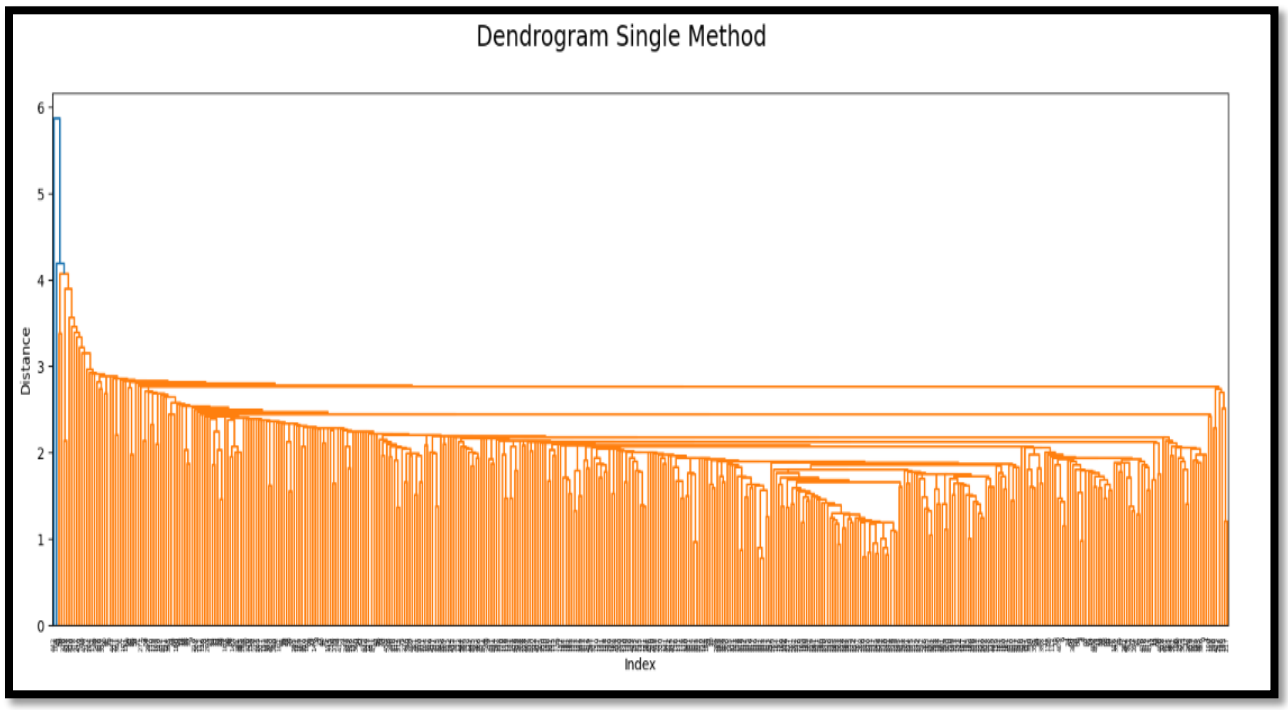
2.1: Unsupervised Learning Algorithms

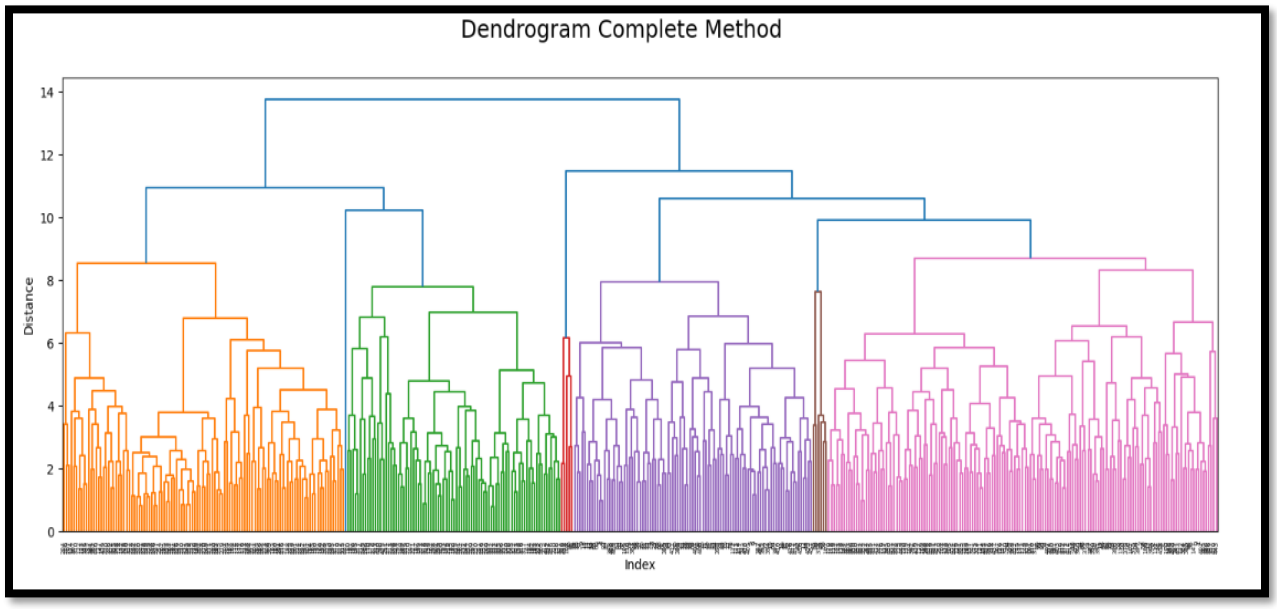
Innocent Bayai

5 December 2024

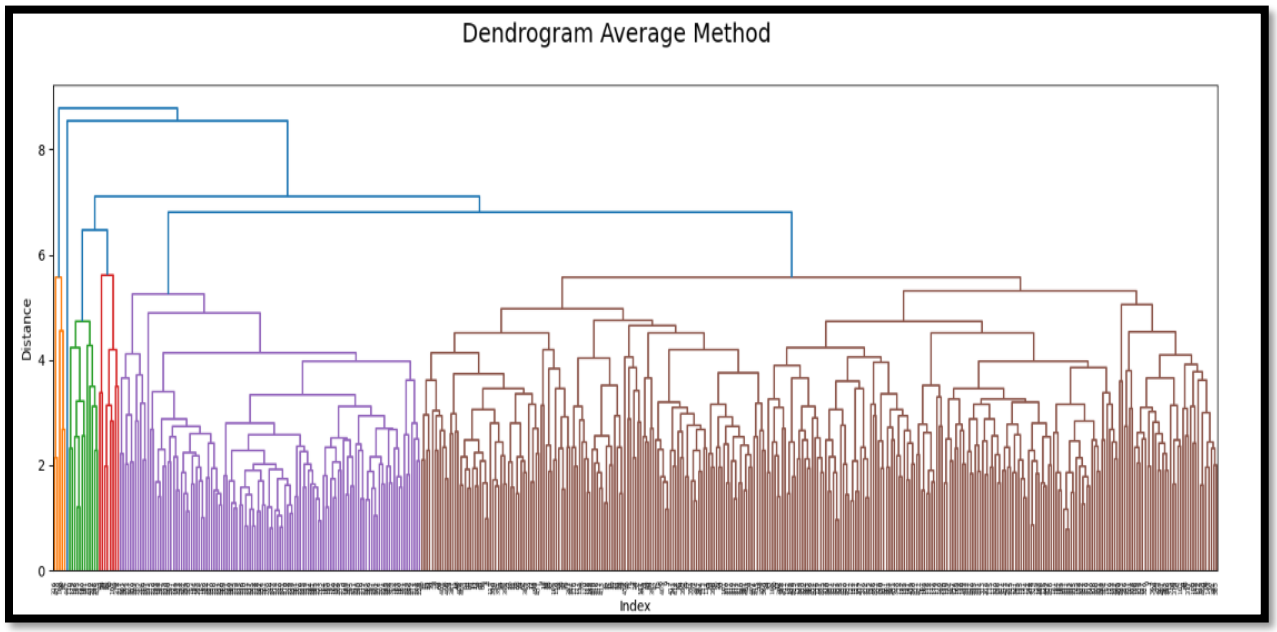
Single Method Dendrogram



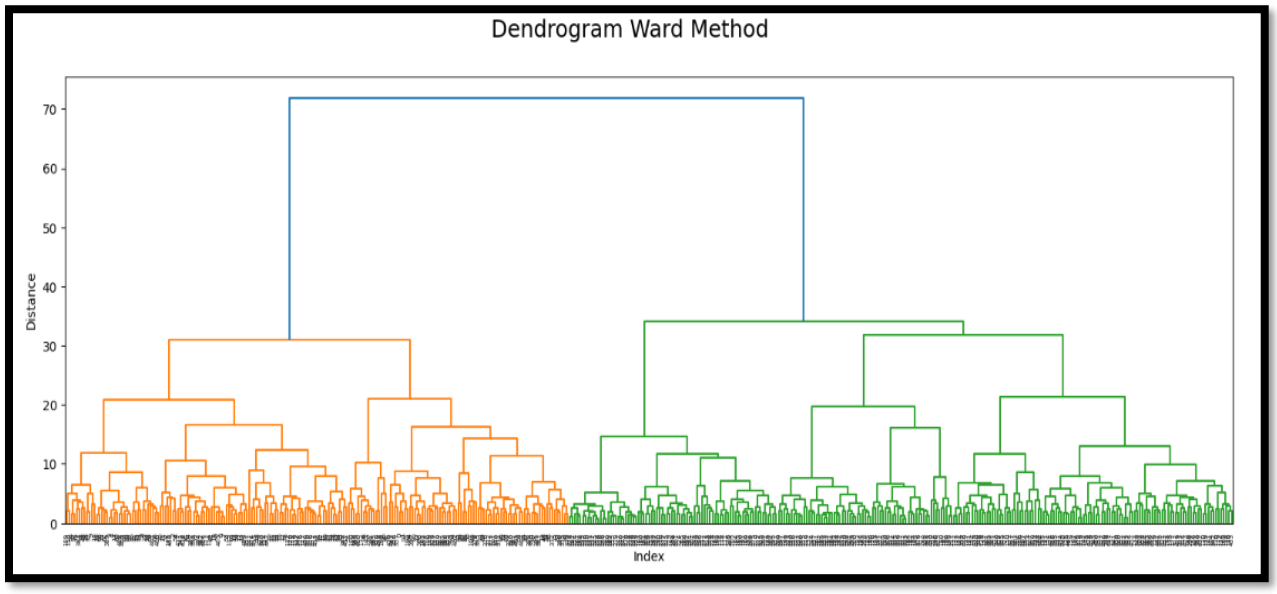
Comple Method Dendrogram



Average Method Dendrogram



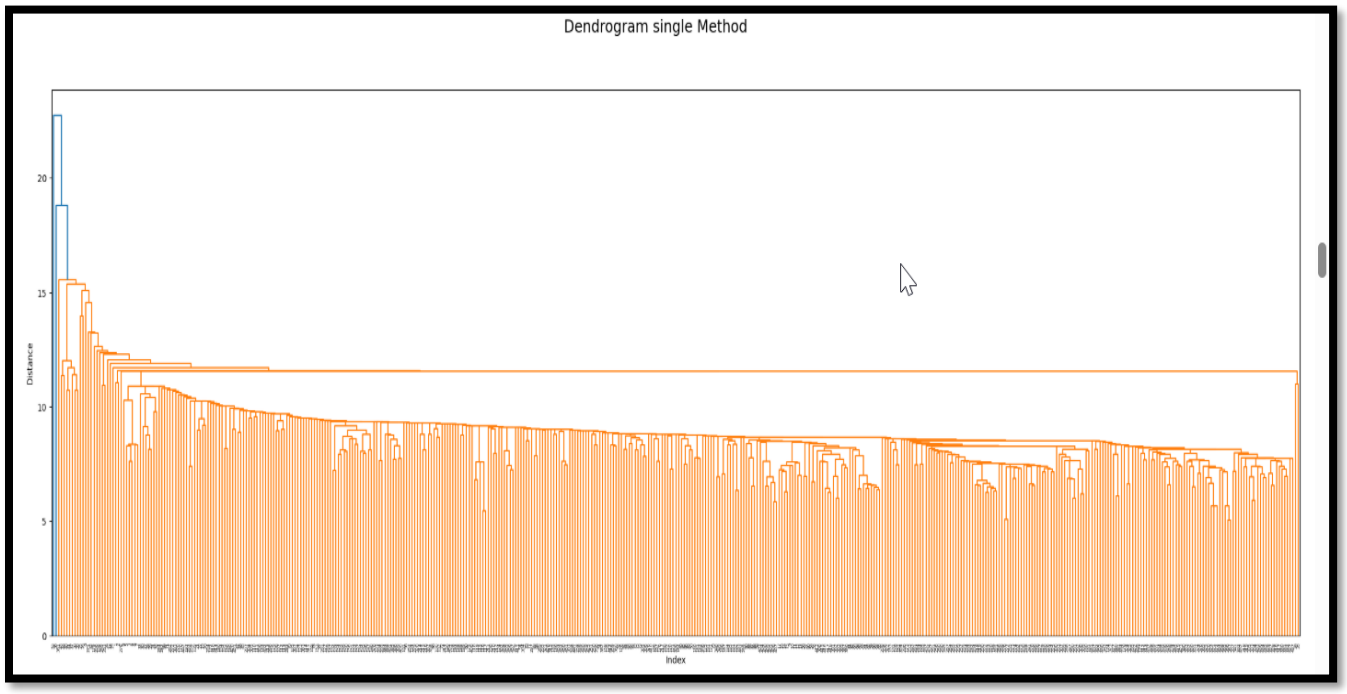
Ward Method Dendrogram



Comments:

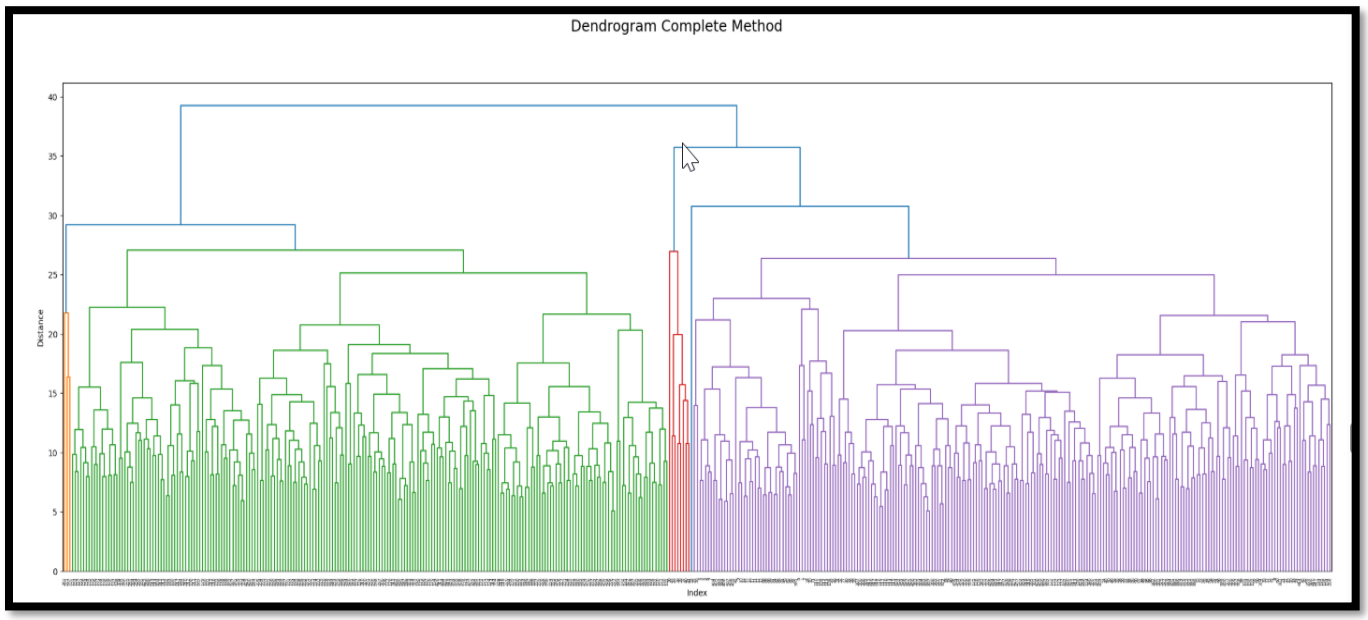
I note that the complete method produces a dendrogram with 4 distinct and well-distributed clusters unlike the other methods whose dendrograms seem not to show any meaningful distributed clusters.

Comparing all stations: Single method dendrogram



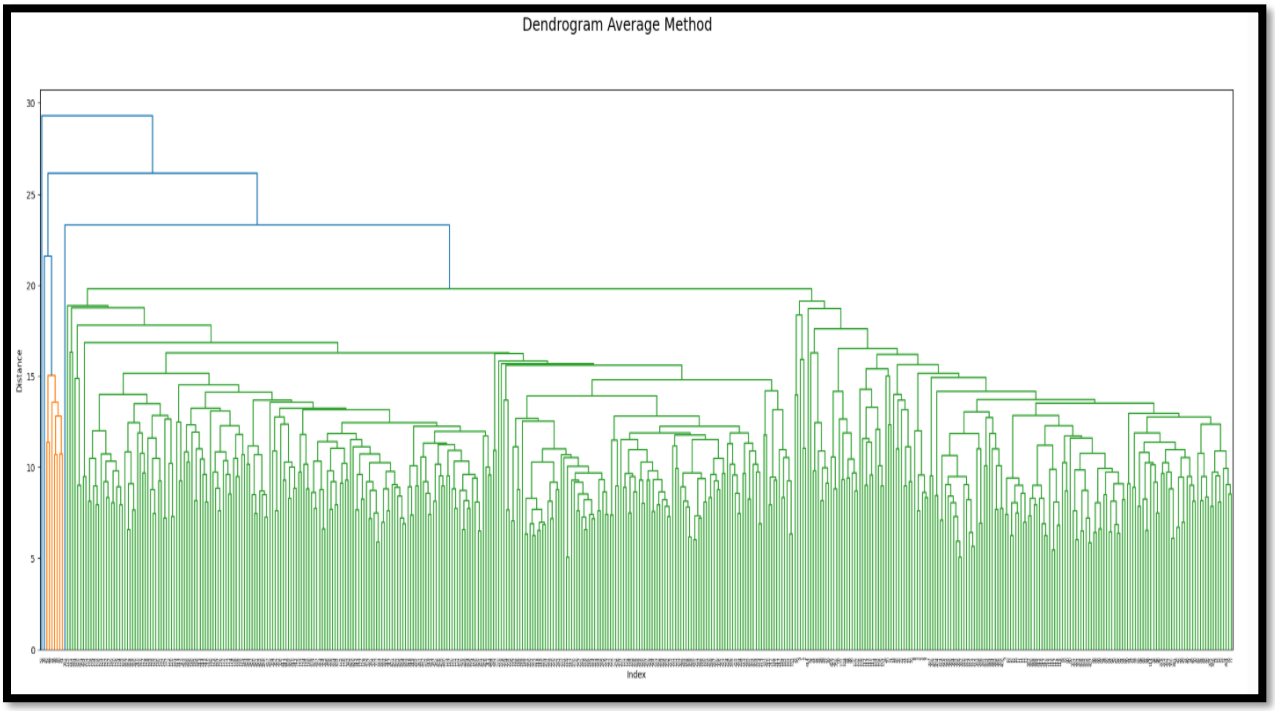
There are 2 clusters although the blue one is almost un-noticeable owing to the dominance of the orange cluster. The Orange one seems not to have any clear sub-clusters discounting its usefulness.

Complete Method Dendrogram for all Stations



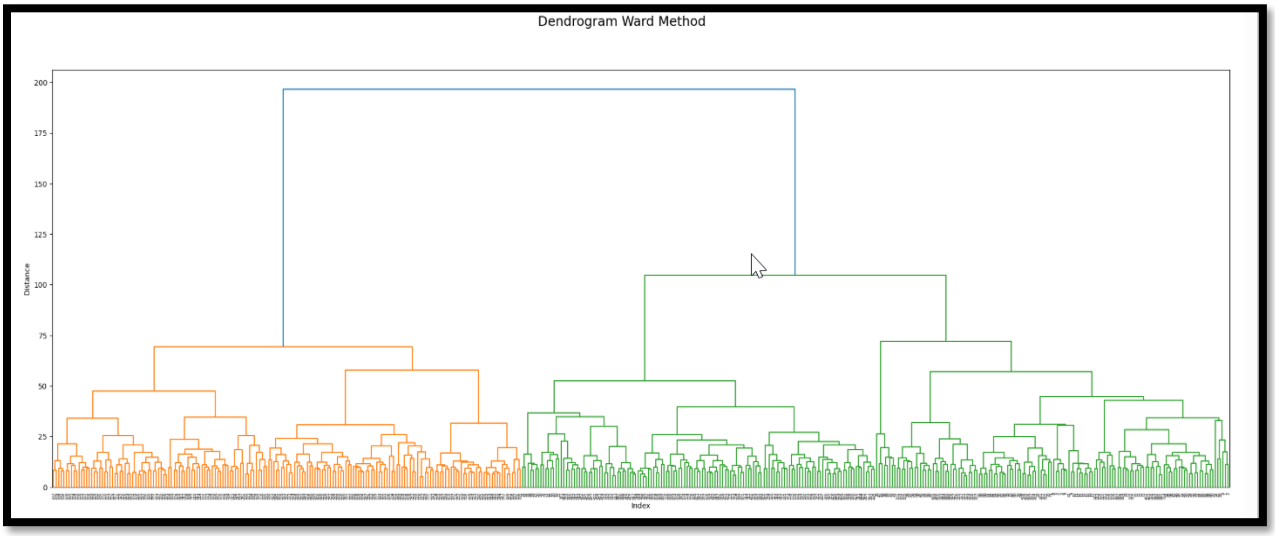
There are four clusters (orange, green, red, and purple). However, the only dominant ones are green and purple whilst orange and red are minor.

Average Method for all Stations



There are 2 clusters (orange and green) but the green one is the dominant one although its sub-cluster not easy to observe.

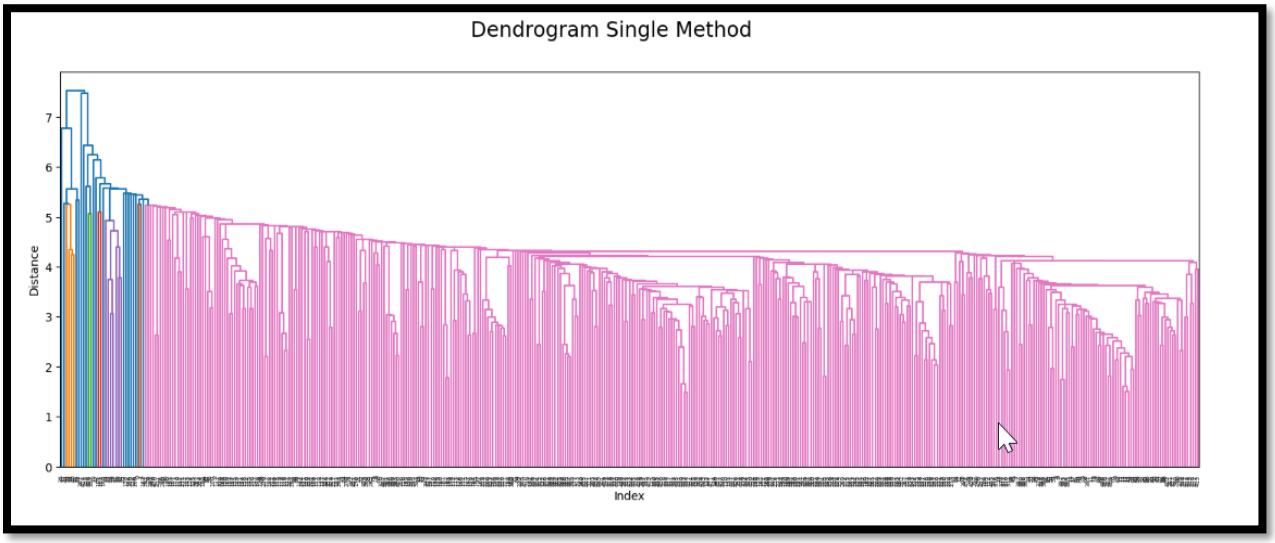
Ward method for all Stations



The dendrogram has only 2 major clusters and several sub clusters whose meaning is not easy to deduce.

Dendrograms from Reduced Dataset

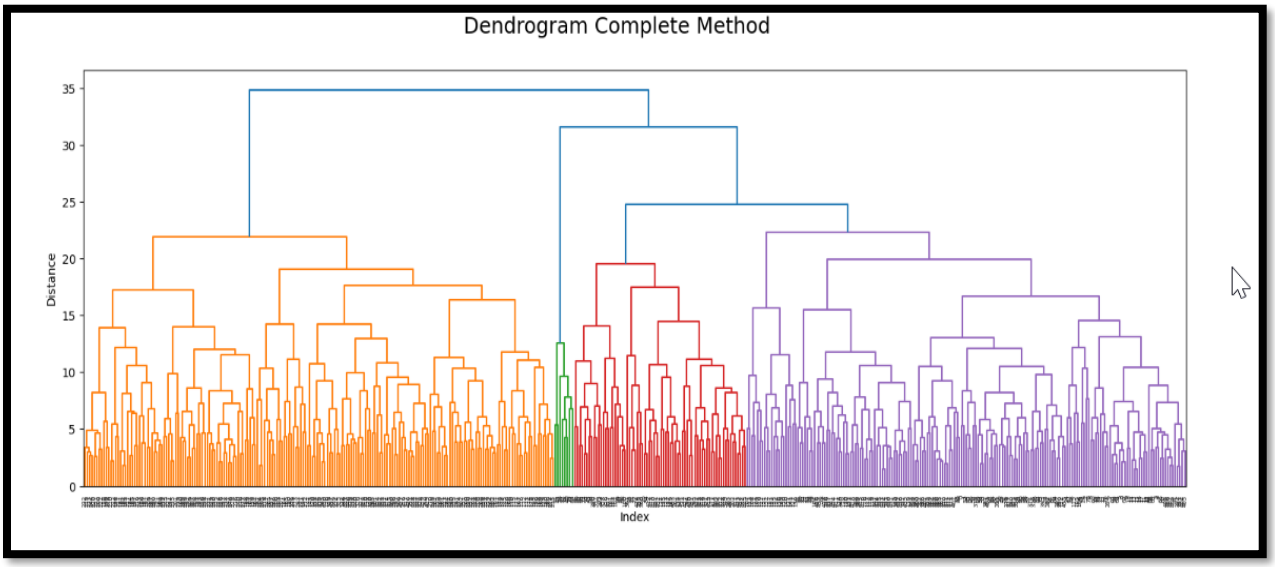
Single Method



Comment

Whilst there are many clusters, only one is dominant.

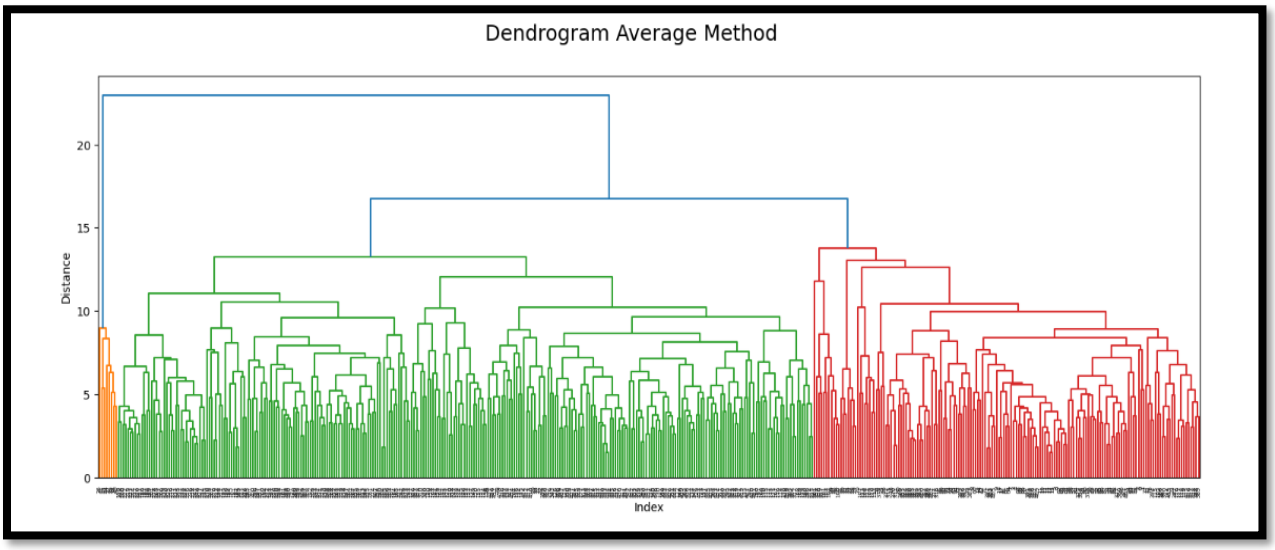
Complete method



Comment:

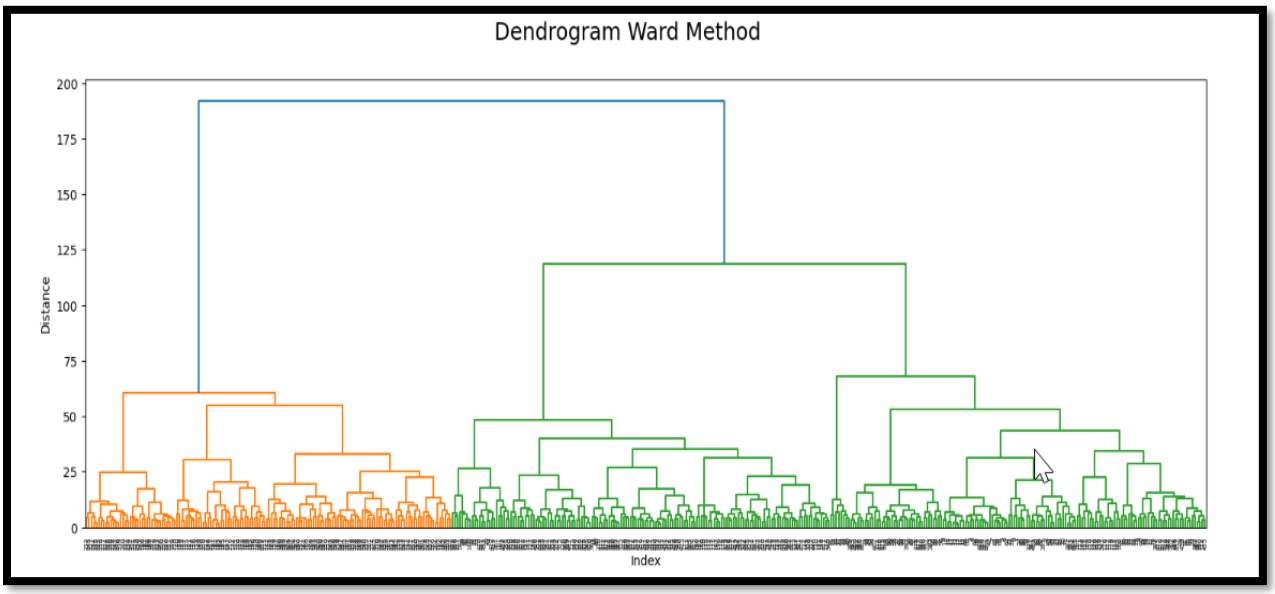
Four clusters are observable although they are not evenly distributed, and the green and red clusters are a minority compared to the orange and purple clusters.

Average Method



Comment:
The method gives 3 clusters that are not evenly distributed, and sub-clusters are innumerable.

Ward Method



Comment:

This method produces 3 distinct clusters. This is seemingly a better distribution using the PCA data.