After cleaning and wrangling the data set for my capstone project, I started to analyze the data based on the outcome of each month. There are 12 variables for Outcomes outlining sales months based on 751 products. I calculated the minimum, maximum, median, and mean for each sales month. I found the lowest and highest outcomes in each sales month. Based on these calculations, I filtered the products into 2 categories. First category was list of products that were sold at less than or equal to the median of each sales month. Second category was list of products that were sold at greater than or equal to the median of each sales month. From this, I found 182 different products were sold below median of each sales month and 117 products were sold above median of each sales month. It was very interesting to find out how many products were high selling products out of 751 products in the dataset.

Since there are 751 products in each sales month, it was very challenging to analyze the dataset. I made stacked bar plot to see the trends of the sales throughout the year. Having 751 products plotted in stacked bar plot (12 in one bar), it was very tough to visualize the plots. Therefore, I made scatterplot Month 1 through 12 to show the yearly sales. I made comparison between stacked bar plot versus scatterplot. From the comparison, scatterplot was better fit for my dataset. From scatterplot, it was easy to understand the trends of the sales. By visualizing the plot, I decided to focus more on high outcome products only. There were also outliers that were sold very high compared to other products throughout the year.

There are also Date\_1 and Date\_2 variables available in my dataset. Date\_1 is the day number the major advertising campaign began and the product launched. Date\_2 is the day number the product was announced and a pre-release advertising campaign began. For these two variables, I also calculated their minimum, maximum, median, and mean to see how many days products has been advertised throughout the year. I also used scatterplot see the trends of pre-advertising and major advertising days.