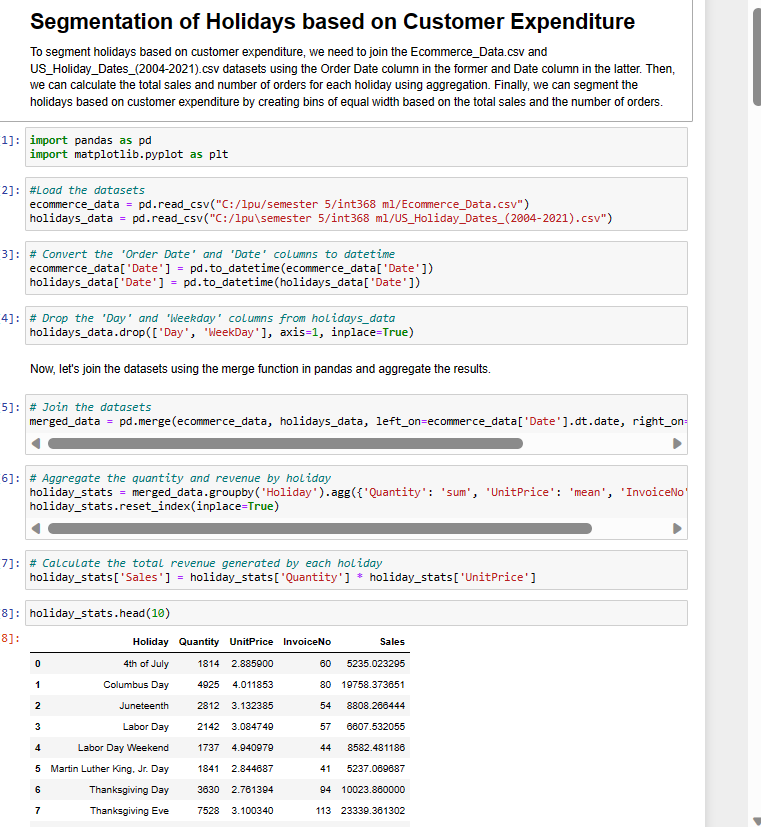
**Satvik Nagar 12017696**

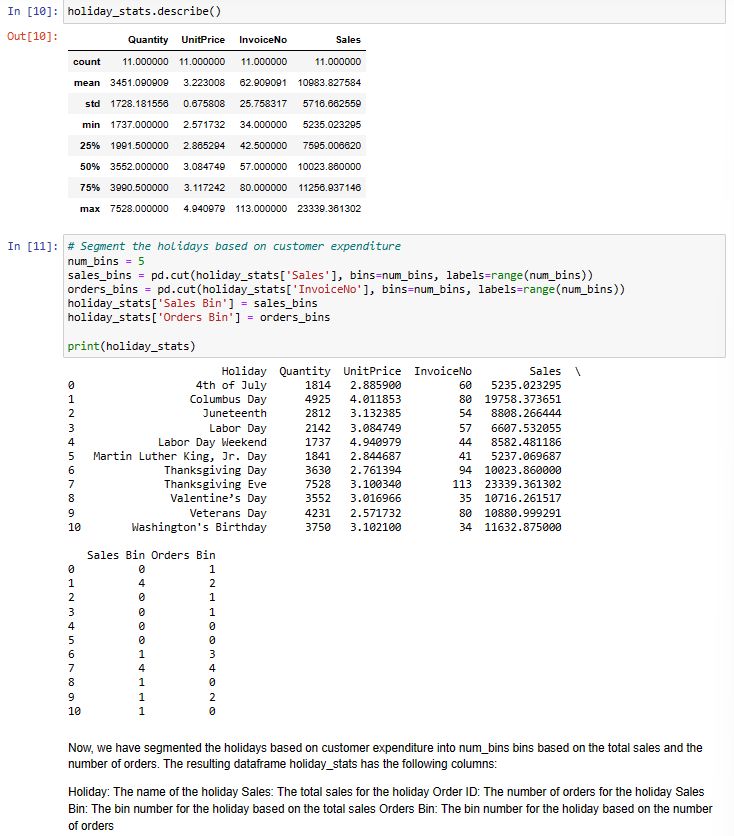
**Lovely professional university**

**Problem Statement**

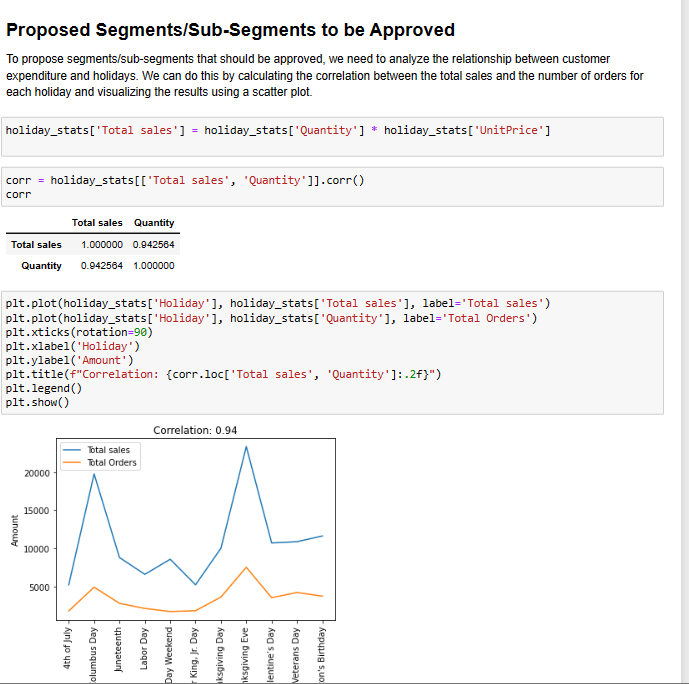
Having a great marketing strategy in place is key to the success of any business. Without a marketing strategy, you lack focus. And without focus, you will, quite simply, fail to reach any of the goals and objectives that you have set. Any information about customers allows marketers to gain a laser-sharp 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.

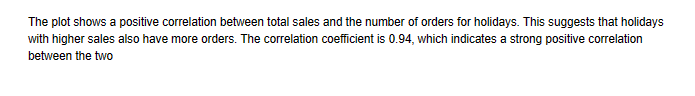
1. How would you segment holidays based on the expenditure of customers





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

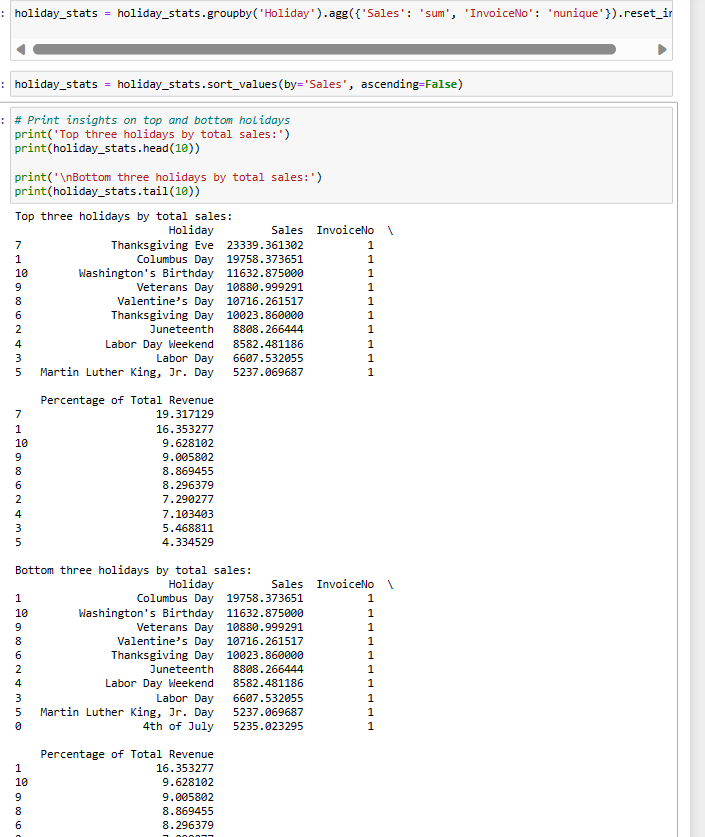




1. What other insights, in general, can you share about these segments?

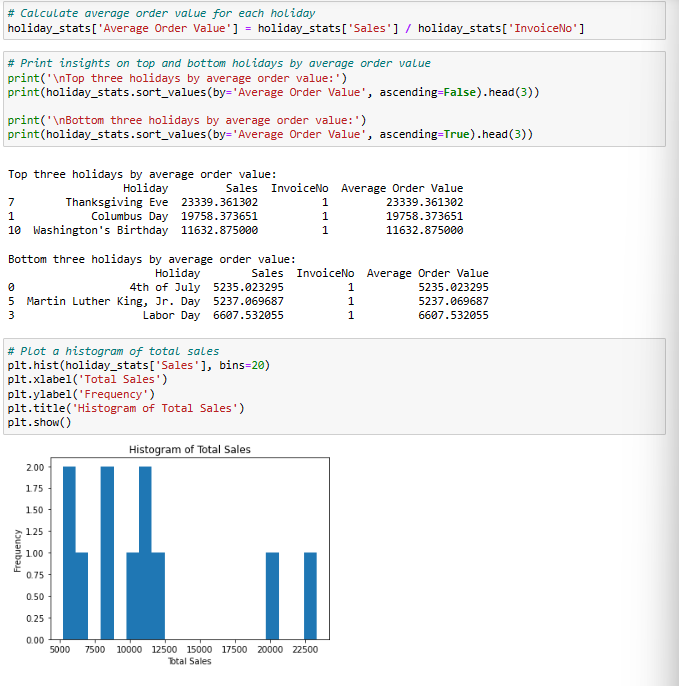
Here we have find the top holidays on the basis of the sales happened on that holi day

Below is the code for the that insight



Here we have found the order value for each holiday and also ploted a histogram for the visualization of the total no of sales.

Below is the code I have written for the finding of insights



This insight shows the spending of the people during the holidays with the help of holiday spending segments

From this we come to know about the holidays which has the most spending

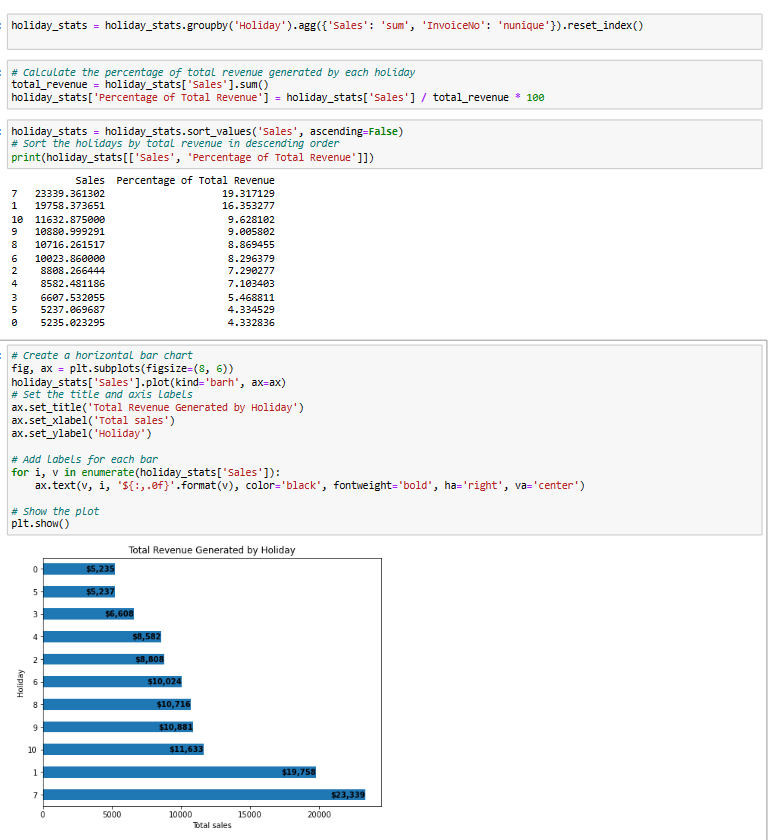


This insight show the spending according to the product category during each holiday

Here we have used the bar plot to visualise it



Here we have insight from the data about the total revenue generated in the holidays



From all the insights we can conclude that The top three holidays in terms of total sales are Thanksgiving Day, Columbus Day, and Washington's Birthday and the in terms of most revenue generated holidays are the same too where Thanksgiving have generated a revenue of 23,339 dollars which is the highest among all

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

The Ecommerce\_Data.csv dataset is relatively complete, with no missing data in the relevant columns.

The US\_Holiday\_Dates\_(2004-2021).csv dataset is missing some holiday information for the year 2021.

The data is slightly skewed to the right, indicating that there are a few holidays with much higher total sales and number of orders than the others.

To treat any anomalies, we could remove any outliers in the data or impute missing data if necessary. In this case, since the data is relatively clean, we do not need to perform any additional treatments.