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:

	MerchantID	Merchant	Industry Segment
1	M0	Apollo Pharmacy	Mid sized
2	M1	MedPlus	Mid sized
3	M2	BigBasket	Startup
4	M3	Grofers	Startup
5	M4	Max Fashion	Mid sized
6	M5	Reliance Trends	MNC
7	M6	Croma	Mid sized
8	M7	Reliance Digital	MNC
9	M8	Ajio	Startup
10	M100	Dmart	Mid sized
11	M125	Apollo Pharmacy	Mid sized
12	M17	Max Fashion	Mid sized
13	M50	Dmart	Mid sized

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:

Date	Year	Month Name	Day Name
01-01-2022	2022	January	Saturday
02-01-2022	2022	January	Sunday
03-01-2022	2022	January	Monday

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:

	CategoryID	CategoryName
1	C1	Healthcare
2	C2	Groceries
3	C3	Apparel
4	C4	Electronics

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:

	Date	MerchantID	CategoryID	Debit_Amount	CategoryName	Merchant
1	01-01-2022	M4	C3	1162.6	Apparel	Max Fashion
2	02-01-2022	M4	C3	1213.17	Apparel	Max Fashion
3	05-01-2022	M0	C1	1003.87	Healthcare	Apollo Pharmacy
4	01-01-2022	M1	C1	562.53	Healthcare	MedPlus
5	05-01-2022	M1	C1	198.53	Healthcare	MedPlus
6	02-01-2022	M6	C4	511.33	Electronics	Croma

Task7: Adding Conditional Column

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

1.2 Debit_Amount	CategoryName	Merchant	Transaction Category
1162.6	Apparel	Max Fashion	High
1213.17	Apparel	Max Fashion	High
1003.87	Healthcare	Apollo Pharmacy	High
562.53	Healthcare	MedPlus	Low
198.53	Healthcare	MedPlus	Low
511.33	Electronics	Croma	Low
529.55	Electronics	Croma	Low
1125.34	Electronics	Croma	High
122381.9111	Electronics	Croma	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