

Lab 1 Data Import & Transformation Instructions

Lab Objectives

Time: 30 Min

1. Import the Power Cycle data files using the Folder connector for the sales transaction by country and the Excel and csv flat file connector for the supporting data files
2. Normalize the data tables by creating a **Reseller** dimension table using the table duplication and table merging feature in Power Query editor
3. Create an ID name **ResellerID** using an index column that can be used to connect the **Reseller** table in the data model.
4. Perform a table Merge and create a **ResellerID** in the **Sales by Country Table**
5. Using Power Query perform a text transformation splitting the **City** and **State** columns into individual columns in the **Reseller** table.

Lab steps

1) Import Lab data files

Import Lab files using the text/csv connector and the Folder connector for the files in the **Country files** folder you will need to use the folder connector

The screenshot illustrates the steps to import data into Power BI. On the left, the 'Home' tab of the Power BI Desktop ribbon is shown, with the 'Get data' button highlighted. A dropdown menu lists common data sources, including Excel workbook, Power BI datasets, Dataflows, Dataverse, SQL Server, Analysis Services, Text/CSV, Web, OData feed, Blank query, and Power BI Template Apps. The 'More...' button at the bottom of this list is also highlighted. On the right, the 'Get Data' dialog box is open, showing a list of data sources. The 'Folder' connector is highlighted with a red box. Below the list, the 'Country Files' folder is selected, and a list of files within that folder is displayed, including Calendar, Power Cycle Sales Target, Product Category, Product Sub-Category, Product, Sale Region, Sales, Salesperson, and SalespersonRegion.

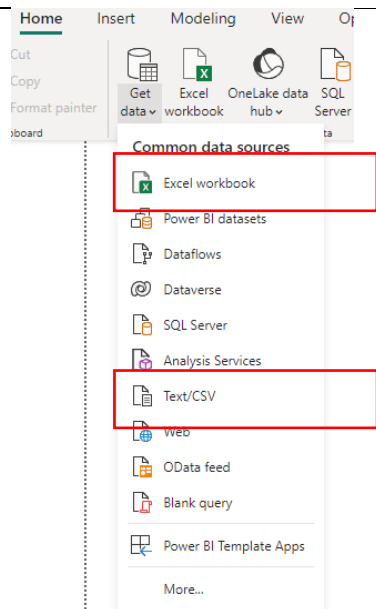
Select combined and load to import all csv files from the folder into power query

Content	Name	Extension	Date accessed	Date modified	Date created	Attributes	Folder Path
Binary	Australia.csv	.csv	8/10/2023 10:11:44 PM	6/21/2023 3:11:09 PM	8/10/2023 10:11:44 PM	Record	C:\Users\shann\OneDrive\Desktop
Binary	Canada.csv	.csv	8/17/2023 1:57:05 PM	6/21/2023 3:14:08 PM	8/10/2023 10:11:44 PM	Record	C:\Users\shann\OneDrive\Desktop
Binary	France.csv	.csv	8/10/2023 10:11:44 PM	6/21/2023 3:15:25 PM	8/10/2023 10:11:44 PM	Record	C:\Users\shann\OneDrive\Desktop
Binary	Germany.csv	.csv	8/10/2023 10:11:44 PM	6/21/2023 3:16:40 PM	8/10/2023 10:11:44 PM	Record	C:\Users\shann\OneDrive\Desktop
Binary	UK.csv	.csv	8/10/2023 10:11:44 PM	6/21/2023 3:17:50 PM	8/10/2023 10:11:44 PM	Record	C:\Users\shann\OneDrive\Desktop
Binary	US.csv	.csv	8/10/2023 10:11:44 PM	6/21/2023 3:19:13 PM	8/10/2023 10:11:44 PM	Record	C:\Users\shann\OneDrive\Desktop

Combine & Transform Data Transform Data Cancel

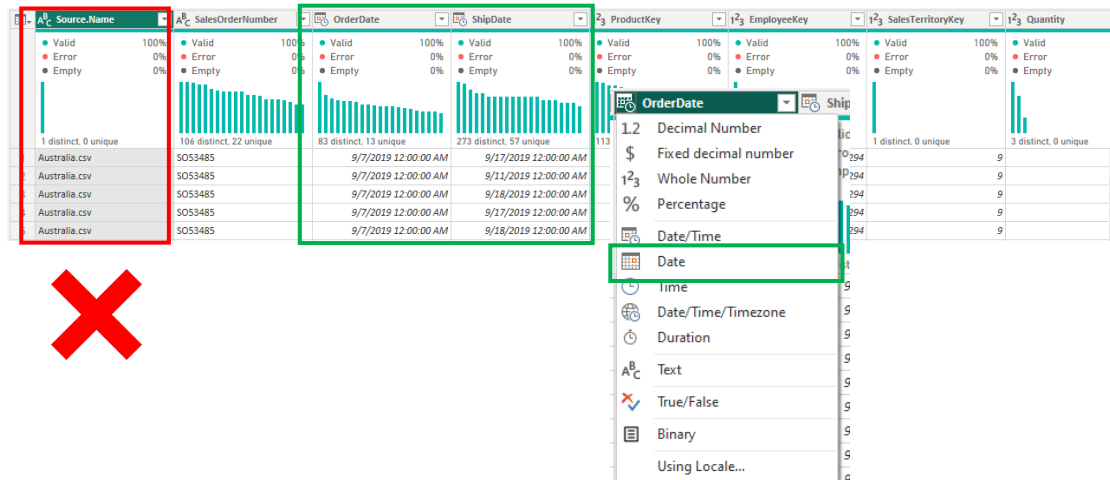
Import the following csv files.

- Product-Category
- Product Sub-Category
- Product
- Sale Region
- Sales By Country Files
- Salesperson
- Power Cycle Sales Target



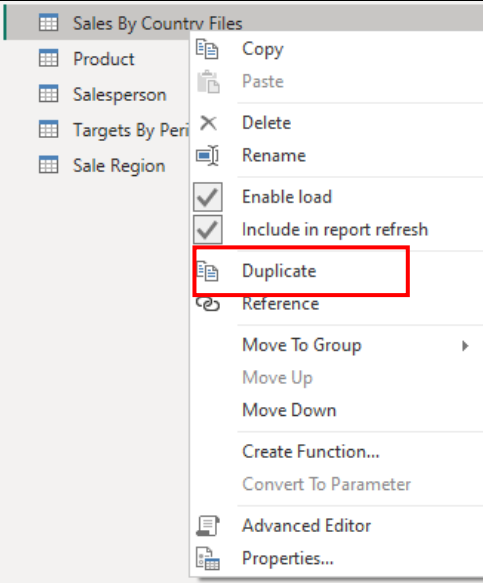
Scan through all tables and ensure that the data fields are classified correctly.

- Change date/time fields to Date only
- Remove the Source name column as it is not a required column for any analysis



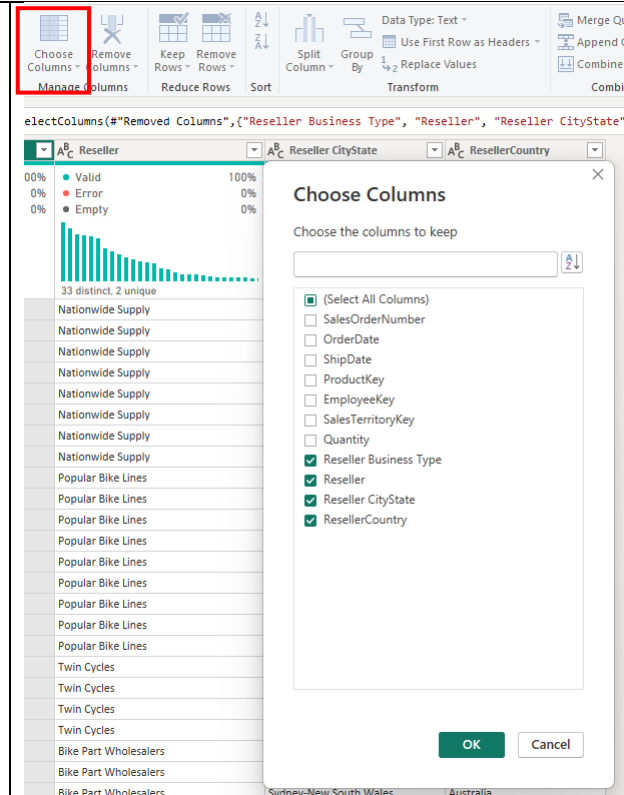
2) Creating a Reseller Dimension table

Create The Reseller table by selecting the **Sales by Country** table then right-click and select duplicate

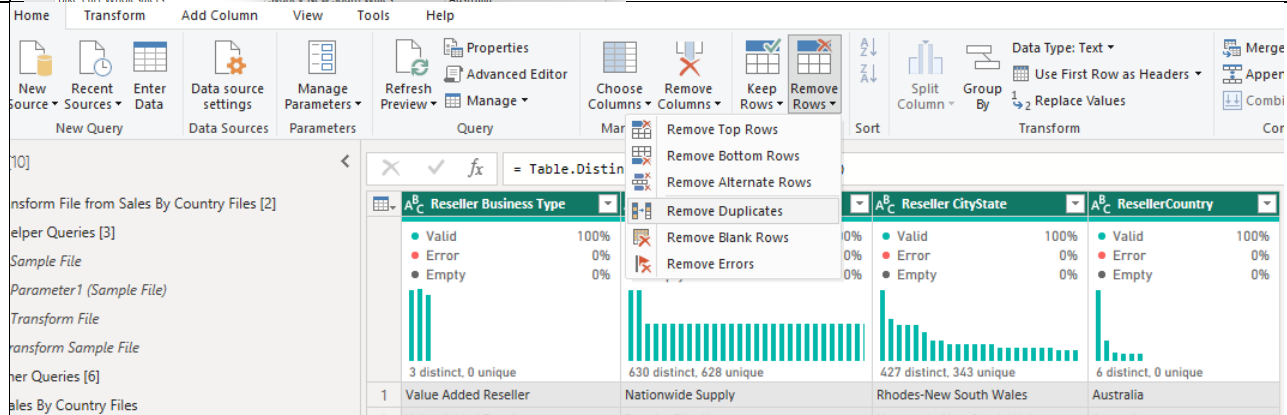


From the **Home Tab** select choose Columns and ensure that the following fields are selected from the list.

1. Reseller Business Type
2. Reseller
3. Reseller CityState
4. ResellerCountry



- On you keyboard select Ctrl+A to select the entire table.
- From the Home Tab select **Remove Rows** and **Remove Duplicated**



3) Create a Reseller ID Index column

- From the **Add Column** tab select **Index Column** with the Index starting from 1
- Right-click on the new Index column and Rename the Index Column to **ResellerID**

Formula Bar: `= Table.AddIndexColumn(#"Removed Duplicates", "Index", 1, 1, Int64.Type)`

Reseller Business Type	Reseller	Reseller CityState	ResellerCountry	Index
Value Added Reseller	Nationwide Supply	Rhodes-New South Wales	Australia	1
Value Added Reseller	Popular Bike Lines	Newcastle-New South Wales	Australia	2

4) Merging Tables

- From the **Home** tab select **Merge Queries** from the dropdown
- Right-click on the new Index column and Rename the Index Column to **Reseller Key**
- The **Sales by Country** table should already be selected at the top and for the table that you are merging from select the **Reseller** table from the dropdown list.
- You will need to ensure that these table are connected by the Reseller name. Select the Reseller Column in both tables

Merge

Select a table and matching columns to create a merged table.

Sales By Country Files

TerritoryKey	Quantity	Reseller Business Type	Reseller	Reseller CityState	ResellerCountry
9	1	Value Added Reseller	Nationwide Supply	Rhodes-New South Wales	Australia
9	1	Value Added Reseller	Nationwide Supply	Rhodes-New South Wales	Australia
9	1	Value Added Reseller	Nationwide Supply	Rhodes-New South Wales	Australia
9	1	Value Added Reseller	Nationwide Supply	Rhodes-New South Wales	Australia

Reseller

Reseller Key	Reseller Business Type	Reseller	Reseller CityState	ResellerCountry
1	Value Added Reseller	Nationwide Supply	Rhodes-New South Wales	Australia
2	Value Added Reseller	Popular Bike Lines	Newcastle-New South Wales	Australia
3	Value Added Reseller	Twin Cycles	Barlinghurst-New South Wales	Australia
4	Value Added Reseller	Bike Part Wholesalers	Sydney-New South Wales	Australia
5	Value Added Reseller	Eastside Cycle Shop	Melbourne-Victoria	Australia

Join Kind

Left Outer (all from first, matching from second)

☐ Use fuzzy matching to perform the merge

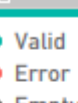
Fuzzy matching options

✓ The selection matches 57851 of 57851 rows from the first table.

OK

Cancel

- From the new **Reseller.1** Column you will now be able to expand the table records by selecting the expansion arrows in the column
- Select only the **ResellerID** and uncheck the “use original column name as prefix option”



Reseller1

Reseller Name	Status	Percentage
Reseller1	Valid	100%
Reseller2	Error	0%
Reseller3	Empty	0%

Table

Table

Table

123 Quantity

Reseller.1

Search Columns to Expand

☒ Expand ☐ Aggregate

☒ (Select All Columns)

☒ ResellerID

☐ Reseller Business Type

☐ Reseller

☐ City

☐ State

☐ ResellerCountry

☐ Use original column name as prefix

OK Cancel

5) Column Splitting

- Going back to the Reseller table you will need to split out the State name from the city so that they are separate columns, respectively.
- From the **Home** tab select **Split Column** option -> **By Delimiter**

AC Reseller CityState

- Valid 100%
- Error 0%
- Empty 0%

427 distinct 363 unique

Rhodes-New South Wales

Newcastle-New South Wales

Darlinghurst-New South Wales

Sydney-New South Wales

Editor

Choose Columns Remove Columns Manage Columns

By Delimiter

Split Column by Delimiter

Split values in the selected column based on the specified delimiter.

100% Valid

0% Error

0% Empty

- From the menu option you will need to ensure that the split at is set to each occurrence
- In the advanced options the number of columns to split should be set to 2 and the split into should be columns
- Re-Name the columns as followed.

Reseller CityState.1 -> City
Reseller CityState.2 -> State

END

Split Column by Delimiter

Specify the delimiter used to split the text column.

Select or enter delimiter

--Custom--

-

Split at

Left-most delimiter

Right-most delimiter

Each occurrence of the delimiter

Advanced options

Split into

Columns

Rows

Number of columns to split into

2

Quote Character

"

☐ Split using special characters

OK

Cancel

Reseller CityState.1

Reseller CityState.2

Valid

Error

Empty

100%

0%

0%

414 distinct, 323 unique

67 distinct, 10 unique

Rhodes

Newcastle

New South Wales

New South Wales