

MERGING MULTIPLE DATASETS

BY

JORDAN GOLDMEIER

LAB 1: MERGING MULTIPLE DATA SETS

You have been tasked to use Power Query to merge multiple datasets together from different systems.

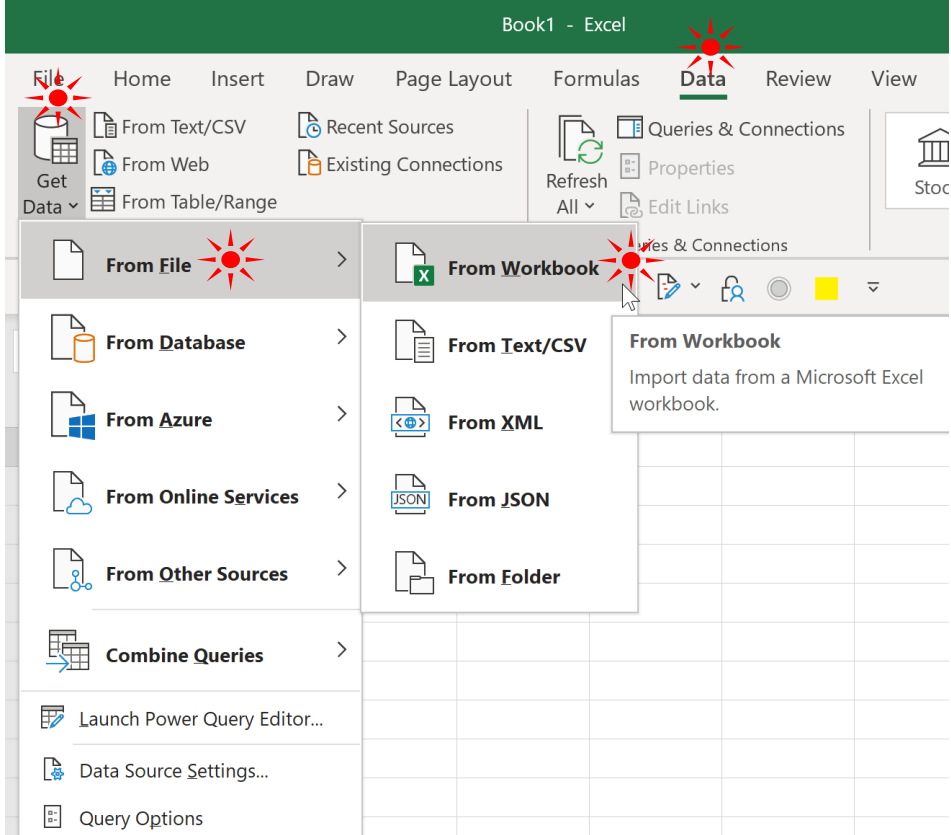
MODULE REQUIREMENTS

- ▶ Pulling in multiple types of data into Power Query
- ▶ Using a merge queries
- ▶ Creating a final table with results from both datasets

HINTS

Watch both videos in the module first before attempting this lab.

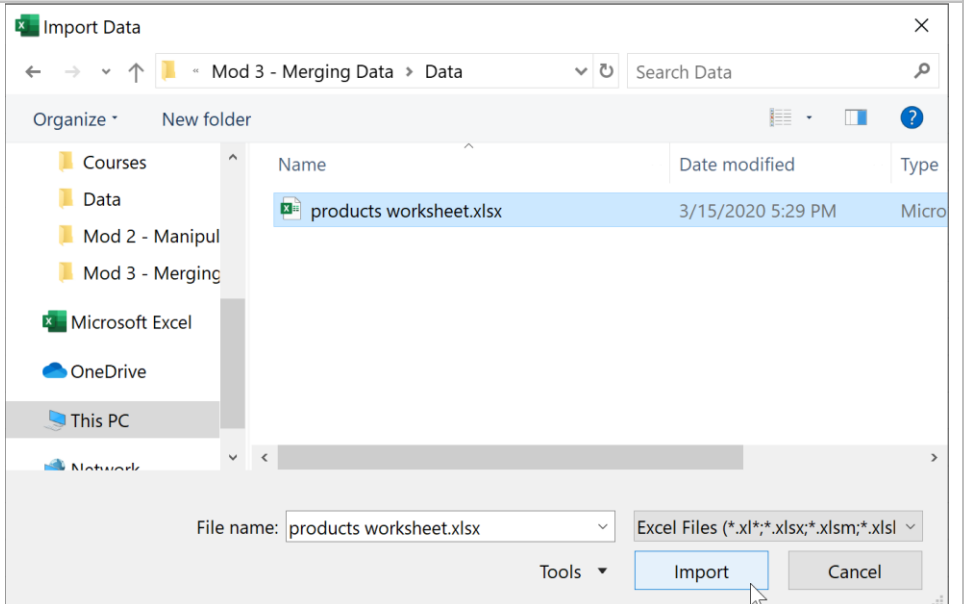
STEP-BY-STEP INSTRUCTIONS

Click Steps	Screen Shots
<ol style="list-style-type: none">1. Create a new fresh Excel File2. Save it Transactions Final or another name of your choice.3. From on the Data tab, select Get Data > From File > From Workbook	 <p>The screenshot shows the Microsoft Excel interface with the 'Data' tab selected. The 'Get Data' menu is open, and the 'From File' > 'From Workbook' path is highlighted. A tooltip for 'From Workbook' is visible, stating 'Import data from a Microsoft Excel workbook.' Red starburst markers indicate the specific clicks described in the instructions: the 'Data' tab, the 'Get Data' icon, the 'From File' option, and the 'From Workbook' option.</p>

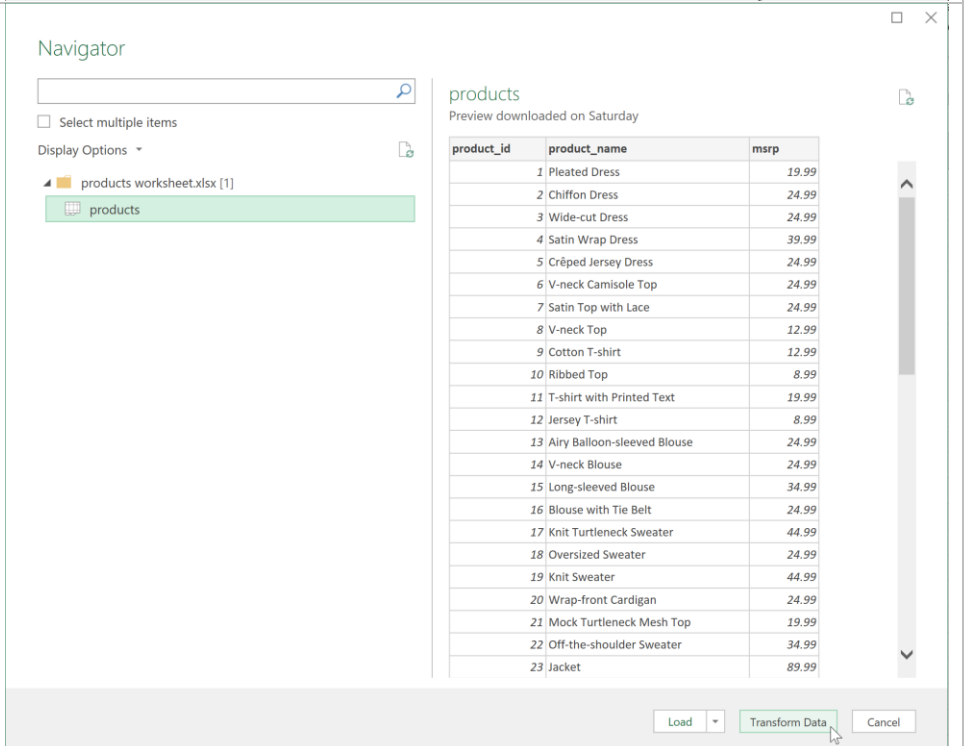
Click Steps

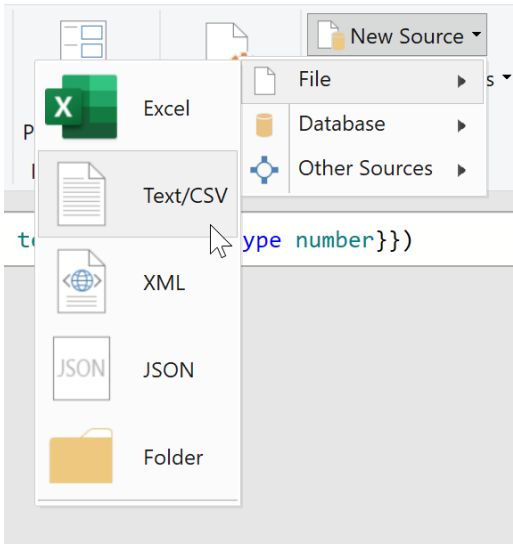
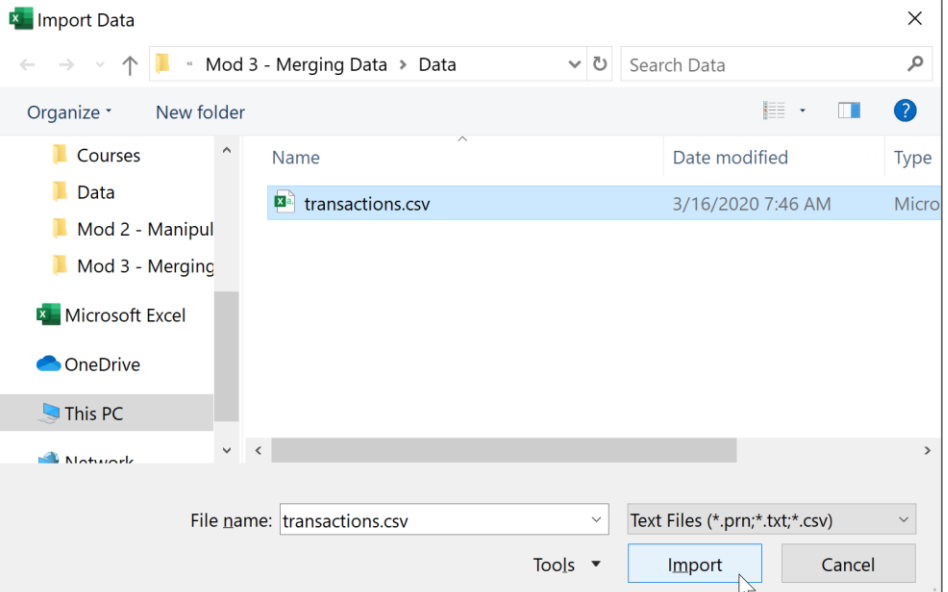
4. From the class files, navigate to ...\\Courses\\Intro to Power Query\\Mod 3 - Merging Data\\Data in the Import Data file selection box.
5. Select products worksheet.xlsx.
6. Press Import.

Screen Shots



7. In the Navigator window, select the products tab worksheet.
8. Then click Transform Data.



Click Steps	Screen Shots
<p>9. Wait until Power Query is finished loading.</p> <p>10. Now it's time to bring the associated CSV file.</p> <p>11. From on the Home tab, select New Source > Text / CSV</p> <p>New Source can be found on the far right of the ribbon tab.</p>	
<p>12. When the Import Data dialog box opens, navigate to ...\Intro to Power Query\Mod 3 - Merging Data\Data from within the project files.</p> <p>13. Select transactions.csv</p> <p>14. Click Import.</p>	

Click Steps

15. Click OK to import the csv with no additional properties set.

Screen Shots

The screenshot shows a dialog box titled "transactions.csv" with the following settings:

- File Origin: 1252: Western European (Windows)
- Delimiter: Comma
- Data Type Detection: Based on first 200 rows

date	time	product_id	qty
9/20/2016	1:43:00 PM	12	5
5/10/2014	3:26:00 PM	55	9
6/14/2013	1:31:00 PM	22	1
11/28/2013	10:36:00 AM	40	2
12/23/2016	12:10:00 PM	50	3
6/15/2016	10:13:00 AM	2	1
6/25/2015	3:29:00 PM	47	6
10/23/2014	2:33:00 PM	51	9
5/26/2013	9:35:00 AM	31	9
11/16/2014	2:35:00 PM	48	3
1/10/2013	10:38:00 AM	26	6
1/5/2015	2:05:00 PM	7	9
3/8/2014	10:04:00 AM	27	6
8/19/2013	2:09:00 PM	48	6
12/25/2016	12:10:00 PM	53	8
2/21/2015	10:42:00 AM	8	10
2/26/2014	9:52:00 AM	44	8
5/24/2014	11:09:00 AM	25	5
8/30/2016	12:41:00 PM	15	4
9/24/2016	4:03:00 PM	12	1

The data in the preview has been truncated due to size limits.

Buttons: OK, Cancel

16. Now it's time to merge both datasets together.

17. From on the Home tab, select the Merge Queries dropdown > Merge Queries as New

The screenshot shows the "Merge Queries" dropdown menu with the following options:

- Merge Queries
- Merge Queries as New
- Combine

Below the menu, the text "Merge Queries as New" is highlighted in blue. Below that, the text "Merge Queries as New" is highlighted in red. Below that, the text "Merge Queries as New" is highlighted in green. Below that, the text "Merge Queries as New" is highlighted in yellow. Below that, the text "Merge Queries as New" is highlighted in purple. Below that, the text "Merge Queries as New" is highlighted in brown. Below that, the text "Merge Queries as New" is highlighted in pink. Below that, the text "Merge Queries as New" is highlighted in gray. Below that, the text "Merge Queries as New" is highlighted in white. Below that, the text "Merge Queries as New" is highlighted in black. Below that, the text "Merge Queries as New" is highlighted in blue. Below that, the text "Merge Queries as New" is highlighted in red. Below that, the text "Merge Queries as New" is highlighted in green. Below that, the text "Merge Queries as New" is highlighted in yellow. Below that, the text "Merge Queries as New" is highlighted in purple. Below that, the text "Merge Queries as New" is highlighted in brown. Below that, the text "Merge Queries as New" is highlighted in pink. Below that, the text "Merge Queries as New" is highlighted in gray. Below that, the text "Merge Queries as New" is highlighted in white. Below that, the text "Merge Queries as New" is highlighted in black.

Click Steps

18. In the merge dialog box, set the following properties:

- transactions in the top dropdown
- product_id selected in the top table
- products selected in lower dropdown
- product_id selected in the bottom table
- Join Kind is set to Left Outer...

19. Click OK

20. When the new table appears, immediately rename it to Final Table

Screen Shots

Transform Combine Parameters Data Sources New Query

Merge

Select tables and matching columns to create a merged table.

transactions

date	time	product_id	qty
9/20/2016	1:43:00 PM	12	5
5/10/2014	3:26:00 PM	55	9
6/14/2013	1:31:00 PM	22	1
11/28/2013	10:36:00 AM	40	2
12/23/2016	12:10:00 PM	50	3

products

product_id	product_name	msrp
1	Pleated Dress	19.99
2	Chiffon Dress	24.99
3	Wide-cut Dress	24.99
4	Satin Wrap Dress	39.99
5	Crêped Jersey Dress	24.99

Join Kind
Left Outer (all from first, matching from second)

Use fuzzy matching to perform the merge

> Fuzzy matching options

✓ The selection matches 34000 of 34000 rows from the first table.

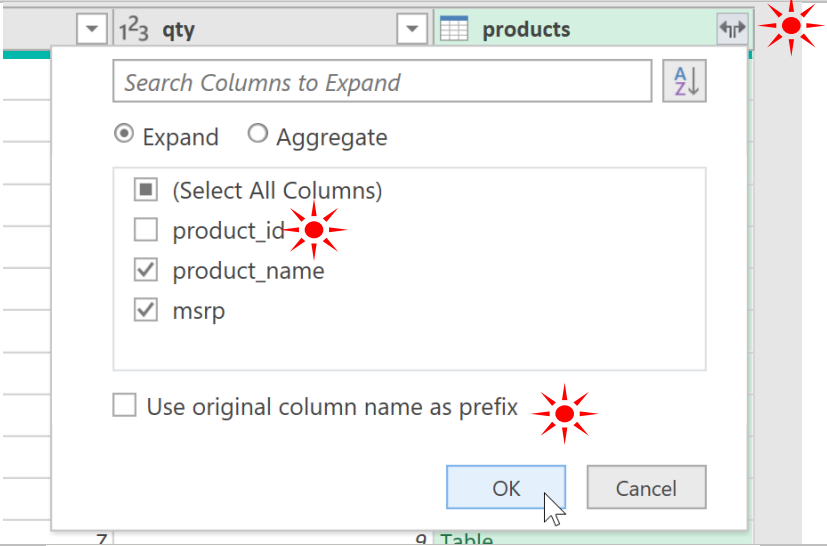
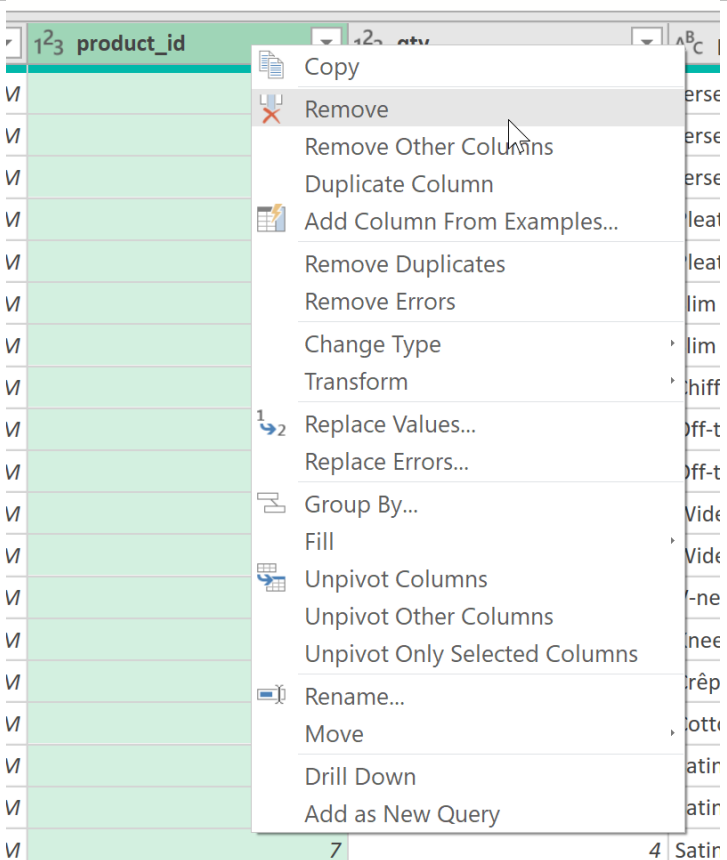
OK Cancel

Query Settings

PROPERTIES

Name
Final Table

All Properties

Click Steps	Screen Shots
<p>21. Click on the expansion icon to the right of the products header</p> <p>22. Uncheck product_id</p> <p>23. Uncheck Use original name as prefix</p> <p>24. Click OK</p>	
<p>25. Remove the product_id column by right-clicking the header and selecting Remove.</p>	

Click Steps

Screen Shots

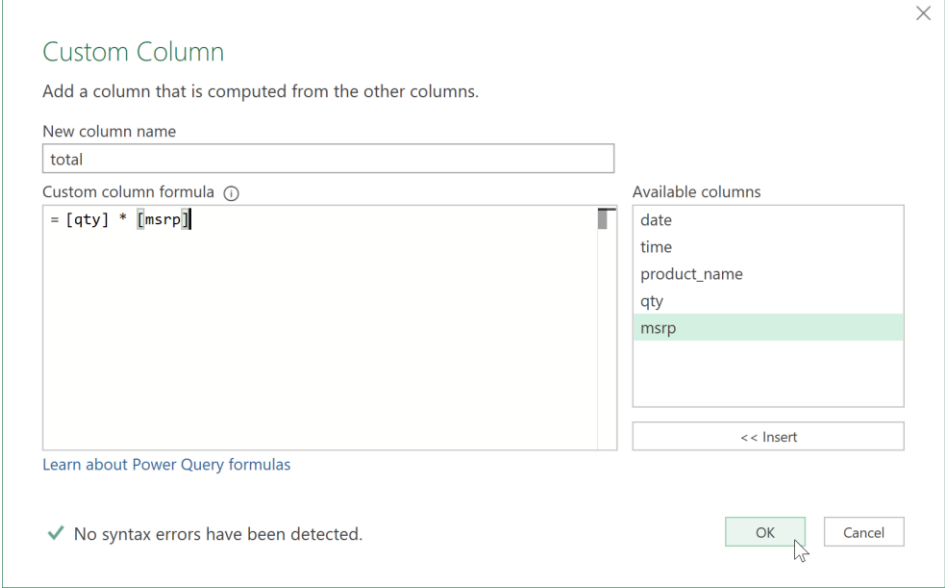
26. Switch qty and product_name by clicking and holding the qty header and then dragging it to the right.

The screenshot shows the Power Query Editor interface. At the top, there are several tabs: Transform, Combine, Parameters, Data Sources, and New Query. Below the tabs, there are various options like 'Data Type: Whole Number', 'Merge Queries', 'Append Queries', 'Combine Files', 'Manage Parameters', 'Data source settings', 'New Source', and 'Recent Sources'. The main area displays a table with the following columns: '123 qty', 'A^c product_name', and 'qty'. The 'qty' column header is highlighted in green, and a mouse cursor is positioned over it, indicating it is being dragged to the right. The table contains 20 rows of data, including columns for time, product name, and price.

	123 qty	A^c product_name	qty	
1:43:00 PM	5	Jersey T-shirt		8.99
4:03:00 PM	1	Jersey T-shirt		8.99
3:44:00 PM	9	Jersey T-shirt		8.99
11:49:00 AM	1	Pleated Dress		19.99
1:33:00 PM	3	Pleated Dress		19.99
3:26:00 PM	9	Slim Jeans		19.99
10:01:00 AM	2	Slim Jeans		19.99
10:13:00 AM	1	Chiffon Dress		24.99
1:31:00 PM	1	Off-the-shoulder Sweater		34.99
1:34:00 PM	6	Off-the-shoulder Sweater		34.99
9:32:00 AM	7	Wide-cut Dress		24.99
10:01:00 AM	6	Wide-cut Dress		24.99
10:36:00 AM	2	V-neck Cotton Sweater		19.99
12:10:00 PM	3	Knee-length Cotton Shorts		24.99
4:32:00 PM	9	Crêped Jersey Dress		24.99
3:29:00 PM	6	Cotton Piqué T-shirt		15.99
2:05:00 PM	9	Satin Top with Lace		24.99
2:56:00 PM	2	Satin Top with Lace		24.99
12:47:00 PM	4	Satin Top with Lace		24.99
2:33:00 PM	9	Sweatshorts with Side Stripes		19.99

27. Create a column that reflects the final sale price.
28. From on the Add Column tab select Custom Column

The screenshot shows the 'Add Column' tab in the Power Query Editor. The 'Custom Column' option is selected, and a dialog box titled 'Custom Column' is open. The dialog box contains the text: 'Create a new column in this table, based on a custom formula.' Below this text, there are two options: 'transactions' and 'Final Table'. The 'Final Table' option is highlighted in green. The background shows the 'Add Column' ribbon with options like 'Column From Examples', 'Custom Column', 'Invoke Custom Function', 'Conditional Column', 'Index Column', and 'Duplicate Column'.

Click Steps	Screen Shots
<p>29. In the Custom Column pop up type total in the New Column name field.</p> <p>30. Click into the custom column formula</p> <p>31. Type =</p> <p>32. Double-click the qty selection from Available columns. This will insert the field into the custom formula.</p> <p>33. Type *</p> <p>34. Double-click msrp to insert he field into the formula.</p> <p>35. Click OK.</p>	
<p>36. Select File > Close & Load To...</p> <p>37. Click Table and New Worksheet. Click OK.</p> <p>38. The updated table now appears on a new worksheet.</p>	