In this final individual milestone, you will submit a starting draft of your final project individual report. See the final project instructions for details and look closely at the rubric. You should describe what you have been working on here at the beginning and where you are thinking to take your analyses in the final weeks of the course. Note that this is not a zero-sum game. You may report on material that you have collaborated on with your partners, and it does not have to be solely your work, but you should be able to describe in detail how you contributed.

(Clustering Analysis Source Code)
My part is focusing on cleaning data and clustering analysis.

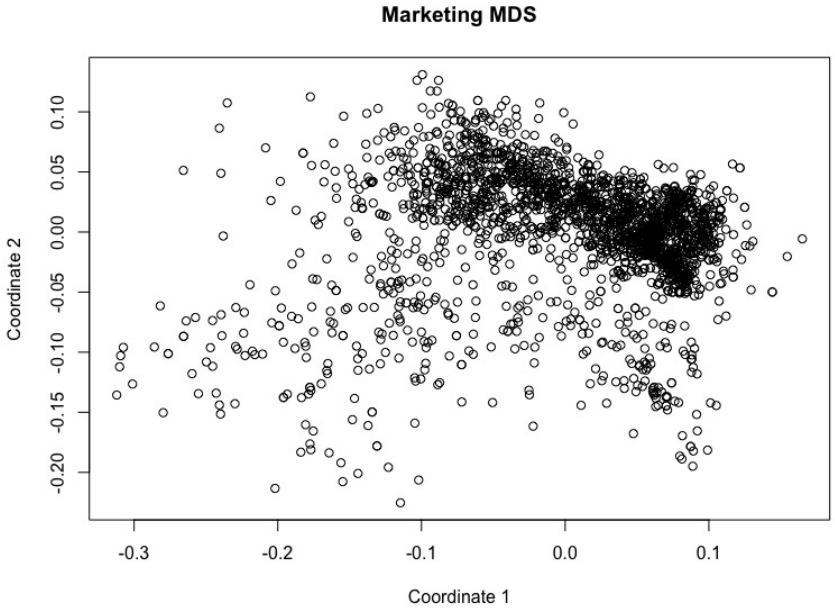

Figure 1: General MDS Plot

Initial we want to find the pattern through general MDS plot, but there is not clue to demonstrate a clear cluster in my graph. Then I try to displays a measure of how close each point in one cluster is to points in the neighboring clusters and thus provides a way to assess parameters like number of clusters visually through silhouette plot, but the silhouette width with less than 0.25 indicate the very little substantial structure has been found. If we try to visualize the clusters from 2 to 8, clearly our clustering plot shows a lot of random pattern and most of them are cross with each other.
Finally, I use elbow method and cluster dendrogram to find the number of optimal clusters are 2 and 3, and they both show very clear clusters in our graphs. In order to find the which variables can be potentially described by 2 or 3 classes, I use histogram to go through all variables, Income in this case is equally separated into two parts, so I use mean value as critical point and label then as "Lower" and "Higher" according to their value compare to its mean. Then I re-plot it again, it shows me the clustering plot has very high correlation with Income, "Higher" are most on the left, "Lower" on right (Figure 9).

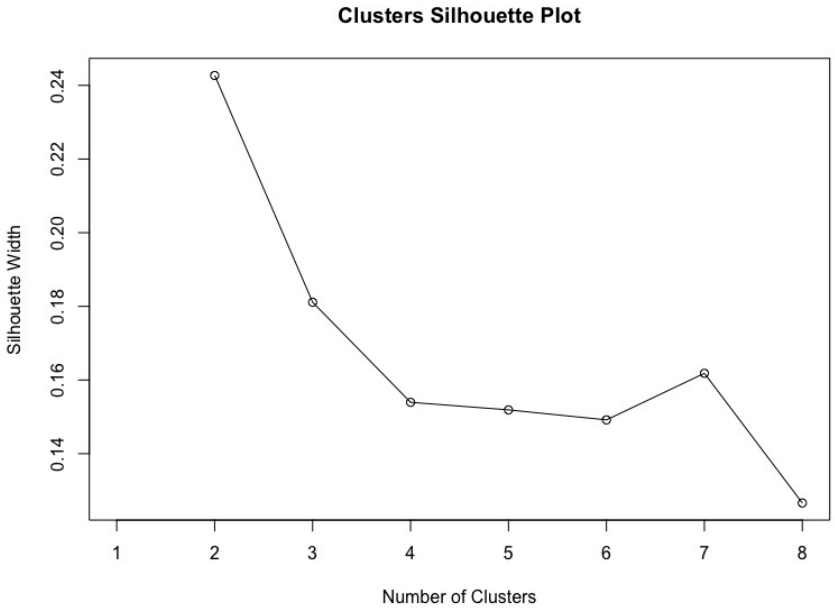

Figure 2: Silhouette Plot

Figure 3: Clustering Plot

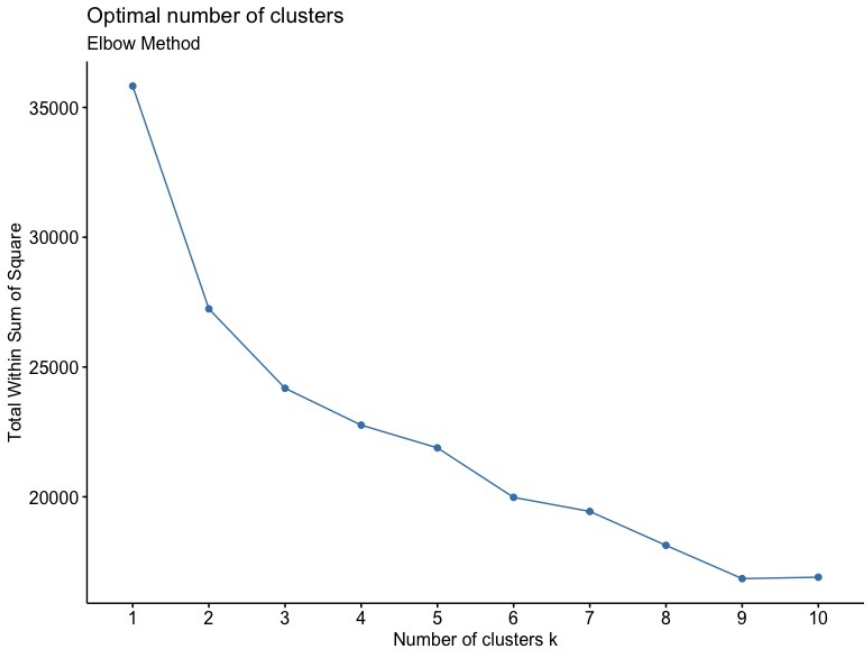

Figure 4: Elbow Method Plot

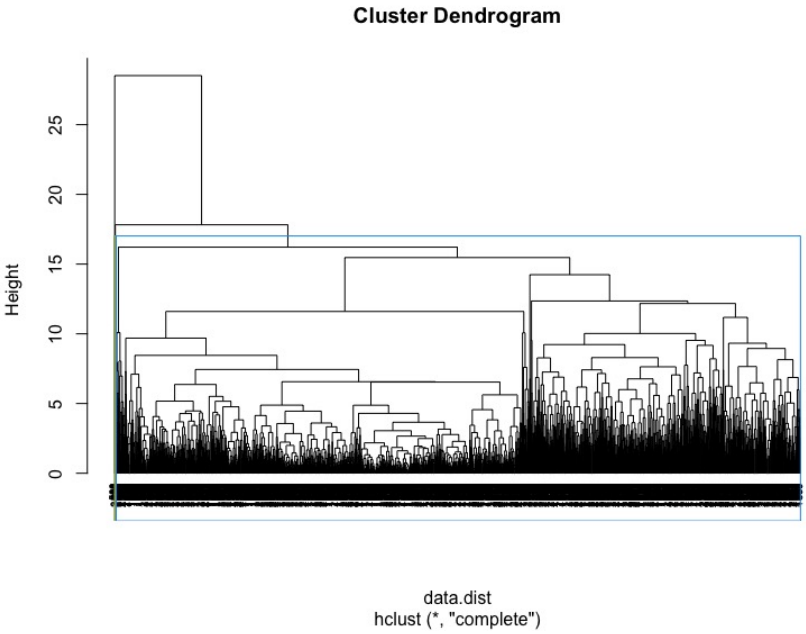

Figure 5: Cluster Dendrogramt

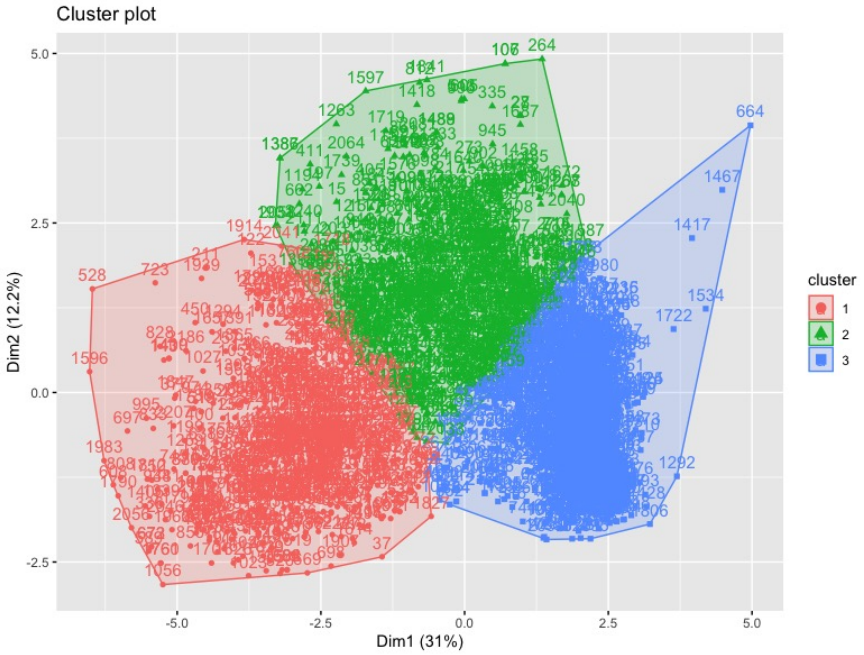

Figure 6: Cluster Plot with 3 clusters

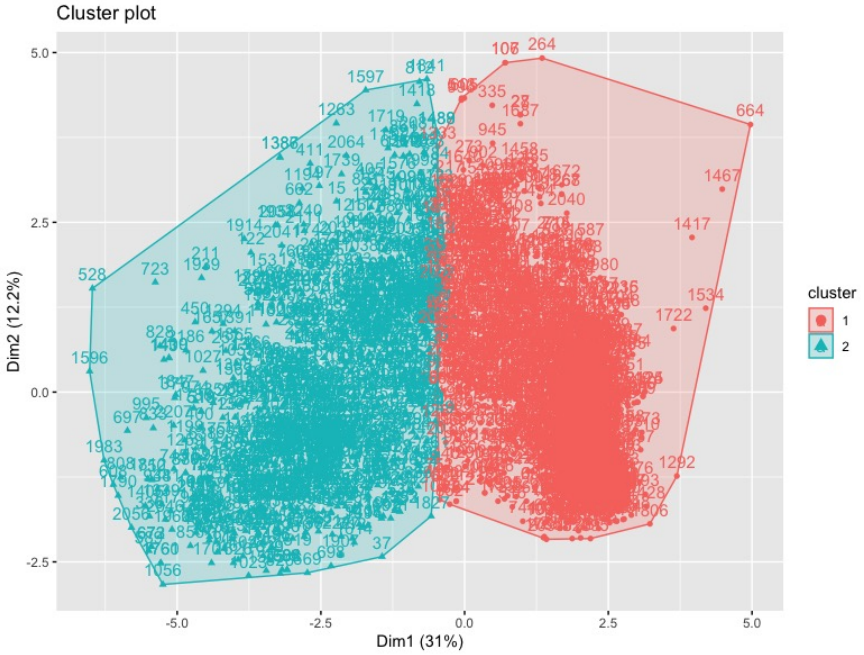

Figure 7: Cluster Plot with 2 clusterst

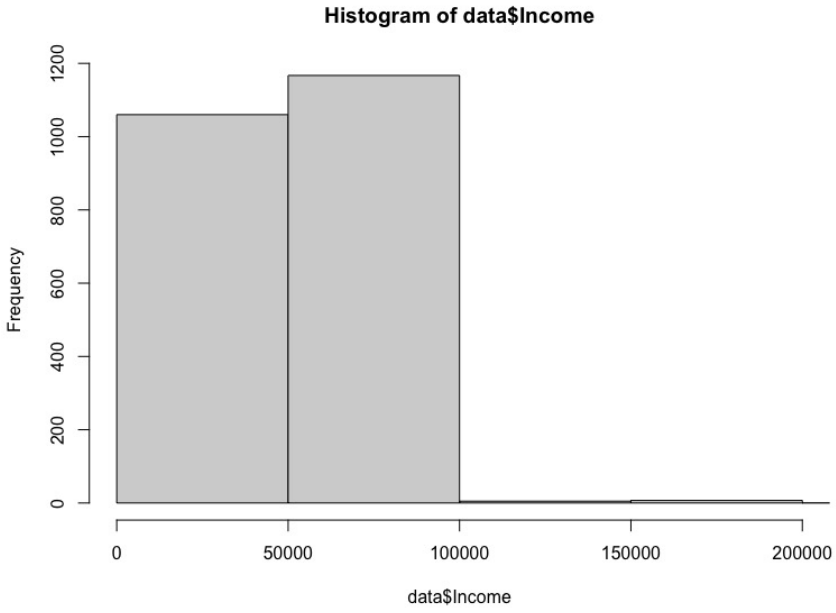

Figure 8: Income Histogram

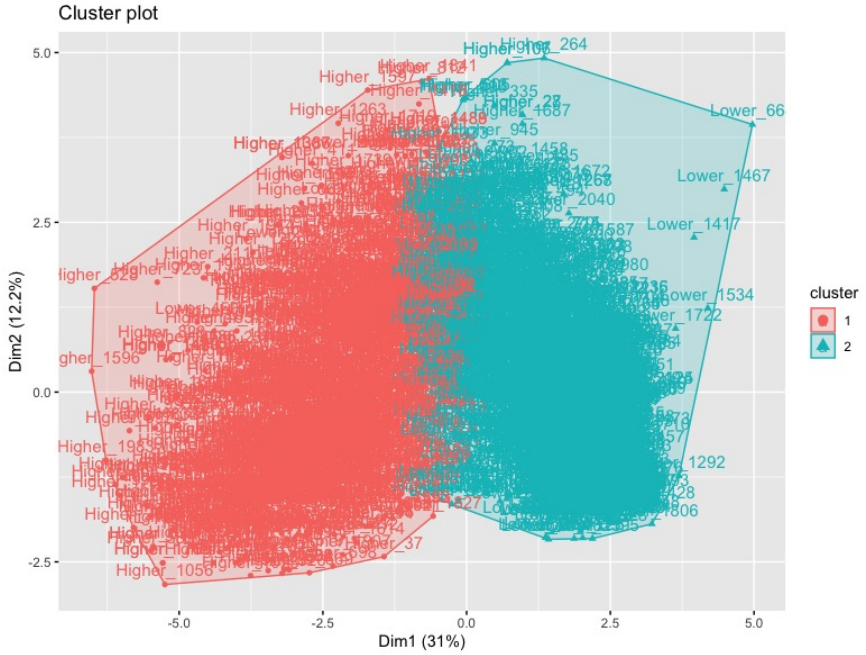

Figure 9: Final Clustering