

## División de la colección

2020-12-02

El objetivo del siguiente trabajo es dividir la colección de los coches del jefe, de manera que seamos capaces de distribuirlos en las 10 residencias, teniendo como restricción un máximo de 15 vehículos por vivienda. A partir de los test realizados, establecemos que el número óptimo de clusters sería 3, pero debido a la restricción de 15 vehículos por vivienda establecemos el óptimo en 5.

Su jefe tiene una semana complicada y le ha pedido que le haga una propuesta de cómo repartir la colección en las distintas residencias. Como Ud. bien sabe, podría repartirlos como máximo en las diez que posee en la actualidad (precisamente está, durante esta semana, cerrando la venta de alguna de ellas, que quizá sustituya por alguna otra), pero, siendo una opción conservadora, quizá no sea la más adecuada, atendiendo a las características que Ud. ya conoce de los vehículos.

El criterio de reparto debe ser consistente, y debe justificar su decisión en un máximo de 4 páginas.



[HTTPS://GITHUB.COM/JLOPEZGALDON/TARD/BLOB/MAIN/03\\_LOS\\_COCHES\\_DEL\\_JEFE/RMD/02 LOS COCHES DEL JEFE.RMD](https://github.com/jlopezgaldon/tard/blob/main/03_los_coches_del_jefe/rmd/02_los_coches_del_jefe.Rmd)

## SELECCIÓN DEL NÚMERO ÓPTIMO DE GRUPOS

Durante este informe el objetivo es ser capaces de segmentar los coches en grupos de máximo 15 vehículos para distribuirlos en 10 viviendas. Para ello, comenzaremos escalando los datos, continuaremos analizando la idoneidad de los clusters, a través del estadístico de Hopkins y métodos visuales y, por último, calcularemos el número óptimo de clusters con distintos algoritmos.

### Métodos de evaluación de la bondad del análisis cluster

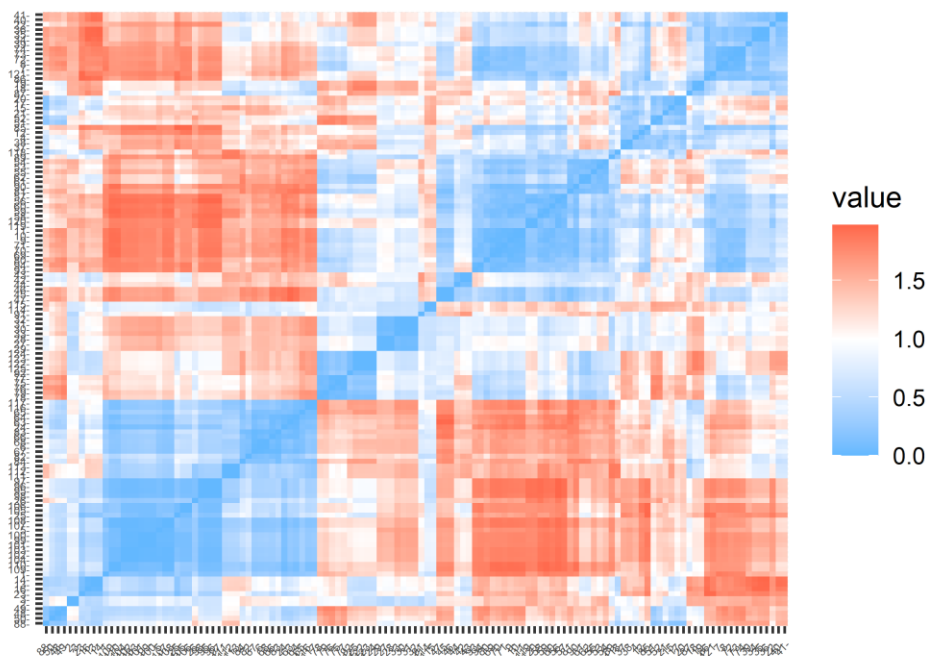
#### *Estadístico Hopkins*

Tras realizar el cálculo del estadístico de Hopkins, podemos rechazar la hipótesis de aleatoriedad por lo que avalamos la existencia de dos o más clusters en el conjunto de observaciones, es decir, existen distancias considerables entre los vecinos, por lo que podemos continuar con nuestro análisis.

#### *Método VAT de evaluación visual*

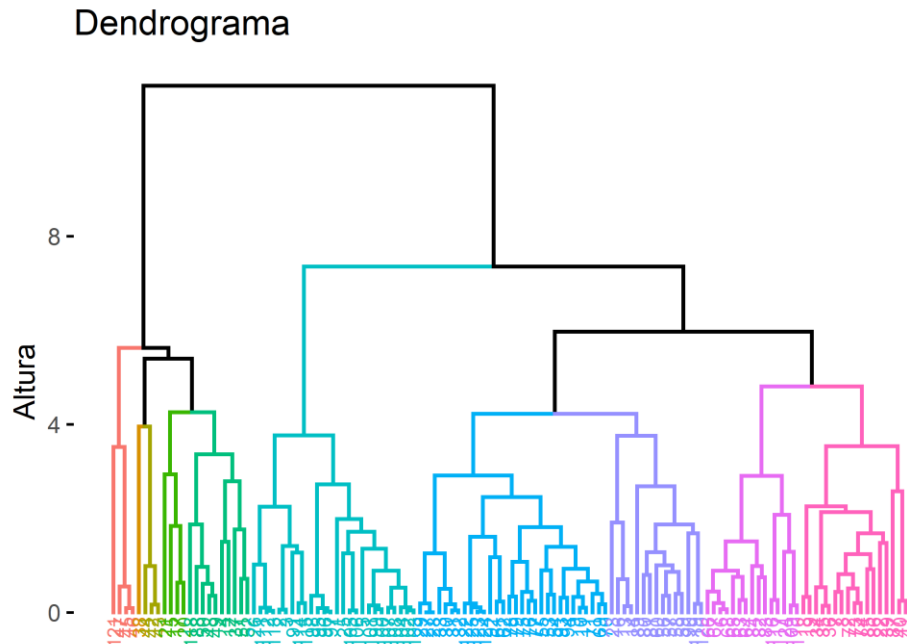
##### Matriz de distancias

Pearson



El gráfico superior representa de manera gráfica la distancia entre los grupos, cuando es azul implica poca distancia entre las observaciones, mientras que si es rojo implica más distancia. En este caso no encontramos gran información, pero si podemos observar cómo existen grupos de observaciones claramente diferenciadas.

Otra manera de visualizar los grupos es mediante un dendrograma:



Análisis cluster aplicando Ward D2

En el gráfico superior, hemos planteado un dendrograma de los vehículos clasificándolos en 10 grupos, siendo este el número máximo de grupos que podemos realizar. Aun así, esto no implica que sea el óptimo, es por ello que realizaremos a continuación el cálculo del número óptimo de grupos.

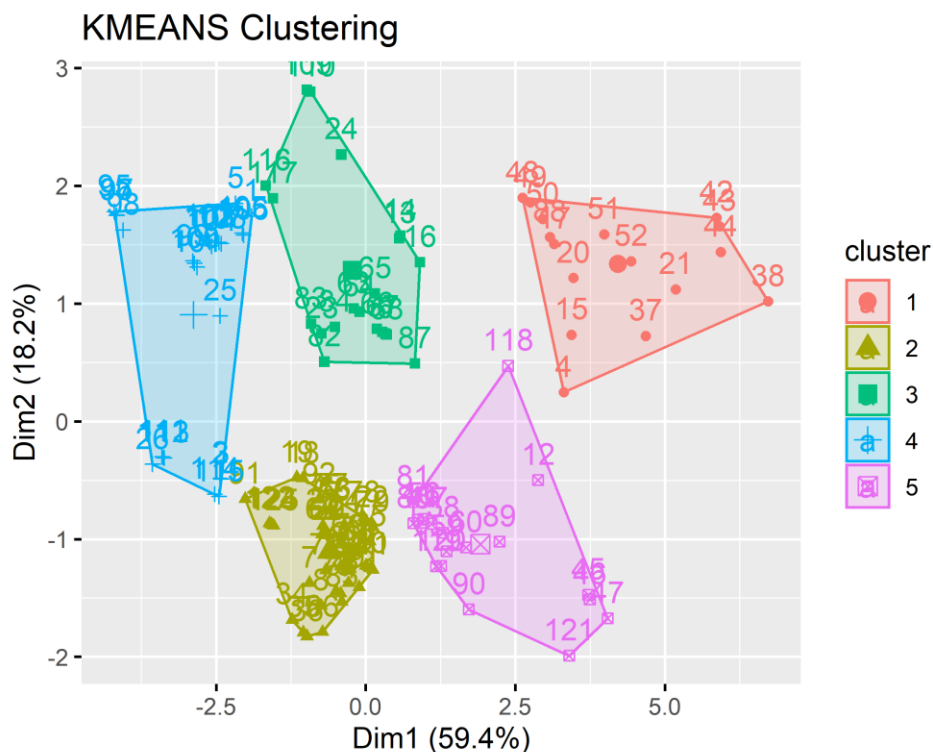
### Determinación del número óptimo de clusters

Tras realizar el análisis obtenemos que el número idóneo de clusters es 3, sin embargo, esto presenta un inconveniente, ya que si dividimos los 125 vehículos en 3 grupos nos sería imposible almacenarlos en los 10 garajes del jefe, ya que la máxima capacidad que poseen es de 15 vehículos por vivienda. Es por ello, por lo que probaremos a graficar los clusters.

$$K = 3$$

Si elegimos 3 grupos salen bien diferenciados y no se producen solapamientos, el inconveniente es que dada la ubicación de las viviendas no consideramos correcto esta segmentación para la situación que nos acontece. Estudiando el mapa de las residencias, podemos observar 5 o 6 zonas donde están distribuidas las viviendas: Paris, La Rochelle, Suiza, Pirineos y Niza, es por ello, por lo que podríamos clasificar los vehículos en 5 grupos.

$$K = 5$$



## CONCLUSIONES

Tras haber realizado analizado el número óptimo de clusters para la colección de 125 TT de nuestro jefe, rechazamos el número óptimo calculado, es decir, 3 grupos, ya que al tener 10 viviendas tan separadas nos imposibilita la tarea de clasificarlos en los garajes de manera óptima. Es por ello, y tras estudiar la ubicación de las viviendas en el mapa otorgado, que establecemos el número óptimo de grupos en 5, los cuales se distribuirán de la siguiente manera:

- **Grupo 1:** 16 vehículos. Distribuidos en la Rochelle. El vehículo que cabe en este garaje pasaría a Córcega, ya que se asemeja a los vehículos del quinto grupo.
- **Grupo 2:** 45 vehículos. Distribuidos en los 3 garajes de la zona de Mónaco-Niza.
- **Grupo 3:** 20 vehículos. Distribuidos en los garajes de Suiza.
- **Grupo 4:** 25 vehículos. Distribuidos en los dos garajes de París.
- **Grupo 5:** 19 vehículos. Dado que algunos se asemejan a los del grupo 2, recomendarías que fuesen a Córcega y el resto a Pirineos.