**Project Overview**

**ETL- SQL and PowerBI:**

* Loaded data (CSV file) into SQL

The file “Sample-Super Store was loaded into SQL Server using SQL Task Manager.

This was main input file was then used to create fact and dimension tables to come up with star schema.

While following the dimensional modeling technique it was ensured that:

* 1 Business Entity relates to 1 Dimension Table
* 1 Business Process or Activity relates to 1 Fact table

It was noticed that for some of the business entities the primary keys were not given e.g postal code. In that case I applied a technique of concatenating these fields which gave me a unique key for each Postal Code (business entity).

**Creation of Dimension Table:**

Creation of all dimension tables followed a very similar approach. I took all the related attributes of a business entity of which a dimension table was created for. After taking all the required columns I took the DISTINCT of those columns.To sequentially number each row, ROW\_Number function was used.

**Except for one dimension table - Date**: Date is very standard dimension table which is why I decided to create a classical date table for it.

**Fact table:**

In this star schema only one fact table was used so it was very straight forward. The fact table is simply a transaction fact table. Only Foreign keys and fact columns were shown in the fact table.

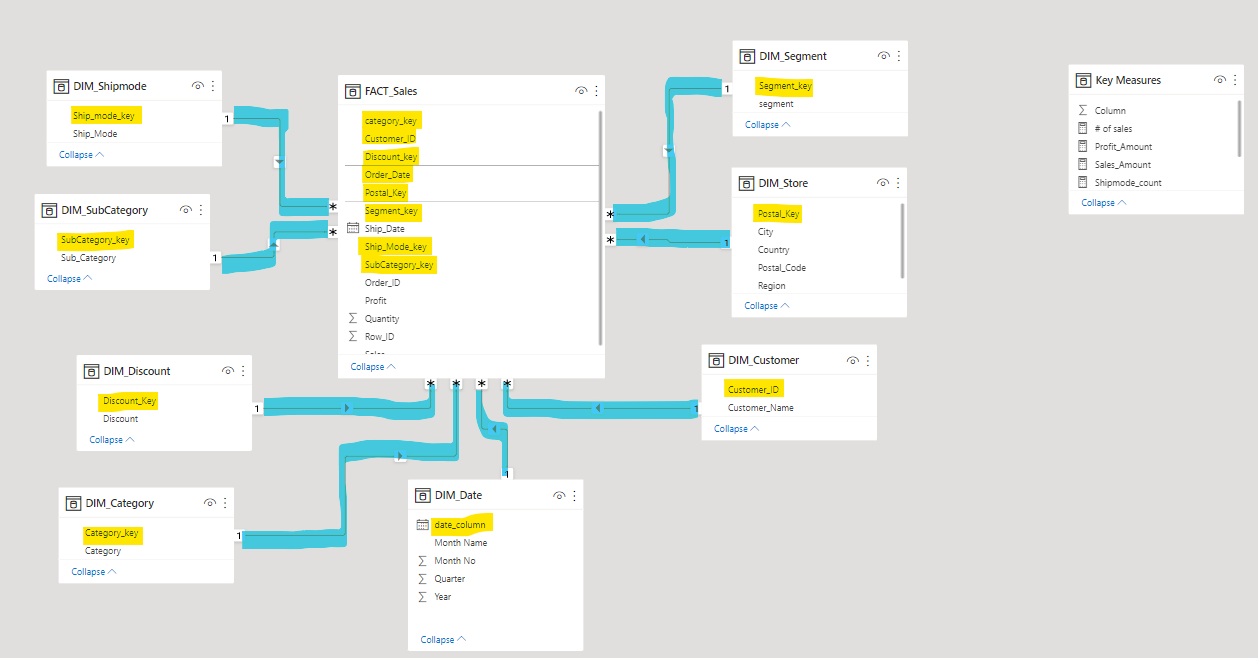
* Loading and Transforming the data

The dimension and fact tables were loaded in PowerBI. Required Transformation were performed on the data to prepare it for visualization.

**Data Model- Start Schema:**

* Dimension Tables (Date, Customer, Store, Category, Sub Category,Segment, Ship Mode, Discount)
* Fact table (Sales, Profit, Discount, Ship Date, Order Date etc. )
* Key Measure: table for measures

As a good data modeling practice the measure table is separately created. This is particularly useful in very large enterprise level data models.



**Data Visualization using PowerBI:**

* Slicers (City, Category, Sub Category, Date(Year, Month)
* Clustered Bar Charts (top 10 Customers with highest sale)
* Line Chart (Sales, Profit for each Month of an year)
* Matrix- Heat maps ( Amount of Sales and Profit for each Sub Category and each Segment)
* Donut Chart (Sales and Profit for each Category)
* Waterfall chart (Sales in each Quarter of an Year)

Dashboard:

[Download](file:///C:\Users\mehak\Downloads\Sample%20-%20Superstore.csv\Project1_SuperStore.pbix) file to open it in Power BI Desktop

[Click here](file:///C:\Users\mehak\Downloads\Sample%20-%20Superstore.csv\Screenshot_pdf.pdf) for Screenshot of dashboard in pdf