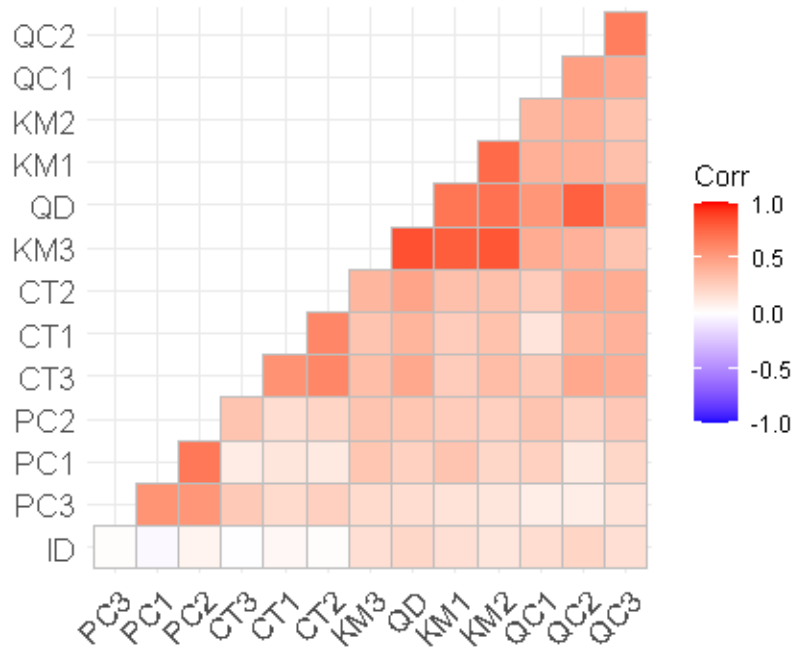


## Practica número 9

### Ejercicio número 1:

Para este ejercicio vamos a utilizar el data set proporcionada por la profesora del link: <https://raw.githubusercontent.com/housecricket/data/main/efa/sample1.csv> por lo que utilizamos el esqueleto del ejercicio visto en clase el día de hoy, por lo que descargamos el archivo con la función de rstudio y lo cargamos al entorno.

Posterior a esto sacamos la correlación de forma grafica



Lo primero es pasarlo por el test de Barlet y el test de KMO el cual arroja los siguientes resultados:

Barlet:

```
> p_esf$p
[1] 5.708066e-125
```

KMO:

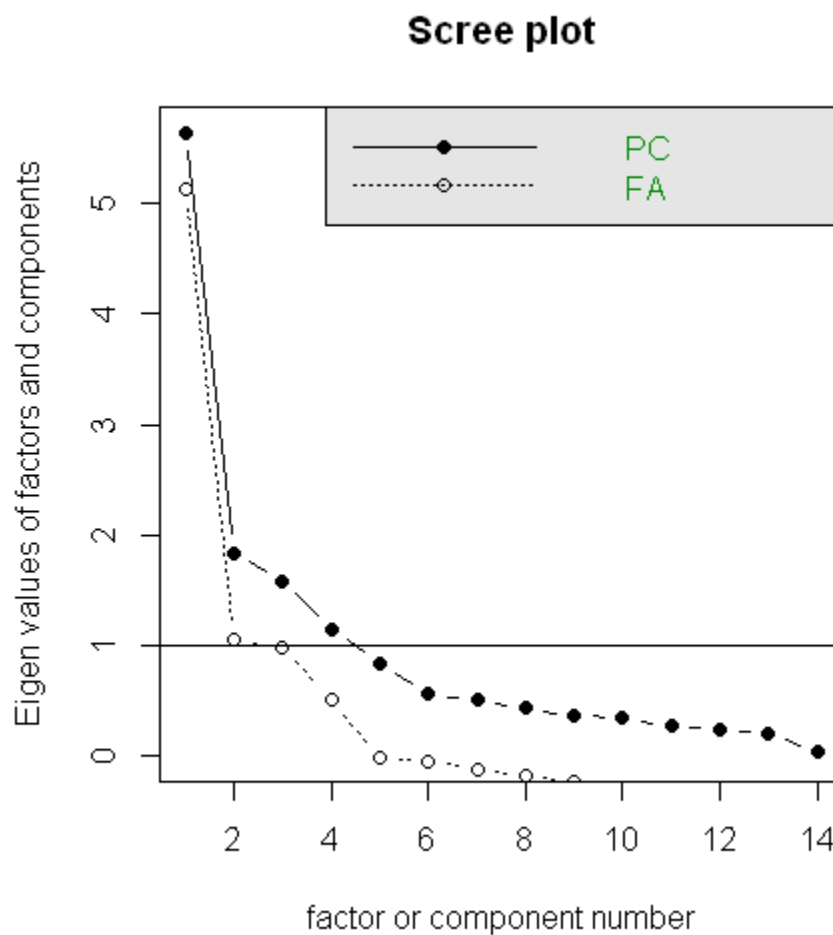
```
> KMO(mat_cor)
Kaiser-Meyer-Olkin factor adequacy
Call: KMO(r = mat_cor)
Overall MSA = 0.8
MSA for each item =
  ID  KM1  KM2  KM3  QC1  QC2  QC3  CT1  CT2  CT3  PC1  PC2  PC3  QD
0.72 0.91 0.91 0.69 0.94 0.67 0.94 0.84 0.86 0.86 0.69 0.78 0.78 0.73
```

Por lo que después del test de KMO nos dice que si podemos continuar con el análisis factorial

Después obtenemos dos factores para realizar la comparación:

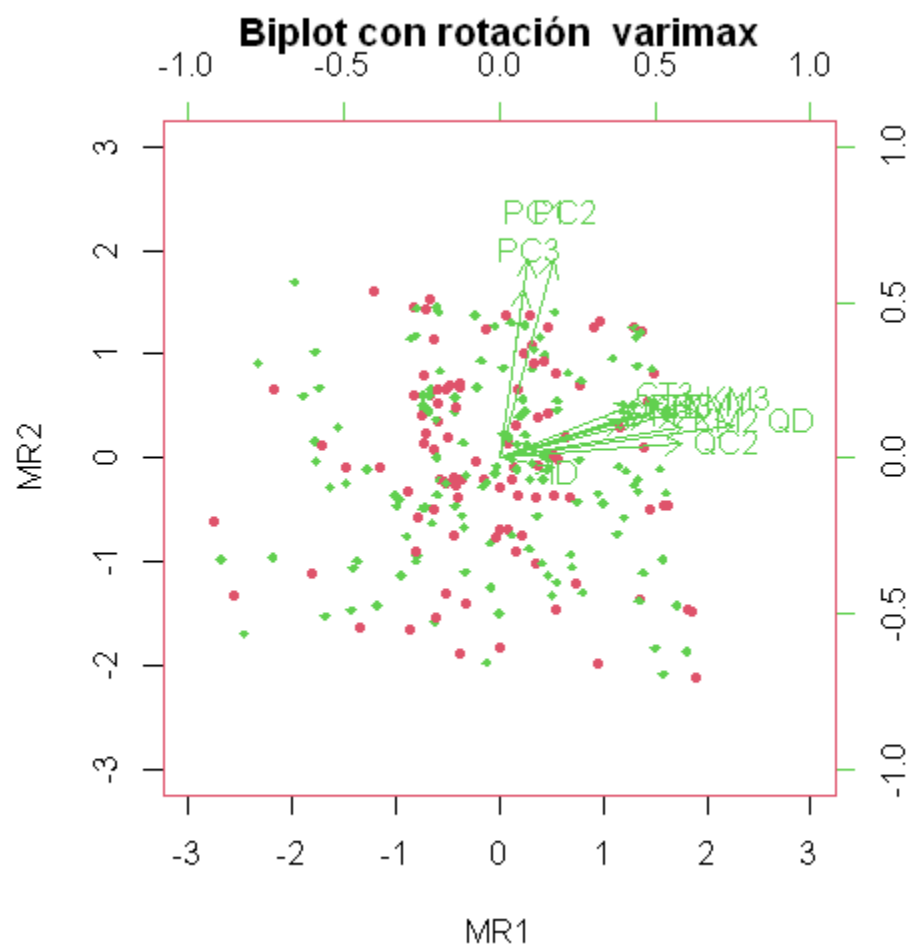
	c1	c2
KM3	0.9950294	0.8962411
QC2	0.9574940	0.8348453
QD	0.9445875	0.7413769
PC2	0.7070931	0.6937535
KM2	0.6785711	0.6746240
KM1	0.6423688	0.6544327

Después realizamos análisis para ver con cuantos factores es factible usar para trabajar con nuestro conjunto de datos:



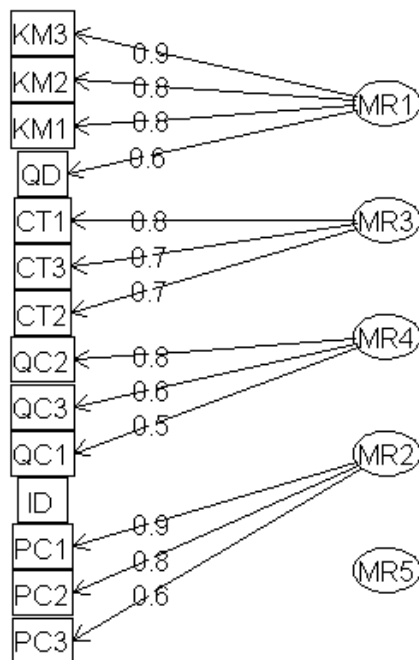
Podemos ver que con al menos dos factores podíamos trabajar

Posterior a esto rotamos las variables, recupero la de “varimax” con el que conseguimos la siguiente grafica.



Posterior obtenemos el grafico de la relación de los factores con las variables

## Factor Analysis



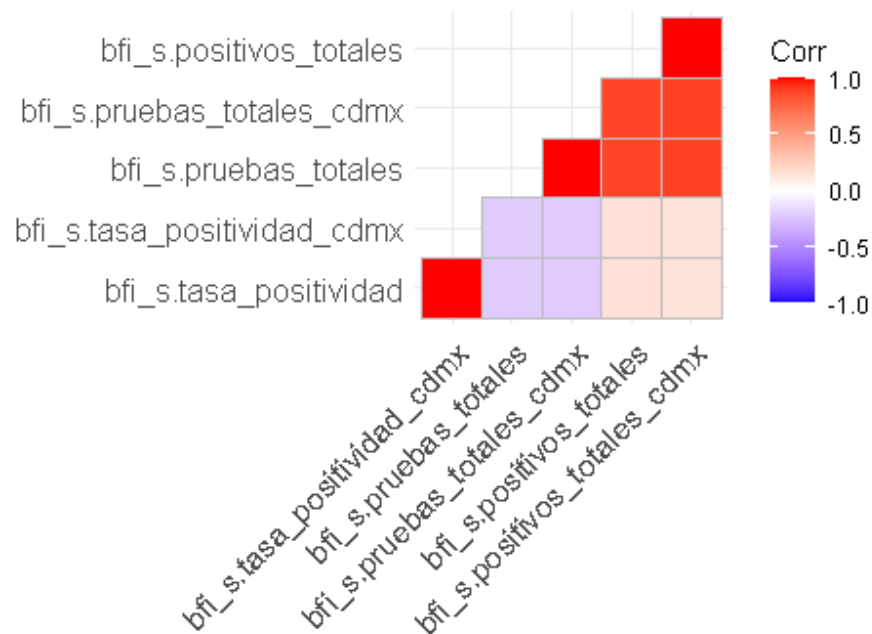
El cual es

## Ejercicio número 2:

Para este ejercicio el data set a escoger es libre y yo he decidido utilizar uno del sitio de datos abiertos de la ciudad de México (link: <https://datos.cdmx.gob.mx/dataset/total-de-pruebas-total-de-positivos-y-tasa-de-positividad/resource/ae2cd306-1aed-45a1-8ee4-3d0b0852ae4b>) se elimina la variable de fecha\_toma\_muestra ya que es de tipo timestamp y ocurría un error al estar trabajando con dicho data set:

Columna	Tipo
fecha_toma_muestra	timestamp
pruebas_totales	numeric
positivos_totales	numeric
pruebas_totales_cdmx	numeric
positivos_totales_cdmx	numeric
tasa_positividad	numeric
tasa_positividad_cdmx	numeric

Obtenemos la gráfica de correlación.



Seguimos trabajando con dicho data set, aplicamos el test de Barlet y el KMO, el cual nos arroja los siguientes resultados:

```
> p_esf$p
[1] 0
> KMO(mat_cor)
Kaiser-Meyer-Olkin factor adequacy
Call: KMO(r = mat_cor)
Overall MSA = 0.51
MSA for each item =
```

	bfi_s.pruebas_totales	bfi_s.positivos_totales	bfi_s.pruebas_totales_cdmx	bfi_s.positivos_totales_cdmx	bfi_s.tasa_positividad	bfi_s.tasa_positividad_cdmx
	0.55	0.52	0.55	0.52	0.43	0.43

Después de la prueba de KMO obtenemos un valor de 0.51 por lo que tenemos que trabajar con estos datos no va a ser tan efectivo, seguimos trabajando y comparamos las comunidades:

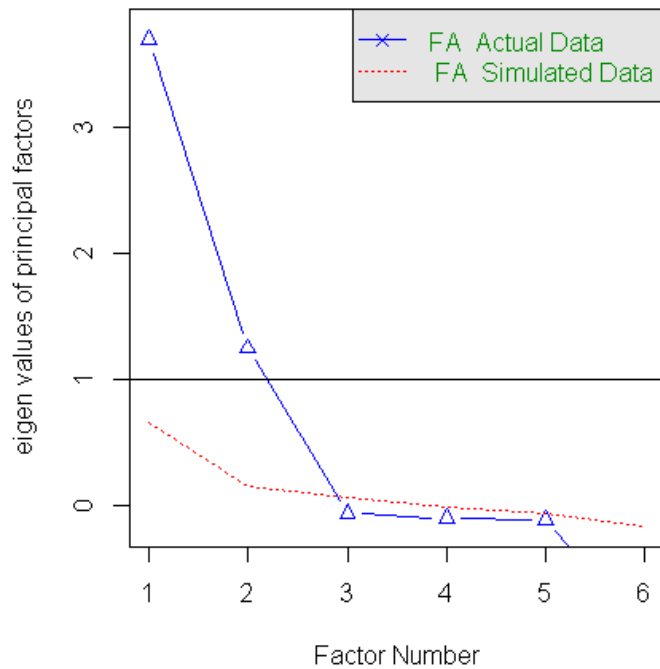
	c1	c2
bfi_s.pruebas_totales_cdmx	0.9973483	0.9973483
bfi_s.pruebas_totales	0.9973127	0.9973127
bfi_s.positivos_totales_cdmx	0.9972355	0.9972355
bfi_s.positivos_totales	0.9971959	0.9971959
bfi_s.tasa_positividad	0.9960498	0.9960498
bfi_s.tasa_positividad_cdmx	0.9960295	0.9960295

Comparamos las unicidades:

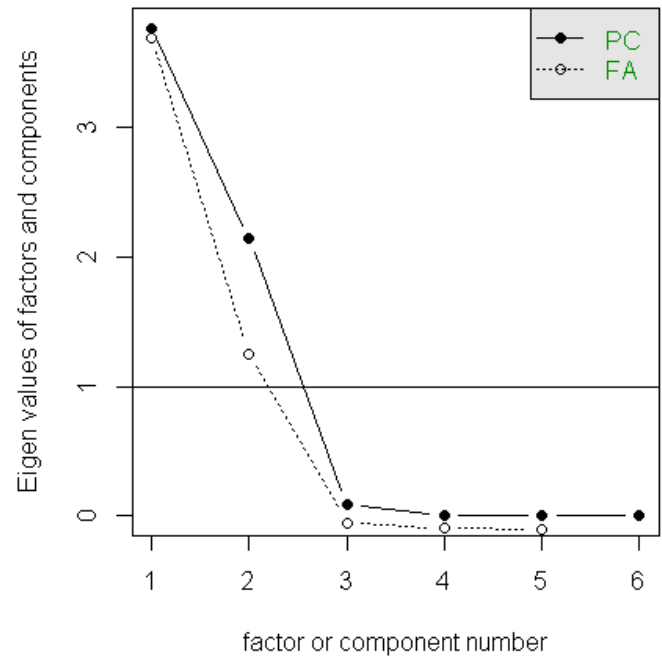
	u1	u2
bfi_s.tasa_positividad_cdmx	0.003970516	0.003970516
bfi_s.tasa_positividad	0.003950244	0.003950244
bfi_s.positivos_totales	0.002804070	0.002804070
bfi_s.positivos_totales_cdmx	0.002764490	0.002764490
bfi_s.pruebas_totales	0.002687309	0.002687309
bfi_s.pruebas_totales_cdmx	0.002651738	0.002651738

Calculamos el número de factores a utilizar:

**Parallel Analysis Scree Plots**



**Scree plot**



El resultado de la prueba nos dice que debemos de trabajar con 3 factores.

Este es el resultado de la gráfica para la interpretación.

## Factor Analysis

