

Exercise-sec23

Mehrab Atighi

12/22/2021

```
rm(list=ls())
Data = data.frame(
  City = c("Atlanta" , "Boston" , "Chicago" ,
           "Dallas" , "Denver" , "Detroit",
           "Hartford" , "Honolulu" , "Houston",
           "Kansas City" , "Los Angeles" ,
           "New Orleans", "New York",
           "Portland" , "Tucson" ,
           "Washington"),
  Murder = c(16.5,4.2,11.6,18.1,6.9,13.0,
             2.5,3.6,16.8,10.8,9.7,10.3,
             9.4,5.0,5.1,12.5),
  Rape = c(24.8,13.3,24.7,34.2,41.5,35.7,
            8.8,12.7,26.6,43.2,51.8,39.7,
            19.4,23.0,22.9,27.6),
  Robbery = c(106,122,340,184,173,477,
              68,42,289,255,286,266,
              522,157,85,524),
  Assault = c(147,90,242,293,191,220,
              103,28,186,226,355,283,
              267,144,148,217),
  Burglary = c(1112,982,808,1668,1534,
               1566,1017,1457,1509,1494,
               1902,1056,1674,1530,1206,
               1496),
  Larceny = c(905,669,609,901,1368,1183,
              724,1102,787,955,1386,
              1036,1392,1281,756,1003),
  AutoThef = c(494,954,645,605,780,
               788,468,637,697,765,862,
               776,848,488,483,793))
```

#a) is with hand writing the distance between Atlanta and Boston:

Handwritten calculations on a piece of paper:

$$\begin{aligned}\text{Atlanta} &= (16.5, 24.8, 106, 147, 1112, 905, 494) \\ \text{Boston} &= (4.2, 13.3, 122, 90, 982, 669, 954) \\ \text{diff} &= (12.3, 11.5, -16, 57, 130, 236, -460) \\ (\text{diff})^2 &= (151.29, 132.25, 256.0, 3249.0, 16900.0, 55696.0, 211600.0) \\ \text{Sum}((\text{diff})^2) &= 287984.5 \xrightarrow{\text{sqrt}} \text{Distance} = \underline{536.6419}\end{aligned}$$

```

#b)

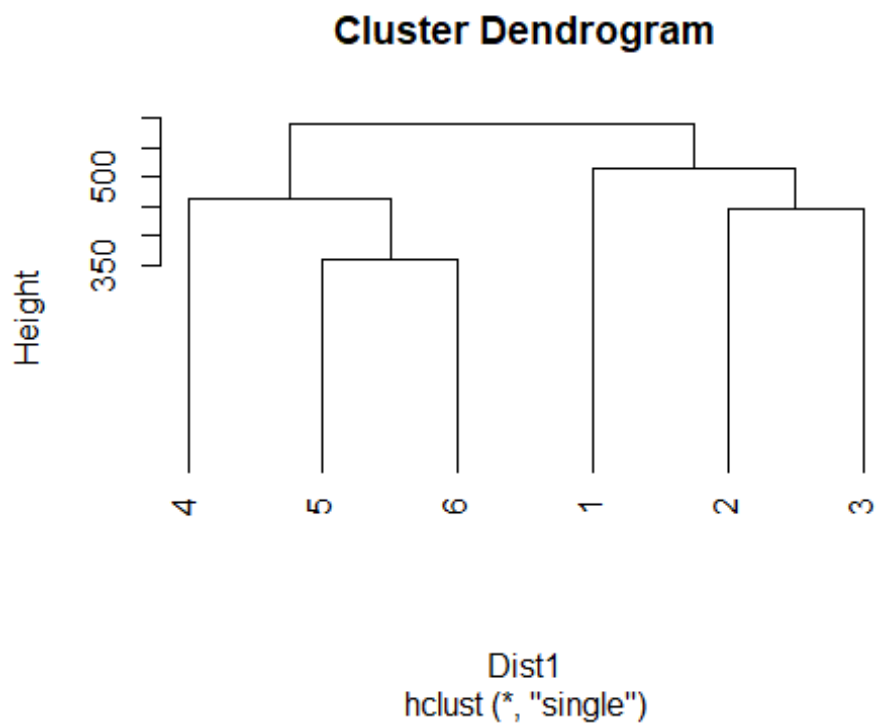
Dist1 = dist(Data[1:6,-1] , method = "euclidean",diag = TRUE , upper = TRUE)

#single method
modell1 = hclust(Dist1 , method = "single")
modell1

##
## Call:
## hclust(d = Dist1, method = "single")
##
## Cluster method   : single
## Distance         : euclidean
## Number of objects: 6

plot( modell1 , hang = -1 )

```



```

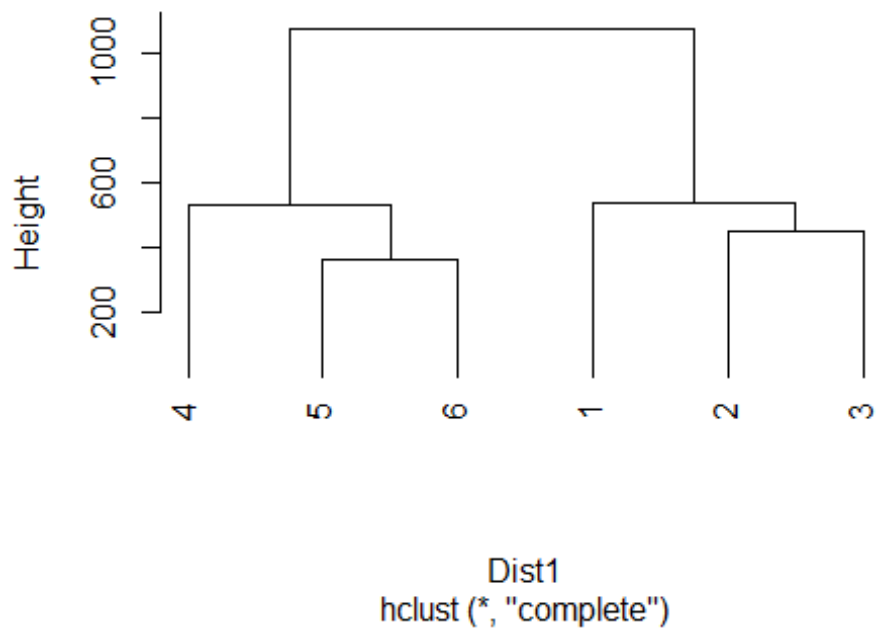
#complete method:
model2 = hclust(Dist1 , method = "complete")
model2

##
## Call:
## hclust(d = Dist1, method = "complete")
##
## Cluster method      : complete
## Distance            : euclidean
## Number of objects: 6

plot( model2 , hang = -1 )

```

Cluster Dendrogram



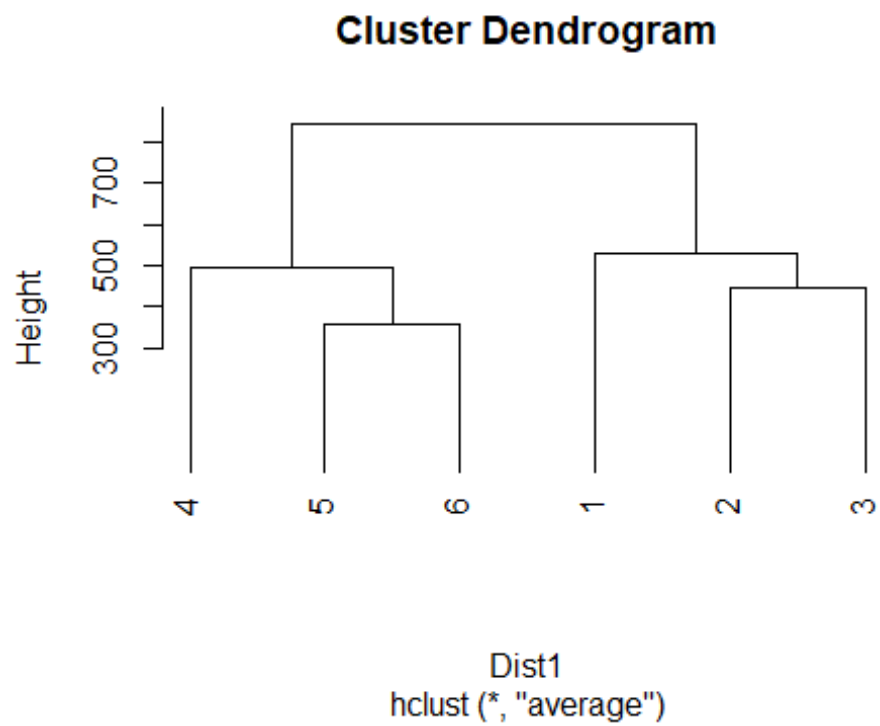
```

#average method:
model3 = hclust(Dist1 , method = "average")
model3

##
## Call:
## hclust(d = Dist1, method = "average")
##
## Cluster method   : average
## Distance         : euclidean
## Number of objects: 6

plot( model3 , hang = -1 )

```



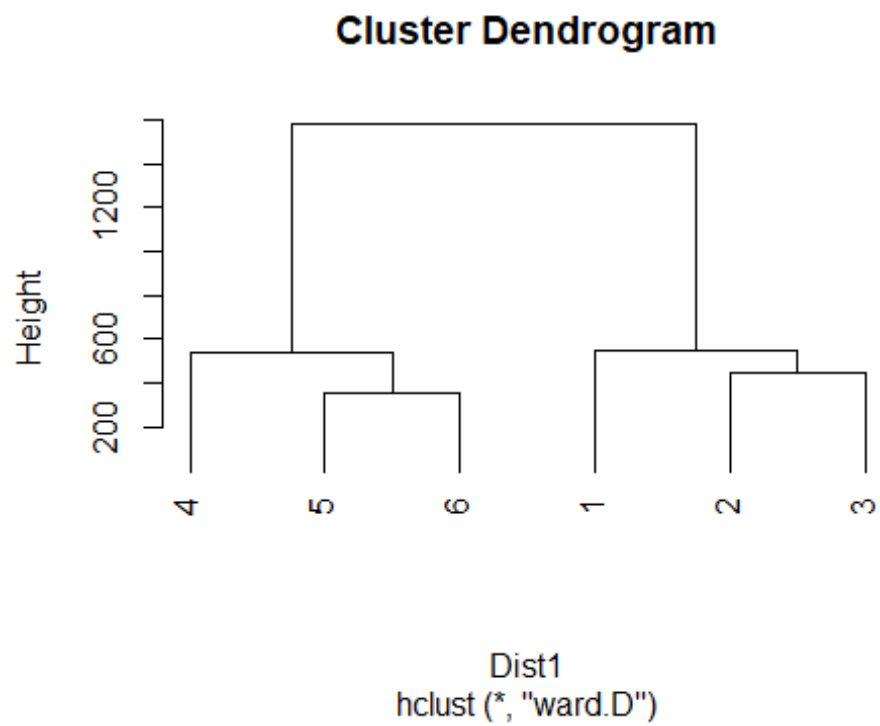
```

#ward method:
model4 = hclust(Dist1 , method = "ward.D")
model4

##
## Call:
## hclust(d = Dist1, method = "ward.D")
##
## Cluster method      : ward.D
## Distance            : euclidean
## Number of objects: 6

plot( model4 , hang = -1 )

```



```

#c)

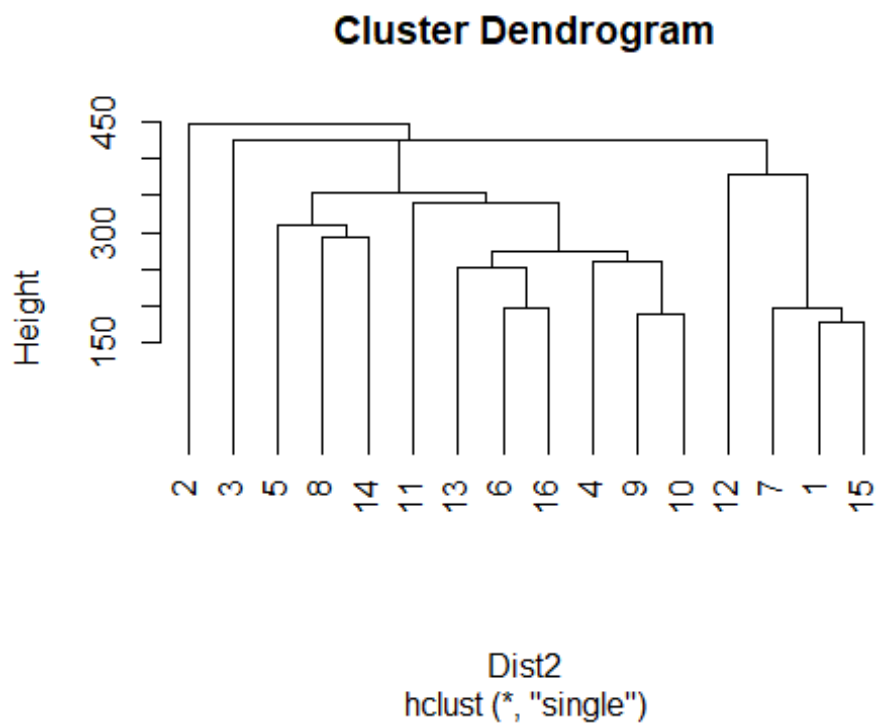
Dist2 = dist(Data[,-1] , method = "euclidean",diag = TRUE , upper = TRUE)

#single method:
model5 = hclust(Dist2 , method = "single")
model5

##
## Call:
## hclust(d = Dist2, method = "single")
##
## Cluster method      : single
## Distance             : euclidean
## Number of objects: 16

plot( model5 , hang = -1 )

```



```

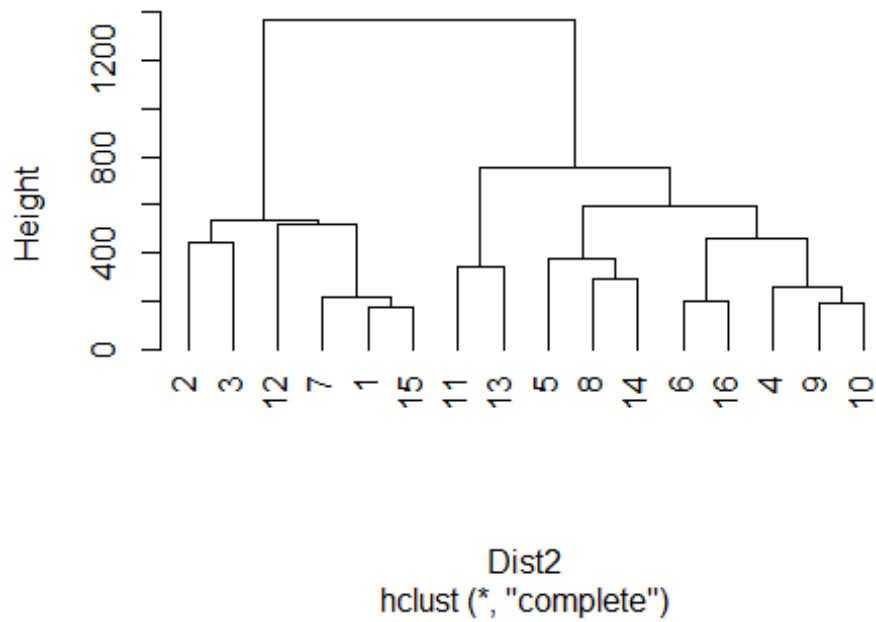
#complete method:
model6 = hclust(Dist2 , method = "complete")
model6

##
## Call:
## hclust(d = Dist2, method = "complete")
##
## Cluster method      : complete
## Distance            : euclidean
## Number of objects: 16

plot( model6 , hang = -1 )

```

Cluster Dendrogram



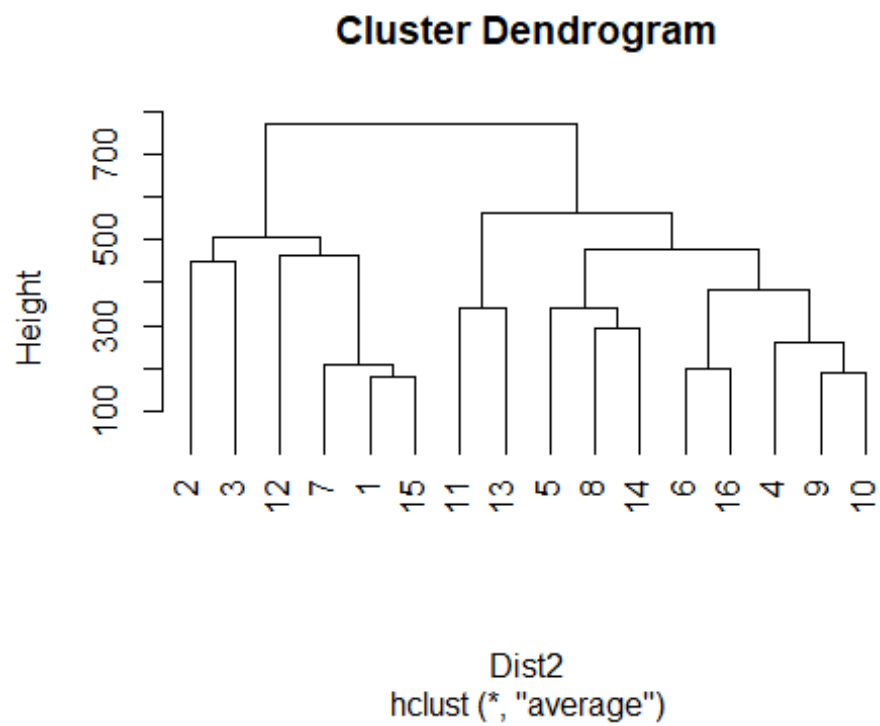

```

#average method:
model7 = hclust(Dist2 , method = "average")
model7

##
## Call:
## hclust(d = Dist2, method = "average")
##
## Cluster method      : average
## Distance            : euclidean
## Number of objects: 16

plot( model7 , hang = -1 )

```



```

#ward method:
model7 = hclust(Dist2 , method = "ward.D")
model7

##
## Call:
## hclust(d = Dist2, method = "ward.D")
##
## Cluster method      : ward.D
## Distance             : euclidean
## Number of objects: 16

plot( model7 , hang = -1 )

```

