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

The **Applied Steps** pane in the Power Query Editor shows a sequential, recorded history of every transformation you have applied to your data. Each step corresponds to a piece of M code. Its purposes are:

- **Audit Trail:** To see exactly what changes were made to the data.
- **Modification:** To edit or delete a specific previous step without redoing all subsequent work.
- **Understanding Data Flow:** To comprehend the logic and sequence of the data cleaning process.

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

1. Select the columns you want to check for duplicates (or select no columns to check the entire row).
2. Go to the **Home** tab (or the **Right-click** menu).
3. Click **"Remove Rows"** > **"Remove Duplicates"**.

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

The **Filter** icon (a funnel shape) in a column header opens a dropdown menu that allows you to sort and filter the data in that column based on its values (e.g., text filters, number filters, date filters). This is the primary way to visually filter rows.

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

Double-click on the "CustID" column header and type "CustomerID". Alternatively, right-click the column header and select **"Rename"**.

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

Clicking **"Close & Apply"** does two things:

1. **Close:** It closes the Power Query Editor window.
2. **Apply:** It saves all the applied transformation steps, executes the final M code, and loads the resulting, cleaned table into the Power BI Data Model. The report canvas then updates to reflect the new data.

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

1. Click the **filter icon** on the "Quantity" column header.
2. Select **"Number Filters"** > **"Greater Than Or Equal To"**.
3. In the dialog box, enter `2` and click **OK**.

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

1. Select the "OrderDate" column (ensure it's a Date/DateTime type).
2. Go to the **Add Column** tab.
3. Click **Date** > **Year** > **Year**. This adds a "Year" column.
4. Repeat for **Month** > **Month** and **Date** > **Day** > **Day**.

(Alternatively, you can use "Split Column" > By Delimiter, but for dates, using the built-in Date functions is more robust).

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

1. Select the "Product" column.
2. Go to the **Transform** tab (or the **Home** tab).
3. Click **Replace Values**.
4. In the dialog box:
    - **Value To Find:** `Mouse`
    - **Replace With:** `Computer Mouse`
5. Click **OK**.

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

1. Click the **filter icon** on the "OrderDate" column header.
2. Select **"Sort Descending"**.

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

Common methods in Power Query:

- **Remove Rows: Home Tab** > **Remove Rows** > **"Remove Empty Rows"** (if the whole row is empty) or filter the column to exclude `(null)`.

- **Replace Values:** Select the column, then **Transform Tab** > **Replace Values**. Leave "Value To Find" blank and enter a value like `0` or the **Average** in "Replace With".
- **Fill Down/Up: Transform Tab** > **Fill** > **Down** (to copy the value from the cell above).

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

In the Power Query Editor:

1. Go to the **Add Column** tab > **Custom Column**.
2. Enter the following:
    - **New column name:** `TotalSpent`
    - **Custom column formula:** `[Quantity] * [Price]`
3. Click **OK**.

The underlying M code added to your step would be:

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

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

1. Select the "CustID" column.
2. Go to the **Transform** tab.
3. Click **Group By**.
4. In the dialog box:
    - **Group by:** `CustID`
    - **New column name:** `TotalSpent`
    - **Operation:** `Sum`
    - **Column:** `TotalSpent` (or `Price` if you haven't created the TotalSpent column yet)
5. Click **OK**.

13. Fix inconsistent date formats in OrderDate.

The most reliable method is to change the column data type.

1. Select the "OrderDate" column.
2. Go to the **Transform** tab or the **Data Type** dropdown on the **Home** tab.

3. Change the data type to **"Date"**. Power Query will automatically interpret and convert the various text formats into a standardized date.

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

1. Go to the **Add Column** tab > **Conditional Column**.
2. Configure the dialog:
    - **New column name:** `OrderLabel`
    - **If:** `[Price]` `is greater than` `100`
    - **Output:** `"High Value"`
    - **Else:** `"Standard"` (or leave as `null`)
3. Click **OK**.

15. Optimize the query to reduce refresh time.

A key optimization is to **remove unnecessary columns and filter rows as early as possible** in the "Applied Steps" sequence. This reduces the amount of data being processed in subsequent steps.

- **Remove Columns Early:** Use **"Choose Columns"** or **"Remove Columns"** immediately after your source step to keep only the columns you need for your report.
- **Filter Rows Early:** Apply any relevant filters (like `[Quantity] >= 2`) right after removing columns.
- **Use Efficient Data Types:** Ensure columns have the correct, most efficient data type (e.g., Whole Number instead of Decimal Number for integers).