Multivariate Analysis for the Behavioral Sciences, Second Edition (Chapman and Hall/CRC, 2019)

Exercises of Chapter 17: Cluster Analysis

19 December 2018

Exercises

Exercise 17.3

Use the matrix lowtemp below, modifying the related R code given in the **Examples of Chapter 17**.

```
lowtemp <- structure(
  c(-8, -7, -44, -12, -27, 4, -25, -8, 53, 12, -22, 23, 30, -17, -6, -23, 17, -26, -16, 24,
    0, -5, 26, 20, -12, 16, 7, 30, -2, 23, 57, 31, 6, 39, 46, 23, 12, 5, 32, 8, 13, 31, 29,
    24, 53, 51, 35, 54, 40, 59, 43, 57, 67, 62, 36, 49, 69, 51, 52, 44, 61, 40, 33, 43, 43,
    55, 28, 25, 5, 28, 17, 29, 3, 25, 64, 33, 11, 41, 51, 26, 28, 13, 34, 15, 8, 34, 28, 29),
.Dim = c(22L, 4L), .Dimnames = list(c("Atlanta", "Baltimore", "Bismark", "Boston", "Chicago",
                   "Dallas", "Denver", "El Paso", "Honolulu", "Houston", "Juneau", "Los Angeles",
                   "Miami", "Nashville", "New York", "Omaha", "Phoenix", "Portland", "Reno",
                   "San Francisco", "Seattle", "Washington"),
                 c("January", "April", "July", "October")))
lowtemp
```

```
##               January April July October
## Atlanta            -8    26   53      28
## Baltimore          -7    20   51      25
## Bismark           -44   -12   35       5
## Boston            -12    16   54      28
## Chicago           -27     7   40      17
## Dallas              4    30   59      29
## Denver            -25    -2   43       3
## El Paso            -8    23   57      25
## Honolulu           53    57   67      64
## Houston            12    31   62      33
## Juneau            -22     6   36      11
## Los Angeles        23    39   49      41
## Miami              30    46   69      51
## Nashville         -17    23   51      26
## New York           -6    12   52      28
## Omaha             -23     5   44      13
## Phoenix            17    32   61      34
## Portland          -26     8   40      15
## Reno              -16    13   33       8
## San Francisco      24    31   43      34
## Seattle             0    29   43      28
## Washington         -5    24   55      29
```

Exercise 17.5

Use the protein data below, modifying the related R code given in the **Examples of Chapter 17**.

```
protein <- read.table("data/protein.txt", sep = '\t', header = TRUE)
protein
```

```
##                Rmeat Wmeat Eggs Milk Fish Cereals Sfoods Pulses Fruitveg
## Albania         10.1   1.4  0.5  8.9  0.2    42.3    0.6    5.5      1.7
## Austria          8.9  14.0  4.3 19.9  2.1    28.0    3.6    1.3      4.3
## Belgium         13.5   9.3  4.1 17.5  4.5    26.6    5.7    2.1      4.0
## Bulgaria         7.8   6.0  1.6  8.3  1.2    56.7    1.1    3.7      4.2
## Czechoslovakia   9.7  11.4  2.8 12.5  2.0    34.3    5.0    1.1      4.0
## Denmark         10.6  10.8  3.7 25.0  9.9    21.9    4.8    0.7      2.4
## E.Germany        8.4  11.6  3.7 11.1  5.4    24.6    6.5    0.8      3.6
## Finland          9.5   4.9  2.7 33.7  5.8    26.3    5.1    1.0      1.4
## France          18.0   9.9  3.3 19.5  5.7    28.1    4.8    2.4      6.5
## Greece          10.2   3.0  2.8 17.6  5.9    41.7    2.2    7.8      6.5
## Hungary          5.3  12.4  2.9  9.7  0.3    40.1    4.0    5.4      4.2
## Ireland         13.9  10.0  4.7 25.8  2.2    24.0    6.2    1.6      2.9
## Italy            9.0   5.1  2.9 13.7  3.4    36.8    2.1    4.3      6.7
## Netherlands      9.5  13.6  3.6 23.4  2.5    22.4    4.2    1.8      3.7
## Norway           9.4   4.7  2.7 23.3  9.7    23.0    4.6    1.6      2.7
## Poland           6.9  10.2  2.7 19.3  3.0    36.1    5.9    2.0      6.6
## Portugal         6.2   3.7  1.1  4.9 14.2    27.0    5.9    4.7      7.9
## Romania          6.2   6.3  1.5 11.1  1.0    49.6    3.1    5.3      2.8
## Spain            7.1   3.4  3.1  8.6  7.0    29.2    5.7    5.9      7.2
## Sweden           9.9   7.8  3.5 24.7  7.5    19.5    3.7    1.4      2.0
## Switzerland     13.1  10.1  3.1 23.8  2.3    25.6    2.8    2.4      4.9
## UK              17.4   5.7  4.7 20.6  4.3    24.3    4.7    3.4      3.3
## USSR             9.3   4.6  2.1 16.6  3.0    43.6    6.4    3.4      2.9
## W.Germany       11.4  12.5  4.1 18.8  3.4    18.6    5.2    1.5      3.8
## Yugoslavia       4.4   5.0  1.2  9.5  0.6    55.9    3.0    5.7      3.2
```

Exercise 17.6

Use the W16F50 data below (the corresponding data of life expectancies at different ages for women), modifying the related R code given in the **Examples of Chapter 17**.

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
W16F50 <- read.csv("data/w16f50.csv")
countries <- W16F50$country
row.names(W16F50) <- countries
W16F50 <- W16F50 %>% select(-country)

W16F50[c("Japan", "Italy", "Spain", "United Kingdom", "Finland", "Cuba", "United States"), ]
```

```
##                birthF age25F age50F age75F age100F
## Japan            87.0   62.4   38.1   15.6     2.5
## Italy            85.0   60.4   36.0   14.0     2.1
## Spain            85.6   61.0   36.6   14.4     3.1
## United Kingdom   82.7   58.2   34.0   12.9     2.3
## Finland          84.1   59.5   35.2   13.4     1.7
## Cuba             80.4   56.2   32.3   12.4     2.1
## United States    81.2   57.0   33.3   13.0     2.3
```

```
var(W16F50)
```

```
##             birthF    age25F    age50F   age75F   age100F
## birthF  10.5635454 9.3913946 8.2995102 5.059122 0.8793225
## age25F   9.3913946 8.4798194 7.5162856 4.615836 0.8733141
## age50F   8.2995102 7.5162856 6.7220415 4.170408 0.8337958
## age75F   5.0591225 4.6158365 4.1704083 2.760918 0.6182040
## age100F  0.8793225 0.8733141 0.8337958 0.618204 0.5184653
```