Assignment 2

Problem Statement

Having a great marketing strategy in place is key to the success of any business. Without a marketing strategy, you lack focus. And without focus, you will, quite simply, fail to reach any of the goals and objectives that you have set. Any information about customers allows marketers to gain a laser-sharp understanding of their target audience. The marketing budget is being set for the year 2023. The marketing director would like to know which holiday brings in the most money so the team can adjust the marketing dollars.

Objective

What holidays should the marketing team invest more marketing dollars in? Also, find out whatever other insights you can from the dataset.

1.How would you segment holidays based on the expenditure of customers?

Segmentation of holidays would be based on the expenditure of customers in the following way:

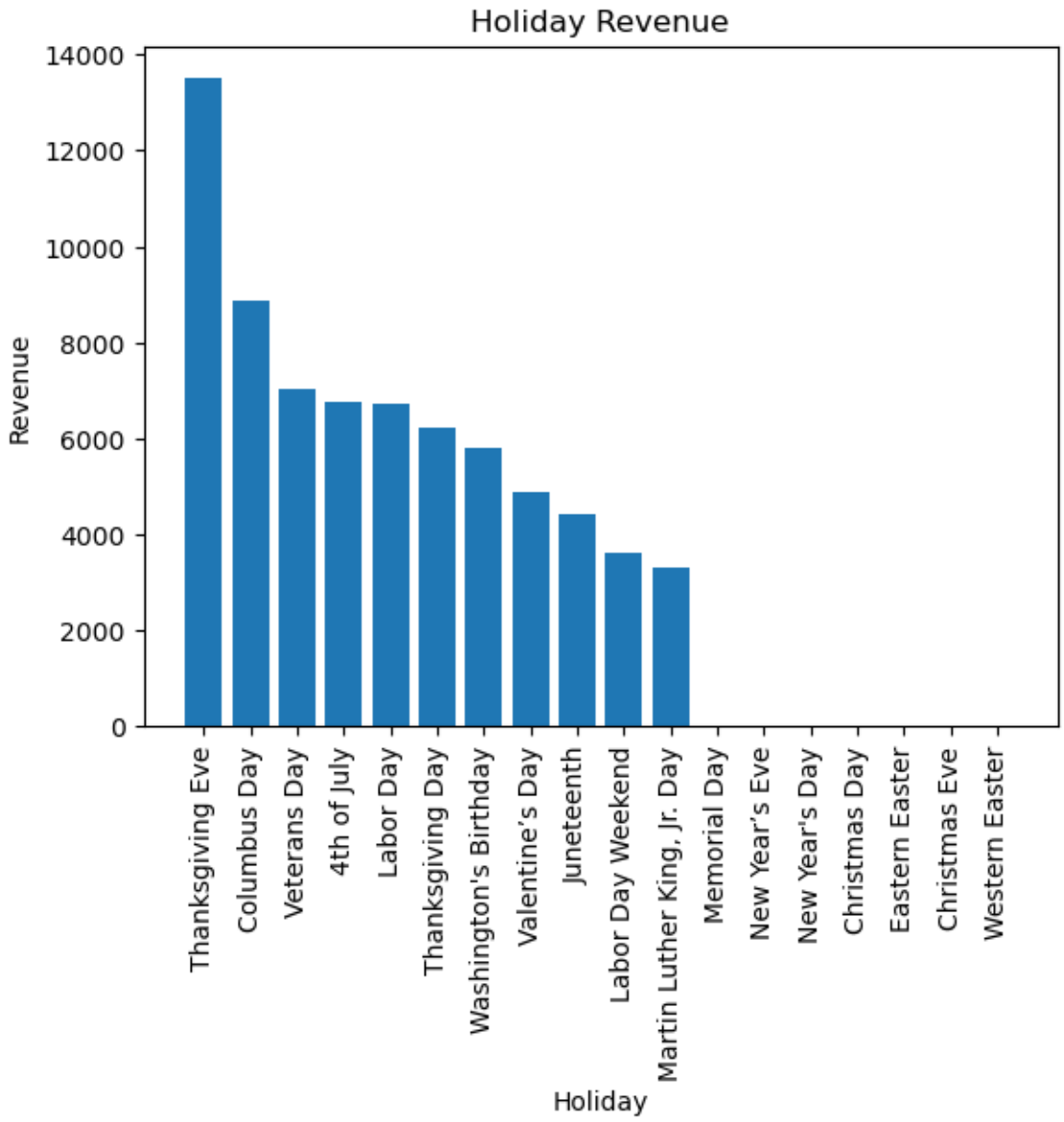
- High-expenditure holidays: These are holidays where customers spend a lot of money. Some examples of high-expenditure holidays include Thanksgiving, Columbus Day and Black Friday.
- Low-expenditure holidays: These are holidays where customers spend a small amount of money. Some examples of low-expenditure holidays include Memorial Day, and Labor Day.
- Total expenditure: This will give me an idea of how much money is being spent on each holiday overall.
- Number of customers: This will give me an idea of how many customers are buying on each holiday.

2.Which of these segments / sub-segments would you propose be approved?

I would propose that the marketing team focus their efforts on high-expenditure holidays. This is because these holidays have the potential to generate the most revenue for the company. The marketing team can do this by creating targeted marketing campaigns that appeal to customers who are likely to spend money on these holidays.

For example, the marketing team could create a Thanksgiving marketing campaign that focuses on food and decorations. By targeting these specific segments of customers, the

marketing team can increase the chances of generating sales during these high-expenditure holidays.



3.What other insights in general can you share about these segments?

Based on the segmentation of holidays by expenditure, we can draw several insights that can help the marketing team optimize their strategy:

Thanksgiving and columbus day are the top two holidays in terms of total sales, followed by Veterans and 4th of july. These holidays should be a priority for the marketing team, and more marketing dollars should be allocated towards them.

Christmas and New Year lower total sales than the top four holidays but still represent a significant opportunity for sales. The marketing team should consider investing more in these holidays to capture more sales.

The duration of the holiday season can affect sales, as we can see in the case of Christmas, which has a longer holiday season compared to other holidays. The marketing team should consider adjusting their strategy based on the duration of the holiday season.

Overall, by understanding which holidays drive the most sales and which segments are most impacted by these holidays, the marketing team can optimize their strategy and allocate their marketing dollars more effectively.

4.Tell us what your observations were on the data itself (completeness, skews) and how you would treat any anomalies (for eg - missing data)

Completeness: The given datasets don't have much null values and we can say it is a complete dataset.
Skews: The ecommerce data contains several outliers in the Quantity and Unit price columns, which could skew our analysis if not handled properly. We may need to remove these outliers or transform the data to reduce their impact on our analysis.

To treat any anomalies or missing data, we can take several steps:

- Imputation: If there are missing values in the data, we can impute them using various techniques, such as mean imputation or regression imputation.

- Outlier removal: If there are outliers in the data, we can remove them using various techniques, such as z-score method or interquartile range (IQR) method.

- Transformation: If the data is heavily skewed, we can transform it using various techniques, such as log transformation or Box-Cox transformation.