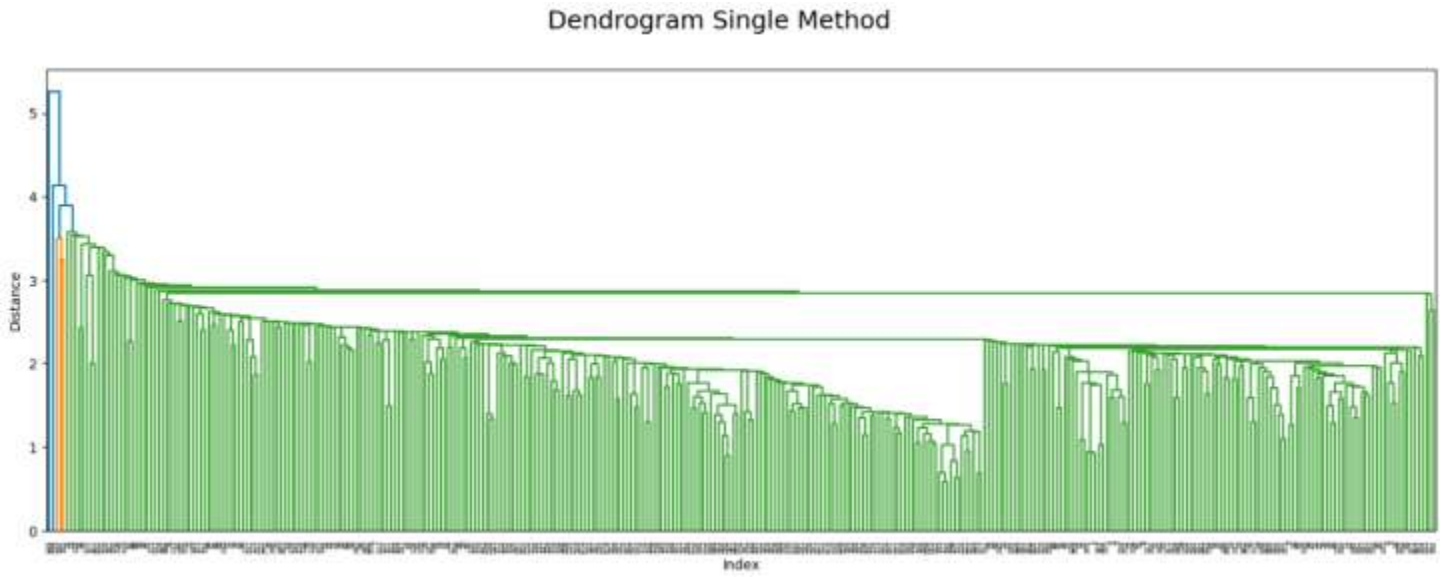
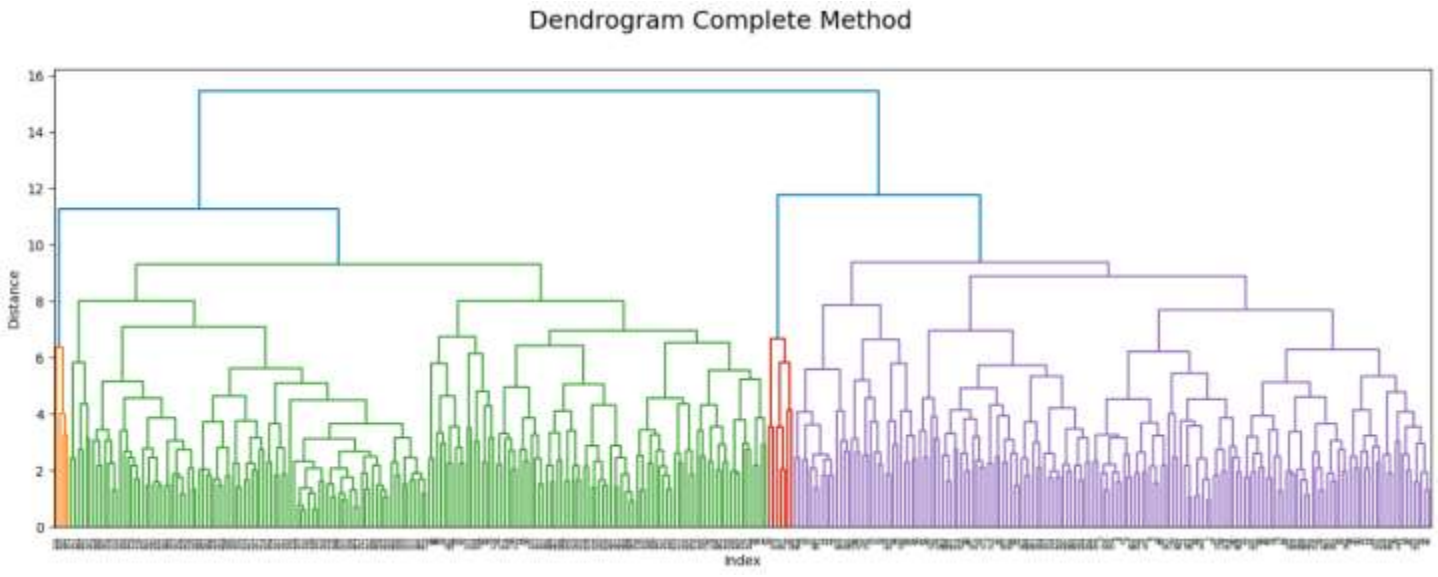
UNSUPERVISED ML ALGORITHMS

CASE STUDY: CLIMATE DATASET
EXERCISE 2-1

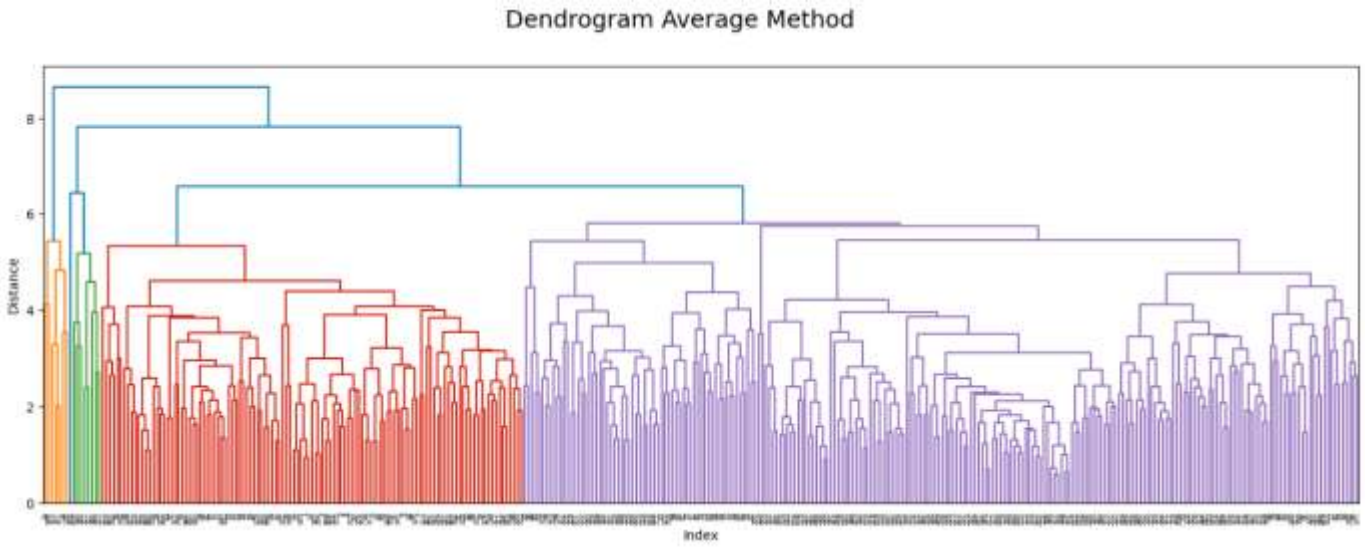
1. Single Method (Data from 1975 for Madrid Vs Belgrade):



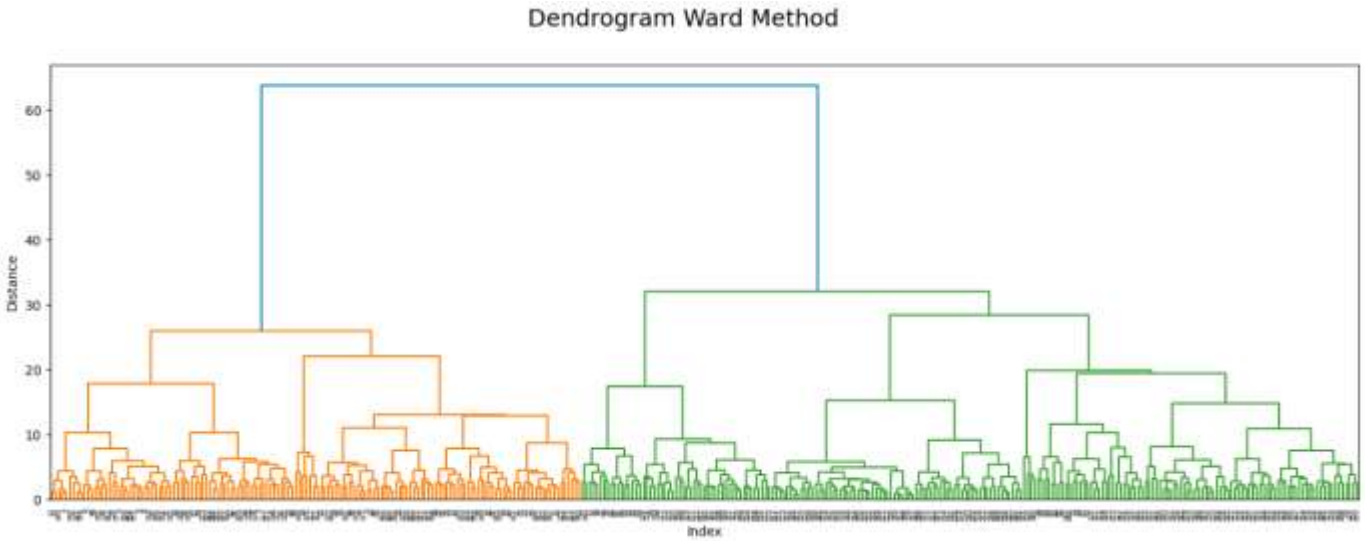
2. Complete Method:



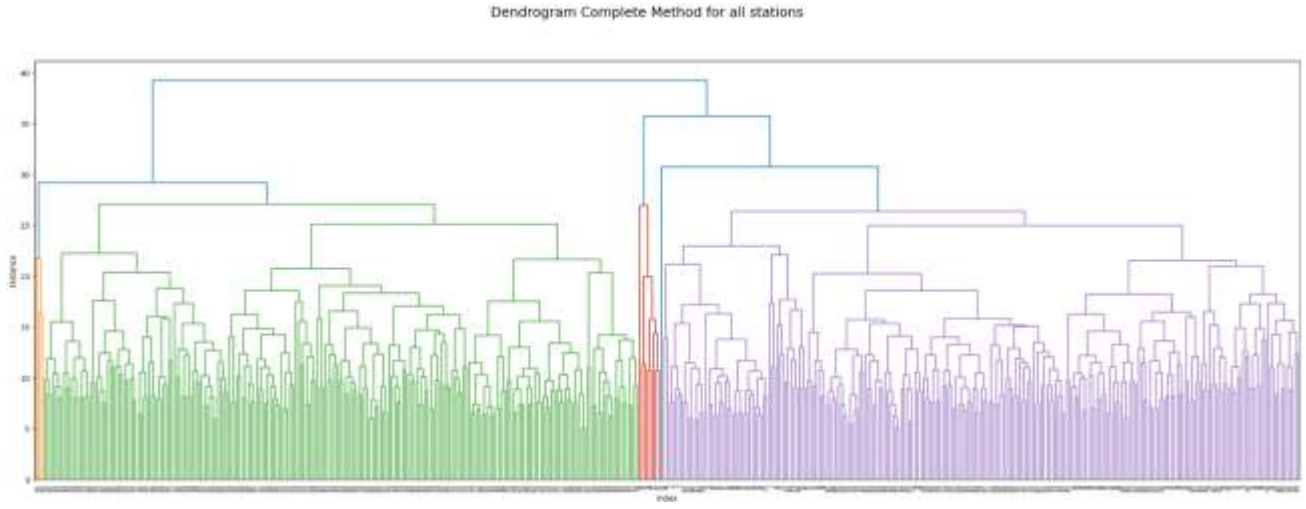
3. Average Method:



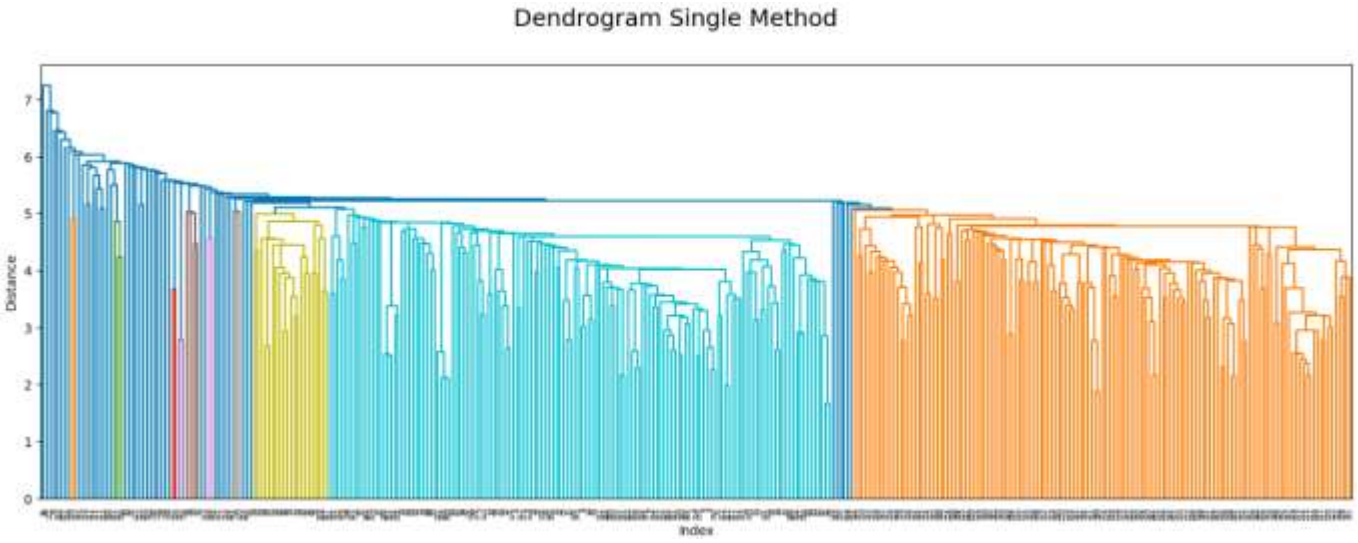
4. Ward Method:



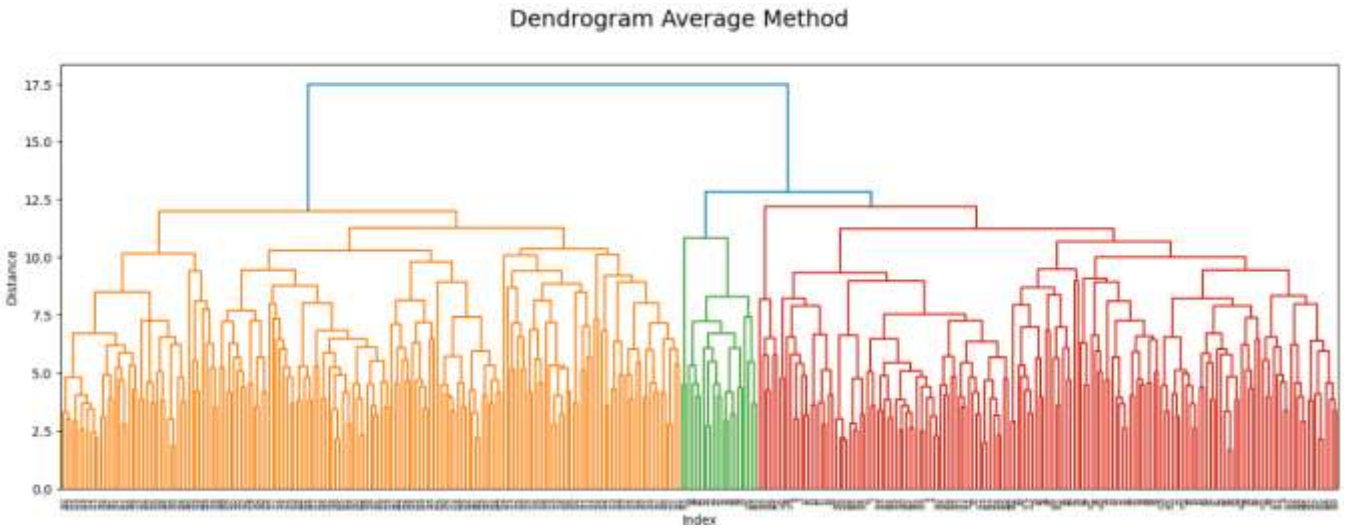
5. Complete Method for all stations:



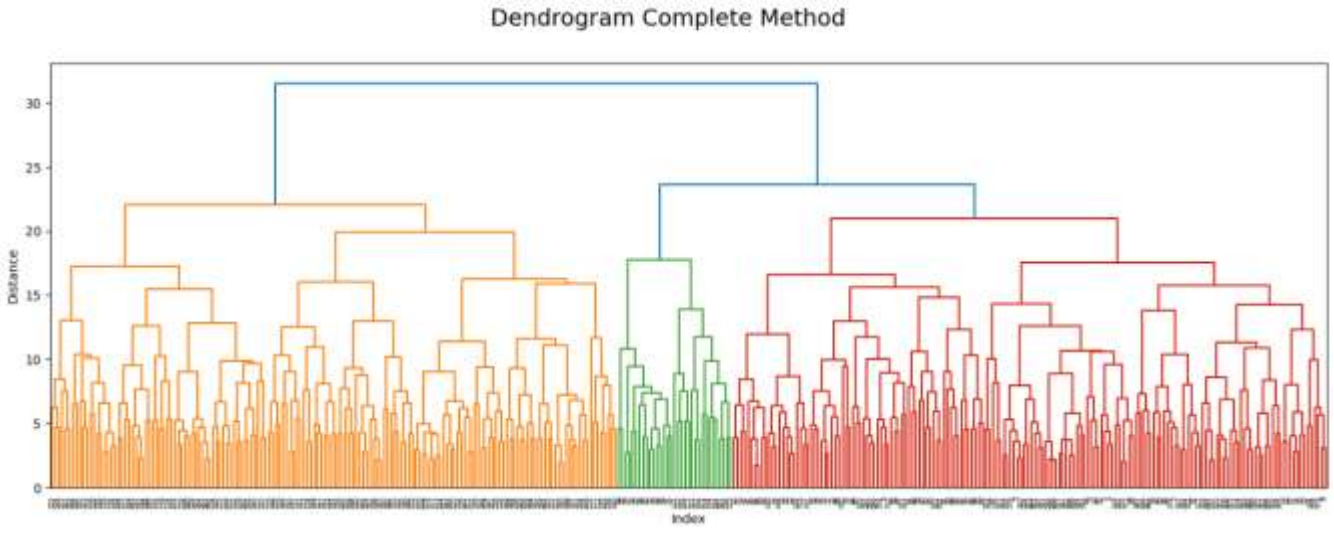
6. Single Method (Reduced Data from 1975):



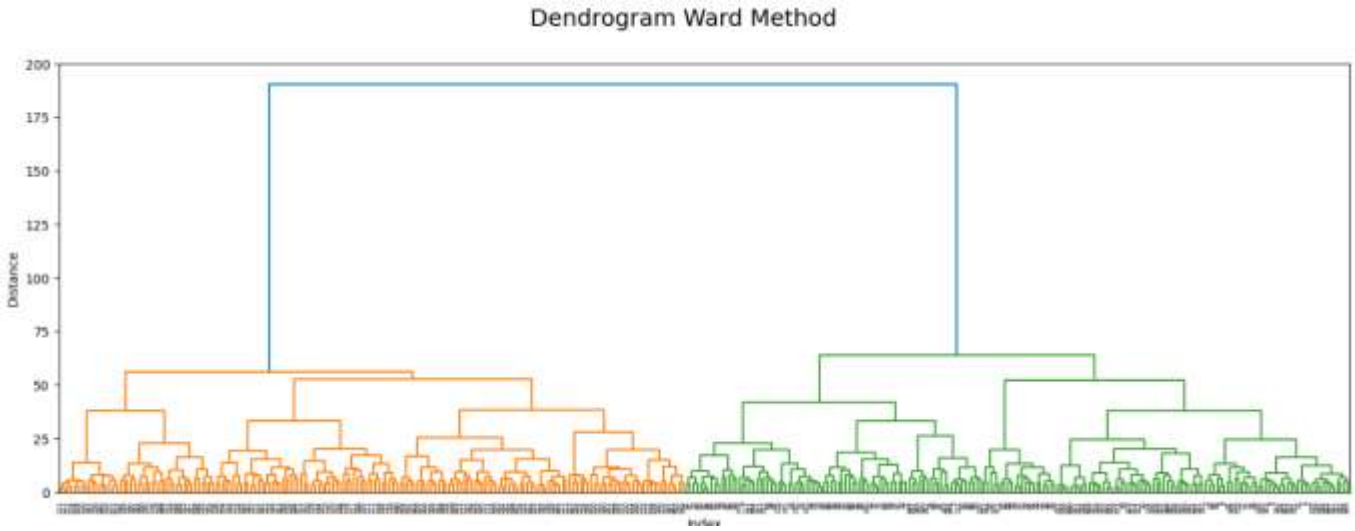
7. Average Method (Reduced Data from 1975):



8. Complete Method (Reduced data from 1975):



9. Ward Method (Reduced Data from 1975):



Observations:

1. Single method is not very useful as it makes one cluster dominate over the others.
2. Complete Method shows a good mix of clusters, which should be the case due to seasonal differences as well as differences between Madrid and Belgrade.
3. Average Method: Mostly shows three distinct categories. Possibly warm weather, cold weather and pleasant days cross over in the middle
4. Ward Method: This shows two clear distinct categories in both datasets. This along with complete method I think are the most valuable in this dataset.