

# Modelos de Soporte No Supervisado

Cluster jerárquico con K-medias

# Jerárquico y K-medias

K-means representa uno de los algoritmos de agrupamiento más populares. Sin embargo, tiene algunas limitaciones:

- Requiere que el usuario especifique el número de clústeres por adelantado y selecciona los centroides iniciales al azar.
- La solución final de clustering k-means es muy sensible a esta selección aleatoria inicial de centros de clusters.
- El resultado puede ser (ligeramente) diferente cada vez que calcule k-means.

# Jerárquico y K-medias

El algoritmo híbrido se resume de la siguiente manera:

1. Calcule la agrupación jerárquica y corte el árbol en k-clusters
2. Calcule el centro (es decir, la media) de cada grupo
3. Calcule k-means utilizando el conjunto de centros de cluster (definido en el paso 2) como centros iniciales de clusters

Tenga en cuenta que el algoritmo k-means mejorará la partición inicial generada en el paso 2 del algoritmo.

# Jerárquico y K-medias

- La función R `hkmeans ()` [en `factoextra`], proporciona una solución fácil para calcular el cluster jerárquico con k-means.
- El formato del resultado es similar al proporcionado por la función `kmeans ()` estándar.

# Jerárquico y K-medias: ejemplo

```
df <- scale(USArrests)
# hierarchical k-means cluster
library(factoextra)
res.hk <- hkmeans(df, 4)
# Resultados de hkmeans()
names(res.hk)
# Imprimimos los resultados
• res.hk
# Gragicamos el dendograma
fviz_dend(res.hk, cex = 0.6, palette = "jco", rect = TRUE, rect_border = "jco",
rect_fill = TRUE)
```

# Jerárquico y K-medias

*# Graficamos el cluster final generado por hkmeans*

```
fviz_cluster(res.hk, palette = "jco", repel = TRUE, ggtheme =  
theme_classic())
```