

Power Query: Data Transformations Assignment

Objective:

This assignment aims to enhance practical skills in using Power Query within Power BI for various data transformation tasks. You will work with bank dataset containing inconsistencies and errors. Your task is to identify these issues and use Power Query to rectify them, ensuring the data is clean, accurate, and well-structured for analysis.

Datasets:

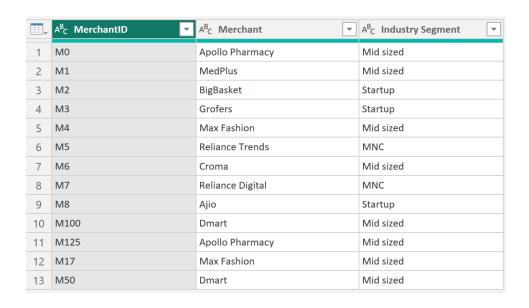
fact_transactions_2022, fact_transactions_2023, dim_merchants, pivoted_dim_category, dim_date

Task1: Data Cleaning in dim_merchants table

- Locate the 'Merchant' column, which contains names that you'll need to split
 using a common delimiter. For instance, if you have a name like "Apollo
 Pharmacy (Mid-sized)," you should split it into two parts: "Apollo Pharmacy"
 and "(Mid-sized)". The new column resulting from this operation should be
 named "Industry Segment".
- In the "Industry Segment" column, remove the blank spaces at the start or end. Use "Trim" function.
- Eliminate the "(" & ")" in the "industry Segment column. Extract values by using "Text Between Delimiters".
- Eliminate duplicate merchants in the table with different IDs.

Desired Format:

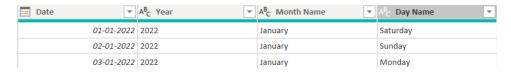




Task2: Extracting Date Information (dim_date)

- Break down the date field in the Date Dimension table into separate components (year, month_name, day_name). Add separate columns for each. Create separate columns for each.

Desired Format:



Task3: Unpivot dim category table

- The data in the dim_category table is not in a suitable format for merging with the fact tables. You should perform a few row transformations and use the Unpivot column option to bring it into the desired format.
- Rename the table to dim_category.

Desired format:





Task4: Filtering unwanted data

- Eliminate transactions with a debit amount below 100 to concentrate on significant transactions. Apply this to both the fact_transactions_2022 and fact_transactions_2023 tables.

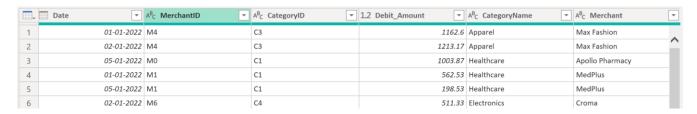
Task5: Append Tables

- Combine data from two years by appending the 2023 transactions to the 2022 transactions table. Take this into separate table named 'fact_transactions' to store this combined data.

Task6: Merge Tables

- Integrate the 'fact_transactions' table with 'dim_category' and 'dim_merchants' to retrieve the corresponding merchant names and category names.

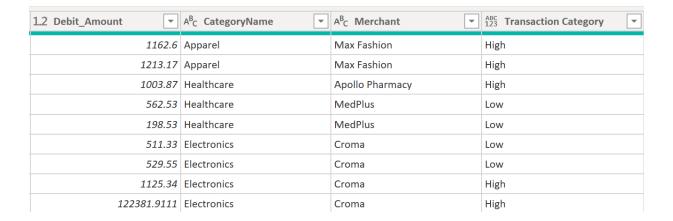
Desired Format:





Task7: Adding Conditional Column

- Categorize debit amount transactions into 'High' and 'Low' using a conditional column. Name the resulting column as 'Transaction Category'.
 - o For amounts below 1000, label as 'Low.'
 - o For amounts above 1000, label as 'High.'



Task8: Sorting and Grouping

 Analyze the total transaction amount per merchant and sort merchants accordingly. (duplicate 'fact_transactions' table and apply the above transformation).

Desired output:

Merchant 🔻	total_transaction_amount
Reliance Trends	2394979.4671
Grofers	2056153.1393
BigBasket	2051215.5436
MedPlus	2017624.6471
Ajio	1867509.1197
Apollo Pharmacy	1744876.0656
Croma	1634962.6198
Reliance Digital	1576590.4082
Max Fashion	1317798.355
Dmart	1002441.3