1. What is Data Cleaning in Power BI? Why is it important?

Answer:

Data cleaning in Power BI involves identifying and correcting inaccurate, inconsistent, or incomplete data to ensure it is reliable for analysis. It’s crucial because messy data can lead to incorrect insights and decisions. For instance, cleaning missing values, removing duplicates, and ensuring uniform formats can make reports more accurate and meaningful.

---

2. How do you open the Power Query Editor in Power BI?

Answer:

To open the Power Query Editor:

1. Load your dataset into Power BI.
2. Go to the "Home" tab on the ribbon and click "Transform Data."
3. The Power Query Editor will open, allowing you to clean and transform your data.

---

3. What are the steps to remove duplicate records in Power Query Editor?

Answer:

1. Select the column or columns where you want to check for duplicates.
2. Navigate to the "Home" tab in Power Query Editor.
3. Click "Remove Duplicates."
   The duplicate rows will be removed, keeping only the first instance.

---

4. How can you deal with missing or null values in Power Query Editor?

Answer:
Missing or null values can be managed by:

- **Replacing Values**: Go to "Transform" > "Replace Values" and specify what to replace nulls with.
- **Removing Nulls**: Use the "Remove Rows" option to delete rows containing null values.
- **Filling Values**: Use the "Fill Down" or "Fill Up" option for propagating non-null values.

---

5. What are transformations, and how are they applied in Power Query Editor?

Answer:

Transformations are modifications to clean or reshape data, such as changing text case, splitting columns, or merging tables.

Example:

- To split a column, right-click the column and choose "Split Column" by a delimiter like a comma or space.

---

6. How do you merge two datasets in Power Query Editor?

Answer:

1. Open both datasets in Power Query Editor.
2. Go to the "Home" tab and select "Merge Queries."
3. Choose a common column from both datasets and select the join type (Inner, Outer, etc.).
4. Click "OK" to merge the datasets into one.

---

7. How can you group data in Power Query Editor?

Answer:

Grouping data allows aggregation based on categories.

1. Select the column to group by.
2. Go to "Transform" > "Group By."

3. In the dialog box, specify the group operation (e.g., Sum, Count).
4. Click "OK" to apply.

---

8. How do you change data types in Power Query Editor? Why is it important?

Answer:

To change a data type:

1. Select the column in Power Query Editor.
2. Choose the appropriate type from the "Data Type" dropdown in the column header (e.g., Text, Date, or Decimal).
   Changing data types ensures the data is treated correctly in calculations and visualizations.

---

9. What is the difference between the "Close & Load" and "Close & Apply" options in Power Query Editor?

Answer:

- **Close & Load**: Loads the cleaned data to the Power BI Data Model.
- **Close & Apply**: Saves changes and applies them to the model but doesn't load the data elsewhere.

---

10. How can you create a new custom column in Power Query Editor?

Answer:

1. Click "Add Column" > "Custom Column."
2. Enter a formula to create the new column (e.g., [Sales] * [Profit Margin]).
3. Name the column and click "OK."
   Custom columns enable calculated fields tailored to specific needs.

---

11. What is the significance of the "Applied Steps" panel in Power Query Editor?

Answer:
The "Applied Steps" panel tracks all changes made to the data. Users can view, edit, reorder, or delete any transformation step. This feature ensures transparency and flexibility in the data-cleaning process.

---

12. How do you split a single column into multiple columns in Power Query Editor?

Answer:

1. Right-click the column and choose "Split Column."
2. Select a delimiter (e.g., comma, space).
3. Choose to split into rows or columns.

   The data is split based on the chosen delimiter.

---

13. What is the function of the "Unpivot Columns" option in Power Query Editor?

Answer:

Unpivot transforms wide tables into long ones by turning column headers into rows, making the data ready for analysis.

- Example: If columns represent months, unpivoting will make "Month" a single column with associated data in rows.

---

14. How do you create a conditional column in Power Query Editor?

Answer:

1. Go to "Add Column" > "Conditional Column."
2. Define conditions, such as If [Sales] > 1000, then "High Sales."

3. Name the column and click "OK."
   Conditional columns classify or segment data based on specified rules.

---

15. How can you add or modify index columns in Power Query Editor?

Answer:

Index columns can be added to provide unique identifiers:

- Navigate to "Add Column" > "Index Column."
- Choose from starting options (0, 1, etc.).
  Indexes are useful for creating sequence numbers or unique keys.

16. How can you remove unwanted characters from a column in Power Query Editor?

Answer:

Use the "Replace Values" option in the Transform tab:

- Select the column and go to "Transform" > "Replace Values."
- Enter the unwanted character (e.g., “#” or special symbols) in the "Value to Find" field and leave the "Replace With" field

blank.

This cleans up data by removing unnecessary characters.

---

17. What is the difference between "Remove Rows" and "Keep Rows" in Power Query?

Answer:

- **Remove Rows**: Deletes specific rows based on conditions like top rows, bottom rows, or duplicates.
- **Keep Rows**: Retains specific rows and removes all others, based on criteria like filtering or index.

  Both are used to target and manage row-level data.

---

18. What are the different types of joins available in Power Query, and when would you use them?

Answer:

Power Query offers six types of joins:

- **Inner Join**: Returns only matching rows from both tables.
- **Left Outer Join**: Keeps all rows from the left table and matching rows from the right.
- **Right Outer Join**: Keeps all rows from the right table and matching rows from the left.

- **Full Outer Join**: Includes all rows from both tables.
- **Anti Join** (Left/Right): Returns rows that do not match in the other table.

  Joins are used for merging tables with common columns.

---

19. How can you extract the year, month, or day from a date column in Power Query Editor?

Answer:

1. Select the date column.
2. Go to "Add Column" > "Date."
3. Choose the desired option: Year, Month, Day, Quarter, etc.

   Power Query creates a new column with extracted date parts for easier analysis.

---

20. What are the benefits of using "Query Parameters" in Power Query?

Answer:

Query parameters allow for dynamic inputs, enabling users to change filter criteria or values without editing the query. For example, a parameter can define a region, and the user can adjust it to display data for different regions dynamically.

---

21. How do you duplicate and reference a query in Power Query? Why would you do this?

Answer:

- **Duplicate a Query**: Makes a copy of the query to independently modify it without affecting the original.
- **Reference a Query**: Creates a linked version that depends on the original query's changes.
  These options are useful when you need multiple views of the same data or incremental transformations.

---

22. What is the use of the "Transform Data" feature in Power Query Editor?

Answer:
"Transform Data" enables applying transformations to clean, reshape, or aggregate data. Examples include modifying text case, replacing null values, filtering rows, splitting columns, and aggregating values for reporting needs.

---

23. How can you dynamically rename multiple column headers in Power Query Editor?

Answer:

Use the "Transform Column Names" option:

1. Go to "Transform" > "Format."
2. Choose an operation, like "Uppercase" or "Lowercase," to apply it dynamically to all column names.

---

24. How can you convert a table into a list in Power Query?

Answer:

1. Select the column or columns to be converted.
2. Right-click and choose "Drill Down" or "Convert to List."
   This conversion is helpful when working with a single column as a list input.

---

25. What are the steps to transpose a table in Power Query? When would this be useful?

Answer:

1. Select the table in Power Query.
2. Go to "Transform" > "Transpose."
   This swaps rows and columns, useful for pivoting data layouts, such as switching between wide and tall datasets.

---

26. How do you apply conditional logic without writing a formula in Power Query Editor?

Answer:

Use the "Conditional Column" option:

1. Go to "Add Column" > "Conditional Column."
2. Define conditions using dropdowns (e.g., IF Sales > 1000, THEN “High Sales”, ELSE “Low Sales.”)
   This creates a logic-based custom column.

---

27. What is the difference between filtering and sorting in Power Query?

Answer:

- **Filtering**: Removes data that does not meet specific criteria. For example, filtering rows where sales are greater than 500.
- **Sorting**: Rearranges data in ascending or descending order without removing any rows.

---

28. How do you create a Pivot Table in Power Query, and how is it different from the Excel Pivot Table?

Answer:

1. Select "Transform" > "Pivot Column."
2. Choose the column to pivot and specify the values to aggregate. Unlike Excel Pivot Tables, Power Query's pivoting is part of the data transformation process and feeds into Power BI visuals.

---

29. How do you use the "Fill Down" and "Fill Up" options in Power Query?

Answer:
"Fill Down" propagates data from the row above into blanks below, while "Fill Up" fills blanks with the data from below. This is particularly useful when hierarchical or grouped data has missing category names.

---

30. How do you use "Remove Errors" in Power Query Editor?

Answer:
"Remove Errors" eliminates rows containing errors from the query. To apply:

- Select the column with errors.
- Go to "Remove Rows" > "Remove Errors."
  This ensures cleaner datasets by eliminating problematic rows.

---

31. How do you troubleshoot performance issues in Power Query transformations?

Answer:

- Avoid using excessive steps or unnecessary transformations.
- Use "Query Diagnostics" to analyze each step’s performance and optimize them.
- Work with summarized data whenever possible to reduce load.