Instituto Politécnico Nacional
Escuela Superior de Computo

Programación para la ciencia de datos.

Cristal Karina Galindo Durán

Practica 9:

Análisis Factorial

Vianey Maravilla Pérez

3AM1

Unidad temática a la que corresponde la práctica. *III. Técnicas y métodos de modelado.*

Objetivo*. Realizar scripts en Lenguaje R que permitan realizar el análisis factorial sobre diversos datos.*

Introducción.

El análisis factorial es un conjunto de técnicas que perteneces a la estadística multivariante. El análisis factorial se divide en dos: Análisis factorial confirmatorio y exploratorio.

El análisis confirmatorio tiene por objetivo la reducción de dimensiones; mientras, el análisis factorial exploratorio busca agrupar e identificar las características que definen a cada grupo.

En esta actividad se incluyen un conjunto de ejercicios que le permiten al discente poner en practica conceptos sobre el análisis factorial en lenguaje R con bibliotecas como: nFactors y GPArotation.

Material o equipo necesario

- *Computadora*
- *Internet*
- *Lenguaje R y R Studio*

Consideraciones: Debemos de tener en cuenta que el numero de las observaciones debe de ser mínimo 3 o 4 veces mayor que el número de variables, los datos deben de tener distribución normal bivariada, los datos o más bien las componentes deben ser cuantitativas.

Ejercicios

1.- Considera los datos del siguiente enlace: https://raw.githubusercontent.com/housecricket/data/main/efa/sample1.csv

a) Aplica el análisis factorial

Procedimiento:

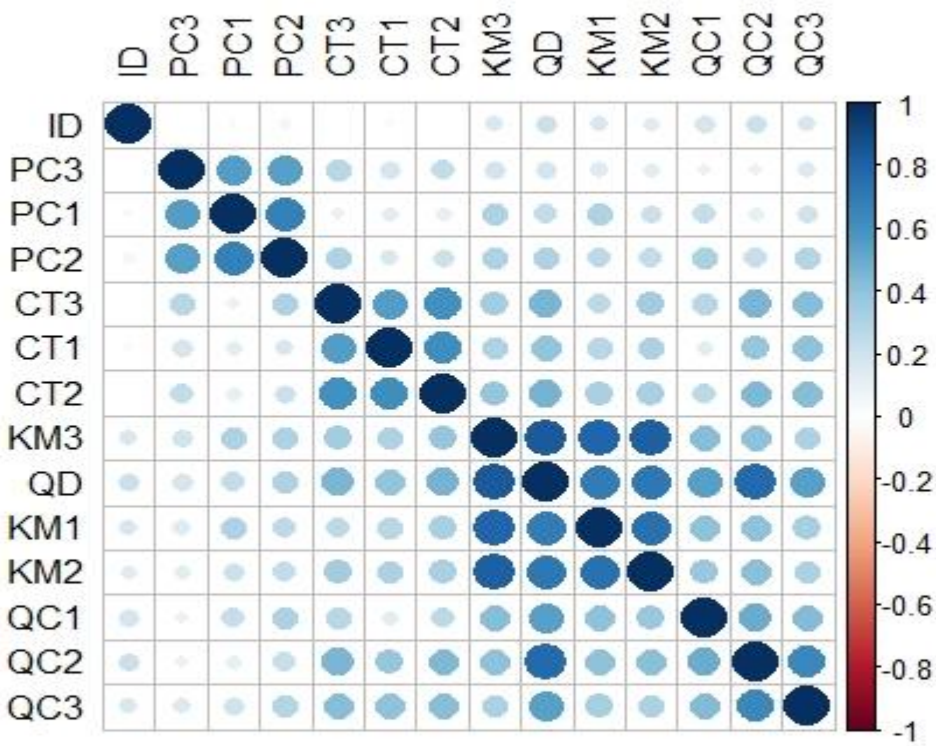
```r
#Ejercicio 1 Practica 9
# Análisis Factorial
# Vianey Maravilla Pérez 3AM1

#Definimos las librerias a utilizar durante todo el proceso
library(openxlsx) #Librería que interactúa con MSExcel
library(corrplot) #Librería para el gráfico de correlaciones
library(corrr) #Otra opción de librería para el cálculo y gráfico de correlaciones
library(psych)
library (ggcorrplot)
library(stats) #Librería del sistema base
library(polycor)
library(GPArotation)

# Lectura de la BDD de acuerdo a su ubicación
archivo <- read.csv('C:/Users/viane/Desktop/ESCOM/3.-TERCER SEMESTRE/PROGRAMACION PARA LAS CIENCIAS DE DATOS/Ejercicio1.csv')

# Visualización de la tabla
archivo

# Obtener matriz de correlación
matriz_correlaciones <- cor(archivo, use = "pairwise.complete.obs")
matriz_correlaciones

# Obtenermos la gráfica de las correlaciones
corrplot(cor(archivo), order = "hclust", tl.col='black', tl.cex=1)

# Calculo de un objeto de correlaciones
archivo_correlaciones <- correlate(archivo)

# Obtener la gráfica de las correlaciones
rplot(archivo_correlaciones, legend = TRUE, colours = c("firebrick1", "black", "darkcyan"), print_cor = TRUE)

# Determinante de la matriz de correlaciones de las variables ingresadas
det(matriz_correlaciones)

```

```r
# Obtener el cálculo de los estimadores del Test de Barlett y KMO
bartlett.test(archivo)

KMO(archivo)

factanal(archivo, factors = 2, rotation = "none")

factanal(archivo, factors = 2, rotation = "none", scores = "regression")$scores

puntuaciones <- factanal(archivo, factors = 2, rotation = "none", scores = "regression")$scores
archivo <- cbind(archivo, puntuaciones)
archivo$Factor1 <- round(((archivo$Factor1 - min(archivo$Factor1))/(max(archivo$Factor1) - min(archivo$Factor1))), 2)
archivo

hist(archivo$Factor1, freq = TRUE, main = "Gráfico de la Distribución del Factor 1",
     xlab = "Factor 1", ylab = "Frecuencia", col = "#009ACD")

#Calcular la matriz de correlación policorica
mat_cor <- hetcor(archivo)$correlations #matriz de correlación policorica
ggcorrplot(mat_cor,type="lower",hc.order = T)

# Obtener el cálculo de los estimadores del Test de Barlett y KMO con nuestra matriz de correlación polocorica
cortest.bartlett(mat_cor)->p_esf
p_esf$p

KMO(mat_cor)


# Escoger un método para extraer los factores
# minres: minimo residuo /
# mle: maxima verosimilitud /
# paf: método de ejes principales /
# alpah: alfa /
# minchi: minimos cuadrados /
# minrak : rango minimo /


modelo1<-fa(mat_cor,
            nfactors = 3,
            rotate = "none",
            fm="mle") # Modelo máxima verosimilitud

modelo2<-fa(mat_cor,
            nfactors = 3,
            rotate = "none",
            fm="minres") # Modelo minimo residuo


# Se comparan las comunalidades
sort(modelo1$communality,decreasing = T)->c1
sort(modelo2$communality,decreasing = T)->c2
head(cbind(c1,c2))

#Se comparan las unicidades
sort(modelo1$uniquenesses,decreasing = T)->u1
sort(modelo2$uniquenesses,decreasing = T)->u2
head(cbind(u1,u2))




# Determinar el número de factores
scree(mat_cor)
fa.parallel(mat_cor,n.obs=200,fa="fa",fm="minres")


# Rotar la matriz
rot<-c("none", "varimax", "quartimax","Promax")
bi_mod<-function(tipo){
  biplot.psych(fa(archivo,nfactors = 2,fm="minres",rotate = tipo),main = paste("Biplot con rotación ",tipo),
               col=c(2,3,4),pch = c(21,18),group = bfi[,"gender"])
}
sapply(rot,bi_mod)


# Interpretación
modelo_varimax<-fa(mat_cor,nfactors = 5,rotate = "varimax",
                   fa="minres")
fa.diagram(modelo_varimax)
```

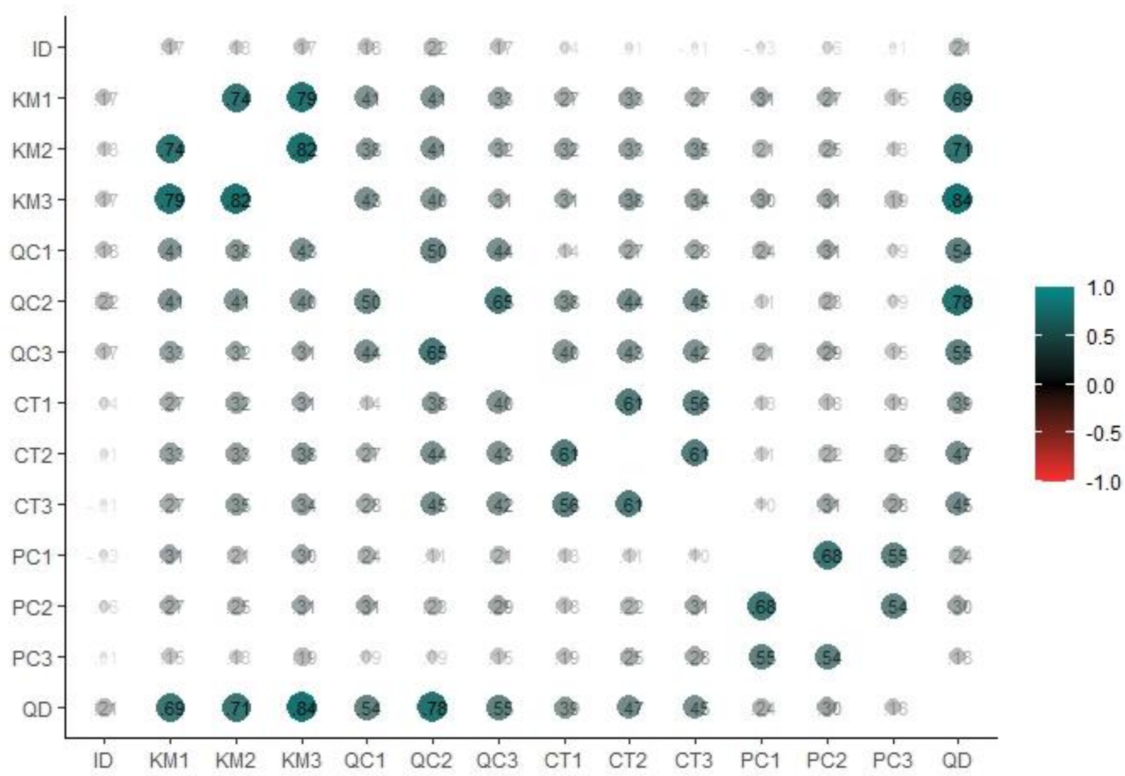
RESULTADOS.

```
> library(openxlsx) #Librería que interactúa con MSExcel
> library(corrplot) #Librería para el gráfico de correlaciones
> library(corrr) #Otra opción de librería para el cálculo y gráfico de correlaciones
> library(psych)
> library (ggcorrplot)
> library(stats) #Librería del sistema base
> library(polycor)
> library(GPArotation)
>
> # Lectura de la BDD de acuerdo a su ubicación
> archivo <- read.csv('C:/Users/viane/Desktop/ESCOM/3.-TERCER SEMESTRE/PROGRAMACION PARA LAS CIENCIAS DE DATOS/Ejercicio1.csv')
>
> # Visualización de la tabla
> archivo
   ID KM1 KM2 KM3 QC1 QC2 QC3 CT1 CT2 CT3 PC1 PC2 PC3 QD
1   1   5   5   5   5   2   1   1   3   1   4   1   3  4
2   2   3   3   3   4   5   3   4   5   4   2   2   2  4
3   3   2   2   2   2   2   1   3   3   3   4   3   5  2
4   4   4   3   3   4   3   4   4   4   4   1   1   3  3
5   5   4   4   4   2   3   4   4   4   4   3   3   5  4
6   6   1   1   1   2   5   3   5   5   5   4   3   5  3
7   7   2   4   4   2   4   2   4   4   4   3   1   4  4
8   8   2   3   3   4   3   3   4   4   4   3   4   3  3
9   9   4   5   5   4   2   5   5   4   5   4   4   5  4
10 10   2   2   2   2   2   2   2   4   2   5   3   5  2
11 11   2   1   2   4   4   4   5   4   4   4   2   4  3
12 12   4   4   4   1   3   5   3   3   3   3   3   5  4
13 13   5   5   5   2   2   3   4   4   3   5   4   5  4
14 14   5   4   5   5   3   3   3   3   4   5   5   5  4
15 15   4   3   4   5   5   4   3   5   3   2   2   2  5
16 16   3   3   4   4   3   3   1   4   3   3   3   3  4
17 17   3   5   3   4   3   2   3   4   3   3   1   4  3
18 18   3   4   2   2   4   4   4   4   2   4   1   4  3
19 19   3   4   3   5   3   3   4   4   3   4   3   3  3
20 20   3   3   3   2   4   3   3   3   5   2   3   4  4
21 21   2   1   1   2   1   1   4   3   4   4   3   4  1
22 22   2   3   3   2   2   2   3   2   3   4   3   4  4
23 23   4   4   4   2   2   2   5   4   4   4   4   4  3
24 24   4   3   4   5   5   4   3   5   3   2   2   2  5
25 25   1   1   1   1   1   1   1   1   1   3   1   3  1
26 26   5   5   5   5   5   5   5   5   5   5   5   5  5
27 27   5   5   5   5   4   4   5   4   5   3   3   4  5
28 28   4   4   4   5   4   5   5   4   5   4   4   4  4
29 29   5   5   5   5   5   5   5   5   5   3   4   3  5
30 30   2   1   1   1   1   3   5   5   3   3   3   3  1
31 31   3   4   2   2   2   2   4   3   4   3   2   3  2
32 32   2   3   2   1   3   3   5   4   3   4   3   5  3
33 33   1   1   1   1   1   2   4   4   5   1   1   3  1
34 34   4   4   3   4   4   1   4   4   4   3   2   2  4
35 35   5   5   5   5   5   1   3   5   5   5   5   5  5
36 36   1   1   1   4   2   3   2   4   4   3   3   3  2
37 37   3   3   3   1   3   3   3   3   3   1   1   1  3
38 38   3   5   5   5   5   5   4   5   5   5   5   5  5
39 39   3   3   3   4   3   3   3   3   3   3   3   3  3
40 40   3   3   3   2   3   2   3   3   1   2   1   1  3
41 41   4   4   2   5   4   5   3   4   2   3   2   5  3
42 42   1   1   2   5   2   1   5   4   4   4   4   4  3
43 43   3   3   3   4   3   3   3   5   3   5   5   5  3
44 44   4   3   4   5   2   3   3   4   2   4   3   4  3
45 45   5   4   5   1   1   3   5   1   2   5   4   5  3
46 46   4   4   5   2   3   3   5   5   1   5   1   3  4
47 47   4   3   3   4   3   4   3   4   3   5   3   3  3
48 48   1   1   3   4   3   3   3   3   3   2   3   3  3
49 49   1   1   1   1   1   5   5   5   5   3   1   5  2
50 50   1   1   1   1   3   2   3   3   3   3   3   4  2
51 51   3   3   3   4   3   3   3   1   3   3   3   1  3
52 52   3   3   3   2   3   2   5   4   3   4   3   4  3
53 53   5   5   5   4   1   1   4   1   1   5   3   2  3
54 54   3   3   3   4   3   3   3   3   3   3   3   3  3
55 55   3   3   3   4   3   3   5   5   5   2   2   3  3
56 56   3   3   3   5   4   4   4   4   4   4   4   4  4
57 57   1   1   1   2   3   5   3   3   3   3   3   5  2
58 58   1   1   1   1   1   1   1   1   1   2   1   1  1
59 59   2   2   4   4   3   3   4   4   4   4   5   4  4
60 60   5   5   5   5   2   3   4   4   2   5   5   5  4
61 61   3   3   3   4   3   3   3   4   5   4   4   4  3
62 62   1   3   3   4   3   2   3   3   3   1   2   1  3
63 63   4   4   4   5   3   4   4   5   4   5   4   4  4
64 64   2   2   2   2   2   2   2   2   2   3   3   2  2
65 65   5   5   5   4   4   3   5   4   3   3   3   5  5
66 66   3   5   2   4   4   4   4   3   2   3   4   2  3
67 67   3   5   2   4   4   4   4   3   2   3   4   2  3
68 68   5   5   5   5   5   5   5   5   5   5   5   5  5
69 69   1   1   1   5   4   3   1   2   1   2   3   3  3
70 70   4   4   4   4   4   4   4   4   4   3   4   3  4
71 71   1   1   1   4   3   1   4   3   1   4   3   5  2
 [ reached 'max' / getOption("max.print") -- omitted 188 rows ]
>
```

```
> matriz_correlaciones <- cor(archivo, use = "pairwise.complete.obs")
> matriz_correlaciones
             ID       KM1       KM2       KM3        QC1        QC2       QC3        CT1        CT2         CT3         PC1        PC2
ID   1.00000000 0.1740073 0.1333292 0.1656654 0.18470059 0.21607543 0.1738844 0.03755648 0.01284096 -0.00724202 -0.03136982 0.05595273
KM1  0.17400730 1.0000000 0.7404031 0.7934378 0.40573706 0.40950984 0.3300700 0.27172555 0.32825097  0.26620029  0.30864695 0.26711101
KM2  0.13332917 0.7404031 1.0000000 0.8164914 0.37794717 0.41408214 0.3169597 0.31938797 0.32894934  0.34834008  0.21196002 0.24705868
KM3  0.16566538 0.7934378 0.8164914 1.0000000 0.42840929 0.40120352 0.3146700 0.30723617 0.38104315  0.34461956  0.30139896 0.30513487
QC1  0.18470059 0.4057371 0.3779472 0.4284093 1.00000000 0.49957287 0.4377126 0.13823971 0.26867358  0.27563691  0.23778751 0.31103404
QC2  0.21607543 0.4095098 0.4140821 0.4012035 0.49957287 1.00000000 0.6462646 0.38416342 0.44324699  0.45450092  0.10973061 0.22926762
QC3  0.17388436 0.3300700 0.3169597 0.3146700 0.43771259 0.64626455 1.0000000 0.40006549 0.42504940  0.42062428  0.20531674 0.29213654
CT1  0.03755648 0.2717256 0.3193880 0.3072362 0.13823971 0.38416342 0.4000655 1.00000000 0.61250104  0.55781308  0.12921991 0.17643667
CT2  0.01284096 0.3282510 0.3289493 0.3810431 0.26867358 0.44324699 0.4250494 0.61250104 1.00000000  0.60915074  0.10747841 0.21824537
CT3 -0.00724202 0.2662003 0.3483401 0.3446196 0.27563691 0.45450092 0.4206243 0.55781308 0.60915074  1.00000000  0.09790339 0.30841476
PC1 -0.03136982 0.3086469 0.2119600 0.3013990 0.23778751 0.10973061 0.2053167 0.12921991 0.10747841  0.09790339  1.00000000 0.67598831
PC2  0.05595273 0.2671110 0.2470587 0.3051349 0.31103404 0.22926762 0.2921365 0.17643667 0.21824537  0.30841476  0.67598831 1.00000000
PC3  0.00702798 0.1465066 0.1331710 0.1922764 0.09303479 0.09426184 0.1510159 0.18646354 0.24966720  0.27969479  0.55058389 0.54455824
QD   0.21279153 0.6914208 0.7132198 0.8366921 0.54456733 0.77775297 0.5475128 0.39044918 0.46630163  0.45085880  0.24267229 0.30193074
            PC3        QD
ID  0.00702798 0.2127915
KM1 0.14650664 0.6914208
KM2 0.13317098 0.7132198
KM3 0.19227642 0.8366921
QC1 0.09303479 0.5445673
QC2 0.09426184 0.7777530
QC3 0.15101592 0.5475128
CT1 0.18646354 0.3904492
CT2 0.24966720 0.4663016
CT3 0.27969479 0.4508588
PC1 0.55058389 0.2426723
PC2 0.54455824 0.3019307
PC3 1.00000000 0.1827814
QD  0.18278142 1.0000000
> corrplot(cor(archivo), order = "hclust", tl.col='black', tl.cex=1)
```



```
> matriz_correlaciones <- cor(archivo, use = "pairwise.complete.obs")
> matriz_correlaciones
             ID       KM1       KM2       KM3        QC1        QC2       QC3        CT1        CT2         CT3         PC1        PC2
ID   1.00000000 0.1740073 0.1333292 0.1656654 0.18470059 0.21607543 0.1738844 0.03755648 0.01284096 -0.00724202 -0.03136982 0.05595273
KM1  0.17400730 1.0000000 0.7404031 0.7934378 0.40573706 0.40950984 0.3300700 0.27172555 0.32825097  0.26620029  0.30864695 0.26711101
KM2  0.13332917 0.7404031 1.0000000 0.8164914 0.37794717 0.41408214 0.3169597 0.31938797 0.32894934  0.34834008  0.21196002 0.24705868
KM3  0.16566538 0.7934378 0.8164914 1.0000000 0.42840929 0.40120352 0.3146700 0.30723617 0.38104315  0.34461956  0.30139896 0.30513487
QC1  0.18470059 0.4057371 0.3779472 0.4284093 1.00000000 0.49957287 0.4377126 0.13823971 0.26867358  0.27563691  0.23778751 0.31103404
QC2  0.21607543 0.4095098 0.4140821 0.4012035 0.49957287 1.00000000 0.6462646 0.38416342 0.44324699  0.45450092  0.10973061 0.22926762
QC3  0.17388436 0.3300700 0.3169597 0.3146700 0.43771259 0.64626455 1.0000000 0.40006549 0.42504940  0.42062428  0.20531674 0.29213654
CT1  0.03755648 0.2717256 0.3193880 0.3072362 0.13823971 0.38416342 0.4000655 1.00000000 0.61250104  0.55781308  0.12921991 0.17643667
CT2  0.01284096 0.3282510 0.3289493 0.3810431 0.26867358 0.44324699 0.4250494 0.61250104 1.00000000  0.60915074  0.10747841 0.21824537
CT3 -0.00724202 0.2662003 0.3483401 0.3446196 0.27563691 0.45450092 0.4206243 0.55781308 0.60915074  1.00000000  0.09790339 0.30841476
PC1 -0.03136982 0.3086469 0.2119600 0.3013990 0.23778751 0.10973061 0.2053167 0.12921991 0.10747841  0.09790339  1.00000000 0.67598831
PC2  0.05595273 0.2671110 0.2470587 0.3051349 0.31103404 0.22926762 0.2921365 0.17643667 0.21824537  0.30841476  0.67598831 1.00000000
PC3  0.00702798 0.1465066 0.1331710 0.1922764 0.09303479 0.09426184 0.1510159 0.18646354 0.24966720  0.27969479  0.55058389 0.54455824
QD   0.21279153 0.6914208 0.7132198 0.8366921 0.54456733 0.77775297 0.5475128 0.39044918 0.46630163  0.45085880  0.24267229 0.30193074
            PC3        QD
ID  0.00702798 0.2127915
KM1 0.14650664 0.6914208
KM2 0.13317098 0.7132198
KM3 0.19227642 0.8366921
QC1 0.09303479 0.5445673
QC2 0.09426184 0.7777530
QC3 0.15101592 0.5475128
CT1 0.18646354 0.3904492
CT2 0.24966720 0.4663016
CT3 0.27969479 0.4508588
PC1 0.55058389 0.2426723
PC2 0.54455824 0.3019307
PC3 1.00000000 0.1827814
QD  0.18278142 1.0000000
> corrplot(cor(archivo), order = "hclust", tl.col='black', tl.cex=1)
```



```
> # Determinante de la matriz de correlaciones de las variables ingresadas
> det(matriz_correlaciones)
[1] 0.0001041695
> # Obtener el cálculo de los estimadores del Test de Barlett y KMO
> bartlett.test(archivo)

	Bartlett test of homogeneity of variances

data:  archivo
Bartlett's K-squared = 18037, df = 13, p-value < 2.2e-16

> KMO(archivo)
Kaiser-Meyer-Olkin factor adequacy
Call: KMO(r = archivo)
Overall MSA =  0.8
MSA for each item = 
  ID  KM1  KM2  KM3  QC1  QC2  QC3  CT1  CT2  CT3  PC1  PC2  PC3   QD 
0.72 0.91 0.91 0.69 0.94 0.67 0.94 0.84 0.86 0.86 0.69 0.78 0.78 0.73 
> factanal(archivo, factors = 2, rotation = "none")

Call:
factanal(x = archivo, factors = 2, rotation = "none")

Uniquenesses:
   ID   KM1   KM2   KM3   QC1   QC2   QC3   CT1   CT2   CT3   PC1   PC2   PC3    QD 
0.947 0.359 0.322 0.005 0.677 0.048 0.550 0.810 0.739 0.745 0.908 0.889 0.961 0.055 
```

```
Loadings:
    Factor1 Factor2
ID   0.178   0.147 
KM1  0.800         
KM2  0.823         
KM3  0.994         
QC1  0.459   0.335 
QC2  0.475   0.853 
QC3  0.364   0.564 
CT1  0.333   0.281 
CT2  0.408   0.307 
CT3  0.375   0.338 
PC1  0.302         
PC2  0.316   0.103 
PC3  0.197         
QD   0.876   0.421 

               Factor1 Factor2
SS loadings      4.323   1.663
Proportion Var   0.309   0.119
Cumulative Var   0.309   0.428

Test of the hypothesis that 2 factors are sufficient.
The chi square statistic is 541.4 on 64 degrees of freedom.
The p-value is 9.57e-77 
> factanal(archivo, factors = 2, rotation = "none", scores = "regression")$scores
         Factor1       Factor2
 [1,]  1.0799382 -1.9185201504
 [2,] -0.1866083  1.4915714441
 [3,] -1.1478011 -0.8533910778
 [4,] -0.3251454 -0.2059202027
 [5,]  0.4356171 -0.4717186484
 [6,] -1.6468111  2.3001225669
 [7,]  0.4503709  0.1324449544
 [8,] -0.3369756 -0.2389319308
 [9,]  1.1009375 -1.6317852111
```

```
[10,] -1.1452401 -0.8294139761
[11,] -0.9940754  1.0679822015
[12,]  0.4300566 -0.4982687587
[13,]  1.0953150 -1.7864066675
[14,]  1.1309996 -1.0561392769
[15,]  0.5730633  1.2770306585
[16,]  0.4134793 -0.5403887974
[17,] -0.3211081 -0.3107588273
[18,] -0.9721515  0.9927908365
[19,] -0.3202374 -0.2383834668
[20,] -0.2308958  0.7399669819
[21,] -1.9419759 -1.2751925758
[22,] -0.3209188 -0.7399131232
[23,]  0.3278172 -1.5548697491
[24,]  0.5731648  1.2778750395
[25,] -1.9713775 -1.4289509484
[26,]  1.2988041  0.8210551155
[27,]  1.2503051  0.0662719116
[28,]  0.4884586  0.3441404853
[29,]  1.2937009  0.8180441229
[30,] -1.9368572 -1.1828420876
[31,] -1.1209677 -0.7798457573
[32,] -1.0244007  0.2812494428
[33,] -1.9521489 -1.2264660912
[34,] -0.2141916  0.7118224635
[35,]  1.2859529  0.6427411804
[36,] -1.8366480 -0.1524741582
[37,] -0.3489797 -0.3430806167
[38,]  1.2817694  0.8030524023
[39,] -0.3358468 -0.2857124435
[40,] -0.3515294 -0.3985404723
[41,] -0.9559093  1.0770639322
[42,] -1.0804686 -0.4069167641
[43,] -0.3248611 -0.2431118812
[44,]  0.3207178 -1.5328123809
[45,]  0.9713996 -2.8827101054
```

```
 [46,]  1.1129964 -1.1100824641
 [47,] -0.3207811 -0.2294532293
 [48,] -0.3692608 -0.2930523992
 [49,] -1.8684824 -0.7486729377
 [50,] -1.8092690  0.4349827153
 [51,] -0.3417800 -0.3212803363
 [52,] -0.3350614 -0.3102583597
 [53,]  0.9749604 -2.9426070008
 [54,] -0.3356776 -0.2843051419
 [55,] -0.3265168 -0.1903565303
 [56,] -0.2154930  0.8475174063
 [57,] -1.7989582  0.5664188038
 [58,] -1.9737061 -1.4263815638
 [59,]  0.4089637 -0.4745885151
 [60,]  1.1019803 -1.7448535677
 [61,] -0.3264516 -0.2290467094
 [62,] -0.3586614 -0.3296801977
 [63,]  0.4483173 -0.3980153148
 [64,] -1.1539193 -0.8595815134
 [65,]  1.2431090 -0.0172332986
 [66,] -0.9601531  1.0264485139
 [67,] -0.9601418  1.0265423340
 [68,]  1.2992781  0.8249955602
 [69,] -1.7023173  1.4876594876
 [70,]  0.4787976  0.2631717553
 [71,] -1.8047184  0.4294216728
 [72,]  1.2246557  0.4860991185
 [73,]  0.4197055 -0.3693025918
 [74,]  1.1354993 -0.9397286590
 [75,] -0.3502928 -0.3454004157
 [76,] -0.5011664 -2.0499926714
 [77,] -1.0352446  0.2925547816
 [78,] -1.9272758 -1.1753808942
 [79,] -0.3251839 -0.2630844826
 [80,]  0.2947581 -2.2689613825
 [81,]  0.5949957  1.3860814520
```

```
 [82,]  0.4160734 -0.0525691517
 [83,]  0.5527663  0.6116539918
 [84,] -1.8243685 -0.0978292076
 [85,]  0.4844481  0.2797769703
 [86,]  0.4772396  0.2979793706
 [87,] -1.1523503 -0.8659481913
 [88,] -0.3352938 -0.2811152581
 [89,] -1.7892046  0.5183983628
 [90,] -0.3352713 -0.2809276179
 [91,] -1.0450236  0.3230379253
 [92,]  1.2995490  0.8272472428
 [93,] -1.2180618 -1.6105930890
 [94,]  1.2995715  0.8274348830
 [95,]  1.2805526  0.8069478046
 [96,] -0.3382813 -0.2474906371
 [97,] -0.3878061 -0.9793113264
 [98,] -0.1831900  1.5143383312
 [99,]  0.4828465  0.2523537648
[100,]  0.4256116 -0.4703453532
[101,] -1.1166292 -0.6296636720
[102,]  0.4505874  0.2242556803
[103,] -1.6431315  2.4269853622
[104,]  1.2759964  0.6152920987
[105,]  1.2418122 -0.0101093419
[106,]  0.4675051  0.2045172741
[107,]  1.2712001  0.7195127335
[108,] -0.9856024  1.0004361486
[109,] -1.0403665  0.2773526112
[110,]  0.4954663  0.3725410524
[111,]  1.1115978 -1.1489756066
[112,] -0.1750734  1.5974738017
[113,]  1.2787889  0.6979679607
[114,]  1.1354603 -1.0863484599
[115,]  0.3136965 -1.6510002248
[116,]  1.2871664  0.8150848046
[117,] -0.3286398 -0.2711874145
```

```
[118,] -1.6844976  1.5605206948
[119,]  0.4535744  0.0030090287
[120,] -1.6161955  2.2291780745
[121,]  0.9928433 -2.7127626063
[122,] -0.9818005  1.0790517971
[123,]  0.6002851  1.4006168854
[124,]  0.4808832  0.2349180720
[125,]  0.4416695 -0.4726874307
[126,] -0.3308030 -0.2257321916
[127,] -0.3285270 -0.2702492134
[128,] -0.3275337 -0.2395052874
[129,] -0.3285044 -0.2700615732
[130,]  0.4220805 -0.5118820621
[131,] -0.3184135 -0.2498288682
[132,] -0.3316339 -0.2733836430
[133,] -0.3316226 -0.2732898229
[134,] -0.3347747 -0.2767995330
[135,] -1.0170150  0.2907650171
[136,] -1.9740721 -1.4179147616
[137,] -0.3267048 -0.2378231500
[138,] -0.3347295 -0.2764242526
[139,] -0.3320172 -0.2758037534
[140,] -0.3315436 -0.2726330822
[141,] -0.3308050 -0.2717015057
[142,]  0.4266607 -0.5329272286
[143,] -1.9710838 -1.4139069955
[144,]  0.4746240  0.2175784761
[145,]  0.4615106  0.1758025092
[146,] -0.3732467 -1.0778685271
[147,]  0.6077395  1.4010845917
[148,] -0.3346167 -0.2754860515
[149,]  1.1234452 -0.9938983810
[150,]  1.2897316  0.7912281127
[151,] -1.8013365  0.5496921191
[152,]  1.2902737  0.8189890077
[153,]  1.2239479 -0.0033416637
```

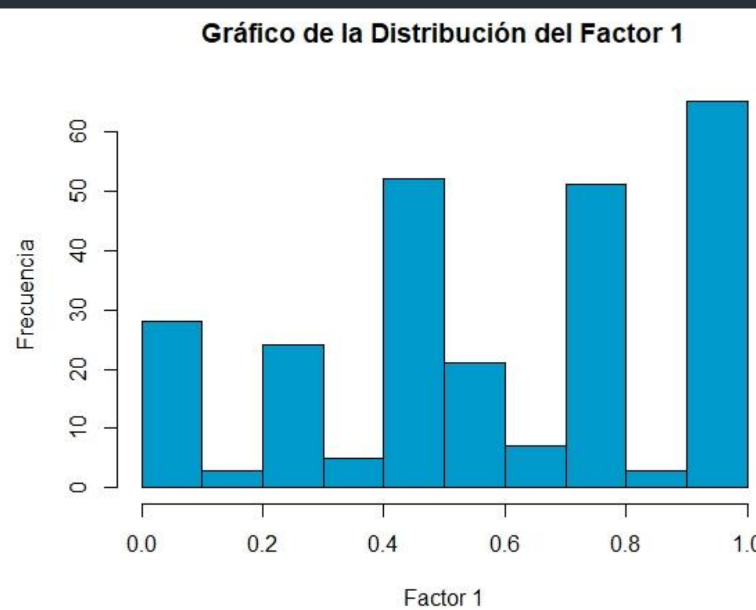
```
[154,] -0.1840847  1.5394228274
[155,] -0.2194497  0.8461374374
[156,] -0.3313630 -0.2711319604
[157,] -0.3287075 -0.2900933211
[158,]  1.3002938  0.8334393701
[159,] -1.1473983 -0.7945604801
[160,]  0.4799029  0.2537086761
[161,]  1.2757920  0.5971950487
[162,]  0.5777874  1.3757519444
[163,]  0.3163628 -1.5650079154
[164,]  1.2901442  0.8223539438
[165,]  0.4567330  0.1859135059
[166,]  0.4745837  0.2155946896
[167,] -0.3233999 -0.2120773290
[168,]  0.4554021  0.2632566441
[169,] -1.6618236  1.6589722495
[170,] -1.8028614  0.5292820654
[171,]  1.2841869  0.7370036820
[172,]  0.4209187 -0.5146349870
[173,] -0.3343345 -0.2731405487
[174,]  0.4263227 -0.5493004524
[175,] -0.2286838  0.7654034006
[176,] -0.3171134 -0.2305254931
[177,] -1.1434596 -0.7448963092
[178,] -0.2390319  0.7884910115
[179,]  0.4403002 -0.4898047242
[180,] -1.1136702 -0.7267364164
[181,]  0.4889957  0.3545803547
[182,] -1.0412504  0.3620555047
[183,] -0.2001950  0.7773924643
[184,]  1.2485183  0.0303205481
[185,] -0.3432386 -0.3280371974
[186,] -1.1568954 -0.8556535976
[187,]  0.4933397  0.2755239634
[188,]  1.2375851 -0.0147868439
[189,]  1.3006437  0.8363477935
```

```
[190,]  0.4443793 -0.4276471011
[191,] -0.2264947  0.9426556507
[192,]  0.5051821  0.3699588569
[193,]  1.2943622  0.8295160136
[194,] -0.3461784 -0.3777275952
[195,] -1.8527449 -0.3468651865
[196,] -0.3277482 -0.2637756258
[197,]  1.3007340  0.8370983544
[198,] -0.2149376  0.8720575636
[199,] -0.3440228 -0.3463928190
[200,]  1.2412759  0.5984965835
[201,] -1.6134783  2.3853365936
[202,]  1.2369554  0.0006587741
[203,]  0.4922948  0.2945325309
[204,]  1.3008130  0.8377550952
[205,]  0.3139157 -1.6130696340
[206,] -0.3181012 -0.3085423612
[207,]  1.2983542  0.8403342227
[208,] -0.3339395 -0.2698568448
[209,] -0.2100139  0.7738277865
[210,]  1.1313219 -1.0500891804
[211,] -1.9523849 -1.3252019353
[212,]  1.3009033  0.8385056561
[213,]  1.1351399 -1.0765700083
[214,] -0.9788905  1.0964988683
[215,] -0.2270874  0.7592984535
[216,] -1.9501684 -1.2908446439
[217,]  0.4705037  0.2581613260
[218,]  1.2336290  0.0374329230
[219,]  1.1335673 -1.0578758224
[220,]  1.1373402 -1.0385973789
[221,] -0.3357098 -0.2733895472
[222,]  1.3010162  0.8394438572
[223,]  1.3010274  0.8395376773
[224,]  1.2930088  0.7908068585
[225,]  0.4075309 -0.5404180839
```

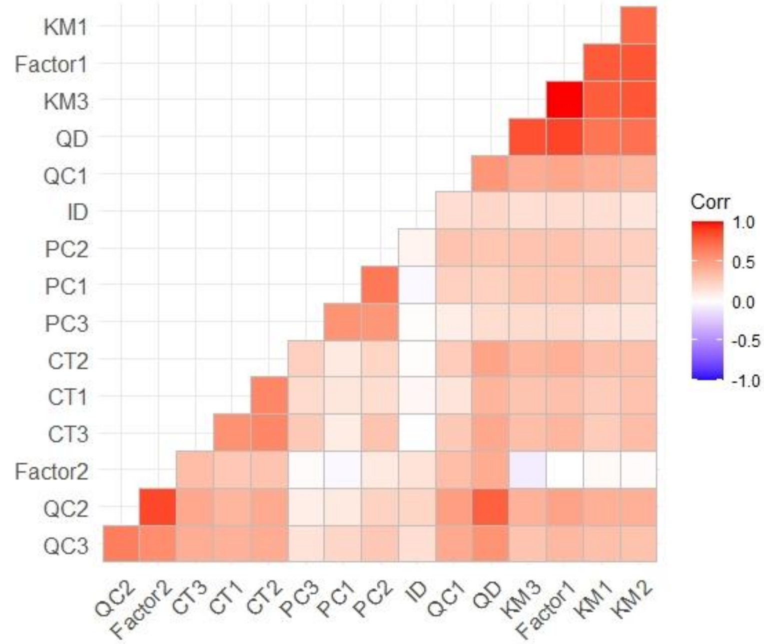
```
[226,] -1.8293575 -0.2187770317
[227,]  1.1423529 -1.0130082319
[228,] -0.3850951 -1.0836818339
[229,]  1.2409926  0.1074598054
[230,]  1.0691162 -1.7964897492
[231,]  1.3011177  0.8402882382
[232,] -0.1818633  1.5501797300
[233,]  0.5950116  1.4072523100
[234,] -1.8014776  0.5484322716
[235,]  1.2914188  0.8296004231
[236,]  1.2899172  0.7493188488
[237,] -1.6239115  2.3563139233
[238,]  0.4774264  0.2959453044
[239,]  1.3012080  0.8410387991
[240,]  0.9918309 -2.8124736041
[241,]  1.2997759  0.8395509266
[242,]  0.9702259 -2.7966240349
[243,] -0.9992655  0.9746873088
[244,] -0.1800419  1.4989737159
[245,] -0.3587449 -0.3091381136
[246,]  0.4786921  0.2888237813
[247,]  0.4738636  0.2604886292
[248,] -1.0432904  0.2377629000
[249,] -1.7990147  0.5374730381
[250,] -0.3047140 -0.1727235800
[251,]  1.3013435  0.8421646404
[252,] -0.3307419 -0.2652020809
[253,]  1.2834538  0.7385246549
[254,]  1.2911599  0.8307977538
[255,] -0.9020613  2.0588826002
[256,]  0.5904953  1.2227663373
[257,]  1.2414862  0.0070771000
[258,]  0.2673795 -2.3263379233
[259,]  0.4564292  0.2717942741
> puntuaciones <- factanal(archivo, factors = 2, rotation = "none", scores = "regression")$scores
> archivo <- cbind(archivo, puntuaciones)
```

```
> archivo$Factor1 <- round(((archivo$Factor1 - min(archivo$Factor1))/(max(archivo$Factor1) - min(archivo$Factor1))), 2)
> archivo
   ID KM1 KM2 KM3 QC1 QC2 QC3 CT1 CT2 CT3 PC1 PC2 PC3 QD Factor1     Factor2
1   1   5   5   5   5   2   1   1   3   1   4   1   3  4    0.93 -1.91852015
2   2   3   3   3   4   5   3   4   5   4   2   2   2  4    0.55  1.49157144
3   3   2   2   2   2   2   1   3   3   3   4   3   5  2    0.25 -0.85339108
4   4   4   3   3   4   3   4   4   4   4   1   1   3  3    0.50 -0.20592020
5   5   4   4   4   2   3   4   4   4   4   3   3   5  4    0.74 -0.47171865
6   6   1   1   1   2   5   3   5   5   5   4   3   5  3    0.10  2.30012257
7   7   2   4   4   2   4   2   4   4   4   3   1   4  4    0.74  0.13244495
8   8   2   3   3   4   3   3   4   4   4   3   4   3  3    0.50 -0.23893193
9   9   4   5   5   4   2   5   5   4   5   4   4   5  4    0.94 -1.63178521
10 10   2   2   2   2   2   2   2   4   2   5   3   5  2    0.25 -0.82941398
11 11   2   1   2   4   4   4   5   4   4   4   2   4  3    0.30  1.06798220
12 12   4   4   4   1   3   5   3   3   3   3   3   5  4    0.73 -0.49826876
13 13   5   5   5   2   2   3   4   4   3   5   4   5  4    0.94 -1.78640667
14 14   5   4   5   5   3   3   3   3   4   5   5   5  4    0.95 -1.05613928
15 15   4   3   4   5   5   4   3   5   3   2   2   2  5    0.78  1.27703066
16 16   3   3   4   4   3   3   1   4   3   3   3   3  4    0.73 -0.54038880
17 17   3   5   3   4   3   2   3   4   3   3   1   4  3    0.50 -0.31075883
18 18   3   4   2   2   4   4   4   4   2   4   1   4  3    0.31  0.99279084
19 19   3   4   3   5   3   3   4   4   3   4   3   3  3    0.50 -0.23838347
20 20   3   3   3   2   4   3   3   3   5   2   3   4  4    0.53  0.73996698
21 21   2   1   1   2   1   1   4   3   4   4   3   4  1    0.01 -1.27519258
22 22   2   3   3   2   2   2   3   2   3   4   3   4  4    0.50 -0.73991312
23 23   4   4   4   2   2   2   5   4   4   4   4   4  3    0.70 -1.55486975
24 24   4   3   4   5   5   4   3   5   3   2   2   2  5    0.78  1.27787504
25 25   1   1   1   1   1   1   1   1   1   3   1   3  1    0.00 -1.42895095
26 26   5   5   5   5   5   5   5   5   5   5   5   5  5    1.00  0.82105512
27 27   5   5   5   5   4   4   5   4   5   3   3   4  5    0.98  0.06627191
28 28   4   4   4   5   4   5   5   4   5   4   4   4  4    0.75  0.34414049
29 29   5   5   5   5   5   5   5   5   5   3   4   3  5    1.00  0.81804412
30 30   2   1   1   1   1   3   5   5   3   3   3   3  1    0.01 -1.18284209
31 31   3   4   2   2   2   2   4   3   4   3   2   3  2    0.26 -0.77984576
32 32   2   3   2   1   3   3   5   4   3   4   3   5  3    0.29  0.28124944
33 33   1   1   1   1   1   2   4   4   5   1   1   3  1    0.01 -1.22646609
```

```
31 31   3   4   2   2   2   2   4   3   4   3   2   3  2    0.26 -0.77984576
32 32   2   3   2   1   3   3   5   4   3   4   3   5  3    0.29  0.28124944
33 33   1   1   1   1   1   2   4   4   5   1   1   3  1    0.01 -1.22646609
34 34   4   4   3   4   4   1   4   4   4   3   2   2  4    0.54  0.71182246
35 35   5   5   5   5   5   1   3   5   5   5   5   5  5    1.00  0.64274118
36 36   1   1   1   4   2   3   2   4   4   3   3   3  2    0.04 -0.15247416
37 37   3   3   3   1   3   3   3   3   3   1   1   1  3    0.50 -0.34308062
38 38   3   5   5   5   5   5   4   5   5   5   5   5  5    0.99  0.80305240
39 39   3   3   3   4   3   3   3   3   3   3   3   3  3    0.50 -0.28571244
40 40   3   3   3   2   3   2   3   3   1   2   1   1  3    0.50 -0.39854047
41 41   4   4   2   5   4   5   3   4   2   3   2   5  3    0.31  1.07706393
42 42   1   1   2   5   2   1   5   4   4   4   4   4  3    0.27 -0.40691676
43 43   3   3   3   4   3   3   3   5   3   5   5   5  3    0.50 -0.24311188
44 44   4   3   4   5   2   3   3   4   2   4   3   4  3    0.70 -1.53281238
45 45   5   4   5   1   1   3   5   1   2   5   4   5  3    0.90 -2.88271011
46 46   4   4   5   2   3   3   5   5   1   5   1   3  4    0.94 -1.11008246
47 47   4   3   3   4   3   4   3   4   3   5   3   3  3    0.50 -0.22945323
48 48   1   1   3   4   3   3   3   3   3   2   3   3  3    0.49 -0.29305240
49 49   1   1   1   1   1   5   5   5   5   3   1   5  2    0.03 -0.74867294
50 50   1   1   1   1   3   2   3   3   3   3   3   4  2    0.05  0.43498272
51 51   3   3   3   4   3   3   3   1   3   3   3   1  3    0.50 -0.32128034
52 52   3   3   3   2   3   2   5   4   3   4   3   4  3    0.50 -0.31025836
53 53   5   5   5   4   1   1   4   1   1   5   3   2  3    0.90 -2.94260700
54 54   3   3   3   4   3   3   3   3   3   3   3   3  3    0.50 -0.28430514
55 55   3   3   3   4   3   3   5   5   5   2   2   3  3    0.50 -0.19035653
56 56   3   3   3   5   4   4   4   4   4   4   4   4  4    0.54  0.84751741
57 57   1   1   1   2   3   5   3   3   3   3   3   5  2    0.05  0.56641880
58 58   1   1   1   1   1   1   1   1   1   2   1   1  1    0.00 -1.42638156
59 59   2   2   4   4   3   3   4   4   4   4   5   4  4    0.73 -0.47458852
60 60   5   5   5   5   2   3   4   4   2   5   5   5  4    0.94 -1.74485357
61 61   3   3   3   4   3   3   3   4   5   4   4   4  3    0.50 -0.22904671
62 62   1   3   3   4   3   2   3   3   3   1   2   1  3    0.49 -0.32968020
 [ reached 'max' / getOption("max.print") -- omitted 197 rows ]
> hist(archivo$Factor1, freq = TRUE, main = "Gráfico de la Distribución del Factor 1",
+      xlab = "Factor 1", ylab = "Frecuencia", col = "#009ACD")
> 
```



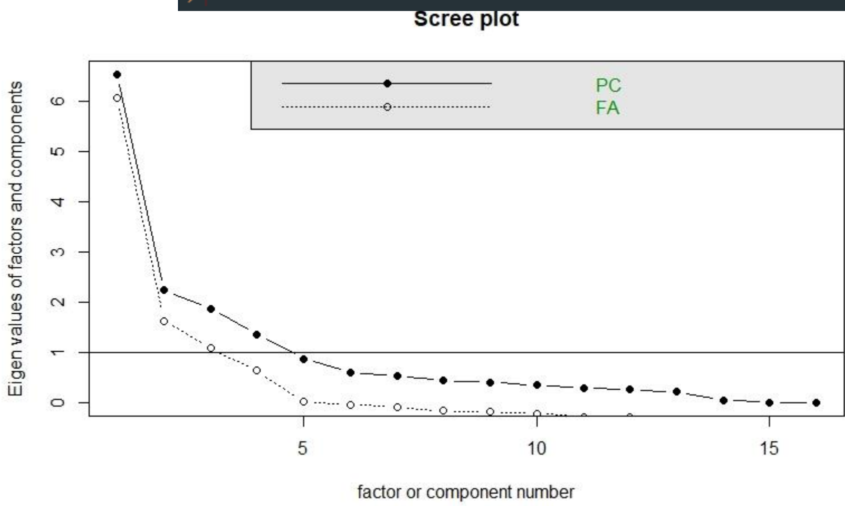
```
> #Calcular la matriz de correlación policorica
> mat_cor <- hetcor(archivo)$correlations #matriz de correlación policorica
Warning message:
In hetcor.data.frame(archivo) :
  the correlation matrix has been adjusted to make it positive-definite
> ggcorrplot(mat_cor,type="lower",hc.order = T)
>
```

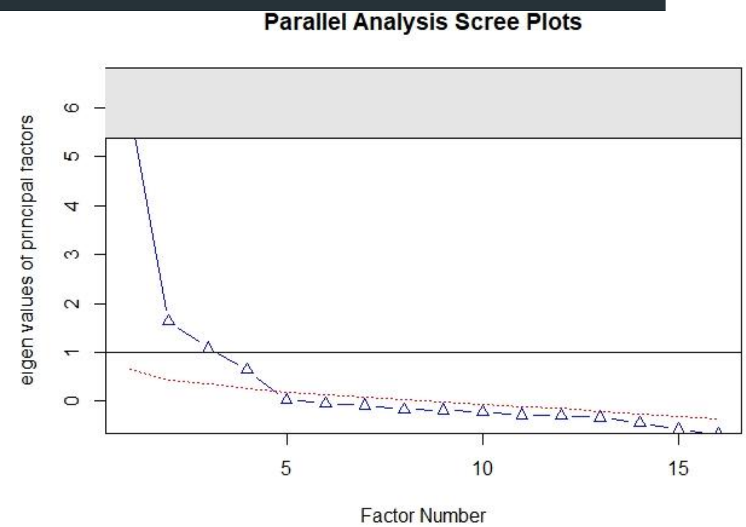



```
> # Obtener el cálculo de los estimadores del Test de Barlett y KMO con nuestra matriz de correlación polocorica
Warning message:
In x$vp : reached elapsed time limit
> cortest.bartlett(mat_cor)->p_esf
Warning message:
In cortest.bartlett(mat_cor) : n not specified, 100 used
> p_esf$p
[1] 0
>
> KMO(mat_cor)
Kaiser-Meyer-Olkin factor adequacy
Call: KMO(r = mat_cor)
Overall MSA =  0.17
MSA for each item =
    ID   KM1   KM2   KM3   QC1   QC2   QC3   CT1   CT2   CT3   PC1   PC2   PC3    QD Factor1 Factor2
  0.02  0.20  0.21  0.23  0.12  0.20  0.15  0.11  0.14  0.13  0.10  0.11  0.08  0.26    0.98    0.11
>
```

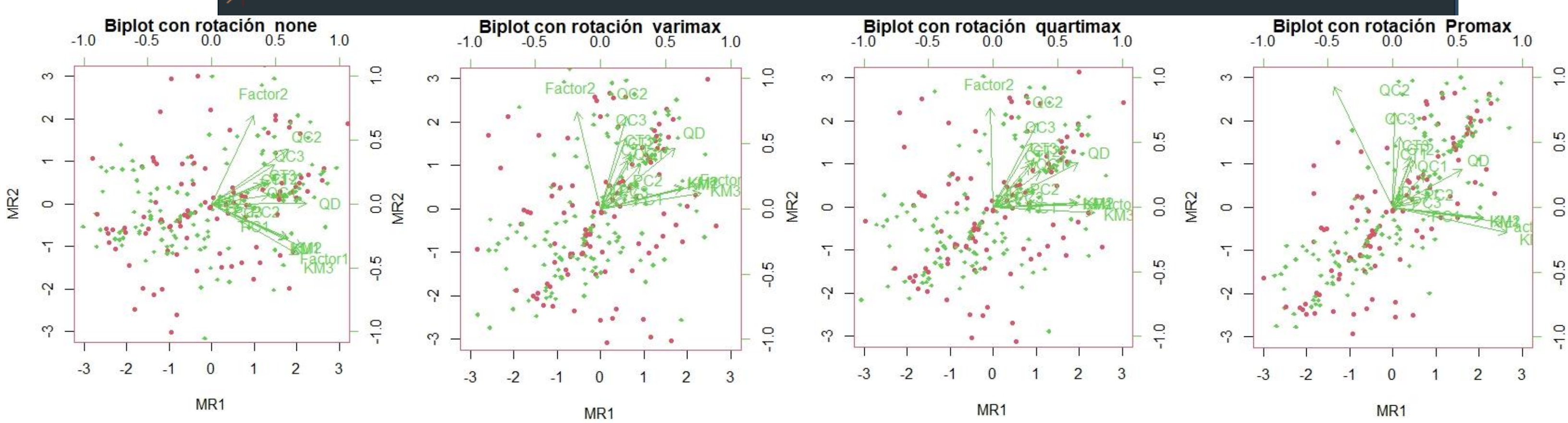
```
> # Escoger un método para extraer los factores
> modelo1<-fa(mat_cor,
+             nfactors = 3,
+             rotate = "none",
+             fm="mle") # Modelo máxima verosimilitud
> modelo2<-fa(mat_cor,
+             nfactors = 3,
+             rotate = "none",
+             fm="minres") # Modelo minimo residuo
Warning message:
In fa.stats(r = r, f = f, phi = phi, n.obs = n.obs, np.obs = np.obs,  :
  The estimated weights for the factor scores are probably incorrect.  Try a different factor score estimation method.
> sort(modelo1$communality,decreasing = T)->c1
> sort(modelo2$communality,decreasing = T)->c2
> head(cbind(c1,c2))
               c1        c2
Factor1 0.9976516 0.9961839
KM3     0.9975269 0.9898527
Factor2 0.9957851 0.9105191
QC2     0.9905528 0.8982413
QD      0.9491866 0.8873357
PC2     0.6999227 0.7085193
> #Se comparan las unicidades
> sort(modelo1$uniquenesses,decreasing = T)->u1
> sort(modelo2$uniquenesses,decreasing = T)->u2
> head(cbind(u1,u2))
           u1        u2
ID  0.9418266 0.9477688
CT1 0.7956501 0.7122305
CT2 0.7230376 0.6795596
CT3 0.7103276 0.6299833
QC1 0.6584258 0.6183254
PC3 0.5407078 0.5114946
>
```

```
> # Determinar el número de factores
> scree(mat_cor)
Warning message:
In fa.stats(r = r, f = f, phi = phi, n.obs = n.obs, np.obs = np.obs,  :
  The estimated weights for the factor scores are probably incorrect.  Try a different factor score estimation method.
> fa.parallel(mat_cor,n.obs=200,fa="fa",fm="minres")
Parallel analysis suggests that the number of factors =  4  and the number of components =  NA
Warning message:
In fa.stats(r = r, f = f, phi = phi, n.obs = n.obs, np.obs = np.obs,  :
  The estimated weights for the factor scores are probably incorrect.  Try a different factor score estimation method.
>
```

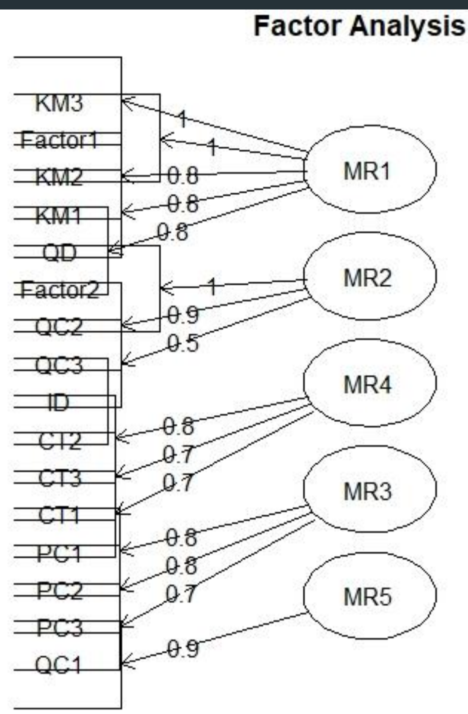





```
> # Rotar la matriz
> rot<-c("none", "varimax", "quartimax","Promax")
> bi_mod<-function(tipo){
+   biplot.psych(fa(archivo,nfactors = 2,fm="minres",rotate = tipo),main = paste("Biplot con rotación ",tipo),
+               col=c(2,3,4),pch = c(21,18),group = bfi[,"gender"])
+ }
> sapply(rot,bi_mod)
In factor.scores, the correlation matrix is singular, an approximation is used
In factor.scores, the correlation matrix is singular, an approximation is used
In factor.scores, the correlation matrix is singular, an approximation is used
In factor.scores, the correlation matrix is singular, an approximation is used
$none
NULL

$varimax
NULL

$quartimax
NULL

$Promax
NULL

There were 24 warnings (use warnings() to see them)
>
```



```
> # Interpretación
> modelo_varimax<-fa(mat_cor,nfactors = 5,rotate = "varimax",
+                     fa="minres")
Warning messages:
1: In fa.stats(r = r, f = f, phi = phi, n.obs = n.obs, np.obs = np.obs,  :
  The estimated weights for the factor scores are probably incorrect.  Try a different factor score estimation method.
2: In fac(r = r, nfactors = nfactors, n.obs = n.obs, rotate = rotate,  :
  An ultra-Heywood case was detected.  Examine the results carefully
> fa.diagram(modelo_varimax)
>
```



Interpretación de los datos:

Ante el análisis factorial completo que se realizó con cada instrucción que vimos anteriormente nos podemos enterar, percatar que todos los datos que teníamos en nuestro poder desde un inicio cumplieron el objetivo de reducirse a más de la mitad de los datos, esto quedando diferentes variables que quedan independientemente entre todos.

La utilización de las gráficas facilitó la interpretación de nuestros datos antes de analizarnos y después de hacerlo.

Ejercicio 2

1.- Elige algún dataset de tu interés

2.- Aplica el Análisis Factorial

Consideraciones:

Se debe considerar que se está utilizando un dataset (el mismo de la practica pasada) que tiene los datos de hospitalizaciones totales de la zona metropolitana de México, esto considerando del Estado de México y CDMX, tomando en cuenta, total de hospitalizaciones, hospitalizaciones de la cdmx, del edomex, así como también sus pacientes que están en una gravedad mayor, es decir, están intubados con oxigeno artificial, por SARS-COV2

Procedimiento:

```
#Ejercicio 2 Practica 9
# Análisis Factorial
# Dataset de hospitalizaciones por SARS-COV2
# Vianey Maravilla Pérez 3AM1

#Definimos las librerias a utilizar durante todo el proceso
library(openxlsx) #Librería que interactúa con MSExcel
library(corrplot) #Librería para el gráfico de correlaciones
library(corrr) #Otra opción de librería para el cálculo y gráfico de correlaciones
library(psych)
library (ggcorrplot)
library(stats) #Librería del sistema base
library(polycor)
library(GPArotation)

# Lectura de la BDD de acuerdo a su ubicación
hospi <- read.csv('C:/Users/viane/Desktop/ESCOM/3.-TERCER SEMESTRE/PROGRAMACION PARA LAS CIENCIAS DE DATOS/Ejercicio2P9.csv')

# Visualización de la tabla
hospi

# Obtener matriz de correlación
matriz_correlaciones <- cor(hospi, use = "pairwise.complete.obs")
matriz_correlaciones

# Obtenermos la gráfica de las correlaciones
corrplot(cor(hospi), order = "hclust", tl.col='black', tl.cex=1)

# Calculo de un objeto de correlaciones
hospi_correlaciones <- correlate(hospi)

# Obtener la gráfica de las correlaciones
rplot(hospi_correlaciones, legend = TRUE, colours = c("firebrick1", "black", "darkcyan"), print_cor = TRUE)

# Determinante de la matriz de correlaciones de las variables ingresadas
det(matriz_correlaciones)

```

```
bartlett.test(hosp)

KMO(hosp)

factanal(hosp, factors = 2, rotation = "none")

factanal(hosp, factors = 2, rotation = "none", scores = "regression")$scores

puntuaciones <- factanal(hosp, factors = 2, rotation = "none", scores = "regression")$scores
archivo <- cbind(hosp, puntuaciones)
archivo$Factor1 <- round(((hosp$Factor1 - min(hosp$Factor1))/(max(hosp$Factor1) - min(hosp$Factor1))), 2)
hosp

hist(hosp$Factor1, freq = TRUE, main = "Gráfico de la Distribución del Factor 1",
     xlab = "Factor 1", ylab = "Frecuencia", col = "#009ACD")

#Calcular la matriz de correlación policorica
mat_cor <- hetcor(hosp)$correlations #matriz de correlación policorica
ggcorrplot(mat_cor,type="lower",hc.order = T)

# Obtener el cálculo de los estimadores del Test de Barlett y KMO con nuestra matriz de correlación polocorica
cortest.bartlett(mat_cor)->p_esf
p_esf$p

KMO(mat_cor)

# Pendiente
# Escoger un método para extraer los factores
# minres: minimo residuo /
# mle: maxima verosimilitud /
# paf: método de ejes principales /
# alpah: alfa /
# minchi: minimos cuadrados /
# minrak : rango minimo /

modelo1<-fa(mat_cor,
           nfactors = 3,
```

```
            nfactors = 3,
            rotate = "none",
            fm="mle") # Modelo máxima verosimilitud

modelo2<-fa(mat_cor,
            nfactors = 3,
            rotate = "none",
            fm="minres") # Modelo minimo residuo


# Se comparan las comunalidades
sort(modelo1$communality,decreasing = T)->c1
sort(modelo2$communality,decreasing = T)->c2
head(cbind(c1,c2))

#Se comparan las unicidades
sort(modelo1$uniquenesses,decreasing = T)->u1
sort(modelo2$uniquenesses,decreasing = T)->u2
head(cbind(u1,u2))

# Termina pendiente

# Determinar el número de factores
scree(mat_cor)
fa.parallel(mat_cor,n.obs=200,fa="fa",fm="minres")


# Rotar la matriz
rot<-c("none", "varimax", "quartimax","Promax")
bi_mod<-function(tipo){
  biplot.psych(fa(hosp,nfactors = 2,fm="minres",rotate = tipo),main = paste("Biplot con rotación ",tipo),
               col=c(2,3,4),pch = c(21,18),group = bfi[,"gender"])
}
sapply(rot,bi_mod)
```


```
            nfactors = 3,
            rotate = "none",
            fm="minres") # Modelo minimo residuo


# Se comparan las comunalidades
sort(modelo1$communality,decreasing = T)->c1
sort(modelo2$communality,decreasing = T)->c2
head(cbind(c1,c2))

#Se comparan las unicidades
sort(modelo1$uniquenesses,decreasing = T)->u1
sort(modelo2$uniquenesses,decreasing = T)->u2
head(cbind(u1,u2))

# Termina pendiente

# Determinar el número de factores
scree(mat_cor)
fa.parallel(mat_cor,n.obs=200,fa="fa",fm="minres")


# Rotar la matriz
rot<-c("none", "varimax", "quartimax","Promax")
bi_mod<-function(tipo){
  biplot.psych(fa(hosp,nfactors = 2,fm="minres",rotate = tipo),main = paste("Biplot con rotación ",tipo),
               col=c(2,3,4),pch = c(21,18),group = bfi[,"gender"])
}
sapply(rot,bi_mod)
```


```
# Interpretación
modelo_varimax<-fa(mat_cor,nfactors = 5,rotate = "varimax",
                   fa="minres")
fa.diagram(modelo_varimax)
```

Resultados:

Se obtuvieron los siguientes resultados en el proceso

```
> #Ejercicio 2 Practica 9
> # Análisis Factorial
> # Dataset de hospitalizaciones por SARS-COV2
> # Vianey Maravilla Pérez 3AM1
>
> #Definimos las librerias a utilizar durante todo el proceso
> library(openxlsx) #Librería que interactúa con MSExcel
> library(corrplot) #Librería para el gráfico de correlaciones
> library(corrr) #Otra opción de librería para el cálculo y gráfico de correlaciones
> library(psych)
> library (ggcorrplot)
> library(stats) #Librería del sistema base
> library(polycor)
> library(GPArotation)
>
> # Lectura de la BDD de acuerdo a su ubicación
> hospi <- read.csv('C:/Users/viane/Desktop/ESCOM/3.-TERCER SEMESTRE/PROGRAMACION PARA LAS CIENCIAS DE DATOS/Ejercicio2P9.csv')
>
> # Visualización de la tabla
> hospi
    X HospTot Hospcdmx HopsEdo Intucdmx Intuedo
1   1       7        6       1        1       1
2   2       5        5       6        6       5
3   3       5        6       5        7       5
4   4      10        8       2        6       6
5   5       7        6       6        7       6
6   6       4        4       6        7       6
7   7       5        5       5        5       6
8   8       5        6       5        5       5
9   9       6        5       7        6       6
10 10       6        5       6        6       6
11 11       6        7       5        6       5
12 12       5        5       4        5       4
13 13       6        6       6        6       5
14 14       8        7       8        8       8
```

```
12 12       5        5       4        5       4
13 13       6        6       6        6       5
14 14       8        7       8        8       8
15 15       6        7       5        6       6
16 16       4        3       4        4       4
17 17       6        4       7        8       7
18 18       6        6       7        7       7
19 19       6        5       4        4       4
20 20       7        7       6        7       6
21 21       4        4       6        7       6
22 22       5        5       5        5       6
23 23       5        6       5        5       5
24 24       6        5       7        6       6
25 25       6        5       6        6       6
>
> # Obtener matriz de correlación
> matriz_correlaciones <- cor(hospi, use = "pairwise.complete.obs")
> matriz_correlaciones
                 X     HospTot   Hospcdmx    HopsEdo   Intucdmx   Intuedo
X         1.0000000 -0.17681614 -0.2051541  0.3649343 0.15480173 0.3115224
HospTot  -0.1768161  1.00000000  0.7415849 -0.1751338 0.09459838 0.1655594
Hospcdmx -0.2051541  0.74158490  1.0000000 -0.1779714 0.10026518 0.1018893
HopsEdo   0.3649343 -0.17513377 -0.1779714  1.0000000 0.77920059 0.7830886
Intucdmx  0.1548017  0.09459838  0.1002652  0.7792006 1.00000000 0.8785392
Intuedo   0.3115224  0.16555941  0.1018893  0.7830886 0.87853924 1.0000000
>
> # Obtenermos la gráfica de las correlaciones
> corrplot(cor(hospi), order = "hclust", tl.col='black', tl.cex=1)
>
> # Calculo de un objeto de correlaciones
> hospi_correlaciones <- correlate(hospi)

Correlation method: 'pearson'
Missing treated using: 'pairwise.complete.obs'
```

```
> # Obtener la gráfica de las correlaciones
> rplot(hospi_correlaciones, legend = TRUE, colours = c("firebrick1", "black", "darkcyan"), print_cor = TRUE)
Don't know how to automatically pick scale for object of type noquote. Defaulting to continuous.
Don't know how to automatically pick scale for object of type noquote. Defaulting to continuous.
>
> # Determinante de la matriz de correlaciones de las variables ingresadas
> det(matriz_correlaciones)
[1] 0.01860279
>
> # Obtener el cálculo de los estimadores del Test de Barlett y KMO
> bartlett.test(hospi)

        Bartlett test of homogeneity of variances

data:  hospi
Bartlett's K-squared = 168.24, df = 5, p-value < 2.2e-16

>
> KMO(hospi)
Kaiser-Meyer-Olkin factor adequacy
Call: KMO(r = hospi)
Overall MSA =  0.65
MSA for each item =
       X  HospTot Hospcdmx  HopsEdo Intucdmx  Intuedo
    0.56     0.51     0.57     0.80     0.66     0.67
>
> factanal(hospi, factors = 2, rotation = "none")

Call:
factanal(x = hospi, factors = 2, rotation = "none")

Uniquenesses:
       X  HospTot Hospcdmx  HopsEdo Intucdmx  Intuedo
   0.849    0.172    0.337    0.199    0.157    0.084
```

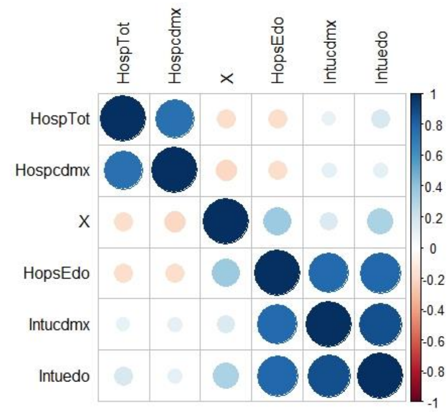
```
Loadings:
         Factor1 Factor2
X         0.294  -0.254
HospTot   0.119   0.902
Hospcdmx          0.810
HopsEdo   0.841  -0.307
Intucdmx  0.918
Intuedo   0.956

               Factor1 Factor2
SS loadings      2.570   1.631
Proportion Var   0.428   0.272
Cumulative Var   0.428   0.700

Test of the hypothesis that 2 factors are sufficient.
The chi square statistic is 3.73 on 4 degrees of freedom.
The p-value is 0.443
>
> factanal(hospi, factors = 2, rotation = "none", scores = "regression")$scores
            Factor1       Factor2
 [1,] -3.3095086868  1.0226143923
 [2,] -0.1332420664 -0.5852703970
 [3,] -0.0606693290 -0.2176096394
 [4,] -0.0848543938  3.1122439831
 [5,]  0.5321255706  0.6728631483
 [6,]  0.4395318620 -1.2858611910
 [7,] -0.0339847109 -0.4370189089
 [8,] -0.4302818309 -0.2439309856
 [9,]  0.4456217496 -0.2031294783
[10,]  0.3180709607 -0.0876907748
[11,] -0.1983019701  0.4771615779
[12,] -0.9690796960 -0.4480558107
[13,] -0.0737448465  0.0955318086
[14,]  1.8536390818  1.2303618707
[15,]  0.2189140662  0.5108033607
```

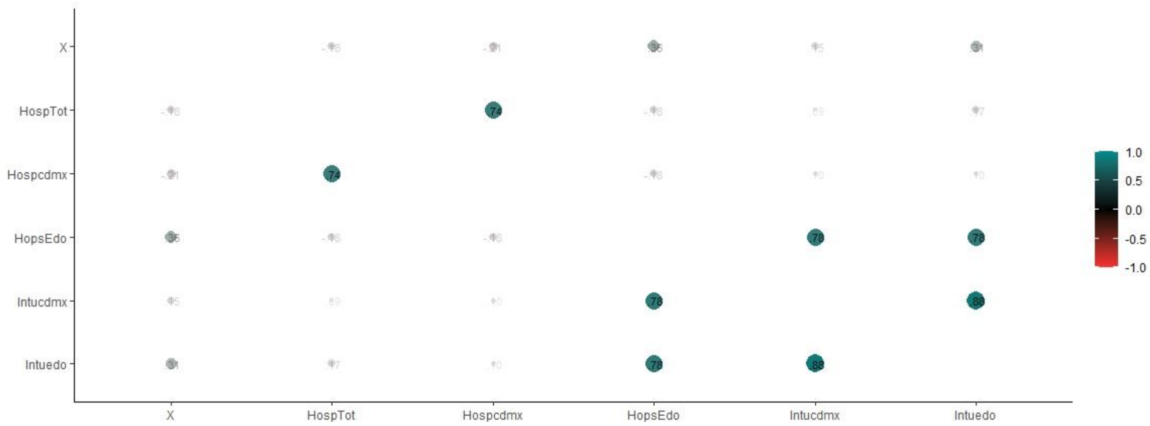
```
[13,] -0.0737448465  0.0955318086
[14,]  1.8536390818  1.2303618707
[15,]  0.2189140662  0.5108033607
[16,] -1.1950905966 -1.4550701199
[17,]  1.2429005608 -0.4389534347
[18,]  1.0741645942  0.0580710090
[19,] -1.1186686279  0.0002550205
[20,]  0.5754510506  0.8502661549
[21,]  0.4731417076 -1.3598512404
[22,] -0.0003748653 -0.5110089583
[23,] -0.3966719853 -0.3179210350
[24,]  0.4792315952 -0.2771195277
[25,]  0.3516808063 -0.1616808242
>
> puntuaciones <- factanal(hospi, factors = 2, rotation = "none", scores = "regression")$scores
> hospi <- cbind(hospi, puntuaciones)
> hospi$Factor1 <- round(((hospi$Factor1 - min(hospi$Factor1))/(max(hospi$Factor1) - min(hospi$Factor1))), 2)
> hospi
    X HospTot Hospcdmx HopsEdo Intucdmx Intuedo Factor1       Factor2
1   1       7        6       1        1       1    0.00  1.0226143923
2   2       5        5       6        6       5    0.62 -0.5852703970
3   3       5        6       5        7       5    0.63 -0.2176096394
4   4      10        8       2        6       6    0.62  3.1122439831
5   5       7        6       6        7       6    0.74  0.6728631483
6   6       4        4       6        7       6    0.73 -1.2858611910
7   7       5        5       5        5       6    0.63 -0.4370189089
8   8       5        6       5        5       5    0.56 -0.2439309856
9   9       6        5       7        6       6    0.73 -0.2031294783
10 10       6        5       6        6       6    0.70 -0.0876907748
11 11       6        7       5        6       5    0.60  0.4771615779
12 12       5        5       4        5       4    0.45 -0.4480558107
13 13       6        6       6        6       5    0.63  0.0955318086
14 14       8        7       8        8       8    1.00  1.2303618707
15 15       6        7       5        6       6    0.68  0.5108033607
16 16       4        3       4        4       4    0.41 -1.4550701199
```

```
10 10       6        5       6        6       6    0.70 -0.0876907748
11 11       6        7       5        6       5    0.60  0.4771615779
12 12       5        5       4        5       4    0.45 -0.4480558107
13 13       6        6       6        6       5    0.63  0.0955318086
14 14       8        7       8        8       8    1.00  1.2303618707
15 15       6        7       5        6       6    0.68  0.5108033607
16 16       4        3       4        4       4    0.41 -1.4550701199
17 17       6        4       7        8       7    0.88 -0.4389534347
18 18       6        6       7        7       7    0.85  0.0580710090
19 19       6        5       4        4       4    0.42  0.0002550205
20 20       7        7       6        7       6    0.75  0.8502661549
21 21       4        4       6        7       6    0.73 -1.3598512404
22 22       5        5       5        5       6    0.64 -0.5110089583
23 23       5        6       5        5       5    0.56 -0.3179210350
24 24       6        5       7        6       6    0.73 -0.2771195277
25 25       6        5       6        6       6    0.71 -0.1616808242
>
> hist(hospi$Factor1, freq = TRUE, main = "Gráfico de la Distribución del Factor 1",
+      xlab = "Factor 1", ylab = "Frecuencia", col = "#009ACD")
>
> #Calcular la matriz de correlación policorica
> mat_cor <- hetcor(hospi)$correlations #matriz de correlación policorica
Warning message:
In hetcor.data.frame(hospi) :
  the correlation matrix has been adjusted to make it positive-definite
> ggcorrplot(mat_cor,type="lower",hc.order = T)
>
> # Obtener el cálculo de los estimadores del Test de Barlett y KMO con nuestra matriz de correlación polocorica
> cortest.bartlett(mat_cor)->p_esf
Warning message:
In cortest.bartlett(mat_cor) : n not specified, 100 used
> p_esf$p
[1] 0
>
> KMO(mat_cor)
```

```
[1] 0
>
> KMO(mat_cor)
Kaiser-Meyer-Olkin factor adequacy
Call: KMO(r = mat_cor)
Overall MSA =  0.32
MSA for each item =
       X  HospTot Hospcdmx  HopsEdo Intucdmx  Intuedo  Factor1  Factor2
    0.09     0.24     0.22     0.31     0.64     0.33     0.74     0.27
>
> # Pendiente
> # Escoger un método para extraer los factores
> # minres: minimo residuo /
> # mle: maxima verosimilitud /
> # paf: método de ejes principales /
> # alpah: alfa /
> # minchi: minimos cuadrados /
> # minrak : rango minimo /
>
> modelo1<-fa(mat_cor,
+             nfactors = 3,
+             rotate = "none",
+             fm="mle") # Modelo máxima verosimilitud
>
> modelo2<-fa(mat_cor,
+             nfactors = 3,
+             rotate = "none",
+             fm="minres") # Modelo minimo residuo
Warning messages:
1: In fa.stats(r = r, f = f, phi = phi, n.obs = n.obs, np.obs = np.obs,  :
  The estimated weights for the factor scores are probably incorrect.  Try a different factor score estimation method.
2: In fac(r = r, nfactors = nfactors, n.obs = n.obs, rotate = rotate,  :
  An ultra-Heywood case was detected.  Examine the results carefully
>
>
```

```
  An ultra-Heywood case was detected.  Examine the results carefully
>
>
> # Se comparan las comunalidades
> sort(modelo1$communality,decreasing = T)->c1
> sort(modelo2$communality,decreasing = T)->c2
> head(cbind(c1,c2))
                c1        c2
Factor1  0.9960170 1.0392072
Factor2  0.9953583 1.0125903
Intucdmx 0.9950678 0.9674924
Intuedo  0.9786363 0.9280627
HospTot  0.9446212 0.8719051
HopsEdo  0.8357561 0.8161935
>
> #Se comparan las unicidades
> sort(modelo1$uniquenesses,decreasing = T)->u1
> sort(modelo2$uniquenesses,decreasing = T)->u2
> head(cbind(u1,u2))
                  u1         u2
X        0.705459374 0.66186292
Hospcdmx 0.250900284 0.31349456
HopsEdo  0.164243907 0.18380654
HospTot  0.055378781 0.12809489
Intuedo  0.021363658 0.07193732
Intucdmx 0.004932196 0.03250760
>
> # Termina pendiente
>
> # Determinar el número de factores
> scree(mat_cor)
Warning messages:
1: In fa.stats(r = r, f = f, phi = phi, n.obs = n.obs, np.obs = np.obs,  :
  The estimated weights for the factor scores are probably incorrect.  Try a different factor score estimation method.
2: In fac(r = r, nfactors = nfactors, n.obs = n.obs, rotate = rotate,  :
```


HospTot
Hospcdmx
X
HopsEdo
Intucdmx
Intuedo
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1


X
HospTot
Hospcdmx
HopsEdo
Intucdmx
Intuedo
1.0
0.5
0.0
-0.5
-1.0

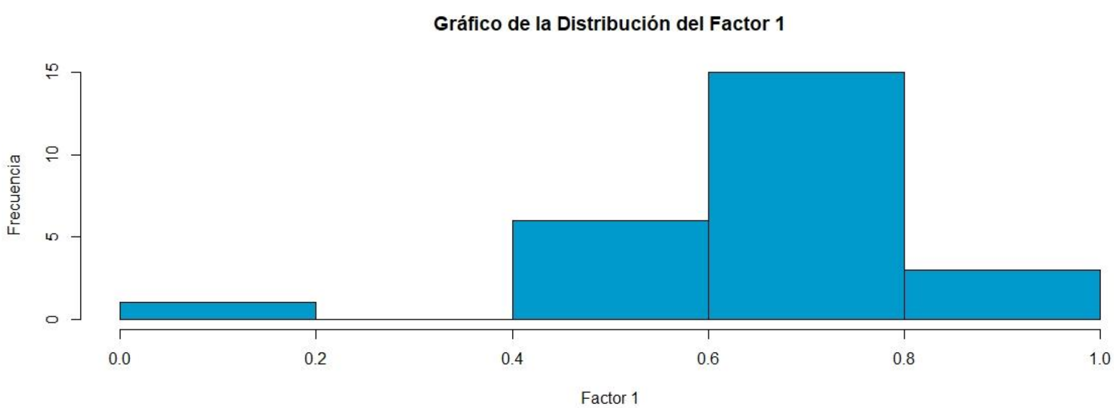
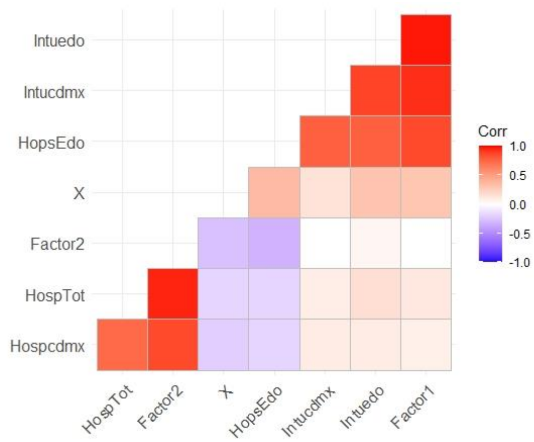
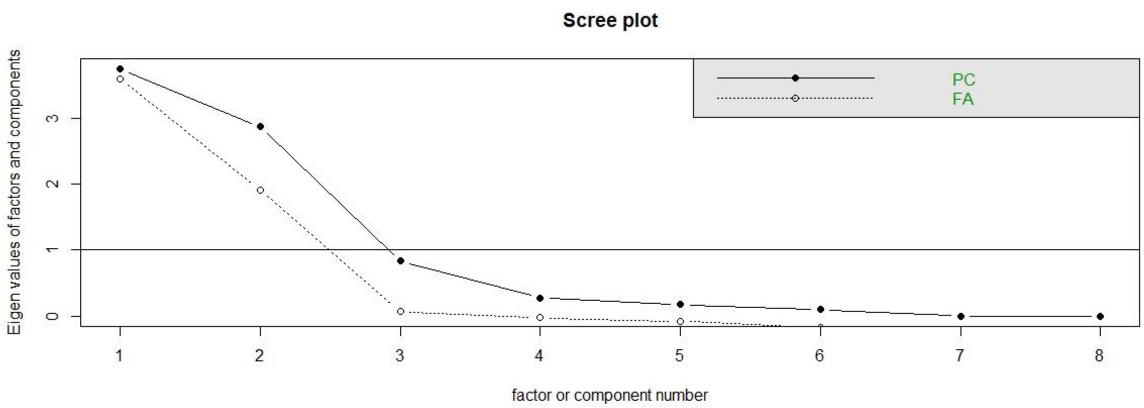
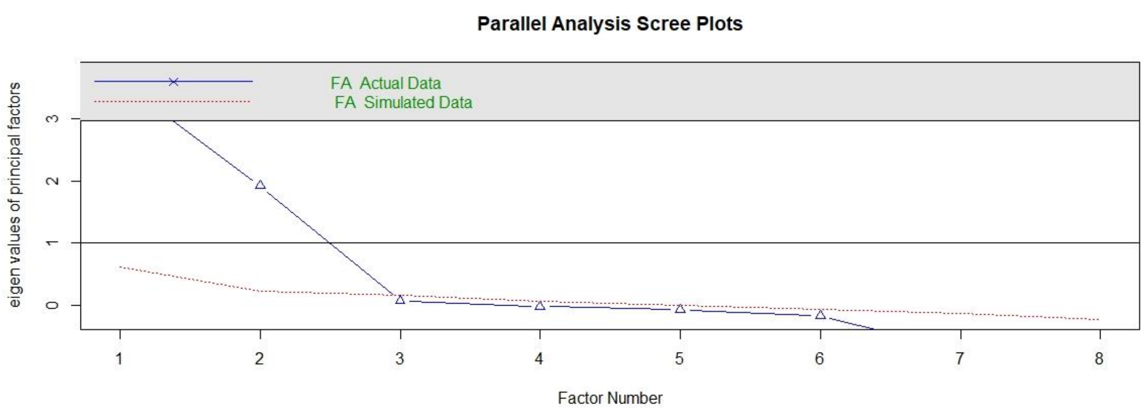


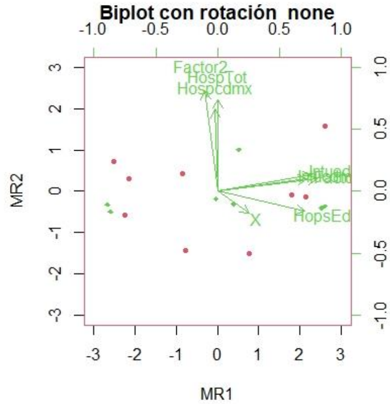
Gráfico de la Distribución del Factor 1
Frecuencia
0
5
10
15
0.0
0.2
0.4
0.6
0.8
1.0
Factor 1


Intuedo
Intucdmx
HopsEdo
X
Factor2
HospTot
Hospcdmx
HospTot
Factor2
X
HopsEdo
Intucdmx
Intuedo
Factor1
Corr
1.0
0.5
0.0
-0.5
-1.0

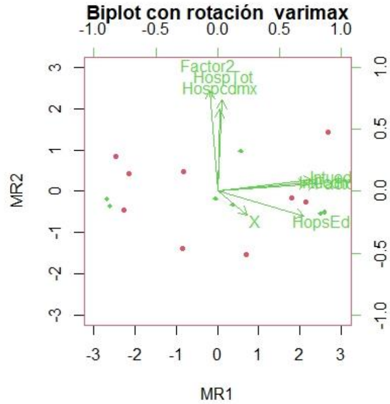

Scree plot
PC
FA
Eigen values of factors and components
factor or component number


Parallel Analysis Scree Plots
FA Actual Data
FA Simulated Data
eigen values of principal factors
Factor Number

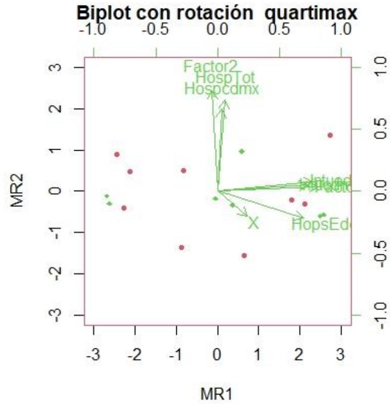
Biplot con rotación none

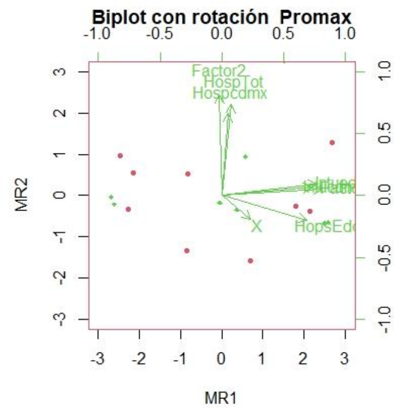


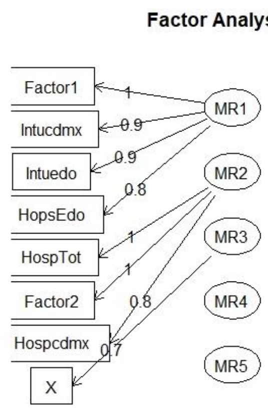
Biplot con rotación varimax



Biplot con rotación quartimax







Interpretación de los datos: *Como podemos darnos cuenta, nuestros datos redujeron aproximadamente más o si no es que ¾ de lo original, dándonos a entender que solo se puede quedar lo más vital e importante para nuestras hospitalizaciones*

Conclusiones: *Al termino de esta práctica se logró emplear el análisis factorial en una seria o conjunto de datos de manera eficiente, además se aprendieron nuevos métodos de rotación y sobre todo de extracción de datos, esto con el fin de tener en mejor organización y manipulación nuestros datos*