

#### Proyecto 04:

Considerar los datos económicos de la siguiente tabla (dados en millones de dólares) de corporaciones industriales:

1. Calcular la distancia de Mahalanobis entre Ford y Exxon
2. Calcular la distancia de Mahalanobis entre General Motors e IBM
3. Calcular la distancia de Mahalanobis entre Philip Morris y Texaco

La tabla de datos se sube como imagen, recuerden crear un CSV para procesar la información.

#### Codificación:

```
1 # Proyecto 04
2 # Vianey Maravilla Pérez
3 # 5AM1
4 # Analítica y Visualización de Datos
5
6 # Considerar los datos económicos de la siguiente tabla (dados en millones de dólares) de corporaciones industriales
7
8
9
10 # 1. Calcular la distancia de Mahalanobis entre Ford y Exxon
11 # 2. Calcular la distancia de Mahalanobis entre General Motors e IBM
12 # 3. Calcular la distancia de Mahalanobis entre Philip Morris y Texaco
13
14
15
16
17 ruta <- file.choose()
18 main <- function(ruta) {
19   print(ruta)
20   datos<- read.csv(ruta,
21                     header = T,
22                     sep = ",")
23   row.names(datos) <- datos[,1]
24   datos <- datos[,2:dim(datos)[2]]
25   View(datos)
26   means_data <- means(datos)
27   print("Means")
28   print(means_data)
29   covars_data <- covars(datos)
30   print("Covars")
31   print(covars_data)
32   print("Distancia de Mahalanobis entre Ford y Exxon")
33   FE_D <- mahalanobis_d(datos["Ford",], datos["Exxon",], covars_data)
34   print(FE_D)
35   print("Distancia de Mahalanobis entre General Motors e IBM")
36   GMIBM_D <- mahalanobis_d(datos["General Motors",], datos["IBM",], covars_data)
37   print(GMIBM_D)
38   print("Distancia de Mahalanobis entre Philip Morris y Texaco")
39   MT_D <- mahalanobis_d(datos["Philip Morris",], datos["Texaco",], covars_data)
40   print(MT_D)
41 }
42
```

```
43 * means <- function(datos) {  
44   col_no <- dim(datos)[2]  
45   row_no <- dim(datos)[1]  
46   means_list <- list()  
47   for(i in 1:col_no) {  
48     suma <- 0  
49     for(j in 1:row_no) {  
50       suma <- suma + datos[j,i]  
51     }  
52     means_list <- c(means_list, suma / row_no)  
53   }  
54   return(means_list)  
55 * }  
56  
57 * mult_mat <- function(mat1, mat2) {  
58   row_no1 <- dim(mat1)[1]  
59   row_no2 <- dim(mat2)[1]  
60   col_no1 <- dim(mat1)[2]  
61   col_no2 <- dim(mat2)[2]  
62   if(col_no1 == row_no2) {  
63     new_mat <- data.frame()  
64     for(a in 1:col_no2) {  
65       rown <- c()  
66       for(i in 1:row_no1) {  
67         suma <- 0  
68         for(j in 1:col_no1) {  
69           suma <- suma + (mat1[i,j] * mat2[j,a])  
70         }  
71         rown <- c(rown, suma)  
72       }  
73       if(dim(new_mat)[1] == 0) {  
74         new_mat <- cbind(rown)  
75       } else {  
76         new_mat <- cbind(new_mat, rown)  
77       }  
78     }  
79     return(new_mat)  
80 * }  
81 * }  
82  
83 * covars <- function(datos) {  
84   cov_mat <- cov(datos)  
85   return(cov_mat)  
86 * }  
87
```

```
87  
88 * mahalanobis_d <- function(vector1, vector2, covars_data) {  
89   dis_vect <- vector1 - vector2  
90   covars_data_inv <- solve(covars_data)  
91   dis_vect_covars_inv <- mult_mat(dis_vect, covars_data_inv)  
92   t_dis_vect <- t(dis_vect)  
93   mahalanobis_dis <- mult_mat(dis_vect_covars_inv, t_dis_vect)  
94   return(mahalanobis_dis)  
95 * }  
96  
97 main(ruta)
```

Resultados:

Vianey Maravilla Pérez  
5AM1  
Análítica y Visualización de datos

```
> main(ruta)
[1] "C:\\Users\\viane\\Music\\ESCOM\\Análisis y visualización de datos\\Datos.csv"
[1] "Means"
[[1]]
[1] 62309100

[[2]]
[1] 2927300

[[3]]
[1] 81248400

[1] "Covars"
      Ventas  Beneficios  Bienes
Ventas 1.000509e+15 2.557560e+13 1.511827e+15
Beneficios 2.557560e+13 1.430020e+12 4.565462e+13
Bienes 1.511827e+15 4.565462e+13 2.980490e+15
[1] "Distancia de Mahalanobis entre Ford y Exxon"
      rown
rown 6.10403
[1] "Distancia de Mahalanobis entre General Motors e IBM"
      rown
rown 5.921156
[1] "Distancia de Mahalanobis entre Philip Morris y Texaco"
      rown
rown 0.214232
```