10. Assignment NMST 539

Frantisek Helebrand

May 2022

1 Introduction

In this assignment we will be analyzing the data from the US, specifically, we will be focusing on the crime data from different US states in 1985. We have observations from 50 different US states with 11 variables.

By using K-algorithm (with 4 clusters) we obtain the non-hierarchic construction of clusters which can be seen in the figure 1. The region is distinguished by the different shapes of points. We can see that Alaska has its own cluster.
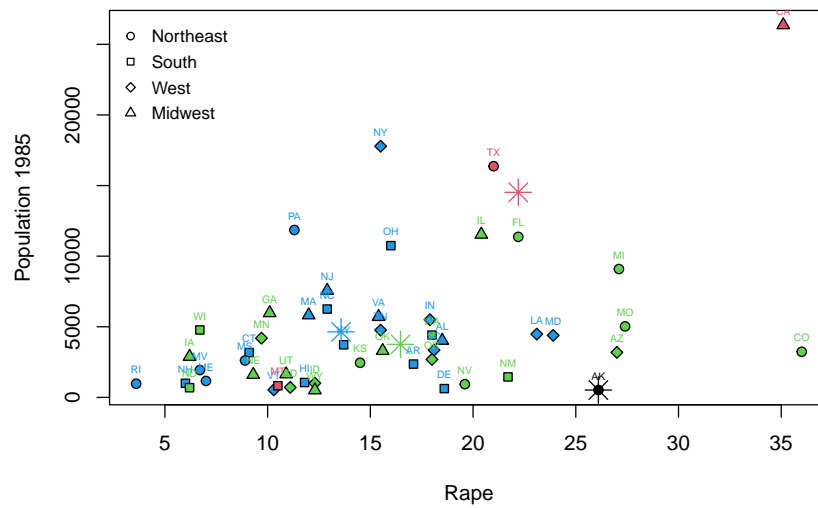

Figure 1: Non-hierarchic construction.

We also compute clusters by hierarchic construction. That can be seen in the figure 2. The clusters are separated by a purple rectangle. The end of

branches refers to a different region. We can see that clusters are similar to region distribution.
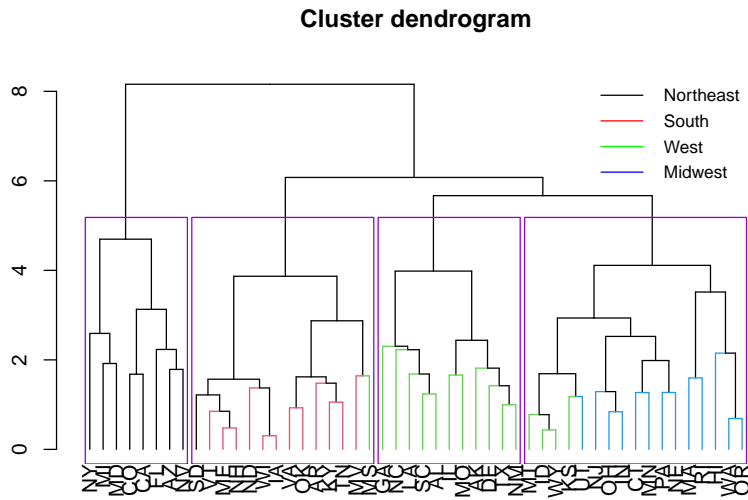
Cluster dendrogram

Figure 2: Hierarchic construction.