

Data Transformation with Power Query (Part 1)

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

It shows the sequence of transformations applied to the data. Each step can be modified, reordered, or deleted.

2. How do you remove duplicate rows in Power Query?

Select the column(s) → **Home** → **Remove Rows** → **Remove Duplicates**.

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

It filters data based on selected conditions (e.g., show only rows where Quantity > 2).

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

Right-click the column header → **Rename** → type "CustomerID".

(M-code: `Table.RenameColumns("#PreviousStep", {"CustID", "CustomerID"})`)

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

The transformed data is loaded into **Power BI Desktop** and applied to the data model.

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

- Filter **Quantity column > 1** using the filter icon.
- M-code:

```
= Table.SelectRows("#PreviousStep", each [Quantity] >= 2)
```

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

Select column → **Transform** → **Date** → **Year/Month/Day** → **Extract**.

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

Right-click column → **Replace Values** → "Mouse" → "Computer Mouse".

9. Sort the table by OrderDate (newest first).

Click the drop-down in **OrderDate** column → **Sort Descending**.

10. How would you handle null values in the Price column?

Options:

- Replace with **0** (Transform → Replace Values).
- Remove rows with null values (Remove Rows → Remove Blank Rows).
- Or fill down values if data is sequential.

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

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

12. Group the table by CustID to show total spending per customer.

- Select **CustID** column → Home → Group By.
- Aggregate with **Sum of TotalSpent**.

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

Select **OrderDate** → Transform → Data Type → Date.

Power Query standardizes all rows into one consistent format.

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

- Add Column → Conditional Column.
- If Price > 100 then "High Value", else "Normal".

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

- Remove unnecessary columns at the start (Choose Columns).
- Filter data early to reduce rows.
- Avoid unnecessary steps.