In todays world, everything is available at the tip of your fingers. Online shopping has created a space for smaller companies to make their product available to the world. This new method of shopping has shaped a new world for physical retail stores. In order for companies to stay profitable and relevant, they need to understand what makes people want to buy certain products. The retail business is of interest to our group because of the joy online shopping brings to our lives. The amount people spend is also an indicator of the economy.

The objective of this project is to look at a few questions that overall help answer how people are spending their money online. This topic was chosen because of the relevance it holds in our lives and it interests us to see how healthy our economy is. To accomplish this, we took an online shopping dataset from Kaggle datasets named “Online Shopping Dataset.” To make this dataset usable, we dropped some columns that were not relevant to our analysis. Rows with null values were also dropped to ensure a clean dataset. To make the analysis easier, we combined like categories of “Nest” since they are the same company but different locations. Finally, we combined New Jersey, New York and DC and renamed it as “Northeast.” This made the three locations within the dataset similar in size to each other. This completed our data cleaning.

Is there a relationship between the cost of a product and the delivery charges people are willing to spend? Are people willing to spend more money on delivery on a larger purchase since they are already spending a large amount? Or are people willing to spend on cheaper items in relation to delivery? According to the following scatter plot, we find no relationship between the item cost and the shipping cost. It brings into question some of the accuracy of the delivery charges. Are delivery charges being used as the source of revenue for some of these companies? This plot raises more questions than it answers.

Since no questions were not answered in the previous plot, a new scatter plot was made to answer if the time of year affected the delivery charges. Again, the outliers are seen throughout the year except a dip in the delivery charges during the warmer months of the year. Are these outliers freight charges? What are people buying during the cooler months that would prompt a willingness to spend more on delivery? Again, more analysis is required to make sense of the data.

The final graph in relation to delivery charges is in relation to the time of year and if there was a difference between gender. Overall, men and women tend to spend the same amount on delivery charges throughout the year. There is a time of year when men will spend more on delivery. This is generally in the summer months. What is this in relation to? Are these outdoor items that require special delivery methods? Or are these last minute anniversary gifts that have to be expedited? Women will spend significantly more just before the “official” shopping season begins. Overall, there don’t seem to be a relationship between any delivery charges and the time of year, gender and product cost.

Divakar, Jackson  R. “🛒 Online Shopping Dataset 📊📉📈.” *Kaggle*, 12 Nov. 2023, www.kaggle.com/datasets/jacksondivakarr/online-shopping-dataset.

“Xpert Learning Assistant.” *Bootcamp Spot*, bootcampspot.com/. Accessed June 2024.