

Data Mining Project with Alliance Sales & Marketing

Prof., Sunny Park | MIS 5560 | Introduction to Data Science

Team: Galactic

**TEAM MEMBERS**

Jalpaben Patel

Uroob Mullick

Jenna Reinert

Ka Lor

Fadi Yousif

Index

1. [Executive Summary](#_Executive_Summary)
2. [Data Mining Summary](#_Data_Mining_Summary)
   1. [Business Understanding](#_Business_Understanding)
   2. [Data Understanding](#_Data_Understanding)
   3. [Data Preparation](#_Data_Preparation)
   4. [Model Building](#_Model_Building)
   5. [Testing and Evaluation](#_Testing_and_Evaluation)
   6. [Deployment](#_Deployment)
3. [References](#_References)

# Executive Summary

Alliance Sales and Marketing is a food broker that helps market consumer packaged goods. Alliance Sales and Marketing collects data to share insights and analytics to their clients. This helps their clients by putting them in a better position to market their products and lead to higher sales. After reviewing the data that was collected we have found that Greek SS Yogurt has the highest sales share at 15.3%. Greek SS Yogurt was number one in sales out of 27 different yogurts followed by Kid’s Yogurt. After looking at the data we noticed that when there is no markdown on the pricing, the sales dollars stay consistent with minor spikes but when there is the markdown, we see a significant increase in sales dollars and units. This type of data helps manufacturers understand what type of yogurt sells the best so they can better position themselves.

# 

# 

# 

# 

# 

# Data Mining Summary

# 

# Business Understanding

Alliance Sales & Marketing is a food broker that essentially helps market Consumer Packaged Goods (CPG) to various retailers for a fee. Their clients are usually smaller to medium size CPG companies who need assistance in advocating and/or marketing their products to retailers in hopes retail stores will help sell their products. Alliance Sales & Marketing is CPG’s strategy to enter the market. This dataset is collected as part of showing clients the marketing effectiveness of their service. After all, no one wants to pay for a service that is not providing any business values. Alliance Sales & Services have captured weekly sales from last year to current year of yogurt products to depict the revenue generated by using their marketing service to CPGs. In other words, Alliance Sales & Marketing have collected data to share insights and analytics to their clients, there is significant value in hiring them to broker their yogurt products. Not only that, this data set provides valuable insights to clients so they can better position themselves in the market to generate more revenue.

# Data Understanding

This data provides insights and analytics to help clients understand where their product sits compared to their competitors. It allows clients to filter and/or slice the data in many ways so they can zero in on the information pertinent to them. For example, this dataset ranks yogurt type by sales revenue and volume. It shows that Greek SS Yogurt has the highest sales share, 15.3%, out of the 27 different yogurts followed by Kid’s Yogurt. This helps manufacturers understand what type of yogurt sells the best so they can better position themselves to enter that market, if they haven’t already. It also shows sales changes from last year to this year, it also shows how pricing, promotions have an impact on sales volume. Data insights and analytics is such a powerful asset and is critical in helping clients better position themselves to increase revenue.

There’s a Manufacturer Overview tab that helps clients compare their data to a competitor’s. The Manufacturing Ranking tab shows where clients rank during that period, a good indicator of how competitive you are, whether you’ve moved up or down the rank. If you’re up the rank then keep doing what you’ve been doing, if down the rank then you have to pivot and re-strategize on how to get on top. The Stores Selling Summary shows which stores sells the most

Client Sales Timeline tab shows dollars sales from last year, dollar sales this year, average price for last year and average price for this year. We noticed when there is no markdown on the pricing, the sales dollars stay consistent with minor spikes but when there is the markdown, we see a significant increase in sales dollars and units. There is a strong correlation in markdowns and increased unit and sales revenue.

# Data Preparation

1. Import Data

1.1. Loaded all CSV files into Power BI

# The first step was to import all 4 CSV files to Power BI.

# There were 4 CSV files provided for data visualization of Alliance Sales and Marketing data which consists of two main commodities: 062 Yogurt and 290 NF Yogurt.

# The first 2 files are for 062 Yogurt which falls under the dairy department and the sales data is available from the year 2019 – 2020. While the second 2 files are for 290 NF Yogurt which falls under the natural foods department and the sales data is available from the year 2019 – 2020.

Below steps were followed:-

1. Get Data > Text/CSV > Navigate to Files > Load

# 

# 1.1.2.

Combine Query (Append)

# We combined all 4 files together by appending the data which is attaching new rows to the bottom of the data.

Below steps were followed:-

1. Import all 4 files separately (see 1.1). all files are now in Power BI, but not yet combined.
2. Open the Power Query Editor by selecting Transform Data in the Queries group on the ribbon
3. In the Combine group > Append Queries > Append Queries as New
4. First Table > Weekly Division w\_ Kiosk
5. Second Table > Weekly Division w\_ Kiosk (1)
6. Third Table > Weekly Division w\_ Kiosk (2)
7. Fourth Table > Weekly Division w\_ Kiosk (3)
8. New table created called All\_Weekly\_Division\_Data\_Append
9. “Close and Apply” [top left]
10. “Apply Changes” on top yellow bar

2. Clean and Verify Data

**Cleaning** data is changing data to make it most usable by Power BI.

**Verifying** data is making sure that changes Power BI makes is done correctly.

2.1. **Remove Metadata:**

# After importing each file we had transformed the data by first removing the top 11 rows, then using the top row as header, next we had changed the data type from number to text values for manufacturing code (MFR\_CD) and UPC.

Below steps were followed:-

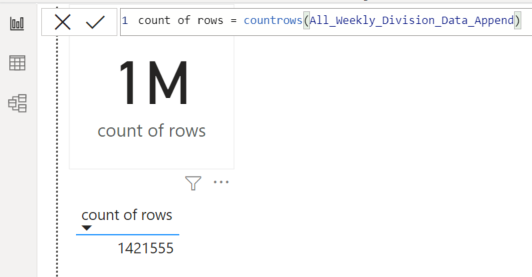
1. In Queries group, select Transform Data to open Power Query Editor
2. Remove Rows > Remove Top Rows > 11 rows [number of rows where metadata is located]
3. Set top row as header : “Use First Row as Header” in Transform group
4. “Close and Apply” [top left]
5. “Apply Changes” on top yellow bar

2.2. **Verify Correct Number of Rows:**

In order to verify that the correct number of rows are appended through the combining query, we use card visualization to display and confirm the count of rows.

Below steps were followed:-

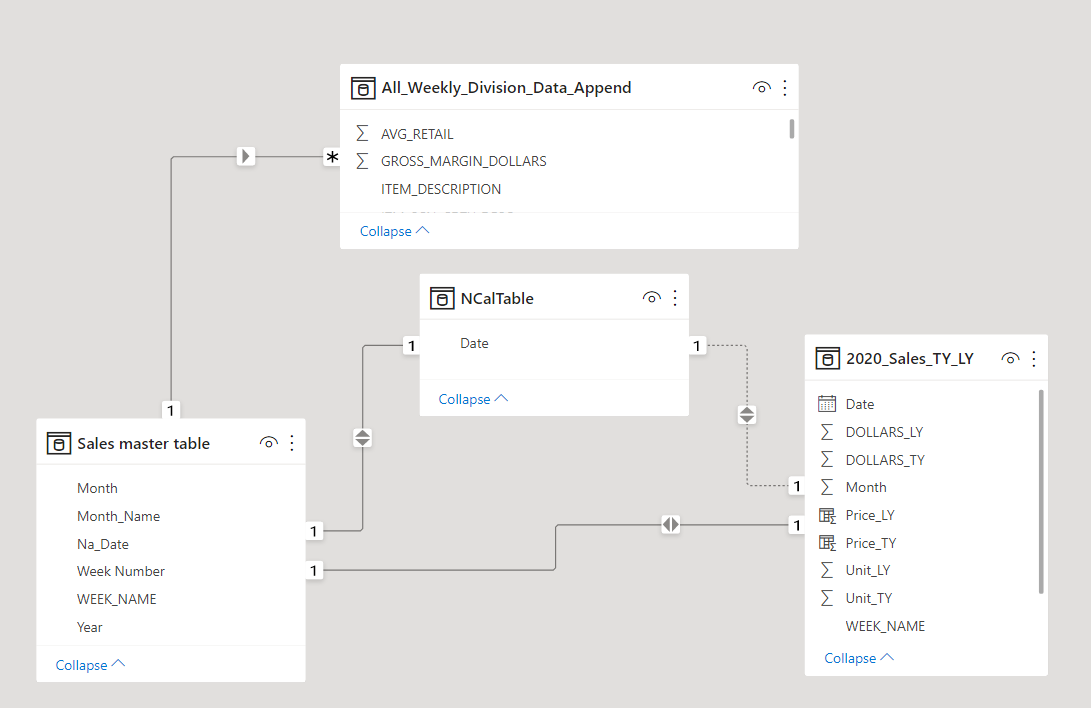
1. Add a “Card” visualization (make sure you de-select the current visualization by clicking in a blank spot in the Canvas before clicking on the Card icon)
2. Select “New Measure”
3. Use the following code: Count of Rows = COUNTROWS(All\_Weekly\_Division\_Data\_Append)



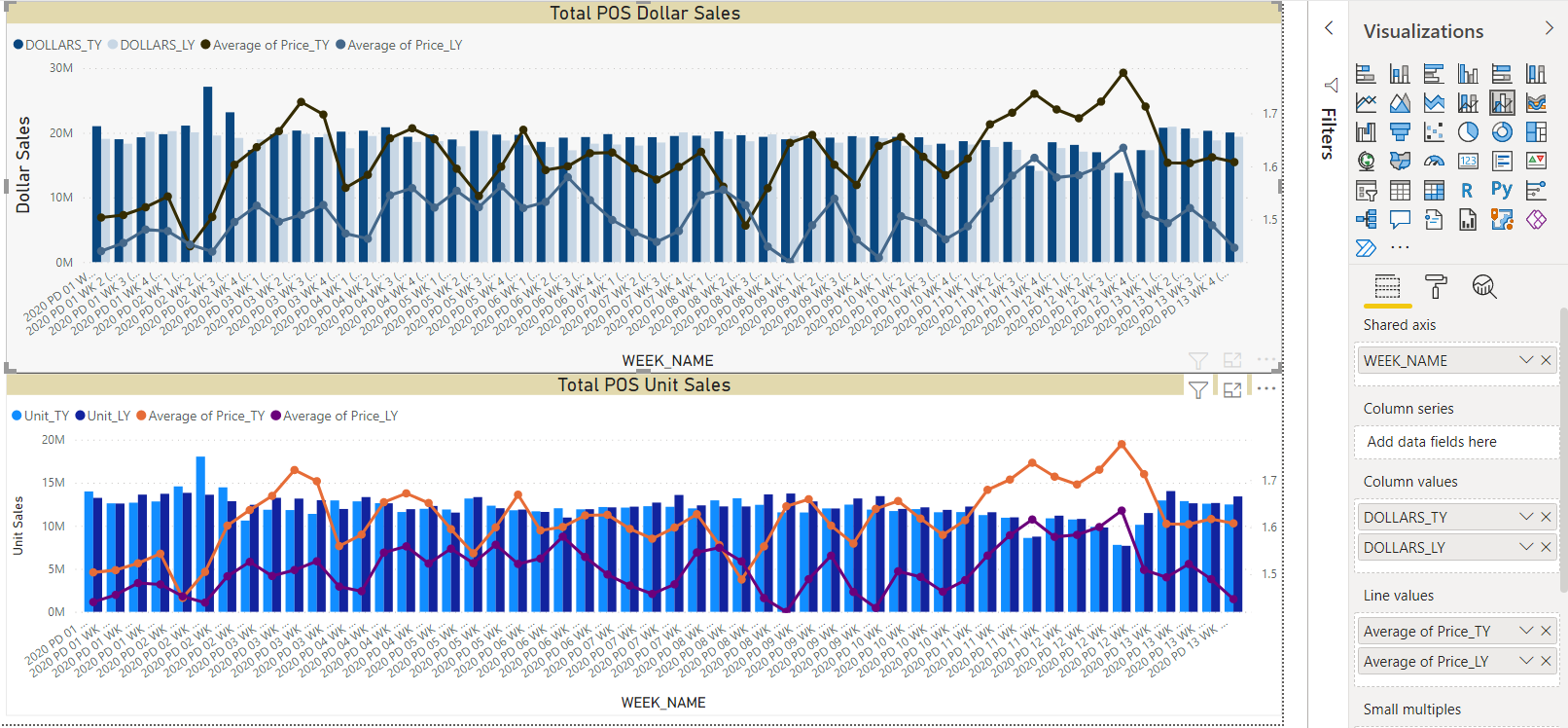
# Model Building

The append table contains the sales data of 2019- 2020. In this report, Power BI tool used for visualization. Which focused on the ‘Manufacturer Sales History’ for given data.

* Create table with calendar date from 1/1/2019 to 12/31/2020, this table called as “NCalTable”
* Then add master table which contain date values including Kroger date and week number
* In order to represent the 2019-2020 sales history, create another table which contains the week number along with “Scanned\_Movements” and “Scanned\_Retail\_Dollars” for both years
* Then, create a new measure called Unit price which is basically the division of Total Sales dollars and Total Units.
* Now establish relationships between all these four tables.



* In Report, use Line and Clustered column charts for data visualization, just drag and drop the data into working space.



* By dragging Sales data under the Values tab and turning it into a stick bar, you’ll see its trends over time.

Below steps were followed for Total POS Dollar Sales visualization:-

1. Place fields in the following sections:
   1. Shared Axis:

i. WEEK\_NAME (Sales master table)

* 1. Column Values

i. DOLLARS\_TY (2020\_Sales\_TY\_LY)

ii. DOLLARS\_LY (2020\_Sales\_TY\_LY)

* 1. Line Values

i. Average of Price\_TY (2020\_Sales\_TY\_LY)

ii. Average of Price\_LY (2020\_Sales\_TY\_LY)

Below steps were followed for Total POS Unit Sales visualization:-

1. Place fields in the following sections:
   1. Shared Axis:

i. WEEK\_NAME (Sales master table)

* 1. Column Values

i. Unit\_TY (2020\_Sales\_TY\_LY)

ii. Unit\_LY (2020\_Sales\_TY\_LY)

* 1. Line Values

i. Average of Price\_TY (2020\_Sales\_TY\_LY)

ii. Average of Price\_LY (2020\_Sales\_TY\_LY)

# Testing and Evaluation

The main benefit of this visualization is the ability to see consumer responses to changes in price. This can be used for sales forecasting, and allow companies to set up strategic promotions to best boost their sales and revenue. On January 7, 2020 for example, there was a sharp decrease in the average price of yogurt, which led to a large increase in sales. Having this visualization for companies to analyze helps them to set a more profitable average price of their product.

We believe that the visualization that was created is accurate, and could be used for a variety of different companies and brands, depending on the data put into it. Since the model we created was based on historical data over two years, the trends that it shows are an accurate representation of consumer behavior; as average price increases, consumers are less willing to spend money on the product. The company can use this to set promotions, determine a fair average price, and continually monitor consumer behavior to see if trends shift.

# Deployment

In order to use this visualization, you need to understand the charts created and the information that they show. The chart titled “Total POS Dollar Sales” shows the data for the total sales for the company over the two years monitored, and displays this with the average price of yogurt for each of the years. As discussed before, this chart can be used to show the relationship between sales and price, and the company can use this to their advantage when coming up with future prices. The second chart, “Total POS Unit Sales” displays similar data, although this chart’s unit of measurement is the total sales, as opposed to the dollar sales. These charts can continuously be updated with new data as it comes in, in order to show emerging trends that consumers may develop, in order to continuously update the business strategy.

In creating these visualizations, there were some challenges faced. When creating the charts mentioned above, we had some difficulty in displaying both years of data, as the chart only shows the dates for this year. Finding a way to display this data and have it make sense for the client was challenging, but we were able to overcome this by using settings within Power BI. Going through the initial data was also challenging at times, but we were able to clear this data and prepare it to create the final visualization.

# References

Alliance Sales and Marketing (2020, 07 09) Retrieved from the Alliance Sales and Marketing Inc.: <https://alliancesalesinc.com/>

Combo chart in Power BI - Power BI | Microsoft Docs (2021, 10 07) Retrieved from Microsoft Docs: <https://docs.microsoft.com/en-us/power-bi/visuals/power-bi-visualization-combo-chart>