

Run the Python code provided and write a brief paragraph regarding the results.

SUMMARY

This assignment was a clustering analysis based upon a Stanford University study regarding 32 personality attributes. There were 240 total responses and this tutorial took the responses and used the K-Means clustering algorithm to help group them. This helps to get equal weightings when comparing the variation in similar responses. It also helps to explain the total overall variation. In order to find the right number of clusters and not base it just on the big five personality grouping, the silhouette coefficients were run. The resulting analysis showed the highest coefficient value occurred at the number of clusters being two. These two clusters were then created into two separate segments and plotted on separate histograms. The data showed that each one had distinct variation and values. These were both based upon the centers of the clusters where each data point would have as accurate as possible equal distance to it.