

Hierarchical-Clustering.R

patriciamaya

2020-11-30

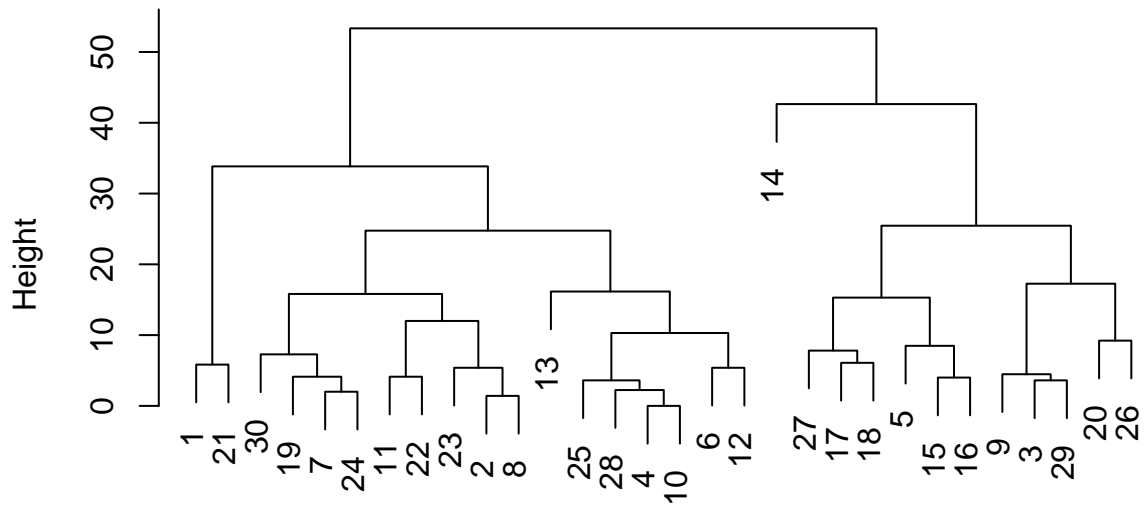
```
#HIERARCHICAL CLUSTERING (unsupervised learning)
library(datasets)
?attitude

data <- attitude[, c(3,4)]
#we are using only 2 variables for learning purposes (see how clusters change)

distance <- dist(data)
clusters <- hclust(distance, method="complete")
#complete link-max distance of all points
clusters

##
## Call:
## hclust(d = distance, method = "complete")
##
## Cluster method   : complete
## Distance         : euclidean
## Number of objects: 30
plot(clusters)
```

Cluster Dendrogram

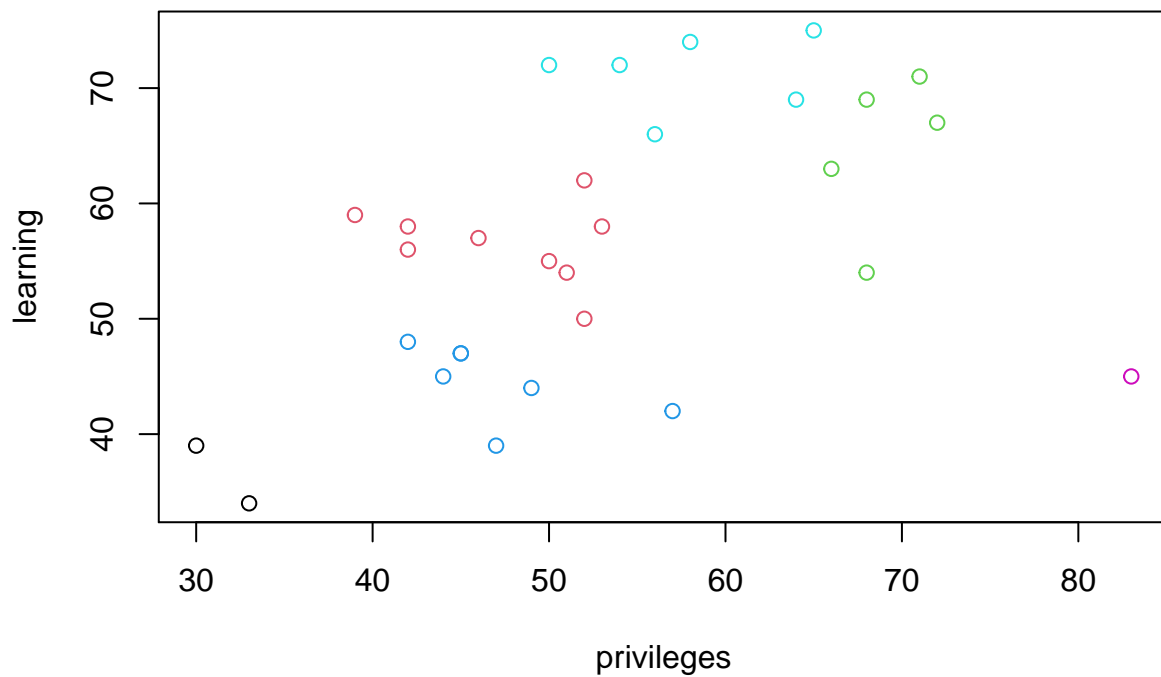


distance
hclust (*, "complete")

```
clusterCut <- cutree(clusters, k=6) #cutting based on num of clusters
clusterCut
```

```
## [1] 1 2 3 4 5 4 2 2 3 4 2 4 4 6 5 5 5 5 2 3 1 2 2 2 4 3 5 4 3 2
```

```
plot(data, col = clusterCut)
```



```
#cutree(clusters, h=20) #cutting based on the height
```

```
#EVALUATING THE MODEL using silhouette
```

```
library(cluster)
```

```
# ?silhouette
```

```
ss<-silhouette(clusterCut, dist(data))
```

```
mean(ss[,3])
```

```
## [1] 0.3839631
```