Practice and Assignment 9

Clustering

Clustering

Use the data file "UtilitiesLab.csv"

1. Agglomerative Hierarchical Clustering

```
agn <- agnes(mydata, stand=TRUE, diss=FALSE, method="complete")
dend_agn <- as.dendrogram(agn)
plot(dend_agn)
```

2. Do Agglomerative clustering and plot dendrogram for "Single" linkage

3. Do Agglomerative clustering and plot dendrogram for "Average" linkage

4. DBSCAN

Using "dbscan" package and its function dbscan()

(a) do density based clustering (into variable Dbscan_cl) for eps=0.45 and MinPts=3

(b) Check the clusters

(c) Plot the clusters

```
pairs(data, col=Dbscan_cl$cluster + 1L)
```

5. Implement k-means using R based on the outline of the algorithm given below (do not using any available k-means function) and apply to the same data file above

arbitrarily choose k objects from D as the initial cluster centers;

repeat

      (re)assign each object to the cluster to which the object is the most similar, based on the mean value of the objects in the cluster;

      update the cluster means, that is, calculate the mean value of the objects for each cluster;

until no change;