**INVESTIGATE CUSTOMER BUYING PATTERNS**

**Executive Summary**

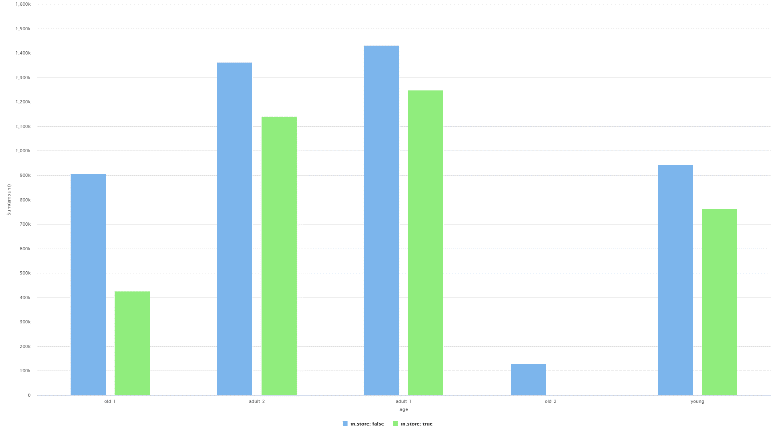
Mr. Goodrich from VP of Sales wants to make some marketing activities and some changes to the website. He hypothesized that (1) customers who shop in the store are older than customers who shop online and that (2) older people spend more money on our products than younger people. Another hypothesis raided by Sherman is: the differences in transactions and customer demographics might be regional. The goal of this analysis is to understand how our customer spend their money in our locations (store/online) in order to maximize the amount of revenue.

An exploratory analysis was done using data sent by Mrs. Sherman. In this database there was information about: age of the customer, number of items purchased, amount spent, region and local (in store/online).

The results show that Goodrich’s hypotheses are partially right. (1) There is no difference in the ages of consumers who used our online platform or visited our stores (Figure 1A), but consumers who used the online platform had a tendency to be older. (2) In the total spent, adult between 29 and 42 years old spent more money than the other age groups (young and elderly), following adults between 43 and 55 years old, young people, elderly between 56 and 74 years old and finally, elderly over 75 years old (Figure 1B). Therefore, we can see that customers who buy our products online tend to be older and that adults between 29 and 42 years old spend more money. Nevertheless, Sherman’s hypothesis seems to be true. The amount spent was much higher in the central region than in the other regions and the age distribution of consumption also varies between regions (Figure 1C). For example, young people and adults between 29 and 42 years old spent more money in the central region, while only the elderly over 75 years old bought our products in the western region. All in all, I recommend expanding our business in the west region, as we did not get good economic results. Also we could provide offers to older people as a target audience.

Please observe these results carefully as I observed some strange behaviors in the data. For example, people over 75 years old only occur in the western region and still in the western region all customers have bought online. For this reason, I suggest we collect more data to better analyze our customers’ consumption behavior. An example of data that we should collect is temporal and spatial data. Finally, I would like to ask you some questions about the dataset I used in this analysis: (1) How long was the data collected? (2) Has online and in-store data collection started at the same time? (3) If there is delay, how long? And (4) if do our stores sell other products besides electronics? Depending on the answers to these questions, the result may change.

**Obs.:** For more results, please check the Appendix. There we can find more information about regional variations in amount spent and age of customers (see topic 1), correlation of age of customers and place of purchase (online/store) (see topic 2), and relationship between number of items purchased and amount spent (see topic 3).



**C**

**B**

**A**

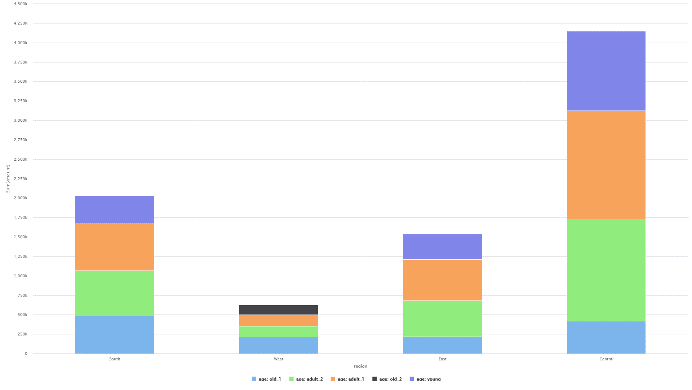
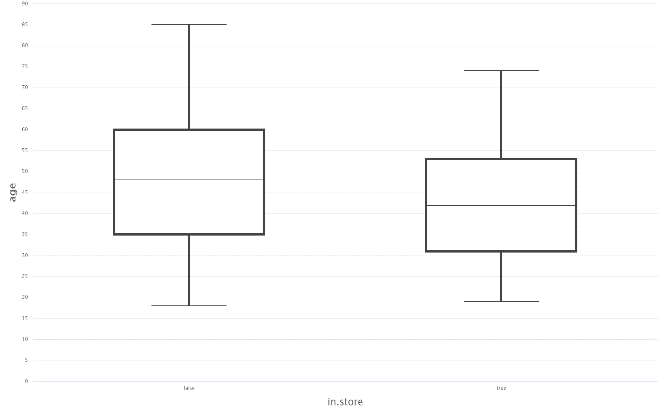


Figure 1. (A) Place of purchase (online and store) is on the x-axis and age is on the y-axis. (B) Age groups (seniors over 75 years old, adults between 43 and 55 years old, adults between 29 and 42 years old, elderly between 56 and 74 years old and young people between 18 and 28 years old, respectively) are on the x axis, sum of the amount spent on the y axis, blue represents online purchases and green represents in-store purchases. (C) Region categories (South, West, East and Central, respectively) are on the x-axis and the amount spent on the y-axis, blue represents seniors over 75 years old, green represents adults between 43 and 55 years old, red represents adults between 29 and 42 years old, black represents elderly between 56 and 74 years old and purple represents young people between 18 and 28 years old.