

Is Your Data Holding You Back? Clean It Up!



Remove Duplicate Rows

- Use the Remove Duplicates feature in Power Query Editor.
- Select the columns that define a unique record.
- Always preview data to confirm which rows will be removed.
- This prevents skewed results and double-counting.
- Consider if all columns are necessary for the check.
- It is a quick way to ensure data integrity.



Correct Data Types

- Ensure numbers are not stored as text.
- Check that dates are recognized correctly by Power BI.
- Use the data type dropdown in the transform tab.
- Incorrect types cause errors in calculations and visuals.
- Look for columns with the ABC icon which indicates text.
- Change decimal numbers to fixed decimal for currency.



Handle Missing Values

- Identify blanks or nulls with filters in Power Query.
- Decide to remove rows with critical missing data.
- Replace nulls with a default value like zero or "N/A".
- Use the fill feature to propagate values up or down.
- Create a calculated column to flag missing entries.
- Analyze if missing data follows a specific pattern.



Trim and Clean Text

- Apply the Trim transformation to remove extra spaces.
- Use Clean to remove non-printable characters.
- Standardize text case using Upper, Lower, or Proper.
- Replace abbreviations with full names for consistency.
- Remove prefixes or suffixes that are not needed.
- This ensures text fields merge and group correctly.



Split Columns Wisely

- Break apart combined fields like full name or address.
- Use the Split Column by Delimiter feature.
- Choose a common delimiter like a comma, space, or tab.
- Split into a fixed number of columns for predictable data.
- Preview the results to avoid creating too many columns.
- Rename new columns immediately for clarity.



Group and Categorize

- Create new columns to group similar text values.
- Use the Replace Values feature for small groupings.
- For larger sets, use a reference table and merge queries.
- This simplifies reports and reduces visual clutter.
- Create bins for numeric data using the grouping feature.
- Categorize dates into fiscal periods or weeks.



Validate with Conditional Columns

- Add a new column based on rules and conditions.
- Flag outliers or unexpected values for review.
- Create categories like "High", "Medium", "Low".
- Validate data entry against expected ranges.
- Use custom column for more complex logic with M code.
- This adds a layer of data quality control.



Profile Your Data

- Use Column Profile in Power Query to see data distribution.
- Check the column quality indicators for errors and empties.
- Identify outliers in numeric data quickly.
- See a sample of the top and bottom values.
- This helps you decide which transformations are needed.
- Understand your data before you start building reports.



Create a Reusable Template

- Save your cleaned query as a Power BI template file (.pbit).
- Document your steps within the Power Query Editor.
- Use parameters for values that might change.
- Build a standard set of cleaning steps for similar data sources.
- This saves time on future projects.
- Ensures consistency across reports and datasets.



Document Your Process

- Use the Advanced Editor to view the underlying M code.
- Add comments to your steps within the formula bar.
- Rename query steps to be clear and descriptive.
- This makes it easier to troubleshoot later.
- Helps others understand the transformations you applied.
- Documentation is key for maintaining reports over time.



FOLLOW FOR MORE



Like this post? Repost!

Share your thoughts in the comments!

