Transforming Data with Ease

Section 1: Learn

What is Data Transformation?

Data transformation is the process of **cleaning, modifying, and restructuring** raw data to make it useful for analysis and visualization. In Power BI, this is done using **Power Query** and the **Query Editor**.

Why is Data Transformation Important?

- **Improves Data Quality**: Removes missing values, duplicates, and inconsistencies.

- **Enhances Analysis**: Converts raw data into a structured format suitable for visualization.

- **Automates Cleaning Tasks**: Saves time by applying repeatable transformations.

- **Ensures Compatibility**: Standardizes different data sources for seamless integration.

How Data Transformation Works in Power BI?

1. **Load Data** → Import data from Excel, SQL, or other sources.

2. **Apply Transformations** → Clean and shape the data (remove duplicates, filter rows, split columns, etc.).

3. **Use Applied Steps** → Each transformation is recorded as a step.

4. **Load Transformed Data** → Apply changes and push data to Power BI reports.

A Brief History

Data transformation has always been an integral part of data analytics. Traditionally, **SQL scripts** were used for transformation, but modern tools like Power BI have introduced **no-code solutions**, making it accessible for everyone.

Section 2: Practice

Basic Data Transformation Techniques

Step 1: Open Query Editor

1. Open **Power BI Desktop**.

2. Click on **Home > Transform Data**.

3. The Query Editor window opens, displaying the loaded dataset.

Step 2: Common Transformations

- **Removing Duplicates**: Select a column → Click **Remove Duplicates**.

- **Filtering Data**: Click the dropdown in a column → Select values to keep.

- **Splitting Columns**: Select a column → Click **Split Column** → Choose a delimiter.

- **Changing Data Types**: Click on a column header → Change to **Text, Number, or Date**.

Example: Cleaning Customer Data

1. Load an Excel file containing customer details.

2. Remove duplicate customer records.

3. Split the "Full Name" column into "First Name" and "Last Name".

4. Change "Date of Birth" column to Date format.

5. Apply filters to remove inactive customers.

6. Click "Close & Apply" to load clean data into Power BI.

Merging Queries (Joining Tables)

1. Click **Home > Merge Queries**.

2. Select two tables and choose a **common column** (e.g., Customer ID).

3. Select the **Join Type** (Inner Join, Left Join, etc.).

4. Click **OK**, and the merged data appears in Query Editor.

Adding a Custom Column

1. Click **Add Column > Custom Column**.

2. Enter a formula like: ([Sales] * 1.1) to increase sales by 10%.

3. Click **OK**, and the new column is added to the dataset.

Section 3: Know More

Frequently Asked Questions

1. Can I undo transformations in Power BI?

Yes! Every transformation is recorded as a **step**, and you can remove or modify it anytime.

2. Does transforming data affect the original dataset?

No! Power BI works on a **copy of the data**, leaving the original dataset unchanged.

3. What happens when the source data changes?

Power BI **automatically updates** the transformed data when refreshed.

4. Can I apply the same transformations to multiple datasets?

Yes! You can create a **Query Template** and apply it to different datasets.

5. Do I need to code for data transformation?

No! Power BI provides a **no-code, drag-and-drop interface**, but advanced users can use **M language** for complex transformations.

6. How do I save changes in Query Editor?

Click **Close & Apply** to save and load the transformed data into Power BI.

These notes will help you **transform data efficiently** in Power BI, making your reports more accurate and insightful. Mastering data transformation ensures better decision-making and business intelligence!