

# Actividad 1.11

Cesar Vazquez

2022-11-24

## Análisis de conglomerados

### Problema 1

A partir de la siguiente tabla de distancias, hallar el dendrograma aplicando dos métodos jerárquicos con el método del promedio.

```
library(ggplot2)
library(factoextra)
```

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

```
Mpre = matrix(c(0,1,5,8.5,7.2, 0,0, 4.5,7.8,6.7,0,0,0,3.6,2.2,0,0,0,0,2,0,0,0,0,0), n
col = 5)
```

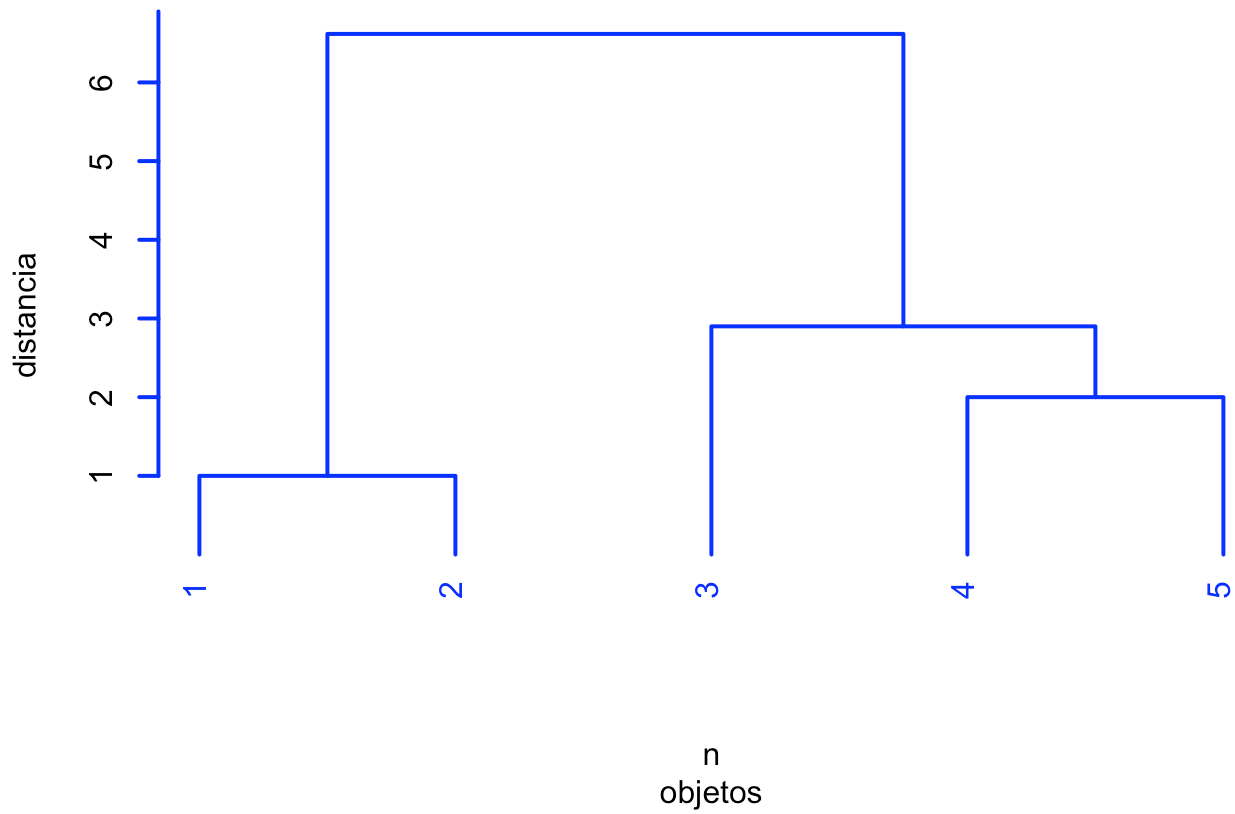
```
M = Mpre + t(Mpre)
M
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]  0.0  1.0  5.0  8.5  7.2
## [2,]  1.0  0.0  4.5  7.8  6.7
## [3,]  5.0  4.5  0.0  3.6  2.2
## [4,]  8.5  7.8  3.6  0.0  2.0
## [5,]  7.2  6.7  2.2  2.0  0.0
```

2. Apliquen las funciones `as.dist`, `hclust` y `plot` para explorar los dendrogramas (que se diferencian por las distancias en el eje Y)

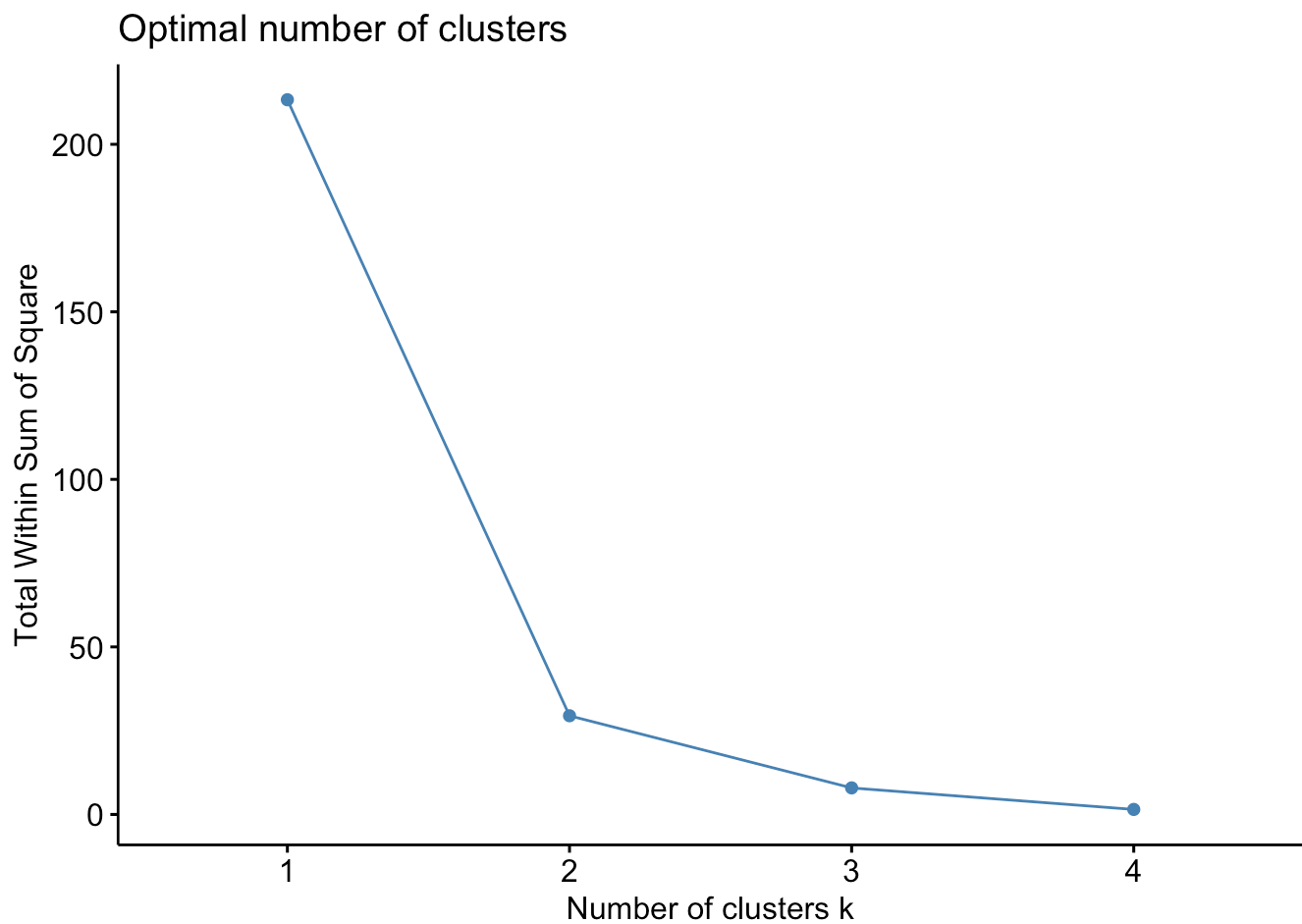
```
d = as.dist(M)
J = hclust(d, method = "average")
plot(J, hang = -1, lwd = 2, col = "blue", main = "Dendrograma de conglomerados", sub
= "objetos", xlab = "n", ylab = c("distancia"))
```

## Dendrograma de conglomerados



3. Para interpretar sobre el número óptimo de clusters puede ayudar la función `fviz_nbclust`

```
fviz_nbclust(M, FUNcluster = kmeans, method = "wss", k.max = 4)
```



4. Hallar la matriz de ultra distancias (las distancias de los objetos según el dendograma)

```
d_pre = matrix(c(0,1,6.616,6.616,6.616, 0,0,6.616,6.616,6.616, 0,0,0,2.9,2.9, 0,0,0,0,2, 0,0,0,0,0), ncol = 5)
d = d_pre + t(d_pre)
d
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,] 0.000 1.000 6.616 6.616 6.616
## [2,] 1.000 0.000 6.616 6.616 6.616
## [3,] 6.616 6.616 0.000 2.900 2.900
## [4,] 6.616 6.616 2.900 0.000 2.000
## [5,] 6.616 6.616 2.900 2.000 0.000
```