**Power Bi Mandatory Project**

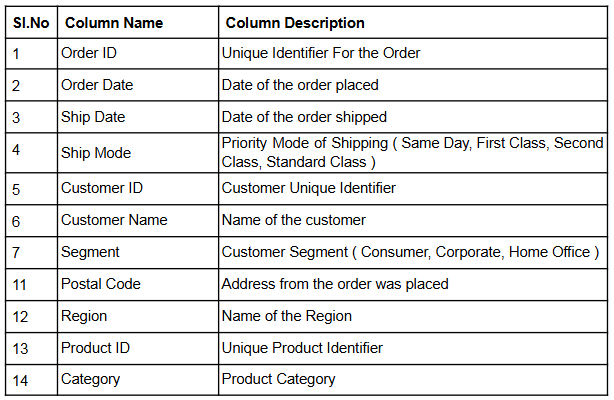
**Case** **Study**

A superstore retail business is a large, multi-department store that sells various products, including groceries, electronics, home goods, clothing, and more. These stores are often designed to be a one-stop shop for customers, offering a wide range of products and services under one roof. Superstores are typically larger than traditional retail stores and may have a larger product selection. Superstores are often part of a larger chain and have multiple locations in a region or country.

**Dataset**

A Superstore dataset typically includes information about the products, customers, and sales associated with a retail store.

Table Description:



# 

Problem Statement: Data Extraction, Cleaning, Loading and Transformation

1. Desk representatives at the stores are not tech savvy hence they directly share the data

in the single excel file. I read the data directly from the excel file.

2. The data coming from the source is in raw form in the flat file; hence I performed cleaning and preparation of data by removing null values .

I deleted the empty columns and rows, change the fields to appropriate data types and split the fields and rename the columns appropriately.

3. I Standardize the values in the column Ship mode . I replaced the values of column into another value such as FC to First class in ship mode column

4. I splitted the address column to City, State, Country and Pincode.

Data Modelling:

1. Tracking sales in the retail business is a weekly task; hence setting up the data model

will be crucial for this. Convert a flat file into STAR schema for better performance of the

analysis. The schema shall have a central Fact table, ‘Orders’ and three dimension

tables, ’Order details’, ‘Customer’ and ‘Product’.

2. Remove duplicate rows from the newly created dimension tables, and ensure there are no empty rows

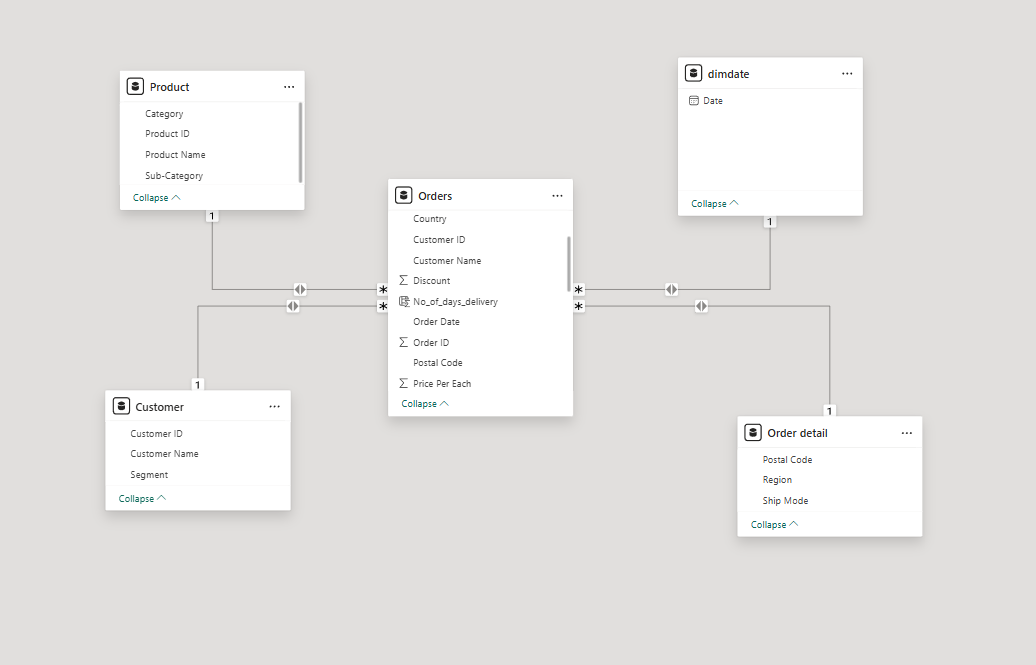
3. Once the tables are created, ensure one- to - many relationships are created between

dimensions and fact table

In Order details Dimension table, I added Ship Mode, Postal Code, Region columns

In Customers dimension table , I added Customer id, Customer name, segment

In Product Dimension table, I added Product ID, Category, Sub-Category, Product Name



Data Analysis:

1. Create a new column ‘Sales’ or ‘Order value’.

I created Sales column by entering below data in power query editor.

Sales = Orders[Quantity]\*Orders[Price Per Each]\*(1-Orders[Discount])

I created a card visual displaying the total sales



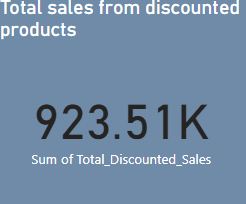
2.Similarly calculate the Sales from discounted products and display the total sales

from discounted products.

I created Total\_Discounted\_ Sales column by entering below data in power query

editor.

Total\_Discounted\_Sales = COALESCE(CALCULATE([Total\_sales],Orders[Discount]>0),0)



3. Since supermarkets sell bulk items, store managers want to know each order's cart

value. Create a column “Cart Value” that categorizes the order value/sales as Low,

medium, high or very high.

Cart Value,

< 1000: Low

<3500: Medium

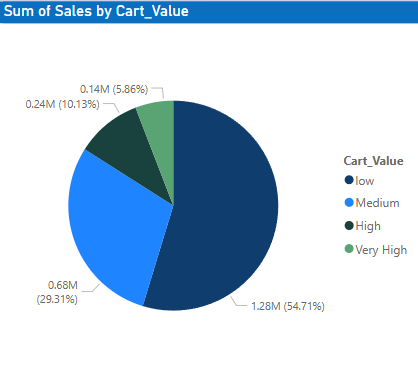
< 10000: High

> 10000: Very High

I created Cart\_Value column by entering below data in power query editor.

Cart\_Value=IF(Orders[Sales]<1000,"low",IF(Orders[Sales]<3500,"Medium",IF(Orders[Sales]<10000,"High","Very High")))

I created a pie chart with Cart value as legend and Order value in Values field



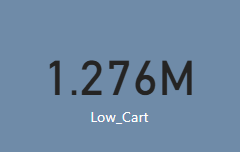
4. Separately visualize the total sales just from the low cart value category (as

mentioned, any value below 1000 can be considered as low value category)

I created Low\_cart Measure by entering below data in power query editor.

Low\_Cart = CALCULATE(SUM(Orders[Sales]),Orders[Cart\_Value]="low")

I created a card visual displaying the Low\_Cart



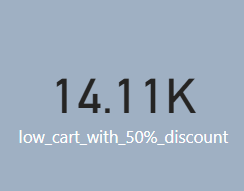
5. Using card visual, track the total sales coming from the low cart category and

discount more than or equal to 50% to find out the contribution and cause.

I created low\_cart\_with\_50%\_discount Measure by entering below data in power

query editor

low\_cart\_with\_50%\_discount = CALCULATE(SUM(Orders[Sales]),Orders[Cart\_Value]="low",Orders[Discount]>=0.5)



6. Find out the number of days it takes to deliver for each shipment type (refer ship

mode) so that delivery issues can be looked at on priority [Hint: No of days to deliver

can be calculated from the difference between order\_date and shipping\_date].

Create a column chart that shows the average number of days it takes to deliver for

each shipment type.

I created No\_of\_days\_delivery column by entering below data in power query editor.

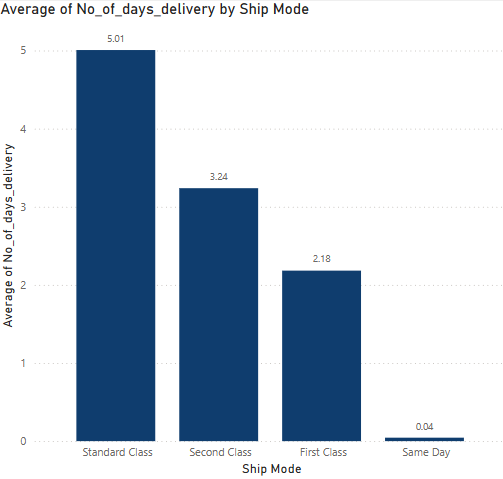
No\_of\_days\_delivery = DATEDIFF(Orders[Order Date],Orders[Ship Date],DAY)

I created a column chart that shows the average number of days it takes to deliver for

each shipment type.

X axis : Ship Mode

Y axis: Average of No\_of\_days\_delivery



7. So far the store manager has managed to see the current snapshot of the sales

based on various criteria. In the Retail business, do we see a spike in sales on

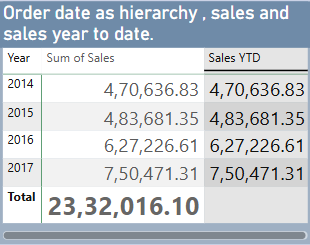
special occasions like festivals? To achieve this, create a matrix visualization that

displays order date as hierarchy , sales and sales year to date.

I created Sales YTD column by entering below data in power query editor.

Sales YTD = TOTALYTD(SUM(Orders[Sales]),DimDate[Date])

I created a matrix visualization that displays order date as hierarchy , sales and sales year to date.



8. Visualize the cumulative sales for each month for all the years to calculate Year on

Year Sales Growth. Calculate YoY growth.

I created Quick measure for Sum of Sales YoY% .And below is dax function for yoy%

Sum of Sales YoY% =

IF(

    ISFILTERED('dimdate'[Date]),

    ERROR("Time intelligence quick measures can only be grouped or filtered by the Power BI-provided date hierarchy or primary date column."),

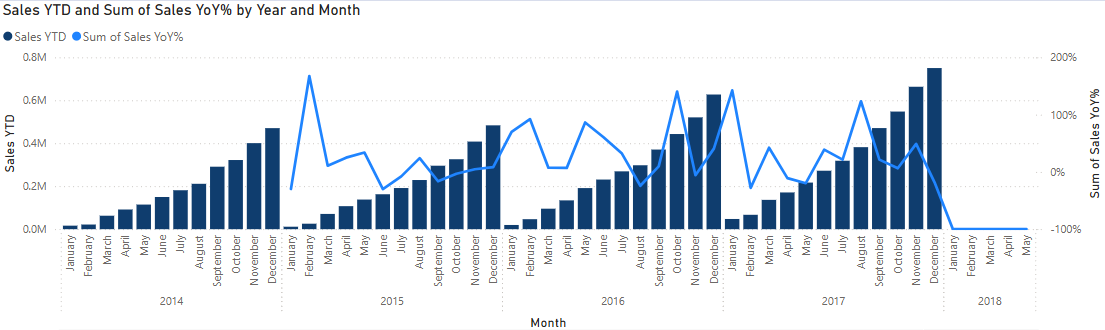
    VAR \_\_PREV\_YEAR = CALCULATE(SUM('Orders'[Sales]), DATEADD('dimdate'[Date].[Date], -1, YEAR))

    RETURN

        DIVIDE(SUM('Orders'[Sales]) - \_\_PREV\_YEAR, \_\_PREV\_YEAR)

)

I used Line and Clustered Column chart for visualization.



And I created line chart and slicer for better analysis

