Power BI Mandatory Project

By- Prince Raj

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.

A new store manager needs your help to better understand his/her Data Operations Team. You are provided with part of the sales data that a Business Intelligence Analyst encounters daily. Design the dashboard to analyze and interpret the data to help provide valuable insights to the store manager.

Dataset Description

Sl. No	Column Name	Column Description
1	Order ID	Unique Identifier For the Order
2	Order Date	Date of the order placed
3	Ship Date	Date of the order shipped
4	Ship Mode	Priority Mode of Shipping
5	Customer ID	Customer Unique Identifier
6	Customer Name	Name of the customer
7	Segment	Customer Segment (Consumer, Corporate, Home Office)
11	Postal Code	Address from the order was placed
12	Region	Name of the Region
13	Product ID	Unique Product Identifier
14	Category	Product Category
15	Sub-Category	Product Sub - Category
16	Product Name	Name of the Product
18	Quantity	Quantity of the product ordered
19	Discount	Discount % on the product
20	Buy Price	Buying price for each item
21	Price Per Each	Selling price for each item

Section- 1: Data Extraction, Cleaning, Loading and Transformation

1) **Read the data directly from the excel file**:
   a) Launch power bi desktop application and select new report.
   b) Click on import data from excel and browse dataset then select it and click open.
   c) There is one table named ‘Orders’, select it and preview.
   d) Click on transform as there is need of data pre-processing and data will load into power query editor.

2) **Delete the empty columns and rows:**
   a) Enable column profile and quality then oversight the data.
   b) Select the “Column 18” in the last and click remove column in manage column ribbon as it contains null only.
   c) There is only one null cell in sub-category column. So, after analysing rest on the feature, product id indicate that it’s belong to subcategory ‘AR’ which stand for art hence, replace the null with ‘Art’ with the help of replace value feature in transform tab.

3) **Change the fields to appropriate data types:**
   a) Click on the calendar icon left to the column name then select date as datatype to change ship date from datetime to date only.
   b) Click on the decimal number left to the column name then select % as datatype to change discount from decimal to percentage.

4) **Split the fields and rename the columns appropriately:**
   a) Select the customer name column and the click on split column tab in text column ribbon, select split on delimiter then select space as delimiter followed by opting split at left-most delimiter then press ok.
   b) Rename both of the newly formed column as first name and last name by double clicking on the column name.
   c) Rename sub-Category column by removing das in-between.

5) **Standardize the values in the column Ship mode:**
   a) Select the column ship mode and replace the value ‘FC’ with ‘First Class’ with the help of replace value tab in any column ribbon.
   b) Click on dropdown arrow right the column name and make sure no FC remains.

6) **Split the address column to City, State, Country and Pin code:**
   a) Select the address column then open split column tab and select split by delimiter followed by opting comma as delimiter and setting it to split on each occurrence then click ok.

b) Rename address.1 as city and address.2 as state by double clicking on column name.
c) Select address.3 column and again go to split column tab and select split by delimiter followed by custom delimiter then provide das (-) as delimiter and press ok.
d) Rename address.3.1 as country and address.3.2 as pin code by double clicking on it.

Section- 2: Data Modeling

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 many- to - many relationships** are created between dimensions and fact table.

- Launch power query editor and duplicate the orders table by right clicking and then selecting duplicate from dropdown.
- Select column while holding ctrl button then choose remove other columns from manage columns ribbon and rename the new table as products.
- Remove duplicates from product id column in product table then remove columns of product table from orders table except Product ID column (FK).
- Repeat the same steps for creating customers and order details table and make sure no duplicates or null value in primary key of all tables.
- Press close and apply then open model view from left most panel of power bi desktop.
- Drag customer id of customers table view to customer id of orders table view and select many to many relations then press ok to establish relationship between both tables.
- Repeat same steps for rest of table with appropriate pk as product id and order id.

Section- 3: Data Analysis

1. **Create a new column ‘Sales’ or ‘Order value’ [Hint: Sales = Qty * *price per qty* (1- % Discount)]. Create a card visual displaying the total sales.**

   - open table view from left most panel and select orders table.
   - opt new column tab from calculation ribbon and write the following expression for new column:

     ```
     Sales = [Quantity] * [Sell price] * (1-[Discount])
     ```

   - from column tool tab format, the decimal number up to two digits in formatting ribbon.

   - Go to report view from left most panel.
   - Select new card view from visualization pane.
   - Tick sales checkbox and adjust the card visual displaying sum of sales.

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

   - Follow the same step instead use the following formula to filter out only discounted products:

   ```
   Discounted Sales = IF([Discount] <> 0, [Quantity]*[Sell price]*(1-[Discount]), 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 and, > 10000: Very High**

   - Open table view and select the orders table.
   - Click on new column in calculation ribbon.
   - Use following expression:

     ```
     Cart Value =
     IF([Sales] < 1000, "Low",
        IF(AND([Sales] >= 1000, [Sales] < 3500), "Medium",
           IF(AND([Sales] >= 3500, [Sales] < 10000), "High",
              IF([Sales] >= 10000, "Very High", BLANK()))))
     ```

   Create a pie chart with Cart value as legend and Order value in Values field.

   - Go to report view and select pie chart from visualization pane.
   - Drag sales in values field and cart values in legend.
   - Adjust pie chart.

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).**

   - Go to model view and select new measure.
   - Use the following expression:

     ```
     low_cart_sales_value =
             CALCULATE(SUM(Orders[Sales]), Orders[Cart Value] = "Low")
     ```
   - Go to report view and select new card from visualization pane.
   - Select newly created measure low_cart_sales_value.
   - Customise visualization accordingly.

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.**

   - Go to report view and select new card from visualization pane.
   - Drag sum of sales in value field.
   - Expand filter pane and add cart value and select low.
   - Add discount and select greater than or equal to in advance filtering and give 0.5 as value in prompt.
   - Tick and button and apply filter then customise the visual card.

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. Create a column chart that shows the average number of days it takes to deliver for each shipment type.**

   - Open model view and select orders table.
   - Click on new column from calculation ribbon and use following formula:

     ```
     days_to_deliver = DATEDIFF(Orders[Order Date], Orders[Ship Date], DAY)
     ```
   - Open report view and select column chart from visulaization pane.
   - Drag ship mode in x-axis for column chart.
   - Drag new calculated column days_to_deliver into value field.
   - Choose average from the dropdown of y-axis and customise visual.

7. **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.**

   - Open model view and select orders table.
   - Click on new measure and use following formula to create sales year to date measure:
     Sales_YTD = TOTALYTD(SUM(Orders[Sales]), Orders[Order Date])
   - Open report view and select matrix chart from visualization pane.
   - Drag order date into rows field of matrix chart.
   - Drag sum of sales and sales YTD into values field.
   - Ensure drill through and also add a slicer for year selection.

8. **Visualize the cumulative sales for each month for all the years to calculate Year on Year Sales Growth. Calculate YoY growth.**

   - Open report view and select line chart from visualization pane.
   - Drag order date into x-axis of line chart.
   - Drag sum of sales into y-axis of line chart and keep only month and year.
   - Drill down to month view and enable pointer and data labels from format visual pane.

   For YoY

   - Go to model view and select oders table.
   - Click on new measure to calculate ThisYearRevenue measure using following formula:

     ThisYearRevenue = CALCULATE(SUM([Sales]))

   - Select orders table and calculate new measure for PreviousYearRevenue using following formula:

     PreviousYearRevenue = CALCULATE(SUM(Orders[Sales]),
     DATEADD(Orders[Order Date].[Date],-1,YEAR))

- Create new measure year of year percent growth (YoY %) using above parameters with following expression:

  YoY % = DIVIDE(([ThisYearRevenue] - [PreviousYearRevenue]), [PreviousYearRevenue])

- Open report view and select line and clustered column chart from visualization pane.
- Drag order date in x-axis of chart and ensure date hierarchy is there.
- Drag this year revenue into y-axis and year over year percent on line field of y-axis.
- Enable trend line from analytics pane and customise visual accordingly.
- Select table chart from visualization pane and drag year of order date, this year revenue, previous year revenue and YoY % into column fields for another visualization of sales growth over year by year.

Note: Open view tab and change theme to innovate from theme ribbon and adjust tiles on report for better visual presentation.

Thanks & Regards

Prince Raj