Maven Market Project Report

Diego Sanoja

2022-09-21

Objective of the Project

Create an interactive report using Microsoft Power BI Desktop for Maven Market, a multi-national grocery chain with locations in Canada, Mexico and the United States to help the company track KPI (Key Performance Indicators), analyze product-level trends, and identify high-value customers.

Data Used

For this project, 8 csv files provided by the company were used, their names are the following:

- MavenMarket_Regions.
- MavenMarket_Products.
- MavenMarket_Customers.
- MavenMarket_Calendar.
- MavenMarket_Returns_1997-1998.
- MavenMarket_Stores.
- MavenMarket_Transactions_1997.
- MavenMarket_Transactions_1998.

Changelog

Version 1.0.0 (09-15-2022)

- Uploaded each of the files to Microsoft Power BI.
- Removed the MavenMarket_ sub string from the name of each table.
- Removed the _1997-1998 sub string from the name of the Returns table.
- Combined the Transactions_1997 and Transactions_1998 files in a folder called Transactions to combine them in the Query Editor.

The following changes were done using the Query Editor:

a. For the Customers file:

- Changed the data type of the customer_acct_num and customer_postal_code columns from numeric to text.
- Added a new column called Full Name which contains the first and last name of the customers using the first_name and last_name columns. Then moved the column next to the ones used to create it.

- Changed the data type of the birthdate and acct_open_date columns from text to date.
- Added a new column called Birth Year which contains the year in which the customer was born.

b. For the Products file:

- Added a new column called product_price_after_discount by multiplying the original retail price by 0.9. Next, The result was rounded to 2 decimal places. Then moved the column next to the one with the original price.
- Replaced all the null values of the recyclable and low_fat columns by 0.

c. For the Stores file:

- Added a new column called full_address which contains the city, region and country of the store by merging the store_city, store_province and store_country using a ',' separator. Then moved it next to the columns that were used to create it.
- Changed the data type of the store_id and region_id columns from string to whole numbers.

d. For the Calendar file:

- Changed the data type of the Date column from text to date.
- Added the following columns using the Date column:

i. Day which contains the number of the day of the month as a number.

ii. Day Name which contains the name of the day of the week as a string.

iii. Start of Week which contains the first day of the week starting from Sunday as a date type.

iv. Month Name which contains the name of the month as a string.

v. Start of Month which contains the first day of the month as a date type.

vi. Quarter which displays the quarter of the year as a whole number.

vii. Year which contains the year as a whole number.

e. For the Returns file:

- Changed the data type of the return_date column from text to date.
- Changed the data type of the product_id, store_id and quantity columns from text to whole numbers.

f. To the combined file of Transactions:

- Changed the data type of the transaction_date and stock_date column to date.
- Changed the data type of the remaining columns to whole numbers.

The following connections were done in the Model view:

i. The Regions and Stores tables were joined using the region_id column.

ii. The Stores Table was joined with the Transactions and Returns tables by the store_id column.

iii. The Calendar table was linked with the Returns table using the Date and return_date columns of the respective tables.

iv. The Calendar table was linked Transactions table by linking the Date column with the transaction_date and the stock_date. The relationship between the Date and transaction_date columns is the only one active of the 2 connections mentioned.

v. The Stores and Transaction tables were linked by the store_id column.

vi. The Products table was linked to the Transactions and Returns tables by the product_id column.

vii. The Customers and Transactions tables were linked by the customer_id column.

Version 1.1.0 (09-15-2022)

The following changes were done in the Data view:

a. For the Products Table:

- Changed the format of the product_retail_price, product_cost and product_price_after_discount from decimals to currency with 2 decimal places.

b. For the Transactions Table:

- Changed all columns of type date to the dd/mm/yyyy format.
- Added a calculated column called Customer Age During Transaction.
- Added a calculated measure called Type of Customer by conditioning the values of the Customer Age During Sale column.
- Added the following DAX measures:

i. Quantity Ordered.

ii. Total Revenue.

iii. Total Cost

iv. Total Profit.

v. Total Number of Orders.

vi. Percentage of Orders.

vii. Previous Month Revenue.

viii. Previous Month Orders.

ix. Previous Month Profit.

x. Target Orders.

xi. Target Revenue.

xii. Target Profit.

- Changed the format of all measures related to profit, cost and revenue to currency with 2 decimal places and the Percent of Items Ordered to percent.

c. For the Return Table:

- Changed the return_date column to the dd/mm/yyyy format.
- Added the following DAX measures:

i. Quantity Returned.

ii. Total Number of Returns.

iii. Percentage of Returns.

iv. Rate of Returns.

v. Previous Month Returns.

vi. Target Returns.

- Changed the format of the Target Returns measure to whole number.

d. For the Calendar Table:

- Added a calculated column called End of Month.
- Added a calculated column called Weekend which has a 'Yes' if the day is either Sunday or Saturday and 'No' otherwise.

e. For the Customers Table:

- Modified the customer_country column category to country.
- Modified the customer_city column category to city.
- Modified the customer_state_province column category to state/province.
- Modified the full_address column category to address.
- Added a calculated column in the Customer table called Customer's Current Age.

f. For the Stores Table:

- Modified the store_country column category to country.
- Modified the store_city column category to city.
- Modified the store_state column category to state/province.

g. For the Regions Table:

- Modified the sales_district column category to city.

Version 1.2.0 (09-16-2022)

The following changes were done in the reporting view:

- Added an extra page for visuals.
- Changed the name of the sheets from page 1 and page 2 to 'Key Metrics' and 'Tables and Graphs' respectively.
- Added the following visuals to the Key Metrics page:

a. A picture with the logo of the Maven Market company.

b. A matrix whose rows are the names of the 50 most profitable product along with the quantity ordered of those products, revenue and profit generated, and percentage of returns for each of these products.

c. A card which displays the name of the customer that has bought more products.

d. A card which displays the amount of items sold to the customer who bought the most products.

e. A card which displays the most profitable brand of products.

f. A Card which displays the amount of profit generated by the most profitable brand.

g. A KPI (Key Performance Indicator) which shows the amount of orders at the end of each month and compares it to the target value for that respective month.

h. A KPI (Key Performance Indicator) which shows the Total Profit at the end of each month and compares it to the target value for that respective month.

i. A KPI (Key Performance Indicator) which shows the Total Revenue at the end of each month and compares it to the target value for that respective month.

j. A KPI (Key Performance Indicator) which shows the amount of returns at the end of each month and compares it to the target value for that respective month.

k. A slicer to filter the data based on the date.

- Added the following visuals to the Tables and Graphs page:

a. A side by side horizontal bar chart which displays the amount of orders by day through each year.

b. A line chart with 2 lines whose y-axis and x-axis are represented by the amount in dollars and last day of the month respectively. The first line displays the total profit and the second line shows the total revenue at the end of each month.

c. A tree map that separates the amount of orders by country and by year.

d. A vertical side by side bar chart that separates the type of customer by the percentage of orders through each year.

e. A horizontal side by side bar chart that separates the number of orders by the country through each year.

f. A donut chart that displays the profit by the store type along with the amount of orders in each one.

Analysis and Insights

For the analysis phase, each image on each page was examined and each provided valuable insights from the data.

Key Metrics page

The first visual examined was the matrix which contains the name of the 50 most profitable products of the company during the time period of 1997-1998.

The matrix was required to reach the following conclusions:

a. In the 1997 year:

b. The most purchased product from the group was the 'Special Wheat Puffs'. During the year, this product was bought 267 times, and generated a total revenue and total profit of 776.97 and 512.64 dollars respectively.

ii. The product that generated both the most revenue and profit from the list was the 'Hermanos Green Pepper'. This product was purchased 239 times and generated a total revenue and total profit of 922.54 and 619.01 dollars respectively.

b. In the 1998 year:

c. The most purchased product from the group was the 'Tell Tale Fresh Lima Beans'. During the year, this product was bought 467 times, and generated a total revenue and total profit of 1480.39 and 901.31 dollars respectively.

ii. The product that generated the most revenue from the list was the 'Great Pumpernickel Bread'. This product was purchased 466 times and generated a total revenue and total profit of 1742.84 and 959.96 dollars respectively.

iii. The product that generated the most profit from the list was the 'Booker Strawberry Yogurt'. This product was purchased 396 times and generated a total revenue and total profit of 1576.08 and 1073.16 dollars respectively.

iv. In the 2-year period of 1997-1998:

v. The most purchased product from the group were the 'Tell Tale Fresh Lima Beans'. During the 2 years, this product was bought 698 times and generated a total revenue and total profit of 2212.66 and 1347.14 dollars respectively.

vi. The product that generated both the most revenue and profit from the list was the 'Hermanos Green Pepper'. This product was purchased 645 times and generated a total revenue and total profit of 2489.70 and 1670.55 dollars respectively.

After analyzing the matrix, the 4 cards at the left bottom of the page were studied and from them it was concluded that:

a. In 1997:

b. Mary Francis Benigar was the customer who bought the most products during the year. The amount of products bought by this customer was of 518.

ii. The Hermanos brand was the brand that generated the most profit. The profit generated by this brand was on 337.13 hundred thousand dollars.

b. In 1998:

c. Scott Littleford was the customer who bought the most products during the year. The amount of products bought by this customer was of 814.

ii. The Hermanos brand was the brand that generated the most profit. The profit generated by this brand was on 715.69 hundred thousand dollars.

iii. In 1997-1998:

iv. Ida Rodriquez was the customer who bought the most products during the 2 years of data provided. The amount of bought products by this customer was of

1021.

ii. The Hermanos brand was the brand that generated the most profit. The profit generated by this brand was on 1.05 million dollars.

To conclude with the visuals on this page, the KPIs were studied.

Each KPI compares the value of the metric (quantity of items ordered, quantity of items returned, total revenue or total profit) to the target value that the company wished to reach by the end of the month. The values shown are the ones of the last month of the 1998.

By looking at them it was concluded that there was a significant increase in each of the metrics after the 1997 year was concluded. The reason for those sudden increases is explained by the visuals of the Tables and Graphs page which will be discussed next.

Tables and Graphs page

The first visual that was analyzed of this page was the tree map which shows the amount of items sold. The tree map was also designed in a way that also drew a line to separate the proportion of the sales of each country by year. By looking at this visual it was observed that during the year 1997, the sales of the company were only done in the USA while in 1998, products were also sold both in Canada and Mexico. This fact also explains the sudden increase in the 4 metrics showed by the the KPIs discussed in the previous section.

The next visual examined was the vertical side by side bar chart of total profit by country which shows that during the USA was the only country that generated profit during 1997 while in 1998 both Mexico and Canada also generated financial gains for the company. This fact also explains the sudden increase in the profit shown by one of the KPIs from the previous section. Furthermore, at the top of each bar, it is possible to see the exact value of profit provided by each country which also helps to conclude that USA was the most profitable sector for the company both in 1997 and 1998.

To continue with the analysis, the vertical side by side bar chart of Percentage of Orders by Type of Customer was studied. This bar chart shows that: adults(customers whose age is equal or higher than 18 and lower than 60 years old) made the greatest percent of sales during the 2-year period. Additionally, during 1998, no teenagers(customers with less than 18 years of age) bought products of Maven Market.

The next visual analyzed was the donut chart which separated the types of stores of the company by the profit each one generated. Furthermore, the amount of items sold at each of those store was included in the chart using tool tips. This charts provided evidence that both the most profitable type of store and the one

where the greatest amount of sales occurred was the supermarket which constitutes a 44.74% of the donut followed by the deluxe supermarket which constitutes a 37.96% of the visual.

The second to last visualization of the page is horizontal side by side bar chart which shows the amount of sales through each day and then is split by the year. By examining this visual, the following conclusions were reached:

a. More sales occurred in 1998 than in 1997.

b. In 1997:

c. The Sunday was the day where more sales occurred through the entire year. A total of 40677 items were sold this day.

ii. The Tuesday was the day were less sales occurred through the entire year. A total of 33953 items were sold this day.

iii. In 1998:

iv. The Thursday was the day where more sales occurred through the entire year. A total of 85678 items were sold this day.

v. The Tuesday was the day were less sales occurred through the entire year. A total of 76472 items were sold this day.

vi. In the 1997-1998 period:

vii. The Thursday was the day where more sales occurred.

viii. The Tuesday was the day were less sales occurred.

e. The difference between the total amount of sales was not so significant between the days of the week in each year.

To conclude with the analysis, the line chart graph was studied. This graph contains 2 lines, one for the total revenue and one for the total profit. The x-axis is represented by the end of the month date. From the line chart, it was discovered that:

a. The month which generated both the greatest profit and revenue for the company was December of 1998 which were 71682.44 and 120160.84 dollars respectively.

b. Through 1998, the amount of profit was above 50 thousand dollars each month and the amount of revenue was above 90 thousand dollars each month.