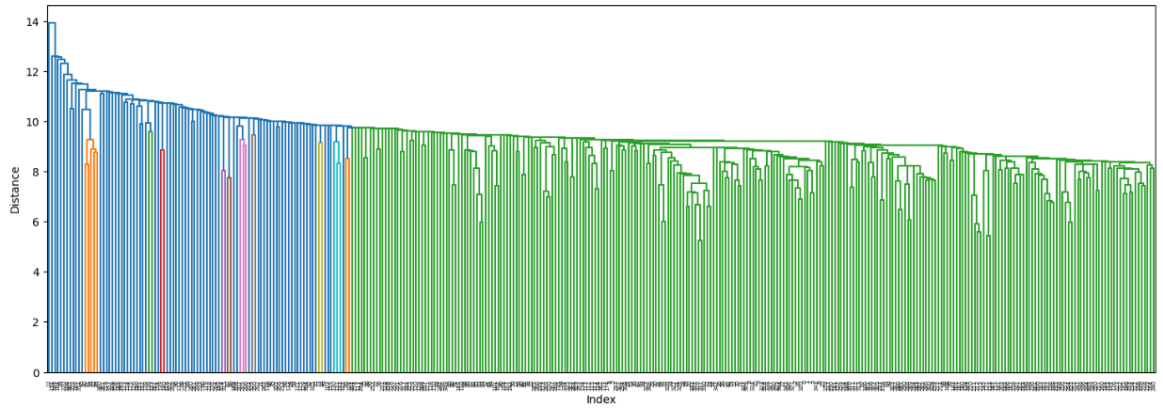
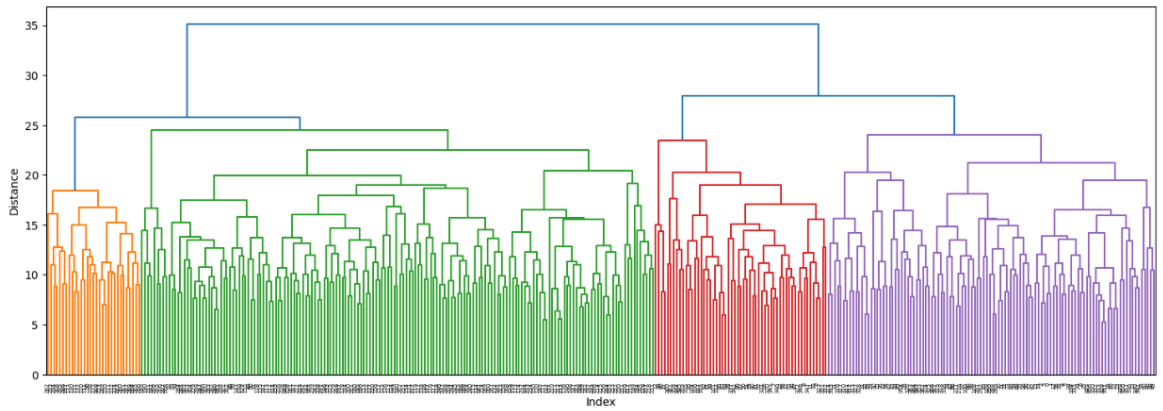
Exercise 2.1: *Unsupervised Learning Algorithms*

RUNNING HIERARCHICAL METHODS ON SCALED DATA

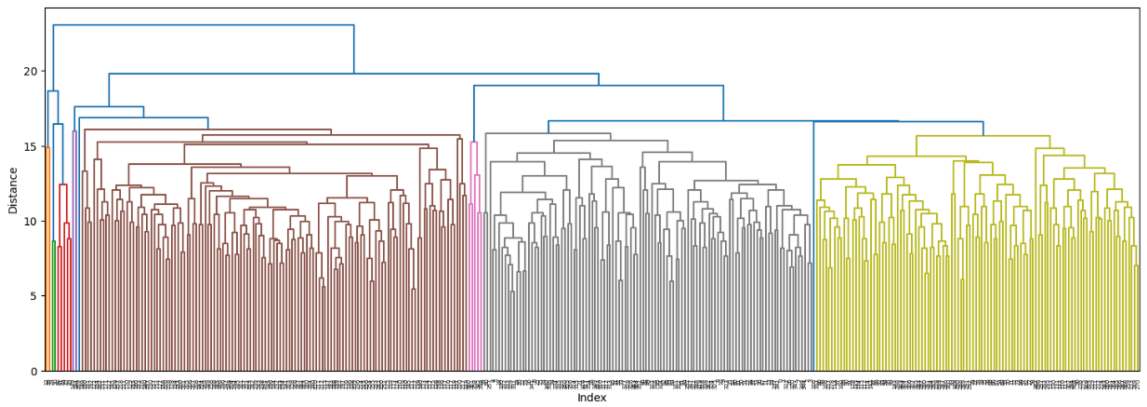
Dendrogram Single Method

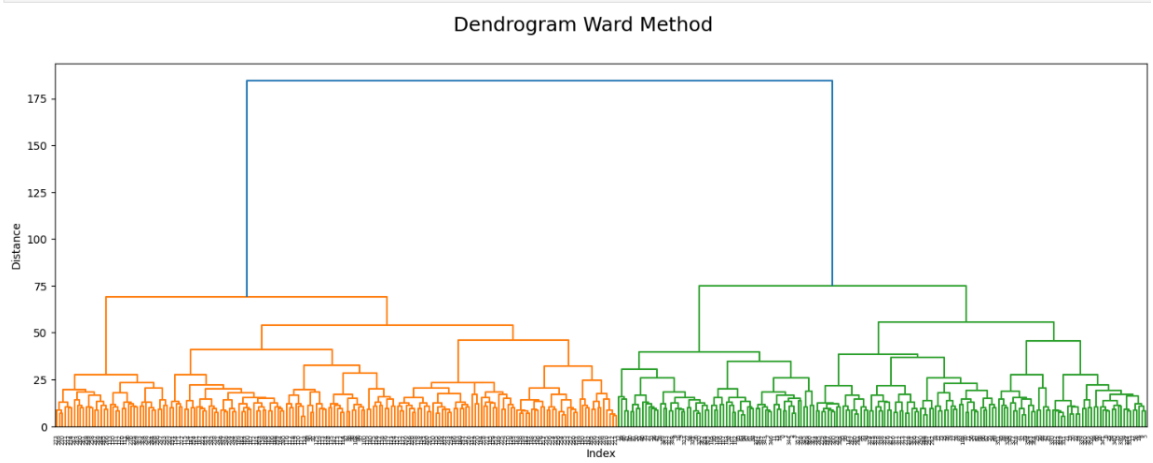


Dendrogram Complete Method



Dendrogram Average Method

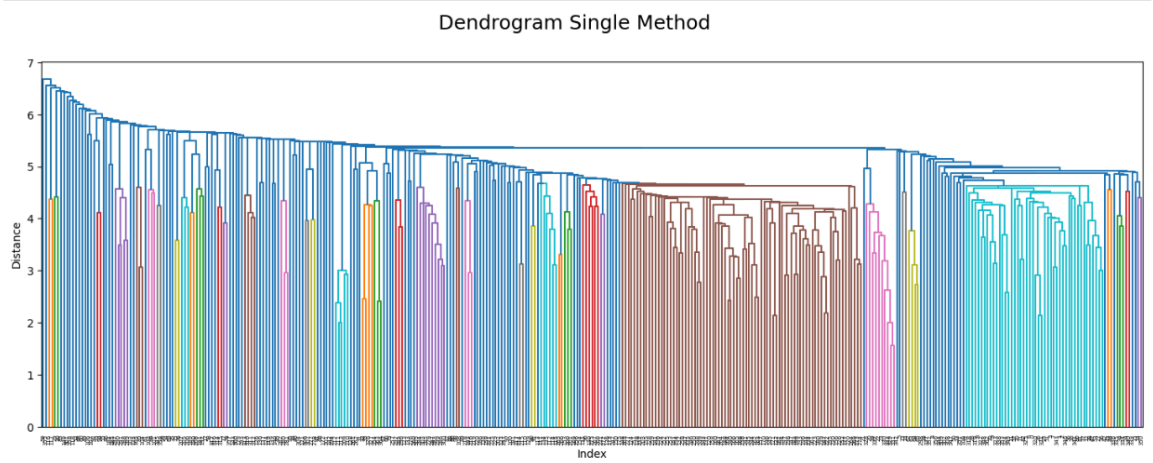


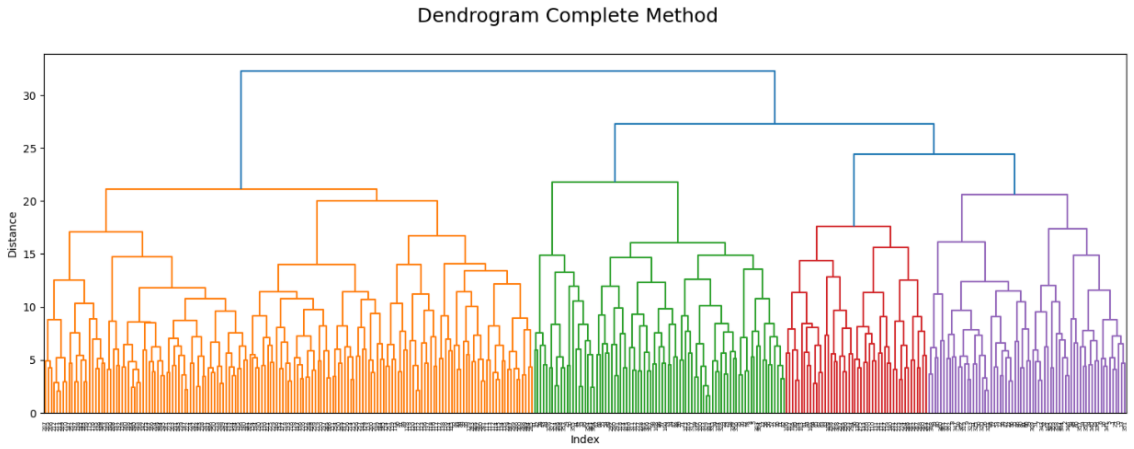


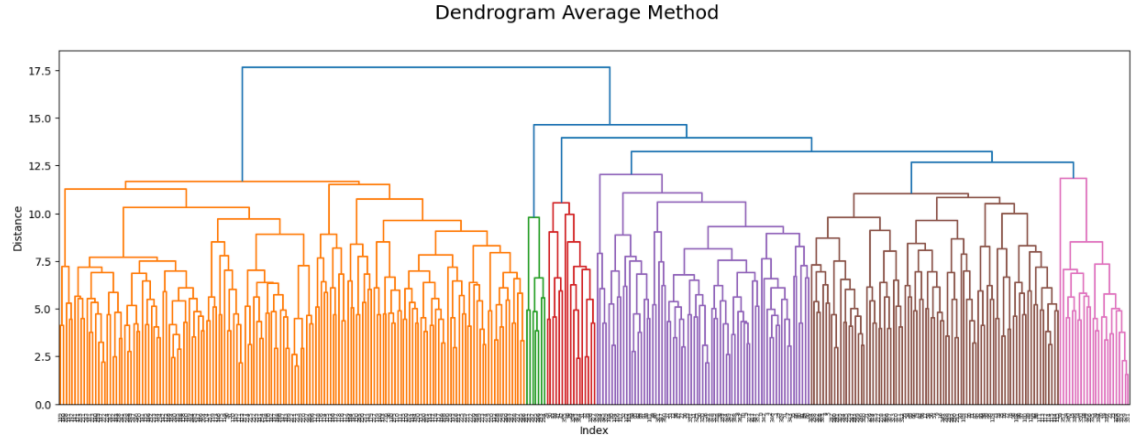
Each of this methods classify data in different ways:

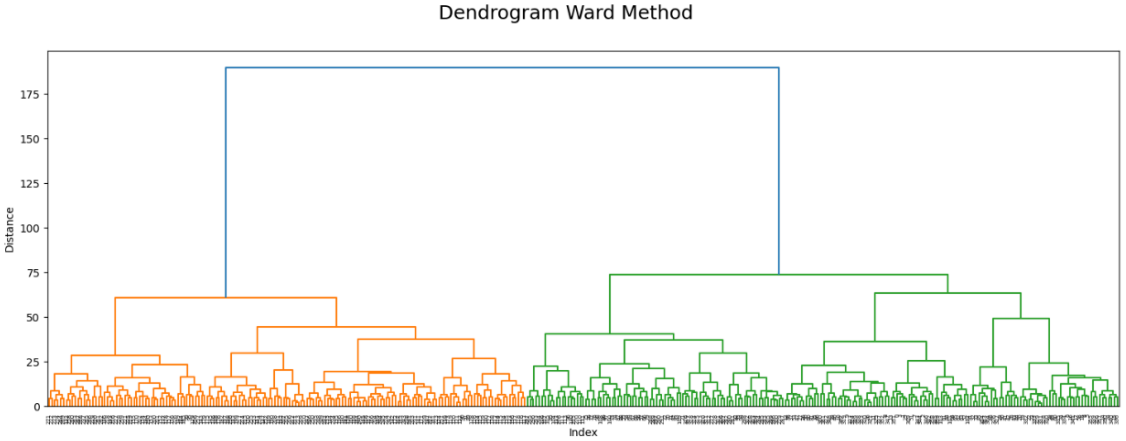
- **Single method**: Shows multiple clusters but most records belonging to the same one, doesn't seem reliable.
- **Complete method**: Shows four different clusters, this might correspond to geographical areas or the 4 seasons.
- **Average method**: Shows 3 main clusters and other minimal ones.
- **Ward method:** Shows two categories very close in terms of size.

RUNNING HIERARCHICAL METHODS ON REDUCED DATA









Again, each method classifies data in different ways:

- **Single method:** Shows multiple a big number of clusters, it doesn't offer any apparent insight.
- **Complete method:** Again, shows four different clusters, this might correspond to geographical areas or the 4 seasons.
- **Average method:** Again, shows 3 main clusters and other minimal ones.
- **Ward method**: Shows two categories very close in terms of size.

I would suggest going for the Complete Method as it looks like the only one that seems to offer insightful information.