Lesson - 3 Homework

Question 1: What is the purpose of the "Applied Steps" pane in Power Query?

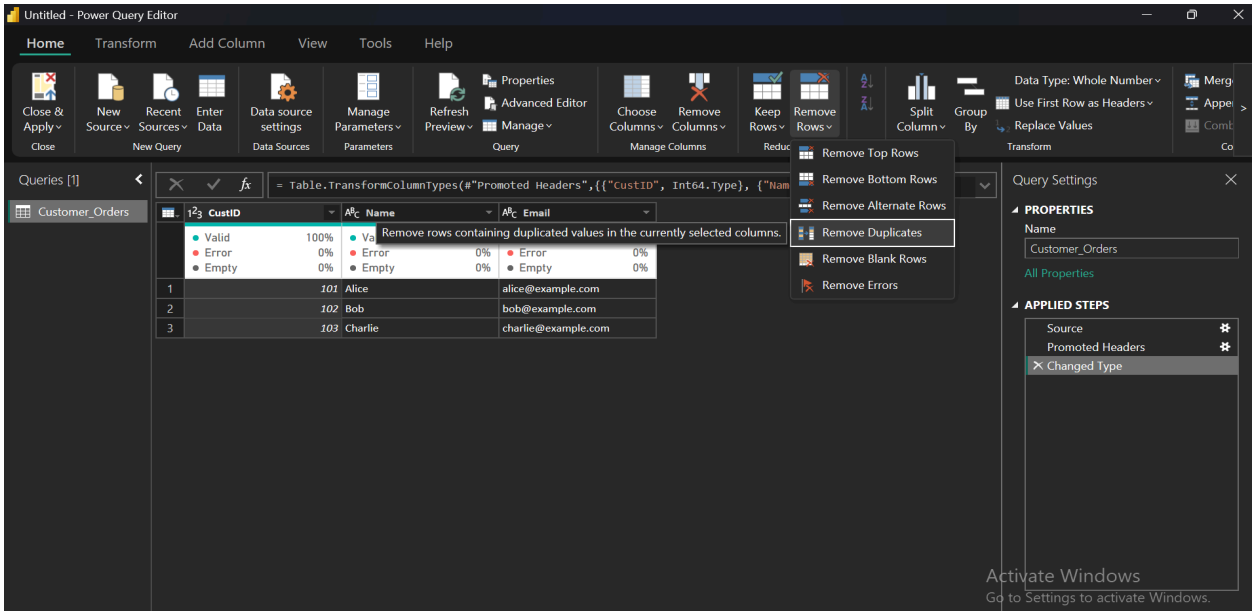
The **Applied Steps** pane in Power Query shows every change you make to your data. It's like a history list: each step is recorded, can be edited or removed, and will run again automatically when you refresh the data. It keeps your process clear, repeatable, and easy to adjust.

Question 2: How do you remove duplicate rows in Power Query?

In Power Query, you remove duplicate rows like this:

1. Select the column(s) you want to check for duplicates.
2. Go to the **Home** tab.
3. Click **Remove Rows** → **Remove Duplicates**.

Done – it keeps only the first occurrence and removes the rest.



Question 3: What does the "Filter" icon do in Power Query?

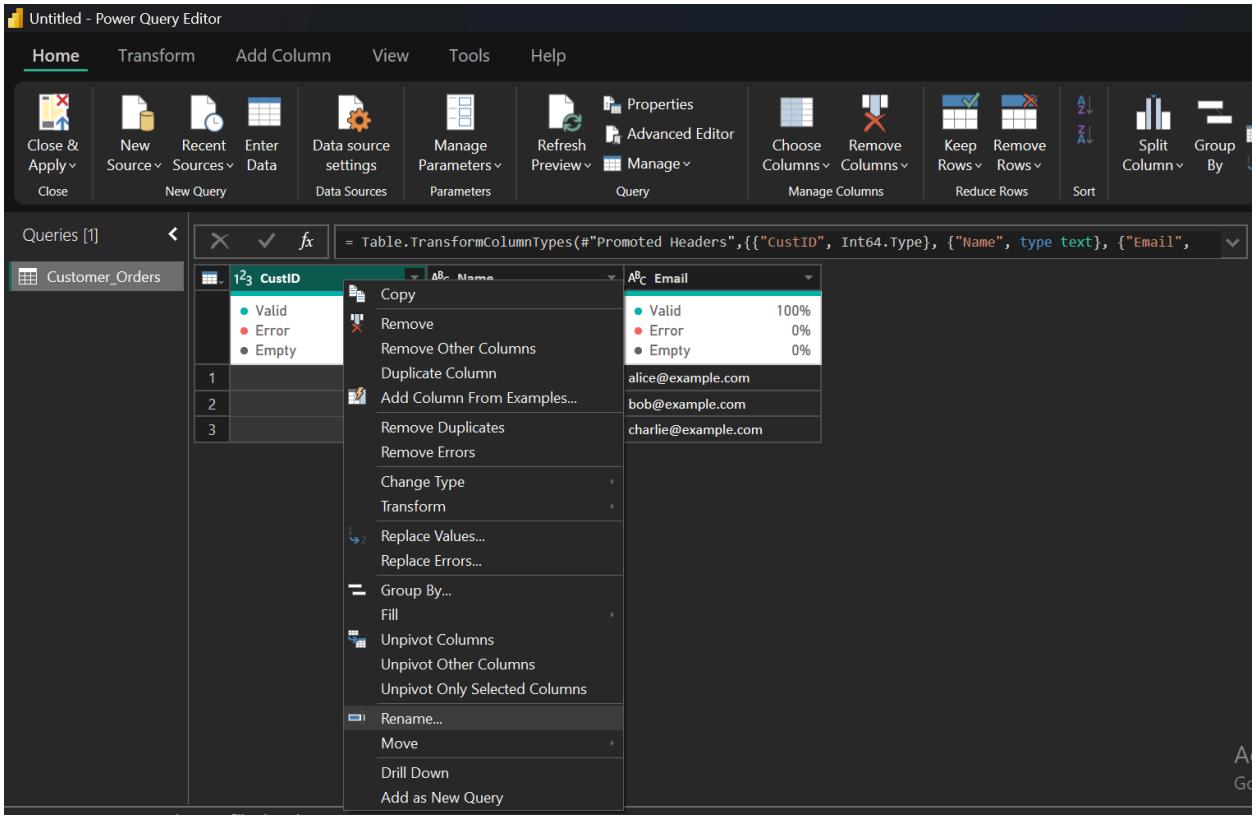
The **Filter** icon in Power Query (the little drop-down arrow on a column) lets you choose which rows to keep or remove. You can filter by values, text, numbers, dates, or even apply advanced conditions (e.g., greater than, contains, before/after). It's like Excel's

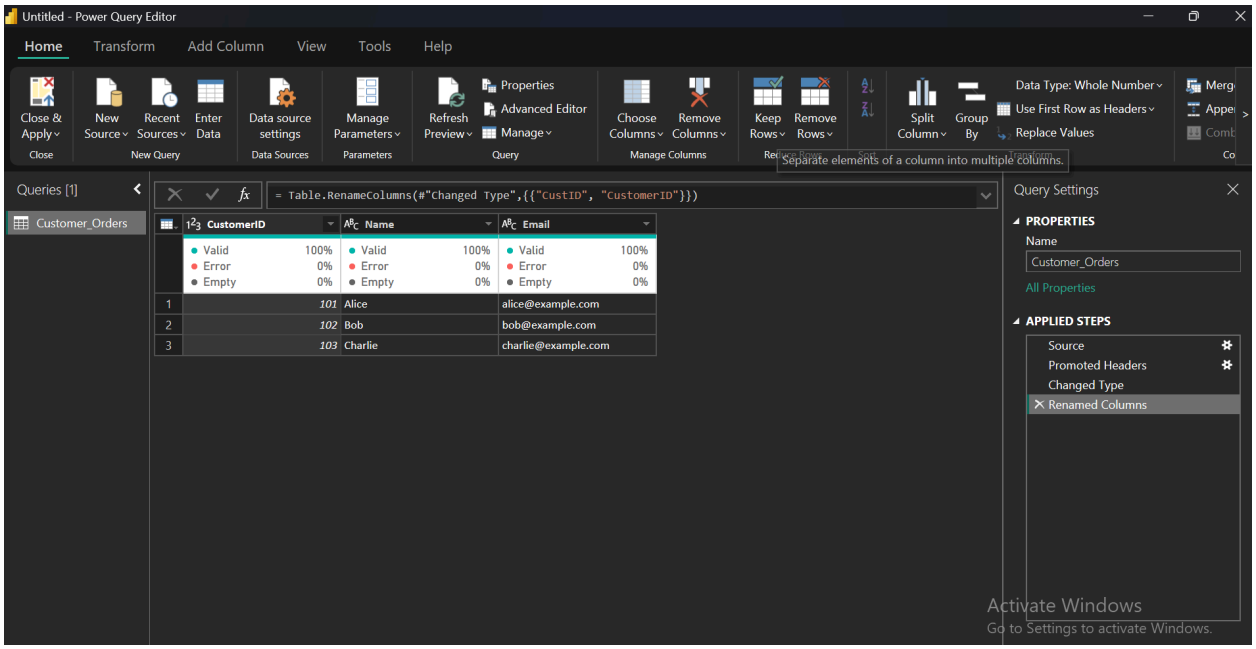
filter, but the step gets saved so it runs automatically on every refresh.

Question 4: How would you rename a column from "CustID" to "CustomerID"?

In Power Query:

1. Right-click the column header **CustID**.
2. Select **Rename**.
3. Type **CustomerID** and press Enter.





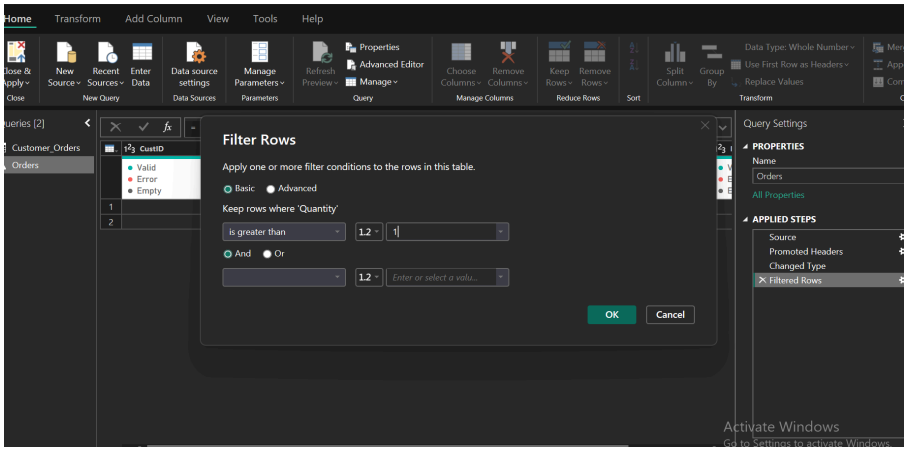
Question 5: What happens if you click "Close & Apply" in Power Query?

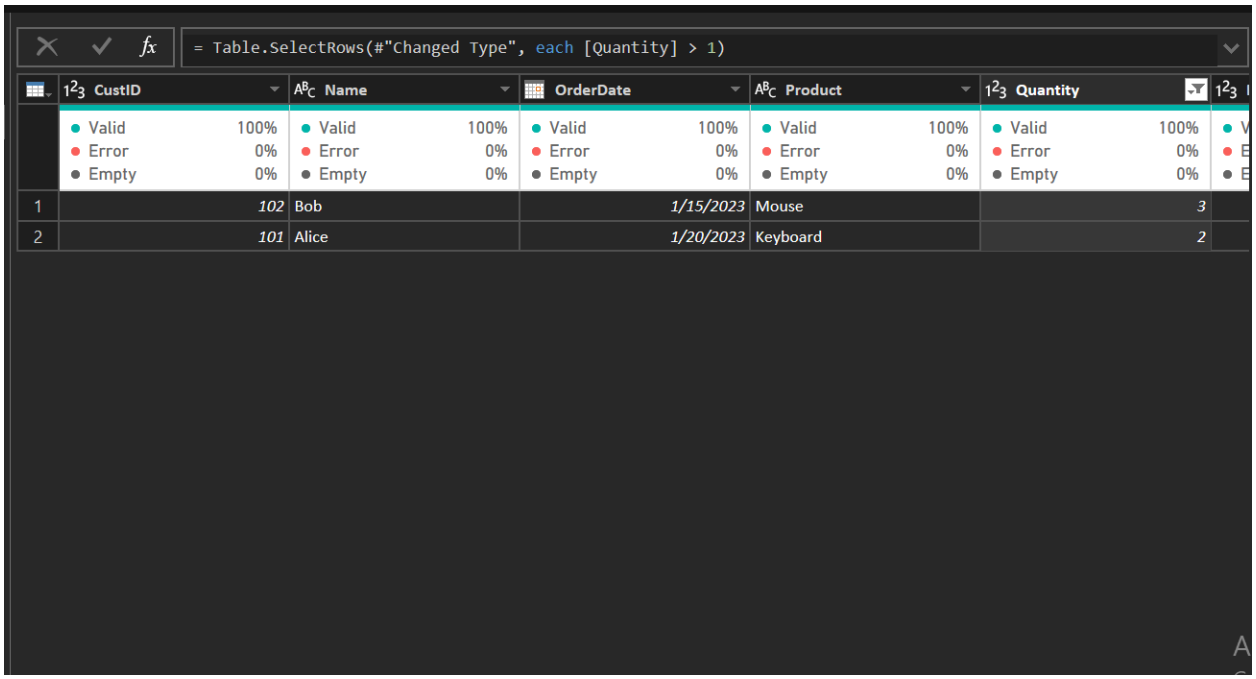
When you click **Close & Apply** in Power Query, it loads the transformed data back into Power BI and applies all the steps you made. After that, the query editor closes.

Question 6: Remove all rows where Quantity is less than 2.

In Power Query:

1. Click the filter icon on the **Quantity** column.
2. Choose **Number Filters → Keep the rows Greater Than 1**



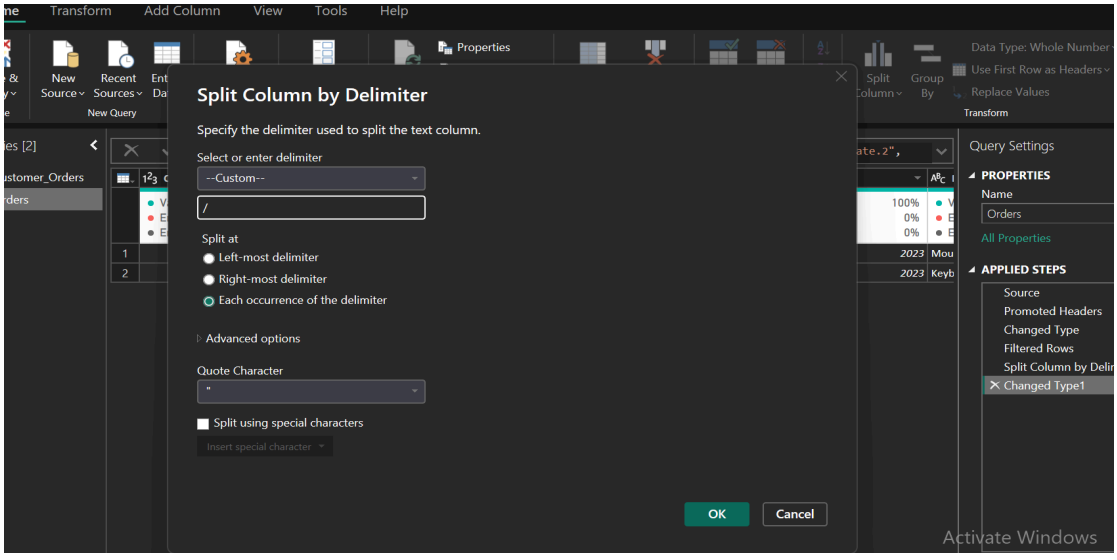


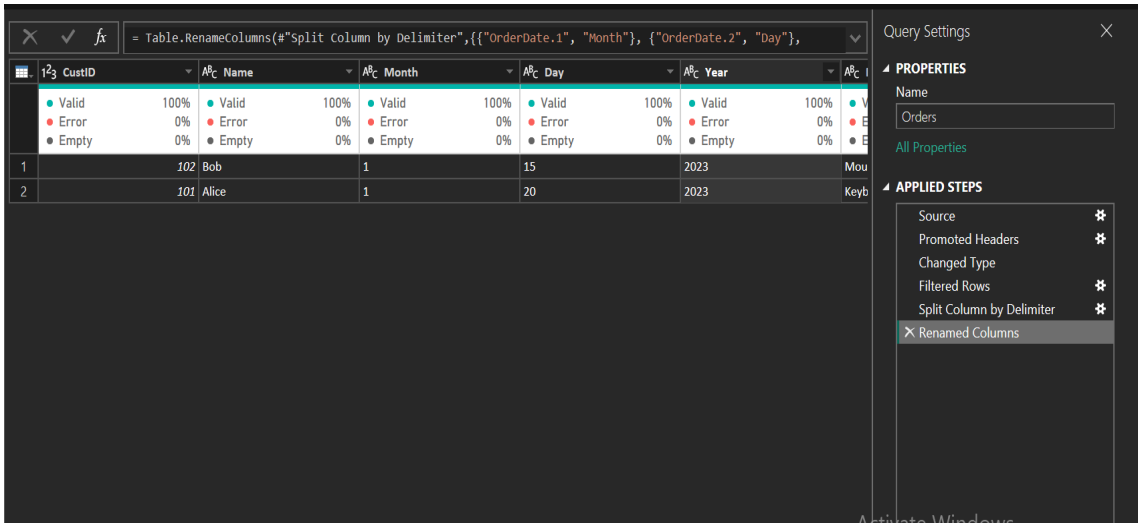
Question 7: Split the OrderDate column into separate "Year," "Month," and "Day" columns.

Click on the **OrderDate** column.

Choose split column, choose split by delimiter and set delimiter (in our case it is “/ ”).

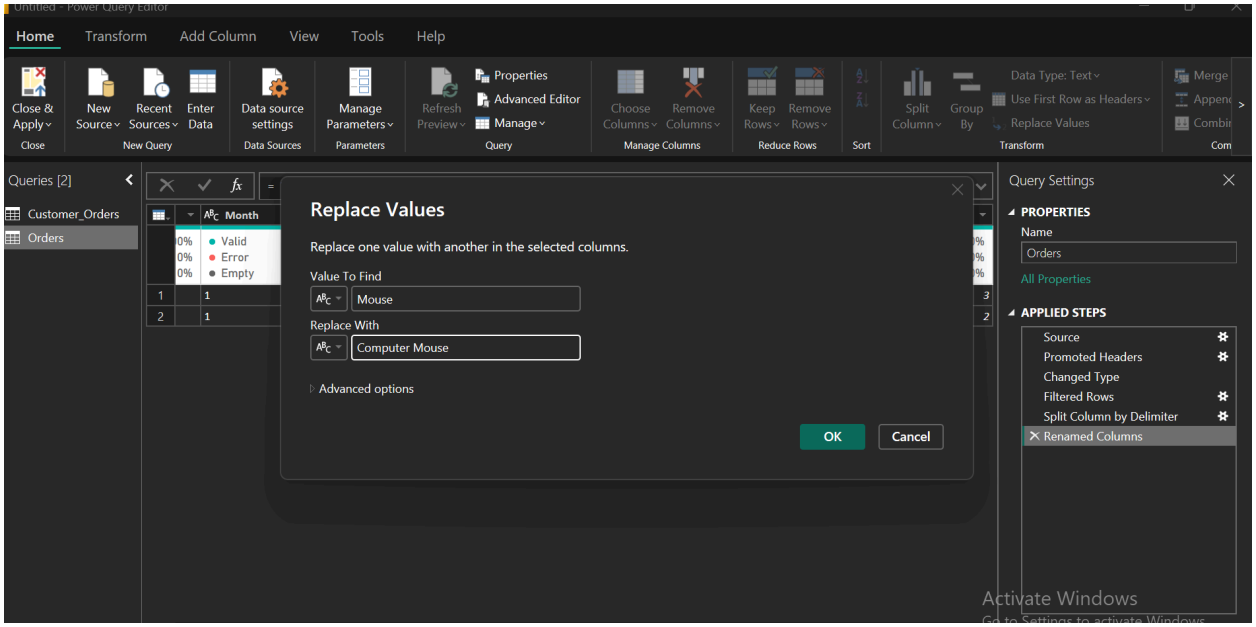
Then it splits into day,month and year (Has to rename those columns respectively).

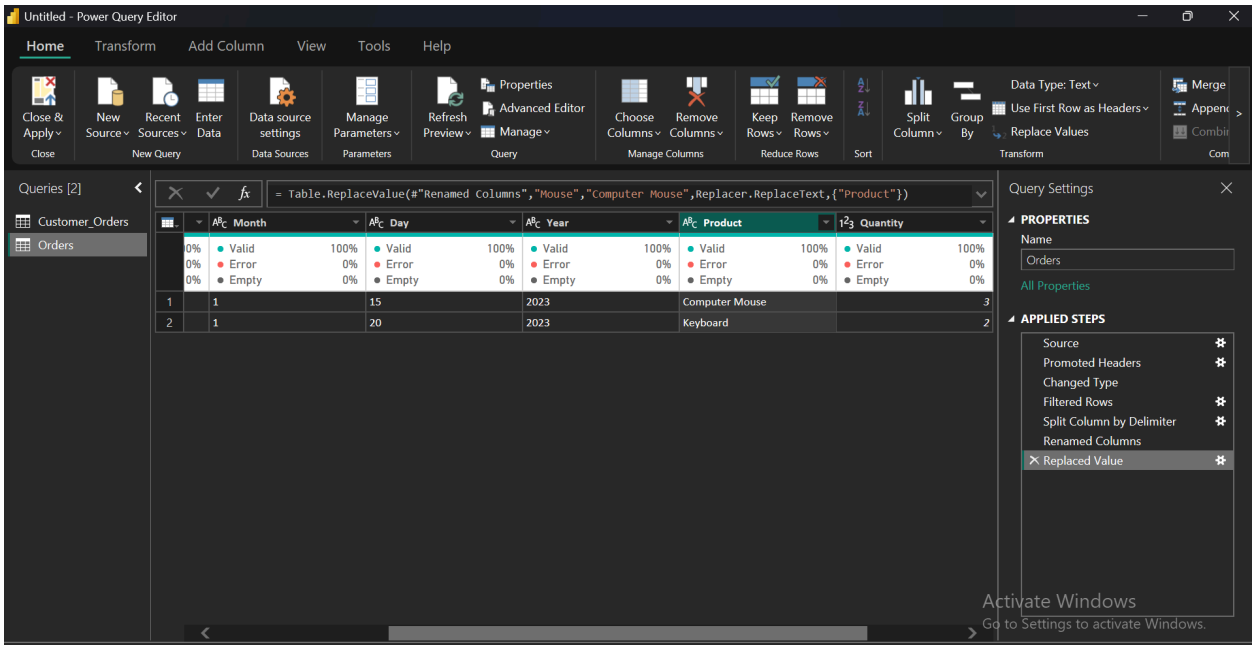




`Question 8:` Replace all "Mouse" entries in the Product column with "Computer Mouse."

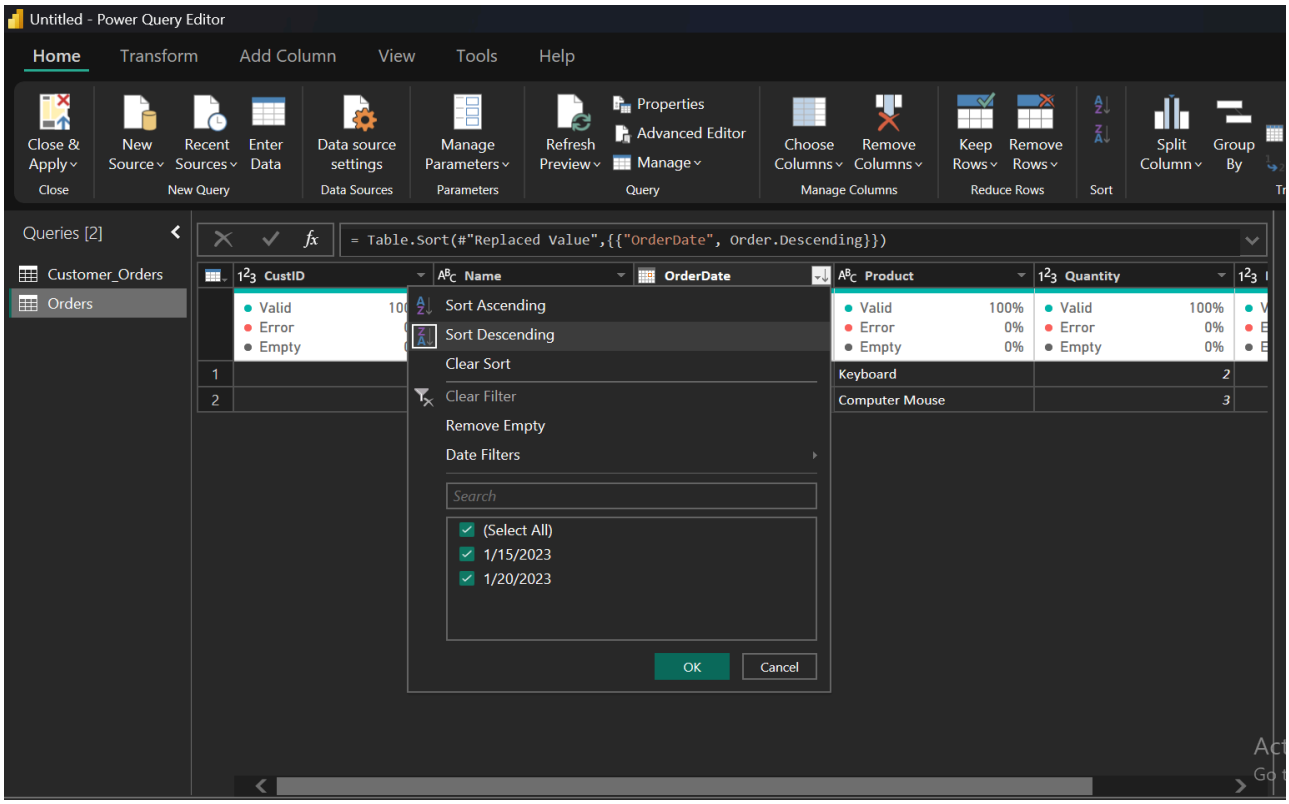
Click the Product column -> Right-click -> Replace Values -> Find: Mouse -> Replace: Computer Mouse -> Click OK.

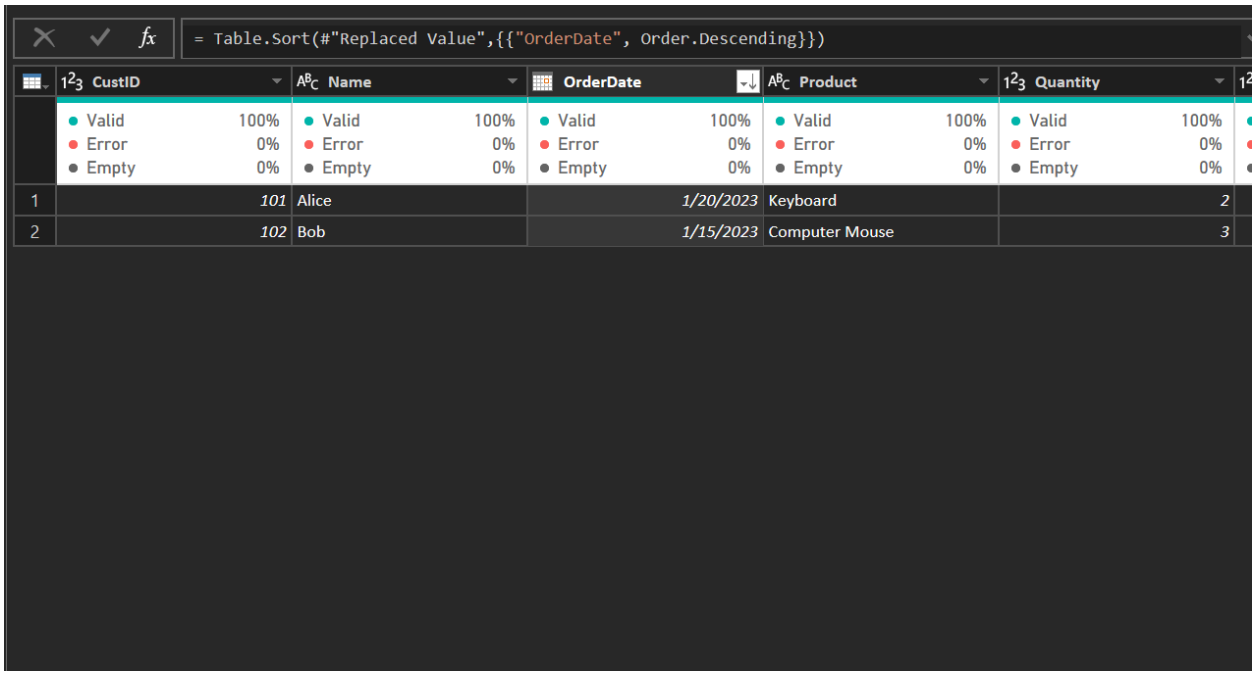




`Question 9:` Sort the table by OrderDate (newest first).

Click the OrderDate column -> Go to the Home tab -> Click Sort Descending





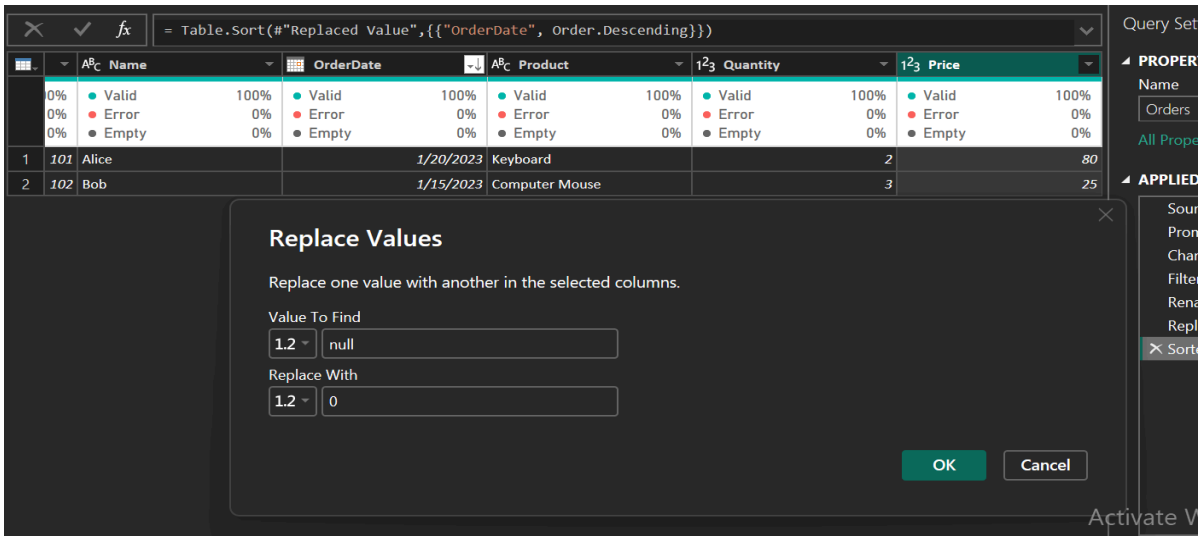
`Question 10:` How would you handle null values in the Price column?

Replace with Zero (Common for price)

Right-click the Price column.

Select Replace Values.

Replace null with 0



`Question 11:` Write custom M-code to add a column calculating TotalSpent = Quantity * Price

= Table.AddColumn(PreviousStepName, "TotalSpent", each [Quantity] * [Price], type number)

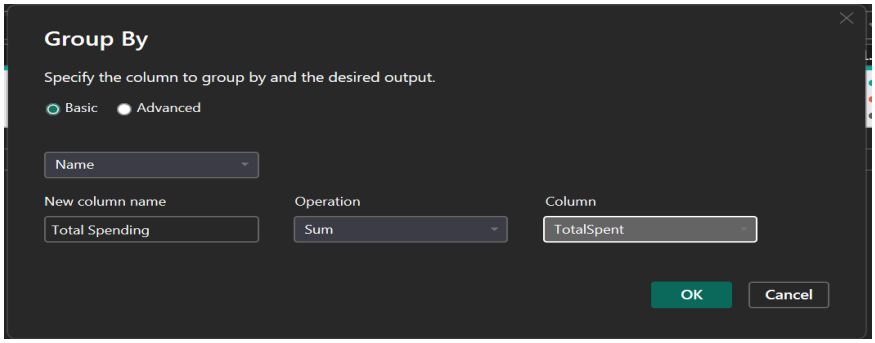
```
let
    Source = Csv.Document(File.Contents("C:\Users\apple servise\Documents\Data_Analytics\Power_BI Classes\bi_class_3\homework\Orders.txt"),[Delimiter=",", Columns=6, Encoding=1252, QuoteStyle=QuoteStyle.None]),
    #"Promoted Headers" = Table.PromoteHeaders(Source, [PromoteAllScalars=true]),
    #"Changed Type" = Table.TransformColumnTypes(#"Promoted Headers",{{"CustID", Int64.Type}, {"Name", type text}, {"OrderDate", type date}, {"Product", type text}, {"Quantity", Int64.Type}, {"Price", Int64.Type}}),
    #"Filtered Rows" = Table.SelectRows(#"Changed Type", each [Quantity] > 1),
    #"Renamed Columns1" = Table.RenameColumns(#"Filtered Rows",{{"OrderDate", "OrderDate"}}),
    #"Replaced Value" = Table.ReplaceValue(#"Renamed Columns1","Mouse","Computer Mouse",Replacer.ReplaceText,{"Product"}),
    #"Sorted Rows" = Table.Sort(#"Replaced Value",{{"OrderDate", Order.Descending}}),
    #"Replaced Value1" = Table.ReplaceValue(#"Sorted Rows",null,0,Replacer.ReplaceValue,{"Price"}),
    #"Added TotalSpent" = Table.AddColumn(#"Replaced Value1", "TotalSpent", each [Quantity] * [Price], type number)
in
    #"Added TotalSpent"
```

```
= Table.AddColumn(#"Replaced Value1", "TotalSpent", each [Quantity] * [Price], type number)
```

	OrderDate	Product	Quantity	Price	TotalSpent
	0% Valid 100%	0% Valid 100%	0% Valid 100%	0% Valid 100%	0% Valid 100%
	0% Error 0%	0% Error 0%	0% Error 0%	0% Error 0%	0% Error 0%
	0% Empty 0%	0% Empty 0%	0% Empty 0%	0% Empty 0%	0% Empty 0%
1	1/20/2023	Keyboard	2	80	160
2	1/15/2023	Computer Mouse	3	25	75

`Question 11:` Group the table by CustID to show total spending per customer.

1. Go to the Transform tab
2. Click Group By
3. Select CustID as the column to group by
4. Under New column name, enter "Total Spending"
5. For Operation, choose Sum
6. For Column, choose TotalSpent



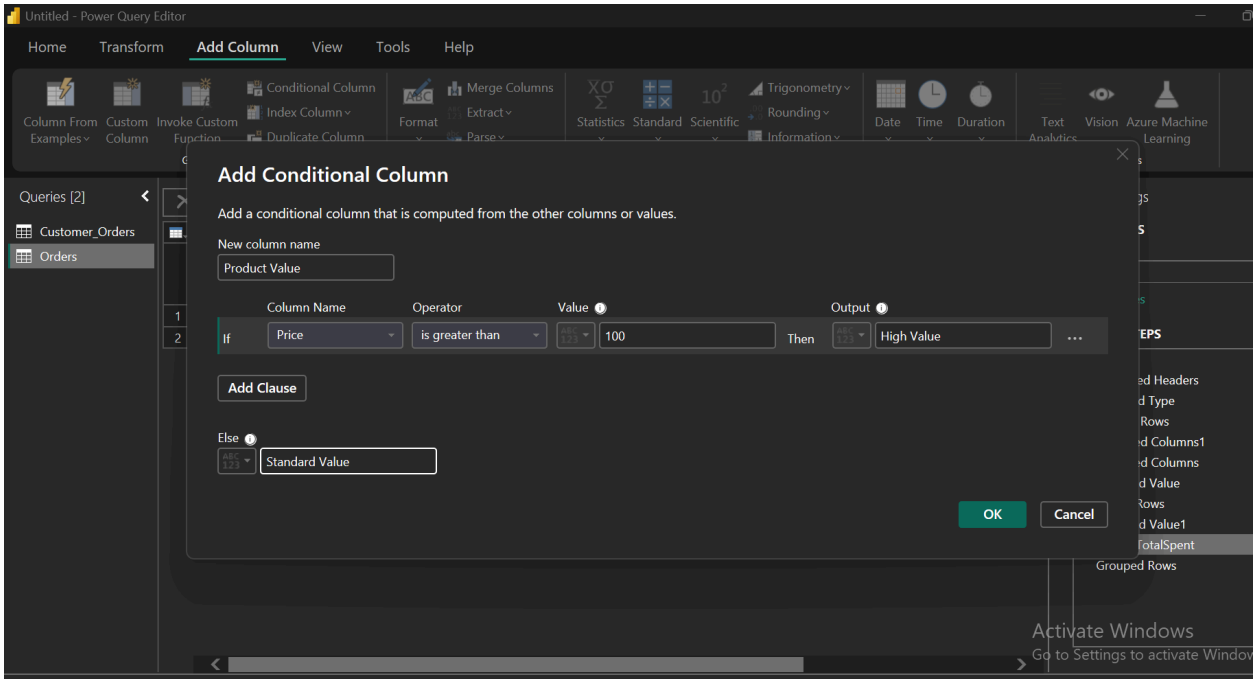
	Name	Total Spending
	Valid 100%	Valid 100%
	Error 0%	Error 0%
	Empty 0%	Empty 0%
1	Alice	160
2	Bob	75

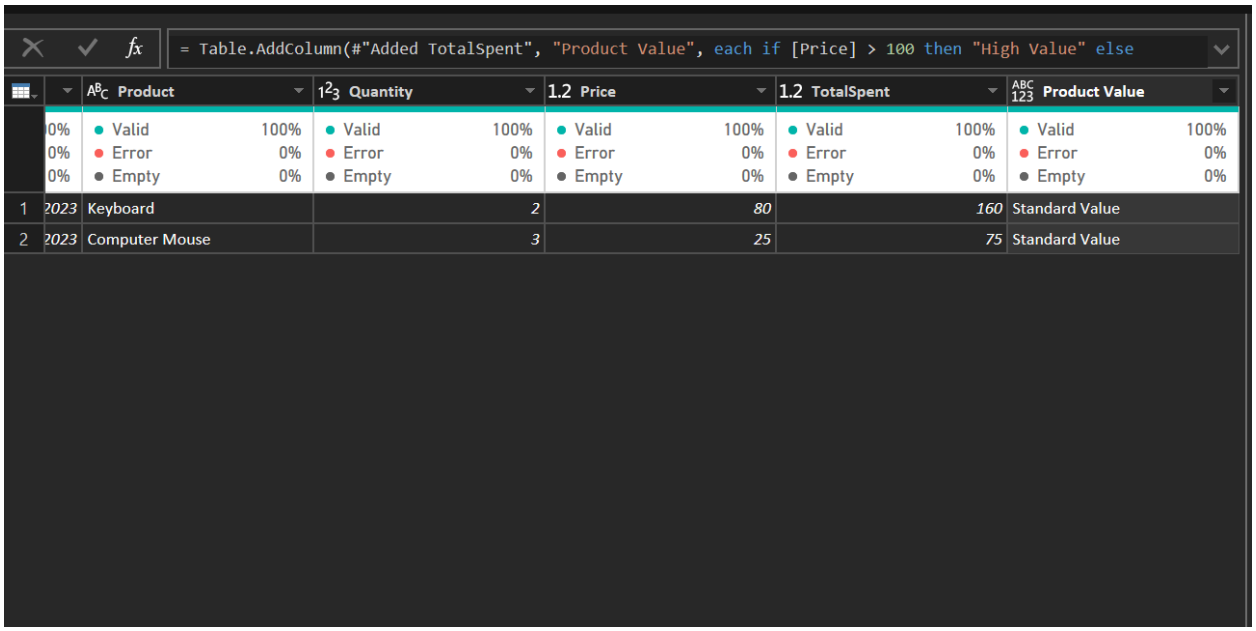
Question 13: Fix inconsistent date formats (e.g., 01/10/2023 vs. 2023-01-10) in OrderDate.

1. In the Power Query window that opens, click on the header of your OrderDate column to select it.
2. Use the "Using Locale" Feature (This is the key step):
3. Go to the "Transform" tab in the ribbon at the top.
4. Click the small dropdown arrow on the "Data Type:" button.
5. From the menu that appears, select "Using Locale...".
6. Configure the Conversion:
7. A new window will pop up.
8. For "Data Type:", choose Date from the dropdown list.
9. For "Locale:", choose English (United States) from the dropdown list.

Question 14: Create a conditional column: Label orders as "High Value" if Price > 100.

1. Go to the Add Column tab.
2. Click Conditional Column.
3. Set it up like this:
4. New column name: Product Value
5. If Price is greater than 100
6. Output High Value
7. Else Standard Value





`Question 15:` Optimize the query to reduce refresh time (e.g., remove unused columns early).

Since we do not have any unused column in our data base i am gonna give theoretical answer:

1. Here's what to do:
2. Open Power Query Editor: Click "Transform data".
3. Remove Unused Columns FIRST:
4. Look at your table and identify which columns you actually use in your reports (e.g., OrderDate, Price, CustomerName).
5. Select the columns you do not need. To select multiple columns, hold Ctrl and click on each column header.
6. Right-click on the selected column headers and choose "Remove Columns".
7. Do this before any other complex steps. This is the most important step for speed.
8. Filter Rows Early:
9. If you only need data from a certain time period (e.g., the last 2 years), apply a filter to the date column early in your steps.
10. Click the filter icon (▼) in the OrderDate column header.
11. Filter out old dates you don't need (e.g., "Date is after... 1/1/2022").
12. Check Applied Steps:
13. Look on the right side in the "Query Settings" pane at your "Applied Steps".
14. If you see any unnecessary steps (like a changed data type on a column you later deleted), you can click the "X" next to that step to delete it and clean up your query.
15. Close & Apply: Click "Close & Apply" to save.