

# Actividad 1.9 Conglomerados jerárquicos

Raúl Correa Ocañas

2024-08-26

## Problema 1

A partir de la siguiente tabla de distancias, hallar el dendrograma aplicando dos métodos jerárquicos con el método del promedio. Indica cuál es el número óptimo de aglomerados que se crean.

Objetos	1	2	3	4	5
1	0				
2	1	0			
3	5	4.5	0		
4	8.5	7.8	3.6	0	
5	7.2	6.7	2.2	2	0

```
dist_matrix <- matrix(c(0, 1, 5, 8.5, 7.2,
                        1, 0, 4.5, 7.8, 6.7,
                        5, 4.5, 0, 3.6, 2.2,
                        8.5, 7.8, 3.6, 0, 2,
                        7.2, 6.7, 2.2, 2, 0),
                      nrow = 5,
                      ncol = 5,
                      byrow = TRUE)

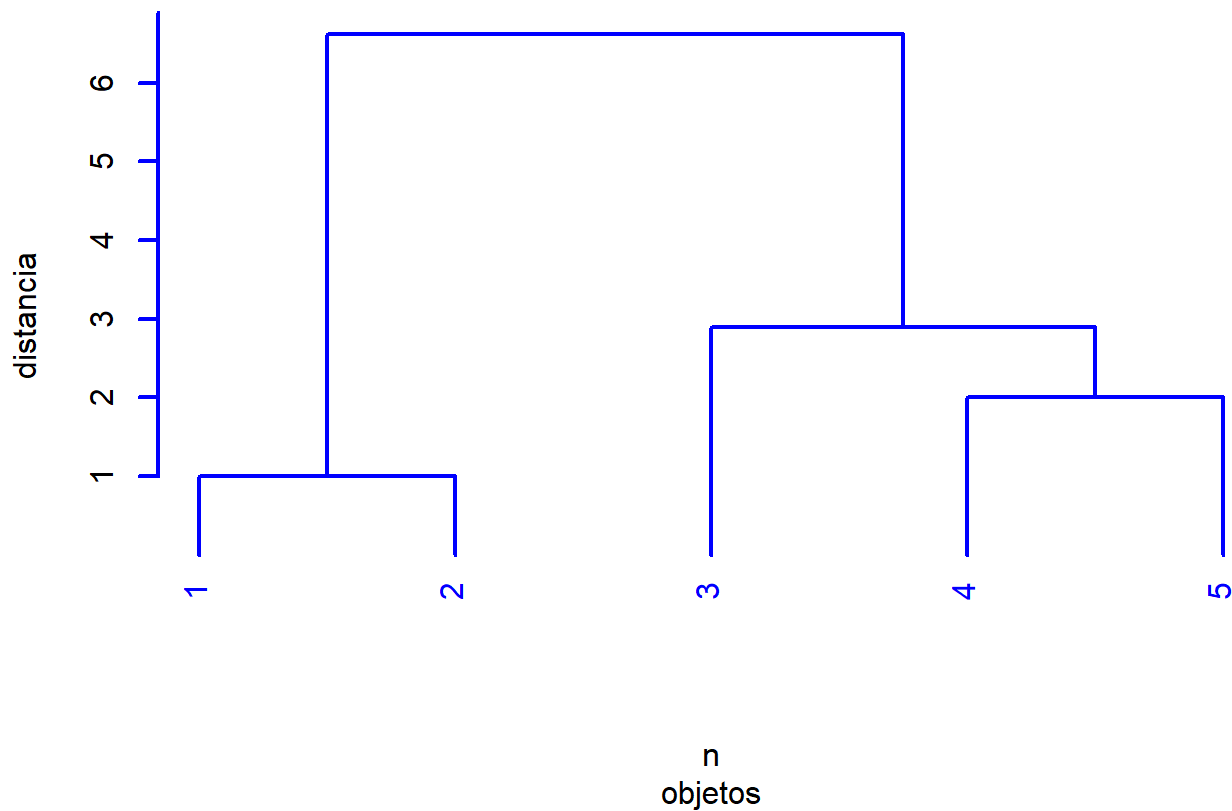
rownames(dist_matrix) <- colnames(dist_matrix) <- c("1", "2", "3", "4", "5")
dist_matrix
```

```
##      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
```

## Conglomerados con Distancia Media

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

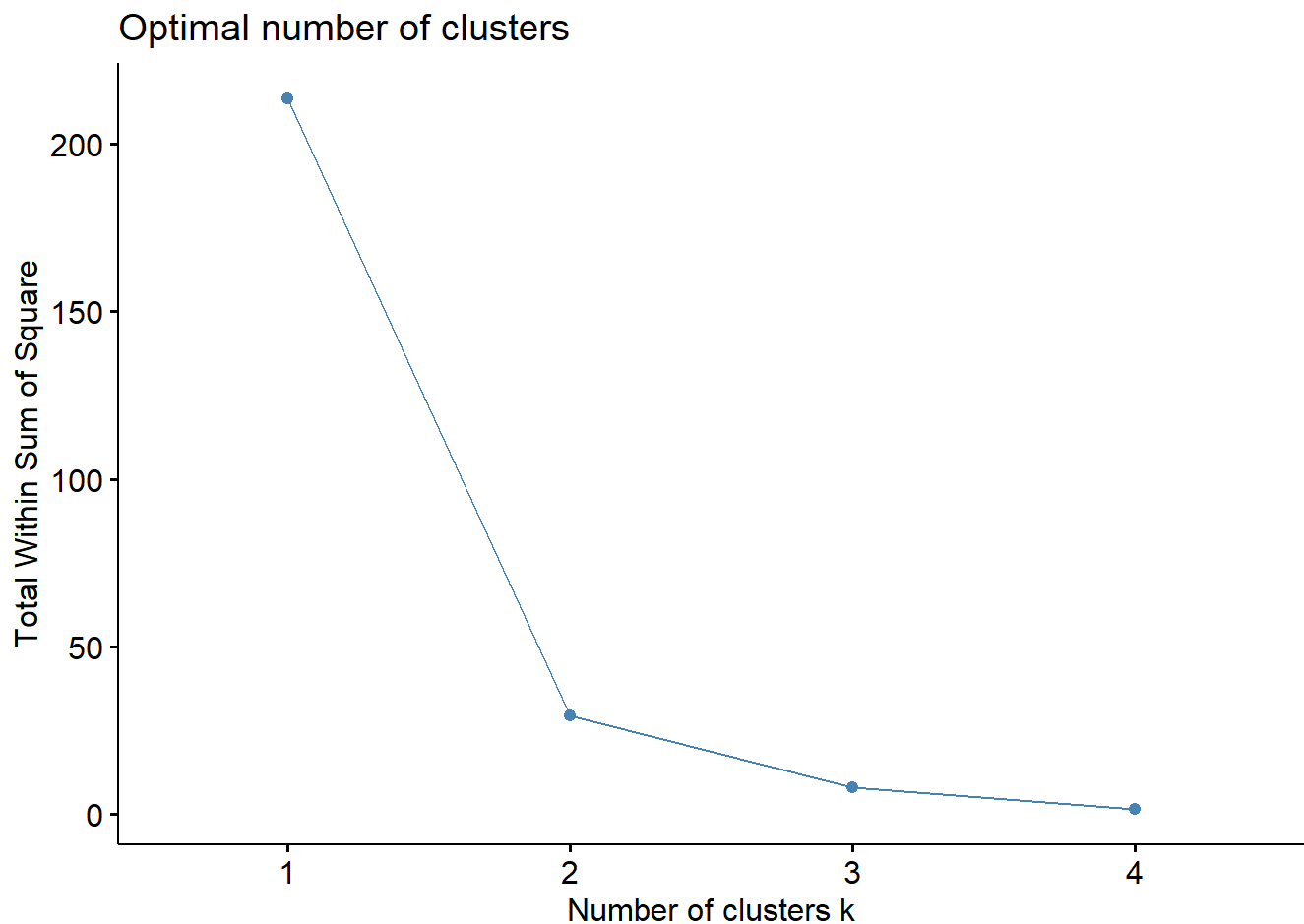
## Dendrograma de conglomerados: Promedios



En el dendrograma utilizando el método de Distancia Media (Promedios), observamos que:

1. Primero se agrupan los objetos 1 y 2.
2. Luego se agrupan los objetos 4 y 5.
3. Posteriormente, el conjunto de los objetos 4 y 5 se agrupa con el objeto 3.
4. Finalmente, el conjunto de los objetos 1, 2, 3, 4 y 5 se agrupan en un solo conglomerado.

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

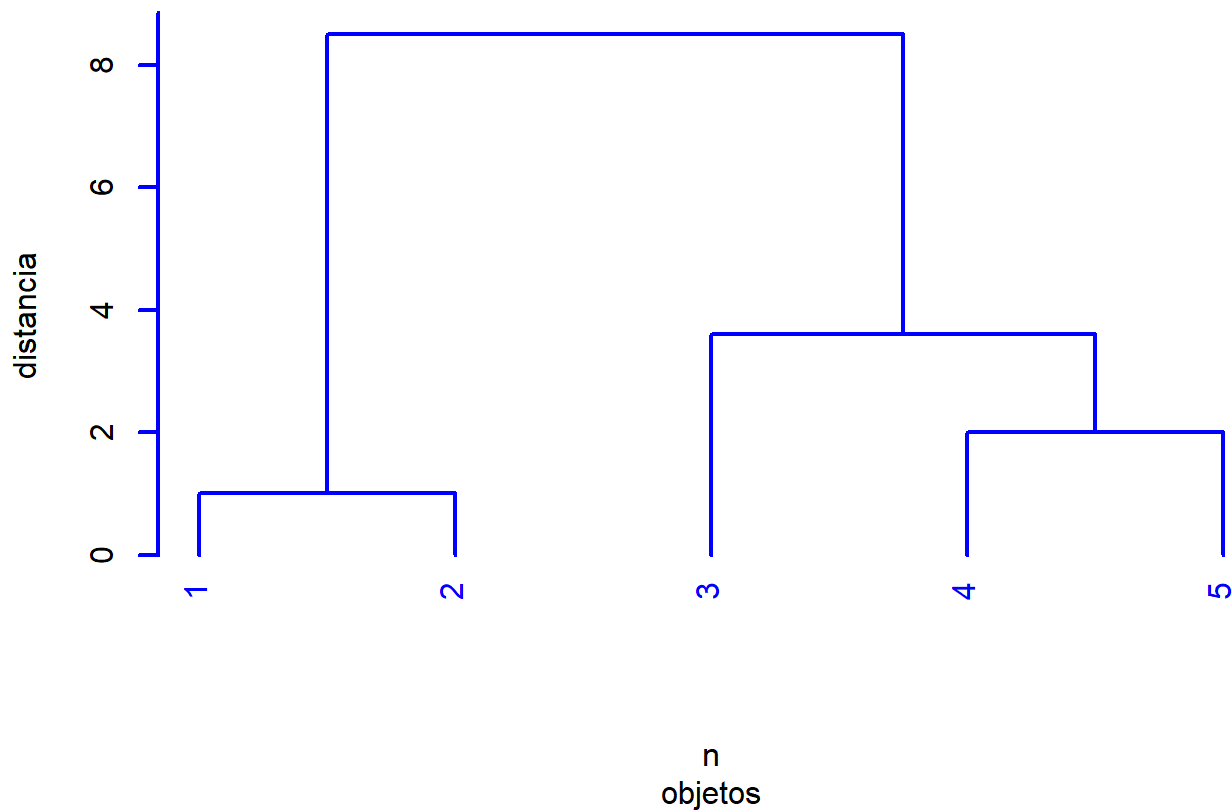


Utilizando el método del codo, se observa que el número óptimo de conglomerados es 2, ya que es en este punto donde la suma total de los errores cuadrados disminuye con menor rapidez.

## Conglomerados con Distancia Minima

```
clustering_complete = hclust(d, method = "complete")
plot(clustering_complete, hang = -1, lwd = 2, col = "blue", main = "Dendrograma de conglomerados: Centroides", sub = "objetos", xlab = "n", ylab = c("distancia"))
```

## Dendrograma de conglomerados: Centroides



Al observar el dendrograma generado con el método de Enlace Completo, notamos que:

- El orden de agrupamiento de los objetos es el mismo que en el caso anterior.
- Sin embargo, las distancias a las que se agrupan los objetos son mayores en comparación con el método de Distancia Media, lo que sugiere que los conglomerados son más compactos.