MAILING LIST PROJECT

Tad Young, Gebil Jibul, Alexia Diaz

March 4, 2023

Table of Contents

Introduction

Business Problem/Hypothesis

Methods/Analysis

Results

Recommendations/Ethical Considerations

Conclusion

References

Appendix

Introduction

For our second and final project we wanted to investigate the different groups of possible customer target groups that would be best to send offers to. The different variables we have in our dataset are genre, age, annual income, and their spending scores. Using a scatter plot and Kmeans clustering, we will look to plot these customers into 4 different groups to determine which would be best to send out our offers to. We don't want to spend extra money sending offers to customers who don't have a high probability of accepting these offers. We need to be strategical and put ourselves in the best position to market.

Along with the clusters, we are also going to put these customers into bins based on the different variables. From low to high we will be able to determine the probability these customers will have based a lot on their spending score, along with the other variables mentioned.

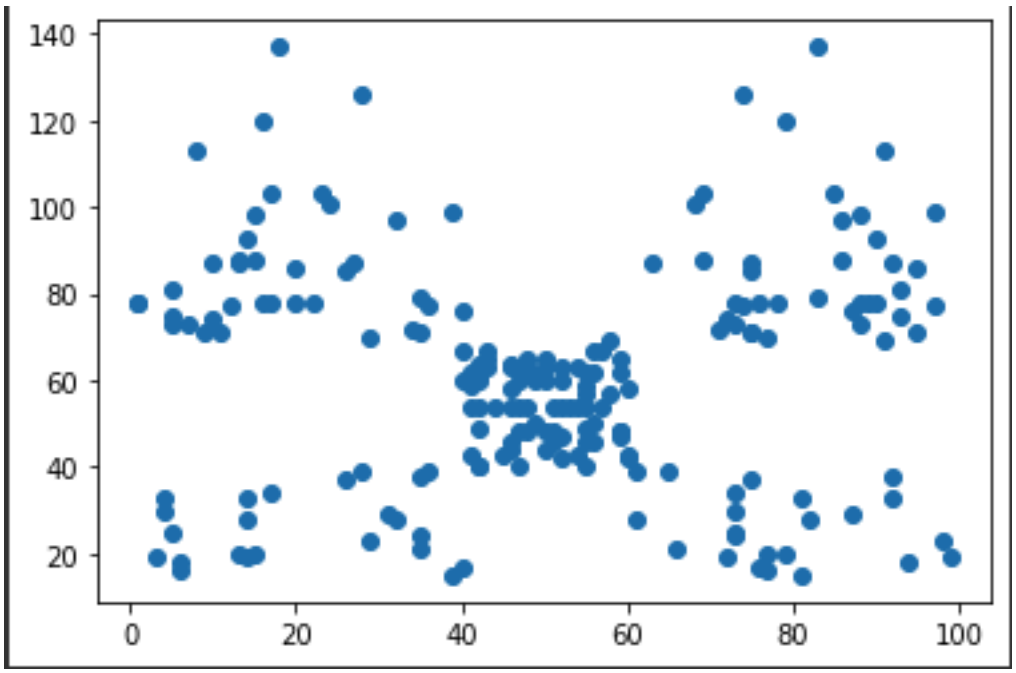
Hypothesis

Our hypothesis is the higher the spending score, along with annual income, will result in customers more likely to accept promotional offers. Factors such as age and gender will most likely be beneficial, however we believe the two most important factors are going to be related to wealth.

Analysis

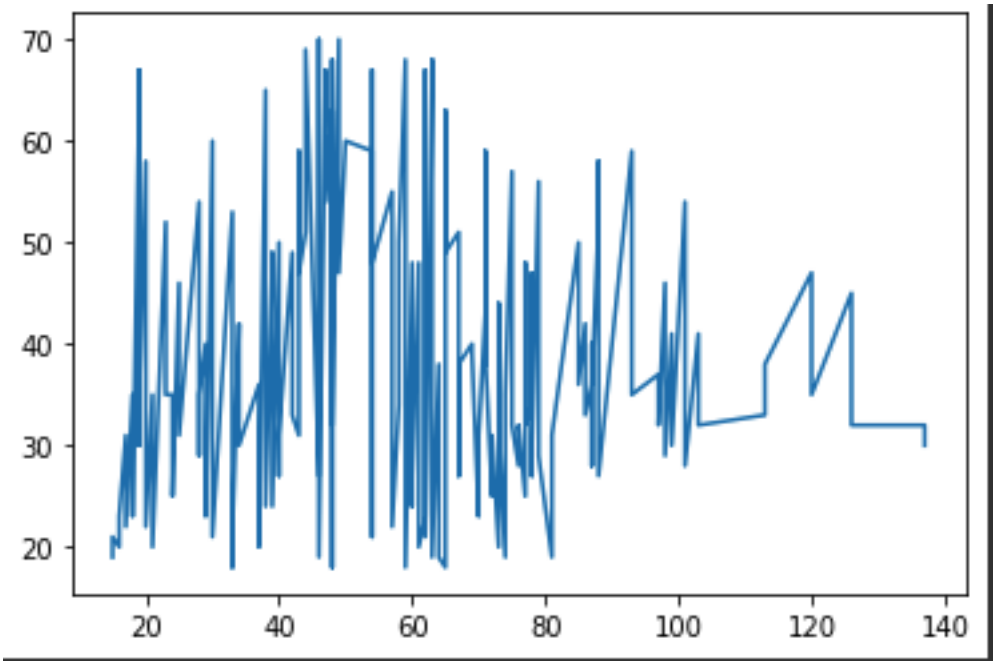
	CustomerID	Genre	Age	Annual Income (k$)	Spending Score (1-100)	Annual Spending	bins
0	1	Male	19	0.000000	0.387755	34.5	medium
1	2	Male	21	0.000000	0.816327	55.5	high
2	3	Female	20	0.008197	0.051020	19.0	low
3	4	Female	23	0.008197	0.775510	54.5	high
4	5	Female	31	0.016393	0.397959	37.0	medium

In the above example is a snapshot of the various variables we will be using to group the different customers. We will be using 4 different clusters, as well as 3 different bins from low-high ratings. We are heavily weighting annual income, along with spending scores, based on the premise that the more money than the more opportunities for the customers to spend it.

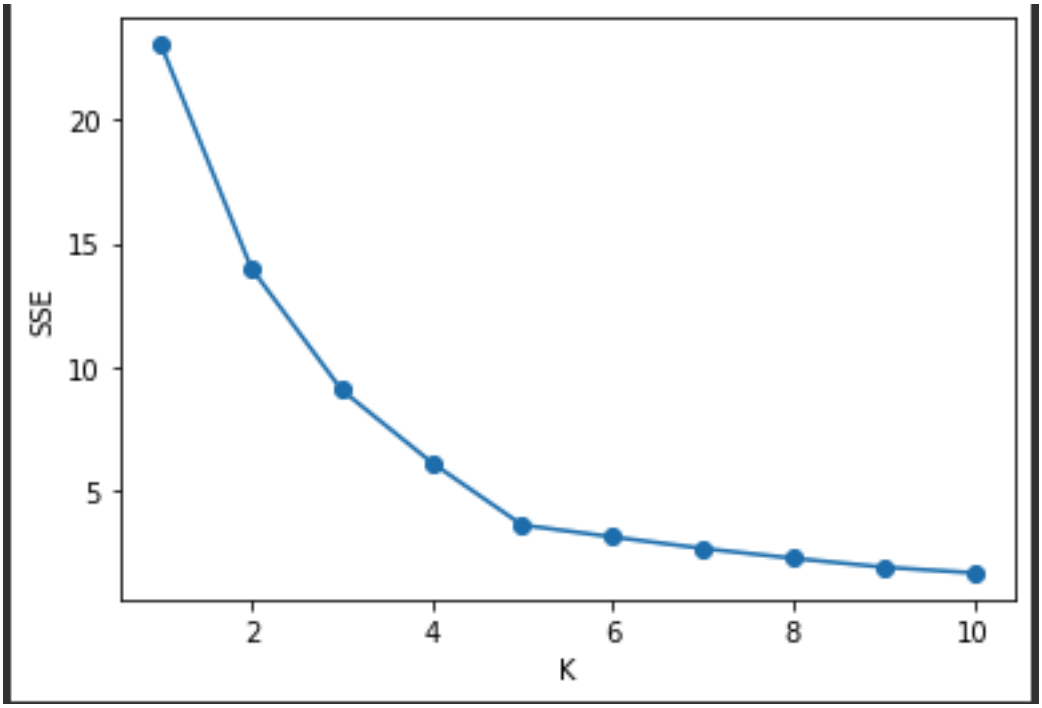


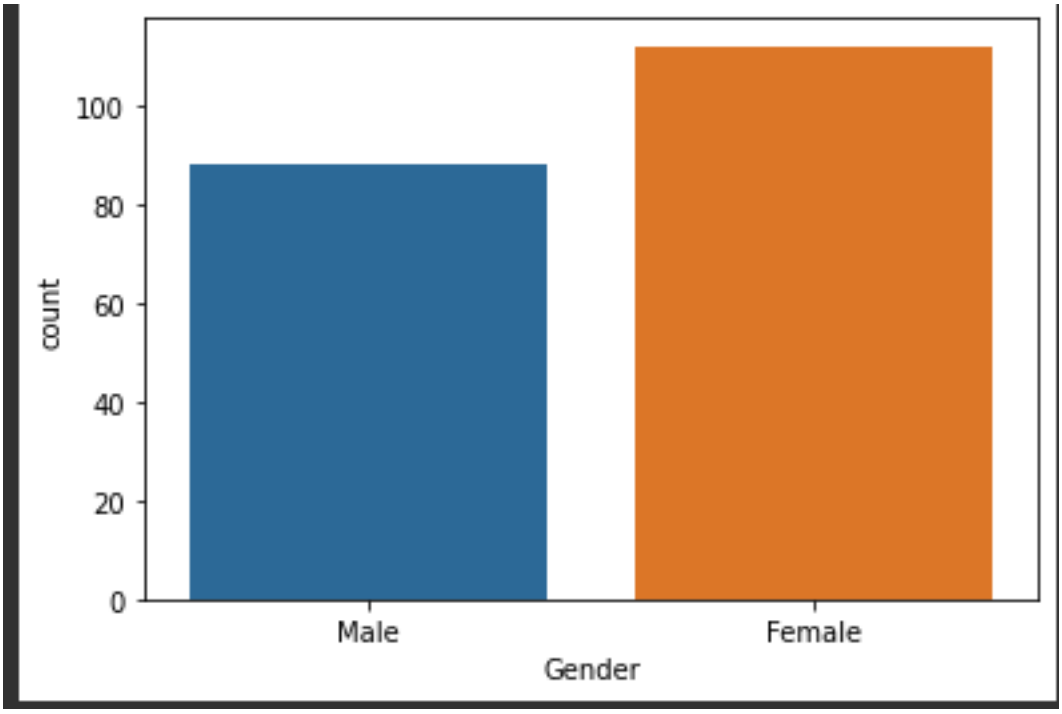
In the visualization above we are plotting the spending score, along with the annual income. The top right will be the wealthier, compared to the bottom left will be lower income customers. We have developed a metric called annual spending which develops the annual income in relation to the spending score. In short that metric is going to look to develop a number off the plotted results above.

Results



The above visualization is going to be age on the y axis, and annual income on the x axis. As you can see, there is quite the variety in relation between the two. The one thing that stands out is that those customers are above the 90k a year range are at least 30 years of age. We decided to focus on that group and no one under 30 in our model.





To finish, the groups of customers we will want to focus on will be between that 4-10 k group. With our results we were able to realize customers above the age of 30 were who we

wanted to first focus on based on the data. Along with this we couldn't find an edge regarding to genre. Income and age were the two variables we were able to focus on. Another factor was customers fitting our annual spending metric. Above 60 is what we were able to distinguish a plus expected value regarding sending offers, as well as over a .70 spending score. This aligns with our hypothesis of the more money than the more opportunities.

We were able to determing the 3 most popular clusters were the 0, 2 ,and 4 clusters. The 2nd cluster is going to be who we are going to focus on as this will be high annual income, high spending score, and is going to be that 30-45 year age range which is exactly what we want to focus on.

Recommendations

A couple things we would change if we were to do this again would be to have a specific item we were selling, for example Life insurance. This would be a lot more efficient when targeting certain customer groups as it is very specific and can broaden our range. With a specific item it would have made our jobs a lot easier to be able to exactly establish who we wanted to target. Another would be additional data that would group in life insurance. That way we can do extra research and analysis to find out what groups buy life insurance, for example older individuals. To finish, overall, we were able to distinguish a group based on the results (30 and above, along with spending score above .70). This will allow our team to spend our marketing dollars wisely, and not to just random customers hoping to strike gold.

References

(LEDU), E. E. (2018, September 12). *Understanding K-means clustering in machine learning*. Medium. Retrieved February 26, 2023, from https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1

6 dead-simple tips to succeed in marketing men VS women. GrowthDot. (2023, February 24). Retrieved February 26, 2023, from https://growthdot.com/6-dead-simple-tips-succeed-marketing-men-vs-women/

Average income by state 2023. (n.d.). Retrieved February 26, 2023, from https://worldpopulationreview.com/state-rankings/average-income-by-state

City and region marketing lists. Business Data List. (2022, May 4). Retrieved February 26, 2023, from https://businessdatalist.com/city-and-region-marketing-lists/

Gong, D. (2023, February 17). *Clustering algorithm for customer segmentation*. Medium. Retrieved February 26, 2023, from https://towardsdatascience.com/clustering-algorithm-for-customer-segmentation-e2d79e28cbc3

Lorincz, N. (2023, February 1). *Gender differences in advertising between men and women: Do gendered advertisements help or hurt?* OptiMonk. Retrieved February 26, 2023, from https://www.optimonk.com/gender-targeting-the-differences-between-men-and-women/

Male & female email & mailing lists. LeadsPlease. (n.d.). Retrieved February 26, 2023, from https://www.leadsplease.com/email-lists/consumer/male-and-female-mailing-lists

Small business guide to direct mail marketing. Business News Daily. (n.d.). Retrieved February 26, 2023, from https://www.businessnewsdaily.com/15776-direct-mail-marketing-guide-for-small-businesses.html

Vaughan, P. (2021, December 7). *How to create marketing offers that don't fall flat*. HubSpot Blog. Retrieved February 26, 2023, from https://blog.hubspot.com/blog/tabid/6307/bid/31252/how-to-create-marketing-offers-that-don-t-fall-flat.aspx

Wikimedia Foundation. (2022, September 28). *List of U.S. states and territories by median wage and mean wage*. Wikipedia. Retrieved February 26, 2023, from https://en.wikipedia.org/wiki/List_of_U.S._states_and_territories_by_median_wage_and_mean_wage

Appendix

The goal of this project was to cluster groups of customers. From these clusters we could determine based on the data which would be best to send offers out to for promotions. Based on the data we were able to distinguish a group we wanted to focus on, which had a direct focus on an age group as well as their income.