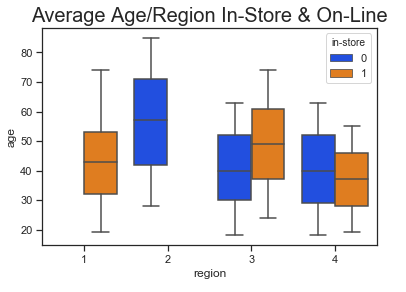
Hello Danielle,

I am pleased that you were able to utilize my last analysis. Per your request, I did a little more investigating and I found some answers that will be very helpful as you build a marketing strategy geared towards your Ecommerce website.

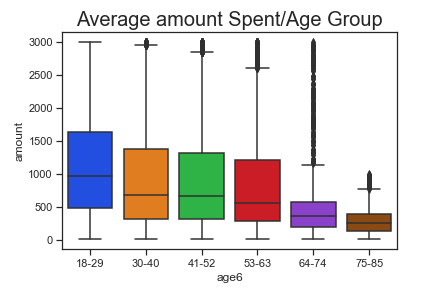
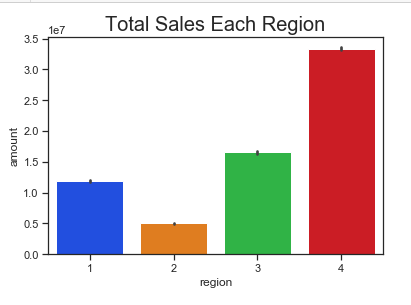
Here are results of my analysis.

This was very surprising. You mentioned that Martin thought that customers who shop in-store would be on average older than those who shop on-line.

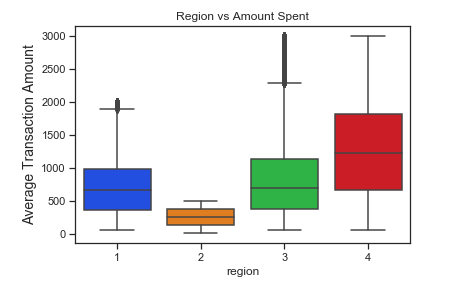


The graphs above show that the average age of the customer who shops online is older than the customer who shops in-store. What is significant is that in region 2, which is comprised of 100% on-line purchases, has the oldest average shoppers. Region 4 which is a mix of on-line and in-store purchases overall has the youngest average aged customers.

However, older customers 64-85 spend quite a bit less on each transaction than younger customers. Customers in the 18-29 age groups spend significantly more on each transaction than those who are 64—85. Region 2 which had the highest average aged customer also had significantly less in total sales.



The best predictor of what region a customer is from is the amount they spent. Region 4 has the largest transactions, the highest total sales and the youngest customer base.



Here are my thoughts. There are fewer older customers 64-85 and they spend less on each purchase of electronics than do the younger demographics. However, older shoppers tend to shop on-line more frequently than in-stores.

Shoppers in youngest age group 18-29 spent the most per/transaction and the oldest 64-85 spent the least/transaction. If Martin wishes to market to older customers, he should focus on attracting more older customers and also focus on increasing the amount they spend/transaction. He may want to market to younger customers in region 2 because younger customers spend more each time they shop.

If you have any other questions, please let me know.

Have a great day,

Sincerely,

Sherri Koski

The algorithtm I chose to use in this exercise was the decision tree classifier. I ran this several times and I may have tried every sort of y imaginable. I used all sorts of age groups(discertized), items, amount(discertized). However, using y =region provided the best accuracy. I achieved my best accuracy 64.91 when I split the amount into 6 bins. I also played around with tree depths and achieved the best results at max\_depth =5.