Customer Buying Patterns Report

# Overview

This analysis is based on 80,000 transactions provided by Danielle Sherman, with 79,979 of those transactions being non-duplicated. The data contained 5 columns:

* The transaction type: online or in-store
* The customer’s age (mean is 45.76 years)
* The number of items purchased in a transaction (mean is 4.5 items)
* The value in dollars of items purchased in a transaction (mean is $835.92)
* The region where the transaction occurred: North, South, East, or West

Within this data, the West had the most transactions (25,988) and the North had the least number of transactions (15,997). The South had 19,994 transactions, and the East had 18,000 transactions. The number of transactions is divided roughly equally between online and in-store.

# Do customers in different regions spend more per transaction? Which regions spend the most/least?

Chart, box and whisker chart

Description automatically generatedA couple of plots show the differences in regional spending. The box plot in Figure 1 shows that the spending patterns are indeed different between the sales regions. The colored areas for each of the plots show the interquartile range, that is, the second and third quartiles for spending in each region. The center line within the colored area shows the mean amount spent in the region per transaction, and the upper and lower lines represent the minimum and maximum values. The diamond points above North and East represent outlier values.

From this graph, we can see that the West region has the highest average amount spent per transaction, with the interquartile range being mostly higher than the mean for every other region.

We can also see that the South region has the lowest average amount spent per transaction, with the interquartile range being lower than the interquartile range of every other region.

Figure : Box Plot showing Amount Spent by Region

Table

Description automatically generatedFigure 2 shows the distribution of amount spent per transaction, grouped by region.Text

Description automatically generated Again, we can see that the South region has the highest density of transactions under $500, with the West and East regions clustered under $1000.

Figure 2: Distribution of Amount Spent by Region

In Figure 3, we can see that the West region has the highest amount spent on average, and the South region has the lowest amount spent on average.

One interesting note is that the South region did hot have any in-store transactions, and the North region did not have any online transactions. That might be worth exploring.

Figure 3: Mean amount spent per region, in-store and online

# Is there a relationship between the number of items purchased and the amount spent?

Chart

Description automatically generatedIn general, a significant relationship between the number of items purchased and the amount spent does not appear to exist. In Figure 4, a scatter plot of 10,000 transactions is displayed by the amount spent and number of items purchased. Also displayed is a linear regression line. We can see that the line is fairly horizontal, suggesting that there is not much of a difference between the number of items purchased and the amount spent in a transaction.

Figure 4: Scatter plot showing amount purchased vs items purchased

# Conclusions

* The West region has the greatest average spending, and the South has the lowest average spending.
* There does not appear to be a relationship between the number of items purchased and the amount purchased.