**Importing and Transforming Data using PowerBI Editor**

Contents

[1. Import United States Sales Data 1](#_Toc180269045)

[2. Import Customer Data 3](#_Toc180269046)

[3. Import Sales Territory data 5](#_Toc180269047)

[4. Import Product data 5](#_Toc180269048)

[5. Import ProductRollup data 6](#_Toc180269049)

[6. Import Other Countries data 6](#_Toc180269050)

[7. Append Sale Queries 7](#_Toc180269051)

[8. Create/Merge Column 7](#_Toc180269052)

[9. Modeling and cleanup 8](#_Toc180269053)

[10. DAX Expressions 8](#_Toc180269054)

# Import United States Sales Data

 **Open Power BI Desktop**

* Go to the **Home tab**.
* Click on **New Source**.
* Select **Text/CSV**.

 **Load Customer Data**

* Navigate to the data folder.
* Select the "customer list" file from the **dimensions** folder.
* Review file parameters (origin, delimiter, data type).
* Click **OK** to load the data.

 **Rename the Query**

* Change the default name from "customer list" to **Customer**.

 **Change Data Type**

* Locate the **House Owner Flag** column.
* Change the data type from **Whole Number** to **Text**.
* Select **Replace Current** to confirm the change.

 **Replace Values**

* Right-click the **House Owner Flag** column.
* Select **Replace Values**.
* Replace **1** with **Yes**.
* Replace **0** with **No**.

 **Create a Conditional Column**

* Go to the **Add Column** tab.
* Click on **Conditional Column**.
* Name the column **Marital Status**.
* Define conditions:
  + If **MS** = "M", then **Married**.
  + If **MS** = "S", then **Single**.
  + Else, **Unknown**.

 **Move Column**

* Select the **Marital Status** column.
* Move it next to the **MS** column for comparison.

 **Remove Unnecessary Columns**

* Select the columns **MS**, **GN**, **Total Children**, and **Number of Children at Home**.
* Click **Remove Columns**.

 **Rename Columns**

* Rename columns for better readability:
  + **Key** → **Customer Key**.
  + **FirstName** → **First Name**.
  + **LastName** → **Last Name**.
  + **B Date** → **Birth Date**.
  + **House Owner Flag** → **Home Owner**.
  + **Number of Cars Owned** → **Cars Owned**.

 **Check and Update Data Types**

* Ensure that data types are correct for all columns:
  + **Customer Key**: Whole Number.
  + **First Name**, **Last Name**, etc.: Text.
  + **Birth Date**: Date.
  + **Marital Status**: Text (change from Any).
  + **Yearly Income**: Fixed Decimal Number.
  + **Cars Owned**: Whole Number.

 **Final Steps**

* Ensure all columns have meaningful names.
* Confirm that unnecessary columns have been removed.
* Check that correct data types are assigned.

# Import Customer Data

 **Open Power BI Desktop**

* Go to the **Home tab**.
* Click on **New Source**.
* Select **Text/CSV**.

 **Load Customer Data**

* Navigate to the data folder.
* Select the "customer list" file from the **dimensions** folder.
* Review file parameters (origin, delimiter, data type).
* Click **OK** to load the data.

 **Rename the Query**

* Change the default name from "customer list" to **Customer**.

 **Change Data Type**

* Locate the **House Owner Flag** column.
* Change the data type from **Whole Number** to **Text**.
* Select **Replace Current** to confirm the change.

 **Replace Values**

* Right-click the **House Owner Flag** column.
* Select **Replace Values**.
* Replace **1** with **Yes**.
* Replace **0** with **No**.

 **Create a Conditional Column**

* Go to the **Add Column** tab.
* Click on **Conditional Column**.
* Name the column **Marital Status**.
* Define conditions:
  + If **MS** = "M", then **Married**.
  + If **MS** = "S", then **Single**.
  + Else, **Unknown**.

 **Move Column**

* Select the **Marital Status** column.
* Move it next to the **MS** column for comparison.

 **Remove Unnecessary Columns**

* Select the columns **MS**, **GN**, **Total Children**, and **Number of Children at Home**.
* Click **Remove Columns**.

 **Rename Columns**

* Rename columns for better readability:
  + **Key** → **Customer Key**.
  + **FirstName** → **First Name**.
  + **LastName** → **Last Name**.
  + **B Date** → **Birth Date**.
  + **House Owner Flag** → **Home Owner**.
  + **Number of Cars Owned** → **Cars Owned**.

 **Check and Update Data Types**

* Ensure that data types are correct for all columns:
  + **Customer Key**: Whole Number.
  + **First Name**, **Last Name**, etc.: Text.
  + **Birth Date**: Date.
  + **Marital Status**: Text (change from Any).
  + **Yearly Income**: Fixed Decimal Number.
  + **Cars Owned**: Whole Number.

 **Final Steps**

* Ensure all columns have meaningful names.
* Confirm that unnecessary columns have been removed.
* Check that correct data types are assigned.

# Import Sales Territory data

* **Connecting to the data source:**  
  Clicked on "New Source" and selected "Excel Workbook".
  + Opened the "Sales Territories" file.
  + Confirmed the data preview and clicked "OK".
  + **Renaming the query:**  
    Renamed the query from "Sheet1" to "Sales Territory".
* **Splitting columns:**  
  Used the "Split Column" feature by delimiter (pipe symbol) to split the "TerritoryRollup" column into two separate columns.
  + **Renaming columns:**  
    Renamed columns to "Sales Territory Key", "Sales Territory Region", and "Sales Territory Country".
  + **Trimming white space:**  
    Selected the "Region" and "Country" columns.
  + Used the "Trim" function to remove leading and trailing white space.

# Import Product data

* **Importing Data**:  
  Go to the Home tab, click on New Source, and select Excel Workbook.
  + Choose the table named "Table1" from the Excel workbook.
  + **Renaming the Query**:  
    Right-click on the query name "Table1" and rename it to "Product".
  + **Removing Columns**:  
    Select the columns to keep: Product Key, Product Code, English Product Subcategory Name, English Product Name, Standard Cost, and Color.
  + Use the "Remove Other Columns" function to remove all other columns except the selected ones.
  + **Changing Column Names**:  
    Rename columns for clarity:  
    EnglishProduct Subcategory to Product Subcategory
    - English Product Name to Product Name
  + **Changing Data Types**:  
    Ensure the data types are correct:  
    Product Key as Whole Number
    - Product Code, Product Subcategory, Product Name, and Color as Text
    - Standard Cost as Fixed Decimal Number
  + **Filtering Rows**:  
    Filter out rows where the Product Subcategory equals "Internal Parts".
  + Standardize the text format to ensure consistency before applying the filter.
  + Goto Tansform->Format->Capitalize each word.

# Import ProductRollup data

* + **Import Data**:  
    Click on "New source" and select "Text/CSV".
  + Select the "product rollup" file and click "Open".
  + **Remove Unnecessary Rows**:  
    Remove the top four rows using the "Remove rows" function.
  + **Set Headers**:  
    Use the first row as headers by selecting "Use first rows as headers".
  + **Standardize Column Data**:  
    Capitalize each word in the "Product Subcategory" column using the "Capitalize each word" function.
  + **Remove Unnecessary Prefix**:  
    Remove the "PC-" prefix from the "Product Category" column using the "Extract" function and selecting "Text after delimiter".
  + **Save and Apply Changes**:  
    Save the work and apply the changes to load the data into the model.

# Import Other Countries data

* **Navigate to Power Query Editor**: Click on "Transform Data" from the Home tab.
* **Import Data**: Use the Folder connector to import data from multiple files in a directory.
* **Combine and Transform Data**: Combine all files into one query and transform the data.
* **Filter Rows**: Filter the data to include only rows from January 1st, 2020, and beyond.
* **Remove Columns**: Remove the source.name column as it has no analytical value.
* **Rename Columns**: Rename PID to Product Key.
* **Change Data Types**:  
  + Change order date from DateTime to Date.
  + Change unit price to Fixed Decimal.
* **Validate Data**: Ensure all files follow the same format to avoid errors during the import process.
* **Remove Unnecessary Data**: Remove data that is not needed for analysis, such as data prior to 2020.

# Append Sale Queries

* **Load Data**: Click "Close and Apply" to load data from sources into the data model.
* **Append Queries**:
  + Go to the Power Query editor.
  + Select the "other country sales" table.
  + Use the "Append" option under the "Combine" menu.
  + Choose "Append as new" to create a new query.
  + Select the tables to append (other country sales and United States sales).
  + Rename the new query to "sales".
* **Inspect Data**: Check columns and data in the new "sales" table.
* **Handle Column Mismatch**:  
  Power Query matches columns based on names.
  + Columns not present in both tables will have null values.
* **Rename Columns**: Rename the new query to "sales".
* **Remove Unwanted Columns**: Delete the "freight" column as it is not needed.
* **Change Data Types**: Adjust data types for columns as needed (e.g., change "unit price" to a fixed decimal number).

# Create/Merge Column

* + **Create a New Column**:  
    In sales table select  Order Quantity and Unit Price columns.
  + Go to the "Add Column" tab.
  + Use the "Standard" function to multiply these columns.
  + **Rename the New Column**:  
    Rename the new column from Multiplication to Line Total Sales.
  + **Check Data Type**:  
    Ensure the new column is set to the correct data type (Fixed Decimal Number).
  + **Goto PowerQuery**
    - Select Product Table and click on Merge Queries
    - Select ProductRollup as second table in merge queries
    - Map Product subcategory from Product table to Product subcategory in ProductRollup table
    - The fields of Product rollup will be appended as table in product out of that select product category column to product table.
  + **Similarly get Sales Territory Key to Sales table from Sales Territory table**

# Modeling and cleanup

* **Load Data**: Click "Close and Apply" to load data into the data model.
* **Inspect Tables**: Check the loaded tables (e.g., customer, other country sales, product, product rollup, sales territory, and United States sales).
* **Disable Tables what we don’t need further**
  + Other Countries, United States Sales and Product Rollup right click on the table names and **uncheck “Enable Load”**
* **In PowerQuery create two groups “Facts” and “Dimension”**
  + Move Sales table under “Fact” group
  + All the remaining three tables customer, product and sales Territory under Dimensional
* **Hierarchy build**
  + Goto Data modelling and select product table
  + In product Table select Product category-> right click and create hierarchy
  + Name it as Product Hierarchy and add   
     Product category, product subcategory and product Name
* Establishing relationship
  + In Data Modelling window drag and drop Product key from Product table to sales table and establish I to many relationship
  + Drag and drop Sales Territory Key from product to Sales Territory key in Sales Territory table.
  + Remove Sales Territory Region column from Sales table

# DAX Expressions

Goto Report view, select the table name right click and select New column and enter the DAX expression.

* **In Sales Table**
  + Calculated Columns
    - Line Total Sales=Sales[UnitPrice]\*Sales[OrderQuantity]
    - Line Product Cost = Sales[Order Quantity]\*Related('Product'[Standard Cost])
    - Line Margin=Sales[Line Total sales]-Sales[Line Product Cost]
    - Line Margin % = DIVIDE(Sales[Line Margin],Sales[Line Total Sales],0)
  + Measures (Aggregate of a column) right click on Sales table and select new measure.
    - Total Sales=SUM(sales[LineTotal Sales])
    - Total Margin=SUM(Sales[Line Margin])
    - Total Margin % = DIVIDE([Total Margin],[Total Sales],0)
    - Total Sales SUMX=SUMX(Sales,(Sales[Unit Price] \* Sales[Order Quantity])) (Returns the sum of the expression in each row in the table)
    - Count of ALL Sales Orders Line Items=COUNTROWS(ALL (Sales))

Target=7000000

* **In Customer Table**
  + Full Name=Customer[First Name] & “ “&Customer[Last Name]
  + Primary Buyer = AND(Customer[Cars Owned]=0,Customer[House Owner] ="No")
  + Secondary Buyer = OR(Customer[Cars Owned]=0, Customer[House Owner]="No")
* **Create in Product**
  + Color Category = If('Product'[Color] IN {"Black","Blue","Red"},"Darker","Lighter")
* Goto Page 1 and drag
  + Line Total Cost,
  + Line Product Cost
  + Line Margin

Quick Measures

Total Sales YoY%=

VAR \_PREV\_YEAR=

Calculate([Total Sales], DATEADD(‘Order Date’[Order Date], -1,YEAR))

RETURN

DIVIDE([Total Sales]-\_PREV\_YEAR, \_PREV\_YEAR)

**Goto DAX formatter website and format it if needed**

**Suggested Charts**

1. **Clustered column chart:**
   1. x-Axis : Sales Territory Country/Continent Y-Axis Total Sales
2. **Clustered Bar Chart:**
   1. Y-Axis: Product -> product hierarchy->product sub category
   2. X-Axis: Toal Margin %
3. **Line Chart**
   1. X-Axis Sales->Order data
   2. Y-Axis Sales-> Total Sale
4. **KPI Visual**
   1. Value: Total Sales
   2. Trend axis: Sales-> Date Hierarchy -> Year
   3. Target: Target from Sales
5. **Card Visual**
   1. Fields: Target
6. **Slicers** 
   1. **Add slicers on country and year**
7. **Filter Pane**
   1. Filter on this Page:-year