Shay Snyder

CSCI 4957 – Data Analytics

Dr. Husari

**Exercise 12 - Cluster Analysis: The Wolf of Wall Street**

As a new data analyst at the world-class merchant and investment banking firm, *WallStock*, my task is to split the ‘PotentialClients.csv’ dataset into the various groups (clusters) that contain similar individuals. The programming language R and the K-means algorithm will be used to accomplish this task. Before I began manipulating the data, I plotted the data to get a better view of the dataset.

A screenshot of a cell phone

Description automatically generated

From this plot, we can see 15 distinct clusters of clients. One could simply move forward with the K-means algorithm from this point, but I decided to calculate the sum of squared distances (WSS) for various numbers of clusters. To build the elbow plot, I used a tutorial on [r-bloggers.com](https://www.r-bloggers.com/finding-optimal-number-of-clusters/). (I am unsure if this step is necessary but I like to give credit to the resources I used to accomplish my programming tasks). My elbow plot was calculated using 50 different starting points with the number of clusters ranging from 1 to 25. I realize that 25 clusters would be excessive, but I wanted to see how the graph changes from the instructional video on D2L.

A screenshot of a cell phone

Description automatically generated

From the elbow plot and my visual assessment, I am going to use 15 clusters for the K-means algorithm. Past 15 clusters, the slope of the elbow plot is basically zero. This means that very little information is lost at 15 clusters. Now, I ran the K-means algorithm with 15 clusters.

A close up of a logo

Description automatically generated

This plot shows the resulting plot from the K-means algorithm using 15 clusters. The plot doesn’t show 15 distinct colors and I think this could by simply fixed by using a different method to plot the clusters. Since our clients have been grouped into 15 distinct groups, my task would be to study each of the 15 groups to gain insight about our customers. With this information and “immense skill” on my part, I have become The Wolf of Wall Street.