Hierarchical Clustering solutions (beginner)

Below are the solutions to [these](#) exercises on hierarchical clustering.

Learn more about clustering in the online courses [Applied Multivariate Analysis with R](#) and [Foundations of strategic business analytics](#)

```
# Prepare dataset
library(ggmap)
capitals <- c("Albania, Tirana", "Andorra, Andorra la Vella", "Armenia, Yerevan",
              "Austria, Vienna", "Azerbaijan, Baku", "Belarus, Minsk",
              "Belgium, Brussels", "Bosnia and Herzegovina, Sarajevo",
              "Bulgaria, Sofia", "Croatia, Zagreb", "Cyprus, Nicosia",
              "Czech Republic, Prague", "Denmark, Copenhagen", "Estonia, Tallinn",
              "Finland, Helsinki", "France, Paris", "Germany, Berlin",
              "Greece, Athens", "Georgia, Tbilisi", "Hungary, Budapest",
              "Iceland, Reykjavik", "Italy, Rome", "Latvia, Riga",
              "Kazakhstan, Astana", "Liechtenstein, Vaduz", "Lithuania, Vilnius",
              "Luxembourg, Luxembourg", "Macedonia, Skopje", "Malta, Valletta",
              "Moldova, Chişinău", "Monaco, Monaco-Ville", "Montenegro, Podgorica",
              "Netherlands, Amsterdam", "Norway, Oslo", "Poland, Warsaw",
              "Portugal, Lisbon", "Republic of Ireland, Dublin",
```

```
                "Romania, Bucharest", "Russia, Moscow", "San
Marino, San Marino",
                "Serbia, Belgrade", "Slovakia, Bratislava",
"Slovenia, Ljubljana",
                "Spain, Madrid", "Sweden, Stockholm",
"Switzerland, Bern",
                "Turkey, Ankara", "Ukraine, Kiev", "United
Kingdom, London",
                "Vatican City, Vatican City"
                )
theData <- geocode(capitals)
rownames(theData) <- capitals


###############
#             #
# Exercise 1  #
#             #
###############

distances <- dist(theData)

###############
#             #
# Exercise 2  #
#             #
###############

dendrogram <- hclust(distances)

###############
#             #
# Exercise 3  #
#             #
###############

plot(dendrogram)

###############
#             #
# Exercise 4  #
#             #
```

```
###############
#
plot(dendrogram, hang=-1)

###############
#             #
# Exercise 5  #
#             #
###############

cutree(dendrogram, k=3)

##                    Albania, Tirana         Andorra, Andorra la
Vella
##                                                              1
2
##                    Armenia, Yerevan                 Austria,
Vienna
##                                                              1
1
##                    Azerbaijan, Baku                 Belarus,
Minsk
##                                                              1
1
##             Belgium, Brussels Bosnia and Herzegovina,
Sarajevo
##                                                              2
1
##                     Bulgaria, Sofia                 Croatia,
Zagreb
##                                                              1
1
##                    Cyprus, Nicosia         Czech Republic,
Prague
##                                                              1
1
##             Denmark, Copenhagen                     Estonia,
Tallinn
##                                                              1
1
##               Finland, Helsinki                     France,
```

```
Paris
##                                                    1
2
##                  Germany, Berlin                   Greece,
Athens
##                                                    1
1
##                  Georgia, Tbilisi               Hungary,
Budapest
##                                                    1
1
##                              Iceland,  Reykjavik
Italy,  Rome
##                                                    2
1
##                  Latvia, Riga              Kazakhstan,
Astana
##                                                    1
3
##          Liechtenstein, Vaduz              Lithuania,
Vilnius
##                                                    1
1
##          Luxembourg, Luxembourg              Macedonia,
Skopje
##                                                    2
1
##                  Malta, Valletta               Moldova,
Chişinău
##                                                    1
1
##          Monaco, Monaco-Ville              Montenegro,
Podgorica
##                                                    1
1
##                          Netherlands,  Amsterdam
Norway,  Oslo
##                                                    2
1
##                  Poland, Warsaw               Portugal,
Lisbon
```

```
##                                                     1
2
##     Republic of Ireland, Dublin               Romania,
Bucharest
##                                                     2
1
##               Russia, Moscow           San Marino, San
Marino
##                                                     1
1
##               Serbia, Belgrade              Slovakia,
Bratislava
##                                                     1
1
##          Slovenia, Ljubljana                   Spain,
Madrid
##                                                     1
2
##                         Sweden,  Stockholm
Switzerland, Bern
##                                                     1
1
##                              Turkey,  Ankara
Ukraine, Kiev
##                                                     1
1
##        United Kingdom, London      Vatican City,
Vatican City
##                                                     2
1


##############
#           #
# Exercise 6  #
#           #
##############

cutree(dendrogram, h=20)


##             Albania, Tirana         Andorra, Andorra la
Vella
```

```
##                                                              1
2
##                    Armenia, Yerevan                    Austria,
Vienna
##                                                              3
1
##                    Azerbaijan, Baku                    Belarus,
Minsk
##                                                              3
4
##                    Belgium, Brussels Bosnia and Herzegovina,
Sarajevo
##                                                              2
1
##                     Bulgaria, Sofia                    Croatia,
Zagreb
##                                                              1
1
##                     Cyprus, Nicosia          Czech Republic,
Prague
##                                                              3
1
##                    Denmark, Copenhagen                  Estonia,
Tallinn
##                                                              5
5
##                    Finland, Helsinki                    France,
Paris
##                                                              5
2
##                     Germany, Berlin                     Greece,
Athens
##                                                              1
1
##                    Georgia, Tbilisi                    Hungary,
Budapest
##                                                              3
1
##                                   Iceland,  Reykjavik
Italy, Rome
##                                                              6
```

```
1
##                          Latvia, Riga                       Kazakhstan,
Astana
##                                                                      5
7
##                    Liechtenstein, Vaduz                       Lithuania,
Vilnius
##                                                                      1
4
##                 Luxembourg, Luxembourg                       Macedonia,
Skopje
##                                                                      2
1
##                       Malta, Valletta                         Moldova,
Chişinău
##                                                                      1
4
##                 Monaco, Monaco-Ville                       Montenegro,
Podgorica
##                                                                      1
1
##                             Netherlands,  Amsterdam
Norway, Oslo
##                                                                      2
5
##                        Poland, Warsaw                         Portugal,
Lisbon
##                                                                      5
2
##      Republic of Ireland, Dublin                            Romania,
Bucharest
##                                                                      2
4
##                       Russia, Moscow                     San Marino, San
Marino
##                                                                      4
1
##                       Serbia, Belgrade                       Slovakia,
Bratislava
##                                                                      1
1
```

```
##            Slovenia, Ljubljana                    Spain,
Madrid
##                                                        1
2
##                              Sweden,   Stockholm
Switzerland, Bern
##                                                        5
1
##                              Turkey,   Ankara
Ukraine, Kiev
##                                                        3
4
##          United Kingdom, London      Vatican City,
Vatican City
##                                                        2
1


plot(dendrogram)
abline(h=20, col="red", lty=2)


###############
#             #
# Exercise 7  #
#             #
###############

plot(dendrogram, hang=-1)
rect.hclust(dendrogram, k=3, border=1:3)


###############
#             #
# Exercise 8  #
#             #
###############

plot(dendrogram, hang=-1)
rect.hclust(dendrogram, k=3, border="red")
rect.hclust(dendrogram, h=20, border="blue")


###############
#             #
```

```
# Exercise 9  #
#               #
###############

# ward.D
plot(hclust(distances, method="ward.D"), main="ward.D")
rect.hclust(hclust(distances,   method="ward.D"),   k=5,
border=1:5)
# ward.D2
plot(hclust(distances, method="ward.D2"), main="ward.D2")
rect.hclust(hclust(distances,   method="ward.D2"),   k=5,
border=1:5)
# single
plot(hclust(distances, method="single"), main="single")
rect.hclust(hclust(distances,   method="single"),   k=5,
border=1:5)
# complete
plot(hclust(distances, method="complete"), main="complete")
rect.hclust(hclust(distances,   method="complete"),   k=5,
border=1:5)
# average
plot(hclust(distances, method="average"), main="average")
rect.hclust(hclust(distances,   method="average"),   k=5,
border=1:5)
# mcquitty
plot(hclust(distances, method="mcquitty"), main="mcquitty")
rect.hclust(hclust(distances,   method="mcquitty"),   k=5,
border=1:5)
# median
plot(hclust(distances, method="median"), main="median")
rect.hclust(hclust(distances,   method="median"),   k=5,
border=1:5)
# centroid
plot(hclust(distances, method="centroid"), main="centroid")
rect.hclust(hclust(distances,   method="centroid"),   k=5,
border=1:5)


################
#               #
# Exercise 10  #
#               #
```

```
################

myVersion <- hclust(distances, method="complete")
groups     <- cutree(myVersion, 7)
plot(theData, cex=0, xlim=c(-30,75))
text(theData, rownames(theData), cex=0.6, col=groups)
```