Power BI ETL Project – Data Leverage

Name: Ashish
Project: Email form Business Analyst Manager

◆ 1. Sources Used

1. Loaded HTML table from web (GDP forecast data from Wikipedia).

2. Loaded 3 Excel files (Sales Jan, Sales_Feb, Sales Mar) from a folder using **Append Queries from Folder**.

3. Loaded Employee dataset from a separate Excel file containing:

   o Employee_ID

   o Name

   o Department

   o Region

   o Join Date

◆ 2. Transformations Applied

✅ **Basic Transformations**

- Removed blank rows and unnecessary columns.

- Promoted first row to headers.

- Renamed columns to meaningful names.

- Changed data types using **Change Type with Locale**.

- Removed duplicates and filtered null values.

✅ **Text Tools Applied**

- Used UPPER() to standardize names.

- Used LOWER() for case conversion.

- Applied TRIM() to remove extra spaces.

- Applied CLEAN() to remove special characters.

- Used REPLACE to correct text values.

- Split Customer Name into First Name and Last Name using delimiter.

✅ Numeric Transformations

- Rounded Revenue column to 2 decimal places.

- Created new Profit column using:

- Profit = Revenue - Cost

✅ Date & Time Transformations

- Extracted Year, Month, Month Name, Quarter, and Day from Order Date.

- Created Fiscal Year column using custom formula.

- Applied proper date data type using Locale settings.

✅ Conditional & Indexing

- Created Sales Category column:

    o High (≥ 10000)

    o Medium (5000–9999)

    o Low (< 5000)

- Added Index columns (0-based and 1-based).

✅ Merging & Appending

- Appended 3 monthly sales files using Folder source.

- Merged Sales data with Employee data using EmployeeID (Left Outer Join).

✅ **Grouping & Aggregation**

- Grouped data by Region.

- Calculated:

  o Total Sales

  o Average Revenue

  o Transaction Count

✅ **Data Profiling**

- Used Column Quality, Column Distribution, and Column Profile tools.

- Verified 100% valid data and removed errors.

✅ **Parameter & Refresh Simulation**

- Configured dynamic folder path using Parameter.

- Simulated adding new monthly file (Sales_Apr.xlsx).

- Verified auto-refresh and transformation maintenance.

◆ **3. Challenges Faced & Solutions**

1. **Null values after merge**
   → Fixed by matching data types (Whole Number) between tables.

2. **Pivot error due to nested columns**
   → Removed nested table columns before applying pivot.

3. **Decimal formatting confusion in Revenue column**
   → Understood Power Query decimal handling and applied correct data type.

4. **Folder path error in parameter**
   → Corrected folder location and updated parameter value.