

Lesson 3 - Data Transformation with Power Query (Part 1)

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

It shows a chronological list of all transformations applied to the data.

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

Use the 'Remove Duplicates' option under the 'Home' tab.

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

It allows you to include or exclude data based on specific values or conditions.

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

Right-click the column header and select 'Rename'.

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

The transformations are applied and the data is loaded into Power BI.

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

Filter the Quantity column with condition 'is greater than or equal to 2'.

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

Use the 'Split Column by Delimiter' or 'Date' transformation tools.

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

Use 'Replace Values' under the 'Transform' tab.

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

Click on the OrderDate column and use the 'Sort Descending' button.

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

Use 'Replace Values' or 'Fill Down/Up' to handle missing data.

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

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

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

Use the 'Group By' feature, select CustID and sum the TotalSpent column.

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

Use 'Change Type' to convert the column to Date format.

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

Use 'Add Column' > 'Conditional Column' and set condition: if Price > 100 then 'High Value'.

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

Remove unnecessary columns and steps at the beginning of the query.