Actividad-5.R

r2623423

2024-05-08

```r
# TC1002S.201 – Equipo 3
# Actividad Evaluable 5: Análisis de datos
# Jesús Ángel Guzmán Ortega A01799257
# Julio Cesar Vivas Medina A01749879

# Cargamos librerías ------------------------------------
library(corrplot)
```

```
## corrplot 0.92 loaded
```
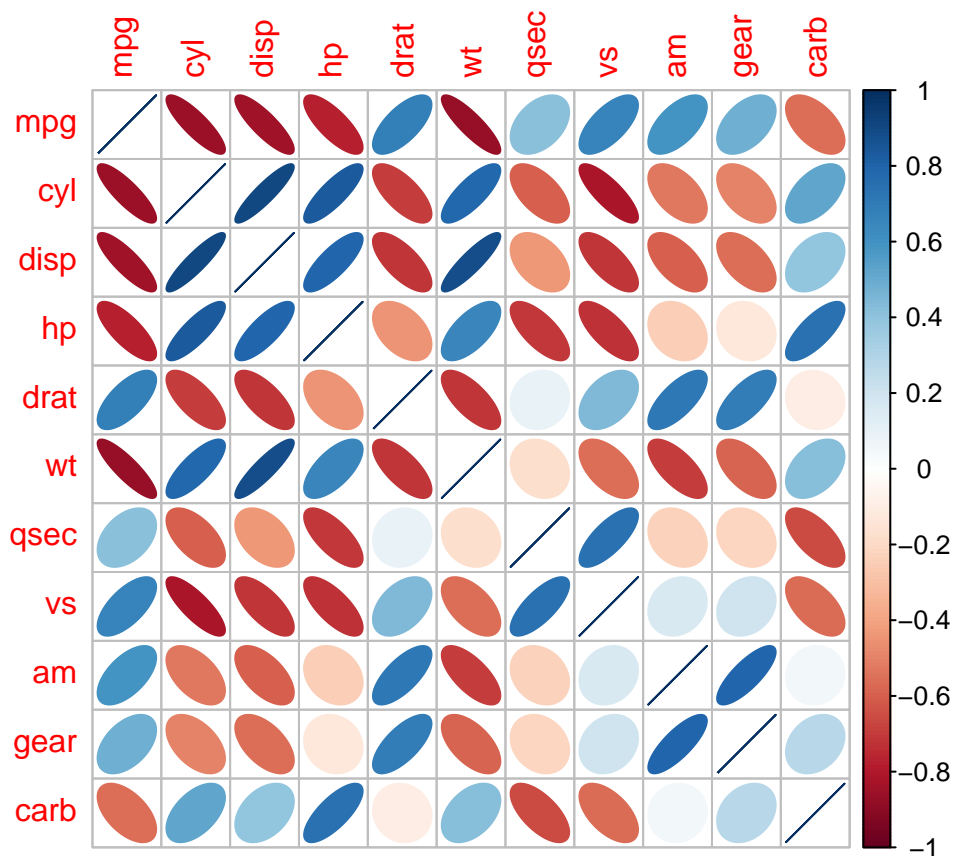
```r
library(factoextra)
```

```
## Loading required package: ggplot2
```

```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

```r
library(ggplot2)

# Preparación de datos ----------------------------------------------
# Cargamos el dataset
data(mtcars)

# Matriz de correlación
correlation <- cor(mtcars)

# Escalado de datos
mtcars_s <- scale(mtcars)

# Análisis de Correlación ----------------------------------------------
corrplot(correlation, method = "ellipse")
```

```
# Análisis de K-means -------------------------------------------------
km2 <- kmeans(mtcars_s, 2)
km2
```
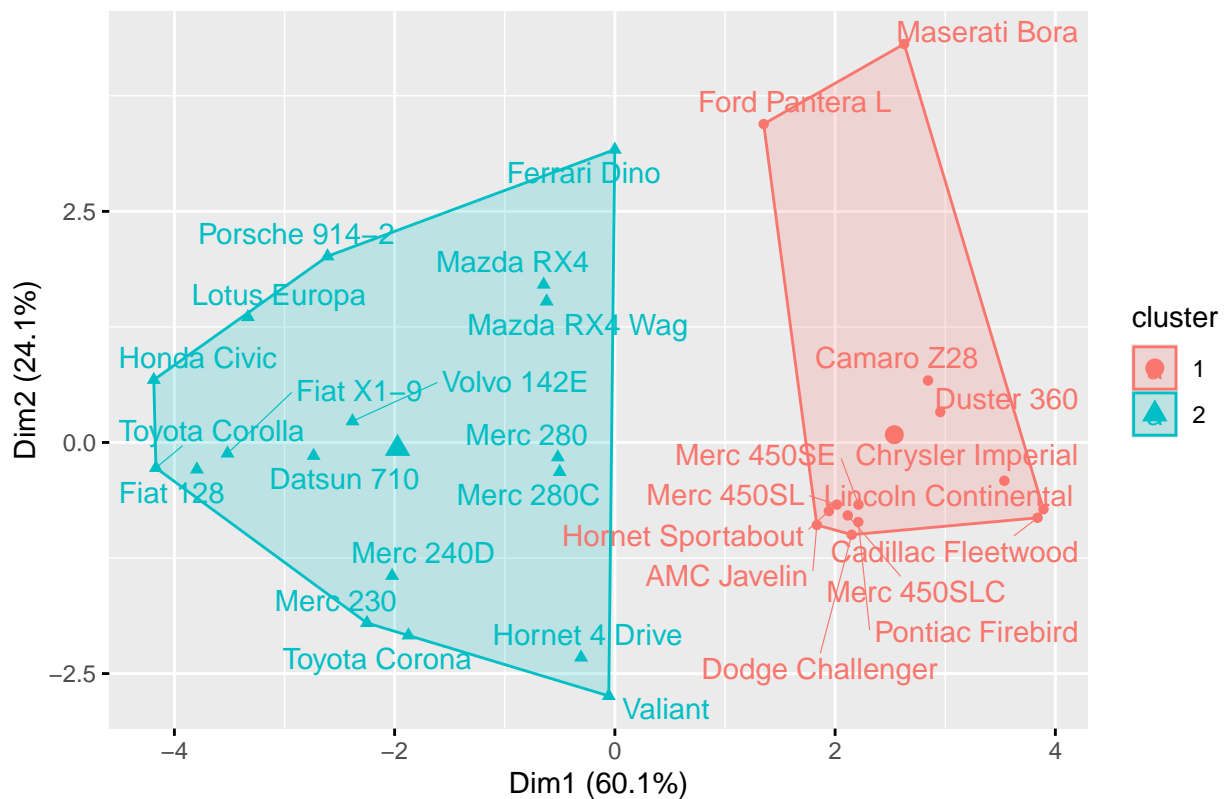
```
## K-means clustering with 2 clusters of sizes 14, 18
##
## Cluster means:
##          mpg        cyl       disp         hp       drat         wt       qsec
## 1 -0.8280518  1.0148821  0.9874085  0.9119628 -0.6869112  0.7991807 -0.6024854
## 2  0.6440403 -0.7893528 -0.7679844 -0.7093044  0.5342642 -0.6215850  0.4685997
##           vs         am       gear       carb
## 1 -0.8680278 -0.5278510 -0.5445697  0.4256439
## 2  0.6751327  0.4105508  0.4235542 -0.3310564
##
## Clustering vector:
##          Mazda RX4        Mazda RX4 Wag          Datsun 710       Hornet 4 Drive
##                  2                    2                   2                    2
##   Hornet Sportabout              Valiant          Duster 360            Merc 240D
##                  1                    2                   1                    2
##           Merc 230             Merc 280           Merc 280C            Merc 450SE
##                  2                    2                   2                    1
##          Merc 450SL          Merc 450SLC   Cadillac Fleetwood  Lincoln Continental
##                  1                    1                   1                    1
##   Chrysler Imperial             Fiat 128          Honda Civic       Toyota Corolla
##                  1                    2                   2                    2
##       Toyota Corona     Dodge Challenger          AMC Javelin           Camaro Z28
##                  2                    1                   1                    1
```

```
##     Pontiac Firebird          Fiat X1-9       Porsche 914-2        Lotus Europa
##                    1                    2                    2                    2
##        Ford Pantera L         Ferrari Dino       Maserati Bora          Volvo 142E
##                    1                    2                    1                    2
##
## Within cluster sum of squares by cluster:
## [1]   64.93441 114.14693
##  (between_SS / total_SS =  47.5 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```r
fviz_cluster(km2, data = mtcars_s, repel = TRUE)
```


Cluster plot

```r
# K-means con 3 clusters
km3 <- kmeans(mtcars_s, 3)
km3
```

```
## K-means clustering with 3 clusters of sizes 14, 8, 10
##
## Cluster means:
##          mpg        cyl       disp         hp        drat         wt
## 1 -0.82805177  1.0148821  0.9874085  0.9119628 -0.68691115  0.7991807
## 2  1.32477913 -1.2248578 -1.1062677 -0.9453003  1.09820619 -1.2008698
## 3  0.09944917 -0.4409488 -0.4973578 -0.5205077  0.08311066 -0.1581571
##        qsec         vs         am       gear        carb
```
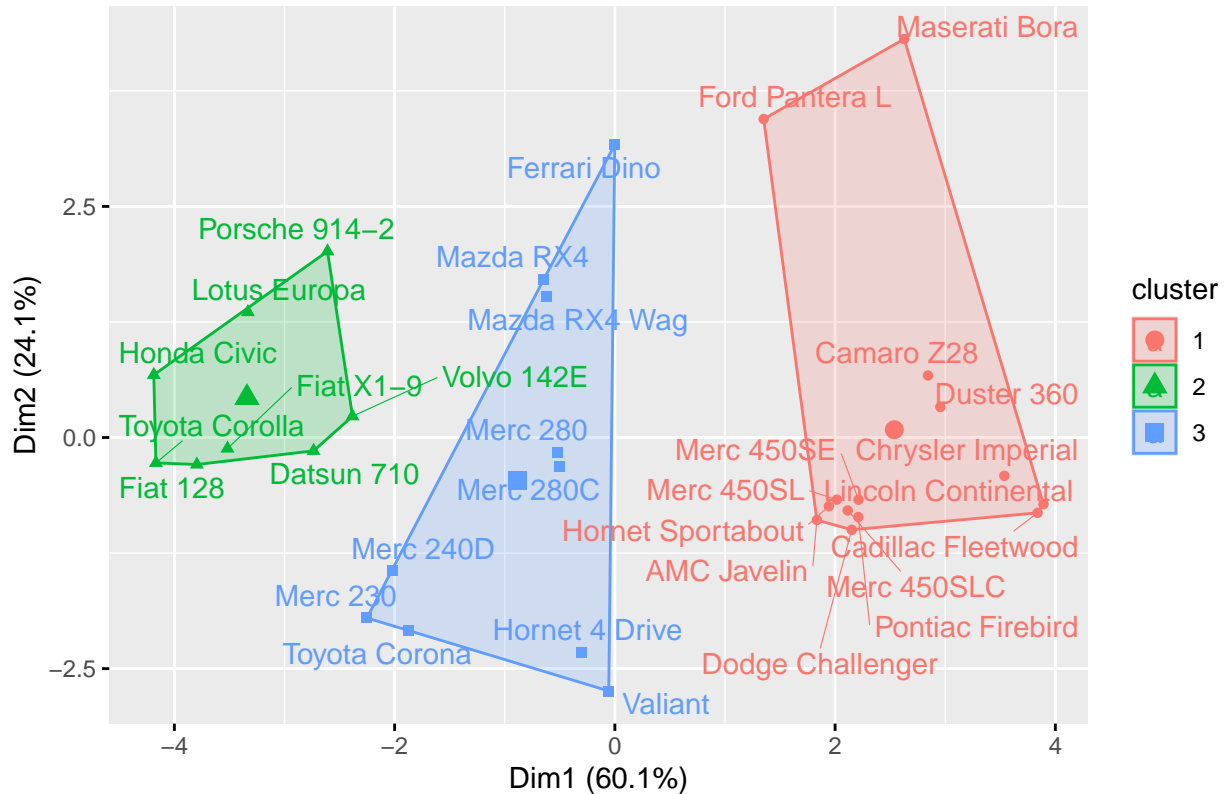
```
## 1 -0.6024854 -0.8680278 -0.5278510 -0.5445697  0.42564389
## 2  0.3364684  0.8680278  1.1899014  0.7623975 -0.81259288
## 3  0.5743048  0.5208167 -0.2129297  0.1524795  0.05417286
##
## Clustering vector:
##          Mazda RX4       Mazda RX4 Wag          Datsun 710       Hornet 4 Drive
##                  3                   3                   2                    3
##   Hornet Sportabout             Valiant          Duster 360            Merc 240D
##                  1                   3                   1                    3
##           Merc 230            Merc 280           Merc 280C            Merc 450SE
##                  3                   3                   3                    1
##          Merc 450SL          Merc 450SLC Cadillac Fleetwood Lincoln Continental
##                  1                   1                   1                    1
##   Chrysler Imperial            Fiat 128         Honda Civic       Toyota Corolla
##                  1                   2                   2                    2
##       Toyota Corona    Dodge Challenger          AMC Javelin           Camaro Z28
##                  3                   1                   1                    1
##     Pontiac Firebird            Fiat X1-9       Porsche 914-2         Lotus Europa
##                  1                   2                   2                    2
##       Ford Pantera L         Ferrari Dino       Maserati Bora            Volvo 142E
##                  1                   3                   1                    2
##
## Within cluster sum of squares by cluster:
## [1] 64.93441 19.04480 59.31024
##  (between_SS / total_SS =  58.0 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```r
fviz_cluster(km3, data = mtcars_s, repel = TRUE)
```

Cluster plot


```r
# K-means con 4 clusters
km4 <- kmeans(mtcars_s, 4)
km4
```

```
## K-means clustering with 4 clusters of sizes 7, 3, 12, 10
##
## Cluster means:
##         mpg        cyl       disp         hp        drat          wt
## 1  0.1082193 -0.5849321 -0.4486701 -0.6496905 -0.04967936 -0.02346989
## 2 -1.3700619  1.0148821  1.8284157  1.0206569 -1.01599172  2.16914561
## 3 -0.6289460  0.9215597  0.6372354  0.8432042 -0.54374552  0.35199964
## 4  1.0900003 -1.0008838 -0.9991381 -0.8632588  0.99206769 -1.05671433
##         qsec         vs         am       gear       carb
## 1  1.18548414  1.1160357 -0.8141431 -0.1573201 -0.4145882
## 2 -0.06085812 -0.8680278 -0.8141431 -0.9318192  0.7352031
## 3 -0.79721801 -0.8680278 -0.3131320 -0.2541325  0.4772371
## 4  0.14508015  0.5208167  1.1899014  0.6946289 -0.5030337
##
## Clustering vector:
##          Mazda RX4       Mazda RX4 Wag          Datsun 710       Hornet 4 Drive
##                  4                   4                   4                    1
##    Hornet Sportabout             Valiant          Duster 360           Merc 240D
##                  3                   1                   3                    1
##           Merc 230            Merc 280           Merc 280C           Merc 450SE
##                  1                   1                   1                    3
##          Merc 450SL          Merc 450SLC  Cadillac Fleetwood  Lincoln Continental
##                  3                   3                   2                    2
```

5

```
##      Chrysler Imperial              Fiat 128         Honda Civic       Toyota Corolla
##                    2                       4                  4                    4
##         Toyota Corona     Dodge Challenger         AMC Javelin          Camaro Z28
##                    1                       3                  3                    3
##      Pontiac Firebird             Fiat X1-9       Porsche 914-2        Lotus Europa
##                    3                       4                  4                    4
##        Ford Pantera L          Ferrari Dino       Maserati Bora          Volvo 142E
##                    3                       3                  3                    4
##
## Within cluster sum of squares by cluster:
## [1] 21.2879804  0.6810421 64.8585007 35.6536100
##  (between_SS / total_SS =  64.1 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```
fviz_cluster(km4, data = mtcars_s, repel = TRUE)
```


Cluster plot