1304	Martin Van Buren
1305	George Washington
1306	John Tyler

Operation 4 :– Verify that the Salesperson field is correct

Click on the "drop-down" next to "Salesperson".

The "Salesperson" field can only contain one of the six salespersons. Browse the data in the "Salesperson" field to verify that all the data is correct. Make any changes that are required.

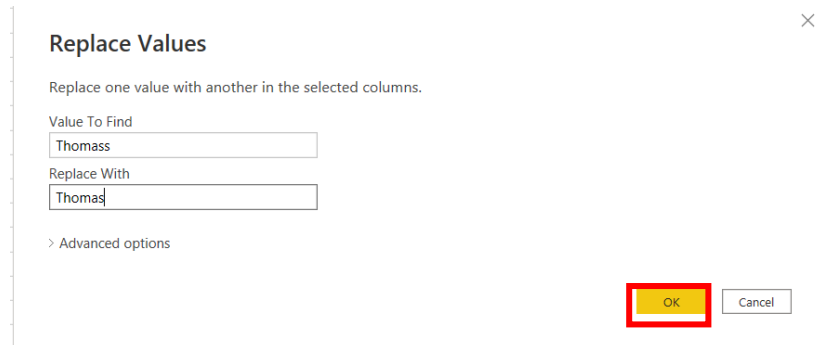"Thomass" must be replaced with "Thomas".
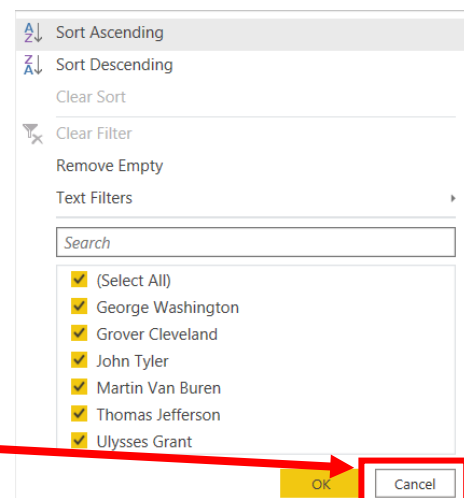

Select Cancel.
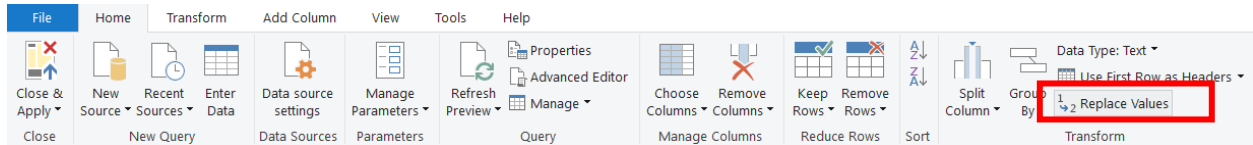
Select Replace Values

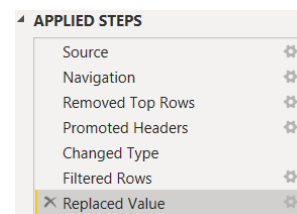Value to find "Thomass"
Replace with "Thomas"

Select OK.

Click on the "drop-down" next to "Salesperson".

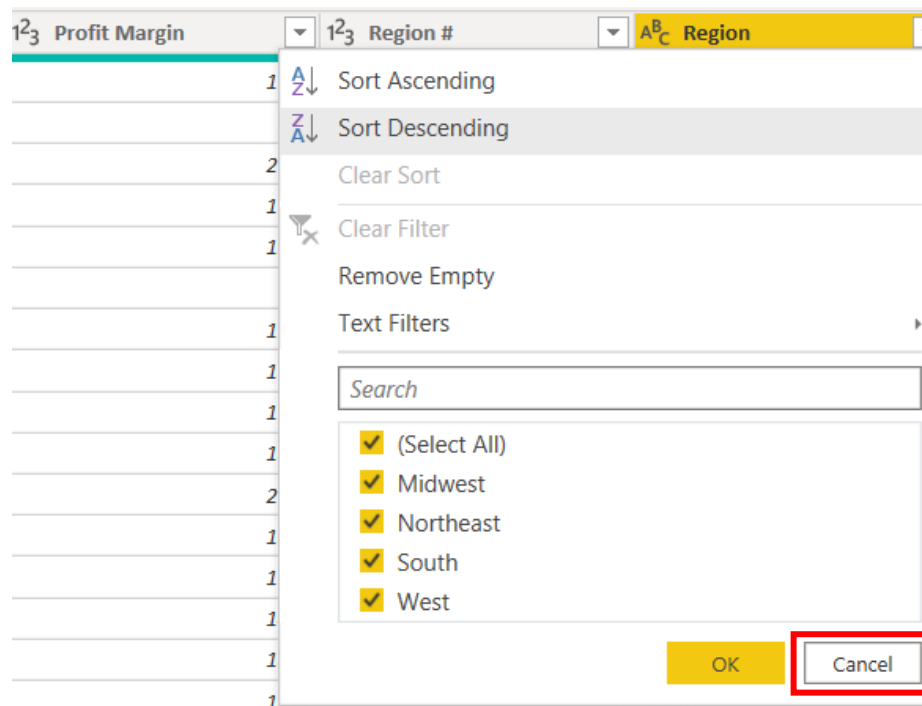Select Cancel.

A step has been added to the Applied Steps.

Operation 5 :– Verify that the Product field is correct

Click on the "drop-down" next to "Region".

Code	Region
1	Midwest
2	Northeast
3	South
4	West

The "Region" field can only contain one of the six products. Browse the data in the "Region" field to verify that all the data is correct. Make any changes that are required**.**
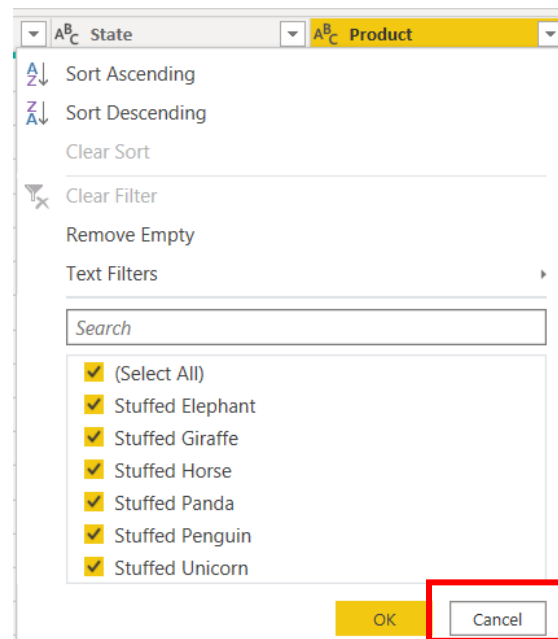
No changes must be made, select Cancel.

Operation 6 :– Verify that the Product field is correct

| Product ID	Product
101	Stuffed Penguin
102	Stuffed Horse
103	Stuffed Unicorn
104	Stuffed Giraffe
105	Stuffed Panda
106	Stuffed Elephant

Click on the "drop-down" next to "Product".

The "Product" field can only contain one of the six products. Browse the data in the "Product" field to verify that all the data is correct. Make any changes that are required.

No changes must be made, select Cancel.

Operation 7: Change the data type. Change the type of all numeric fields that will be used in calculations to "Fixed Decimal". The "# of units should have zero decimal places and dollar values should have two decimal places.

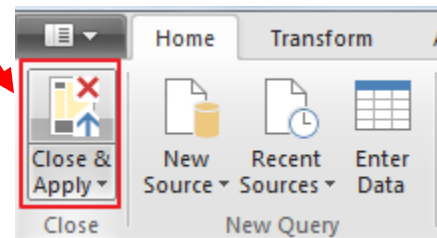Click on the 1²₃ in the left-hand corner of the field "Sales".

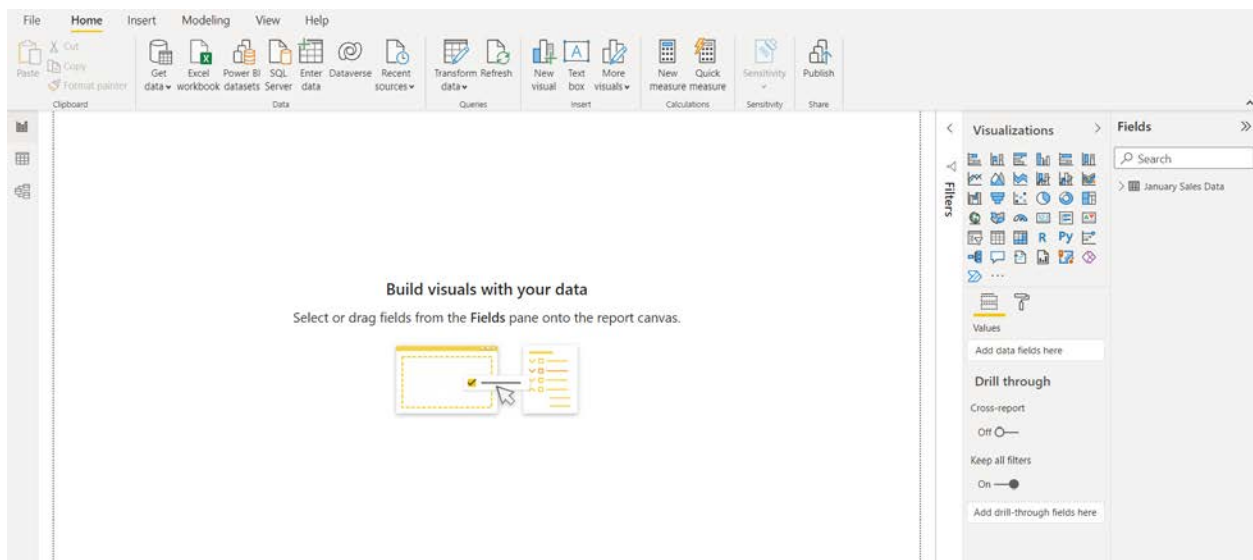Select "$ Fixed decimal number".

Do the same to the "Profit Margin" Field.
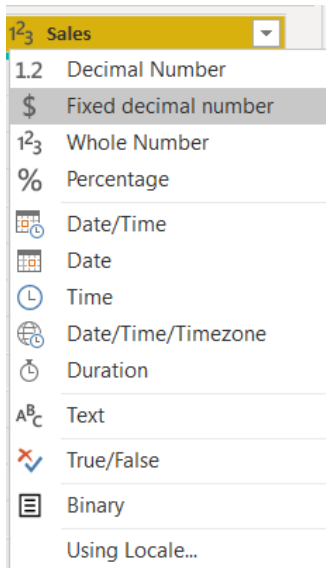
Operation 8: <mark>Load data</mark>. Click on "Close and Apply"

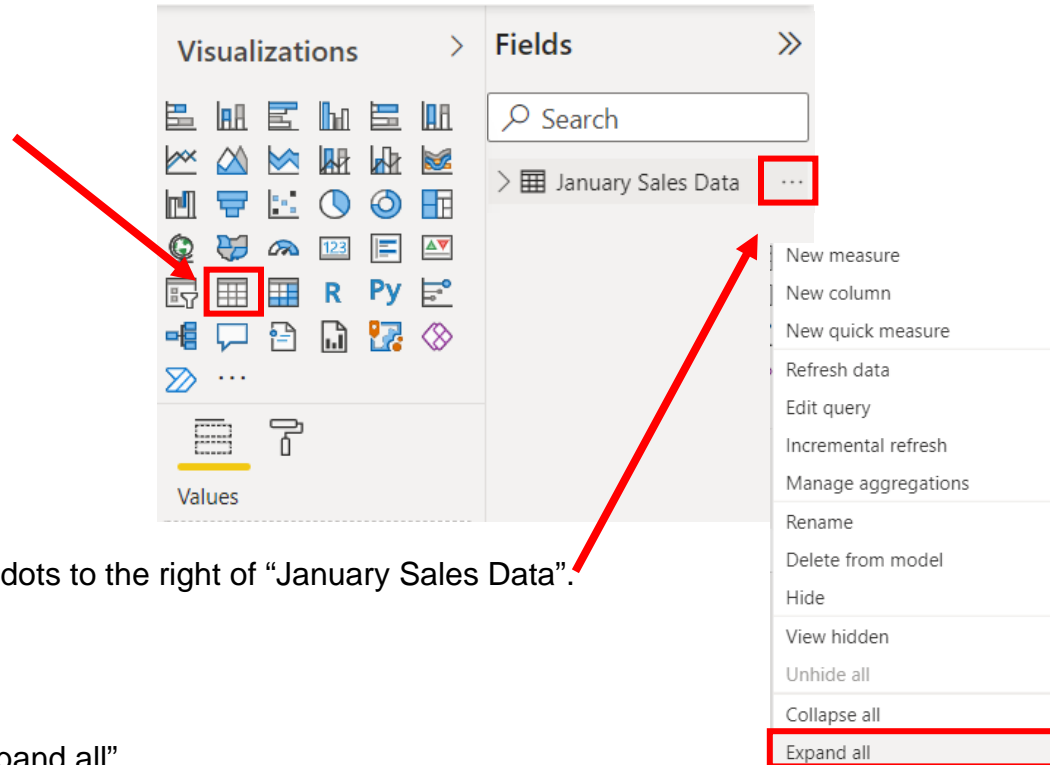This loads the data onto Power BI from Power Query
Editor.

Now you'll be back to the main window of Power BI.

Operation 9: Create a "==Table==".

Click on the Table icon under the Visualization section to add a table.

Select the dots to the right of "January Sales Data".

Select "Expand all".

The order that you select the fields is used to determine the order in which the fields are displayed in the table. Click on the box next to each field in the order that you want the fields to be displayed.

1. Product
2. # of units
3. Sales
4. Profit Margin
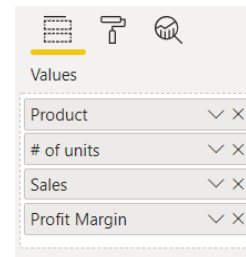
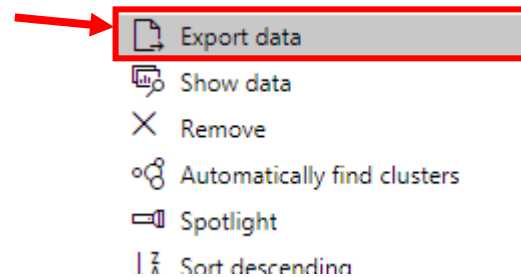| Product	# of units	Sales	Profit Margin
Stuffed Elephant	4392	109800	64782
Stuffed Giraffe	8052	169092	90585
Stuffed Horse	12120	266640	160590
Stuffed Panda	7488	172224	97344
Stuffed Penguin	5016	100320	55176
Stuffed Unicorn	3060	73440	40545
Total	**40128**	**891516**	**509022**

Operation 10: <mark>Exporting data to Excel</mark>

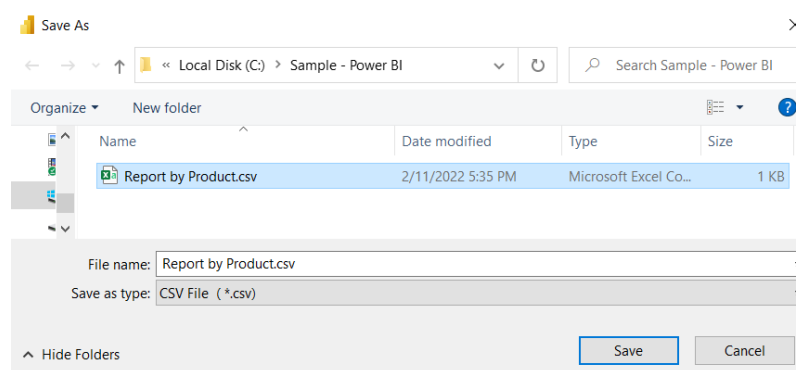Select the table and notice that you get More Options indicated by 3 dots (…) on the top or bottom right of the table.

Product	# of units	Sales	Profit Margin
Stuffed Elephant	4392	109800	64782
Stuffed Giraffe	8052	169092	90585
Stuffed Horse	12120	266640	160590
Stuffed Panda	7488	172224	97344
Stuffed Penguin	5016	100320	55176
Stuffed Unicorn	3060	73440	40545
Total	**40128**	**891516**	**509022**

Click on the More Options

Select "Export Data".

- Export data
- Show data
- ✕ Remove
- Automatically find clusters
- Spotlight
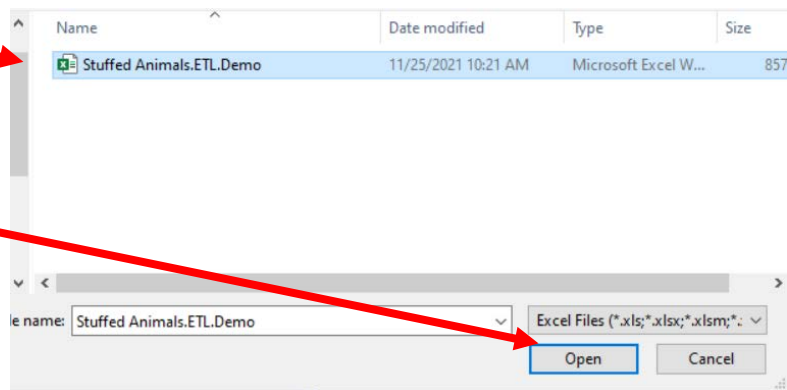- Sort descending

Type the name of the file "Report by Product.csv".

Select Save.
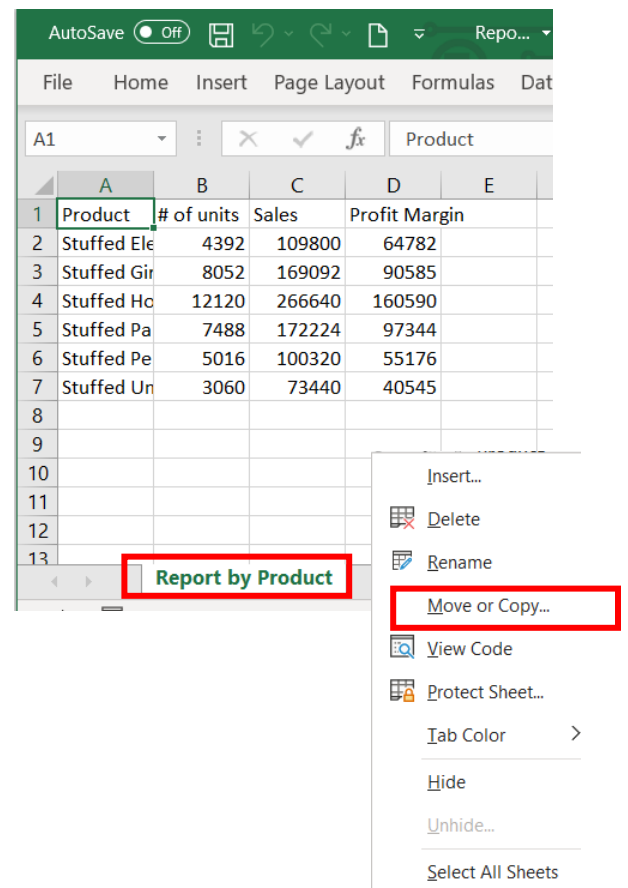
Open the original excel file

Select the file.

Select "Open".

To open the new report, navigate in file manager and double click on the "Report by Product.csv" file.

Copy this tab to the original Excel by right clicking on the tab Report by Product and selecting "Move or Copy".
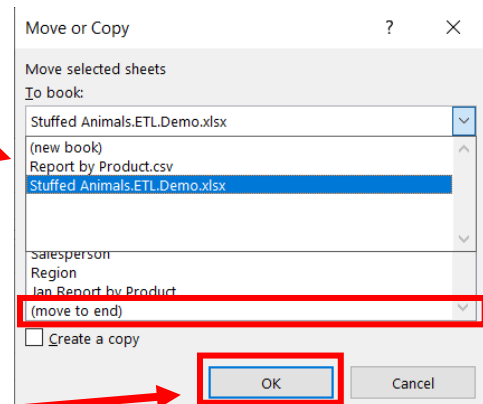
Select the name of the original file.

Then click "OK".

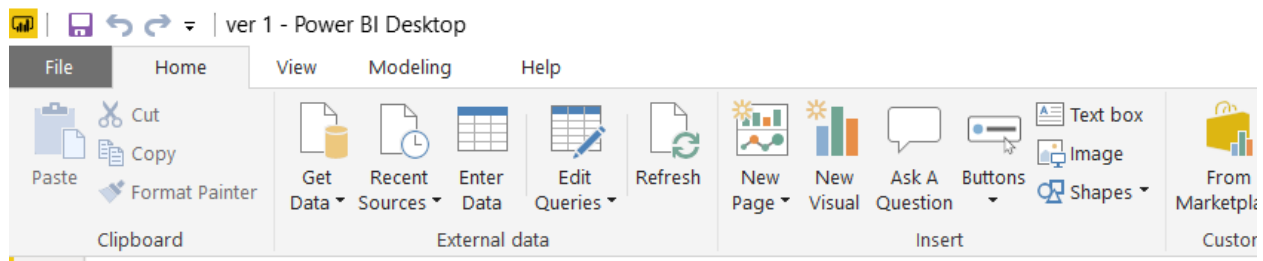Scroll till the end and click on "(move to end)".

Then click "OK".

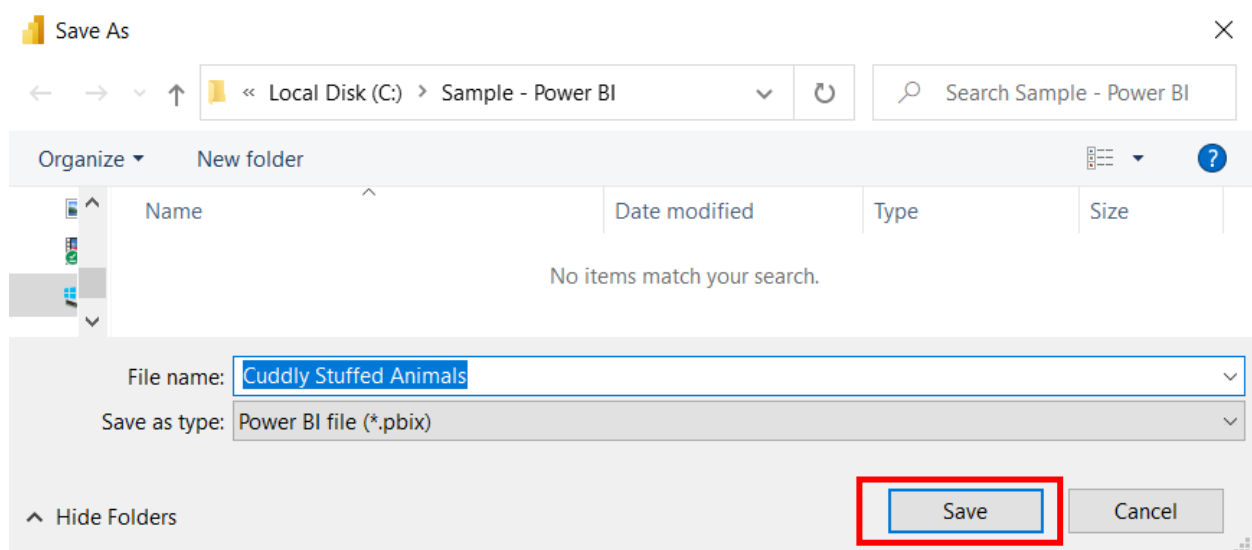Rename the tab as "January Report by Product", format the sheet, save and close the Excel file.

Operation 11: Save the "Flow and Data"

Select "File".

Type the name of the file.

Select Save.

Select Save.

Close Power BI.