

Los coches del jefe, parte 2. Cómo los reparto.

Valentina Díaz Torres

01/12/2020

Introducción

Tras haber realizado un análisis exploratorio en el informe anterior, acerca de las distintas variables que podrían afectar al reparto de los coches del jefe, en 10 lugares diferentes, se eligieron unas variables, las que más aportaban al estudio, en base a diferentes factores.

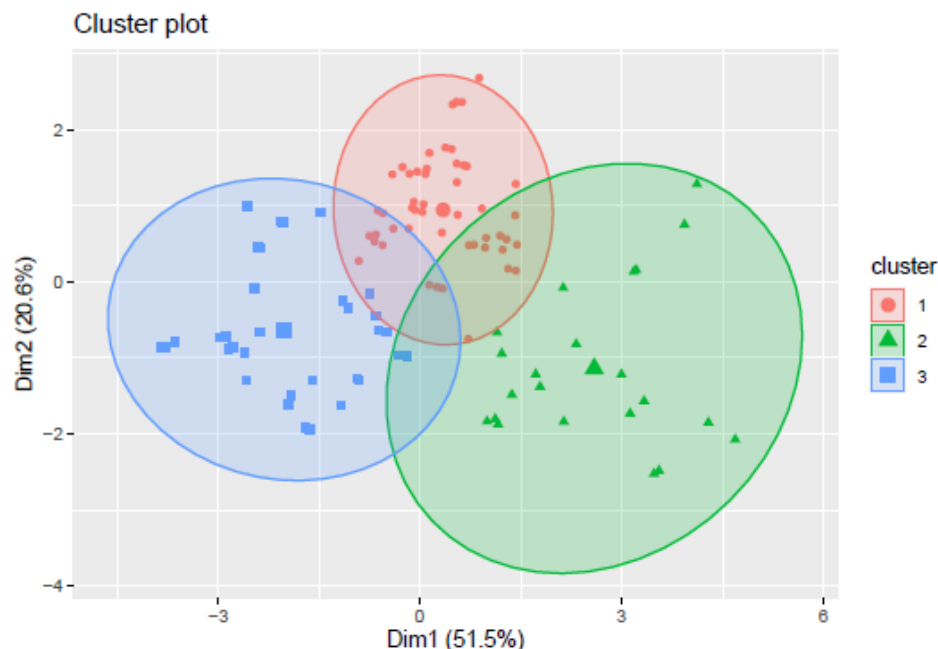
Estas variables son: Cilindrada, revoluciones por minuto, peso, nº de plazas, consumo a 90 km/h, velocidad y precio en Euros.

Una vez que han sido elegidas estas variables, el siguiente objetivo es repartir la colección del coche del jefe en torno a las distintas residencias, que este tiene por el mundo. Esto significa que los coches tienen que ser agrupados en grupo óptimos, en base a sus características.

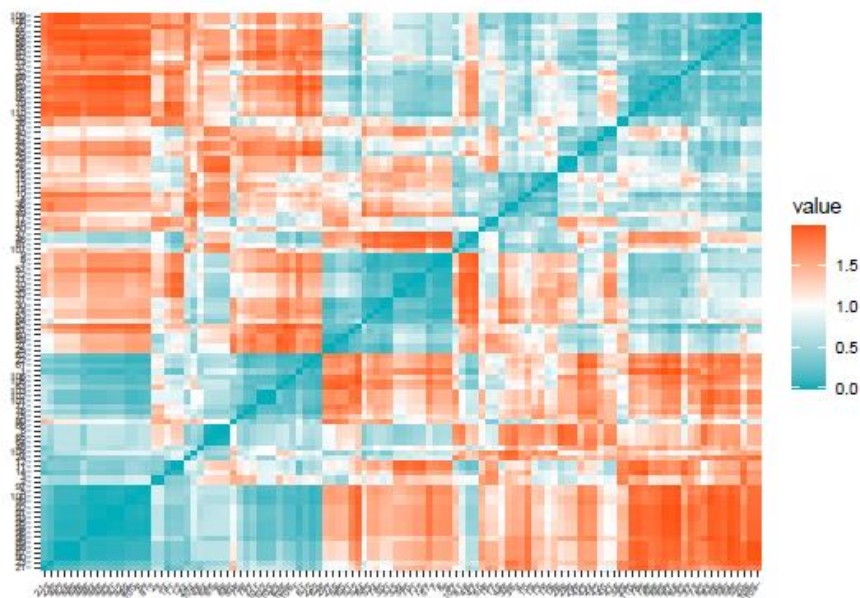
Para decidir en cuántos grupos se dividen los coches habría que tener en cuenta varios factores. En primer lugar hay 125 coches y el Jefe tiene 10 casas repartidas por Europa, en las cuales tiene capacidad máxima de 15 coches. Además, analizando bien la localización de estas casas, se observa que hay tres casas cercanas a Mónaco y próximas a la costa, una en la Isla de Corse y otra en la Rochelle, también costa. Por tanto estas podrían formar un primer grupo, con coches con características parecidas, que normalmente no serán de muchas revoluciones, ni con mucha potencia, por ejemplo. Otro grupo podría estar compuesto por los coches de Andorra y Suiza, que conforman tres casas, ambos necesitarán ser coches de montaña, adaptados al clima y al relieve de las zonas. Por último, hay dos casas en París, donde se necesitarán coches de ciudad, que no consuman ni contaminen demasiado, y que sean ligeros de mover, posiblemente de pocas plazas.

Para comprobar esto, se ha realizado un análisis cluster, para conocer seguro que es óptima esta división.

Para realizar el análisis, se parten de 3 cluster, ya que, a partir de 4, estos empiezan a solaparse. La dimensión 1 explicaría un 51,5% de la información y la dimensión 2 un 20,6%. La representación de 3 cluster en el mapa sería la siguiente:

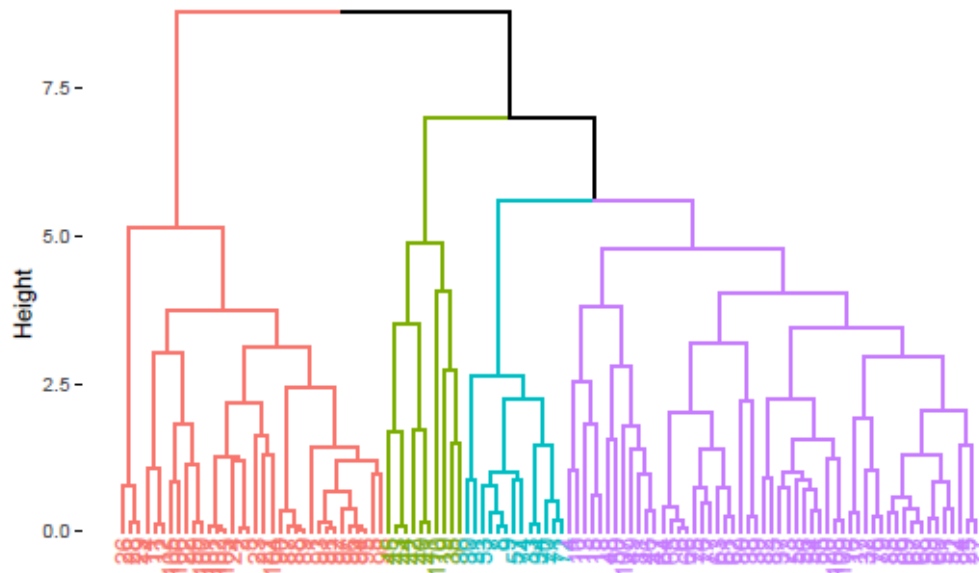


Este gráfico nos muestra que a priori las manchas rojas son las que van a tener más correlación, por lo que se podrían distinguir algunos grupos en los extremos.

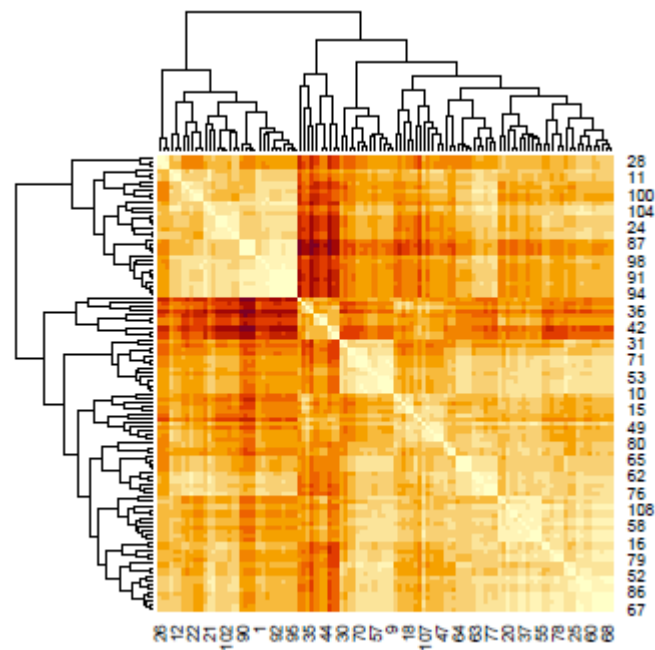


También, se ha realizado un dendrograma, que nos muestra dos grupos claramente diferenciados y un tercero. De estos salen otros subgrupos, por lo que se comprueba que habría tres grupos grandes importantes.

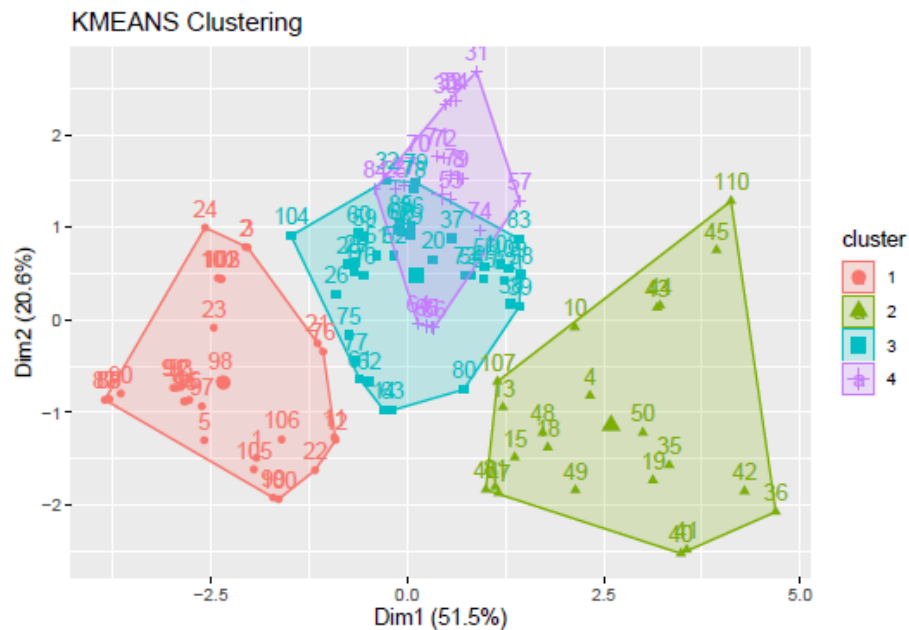
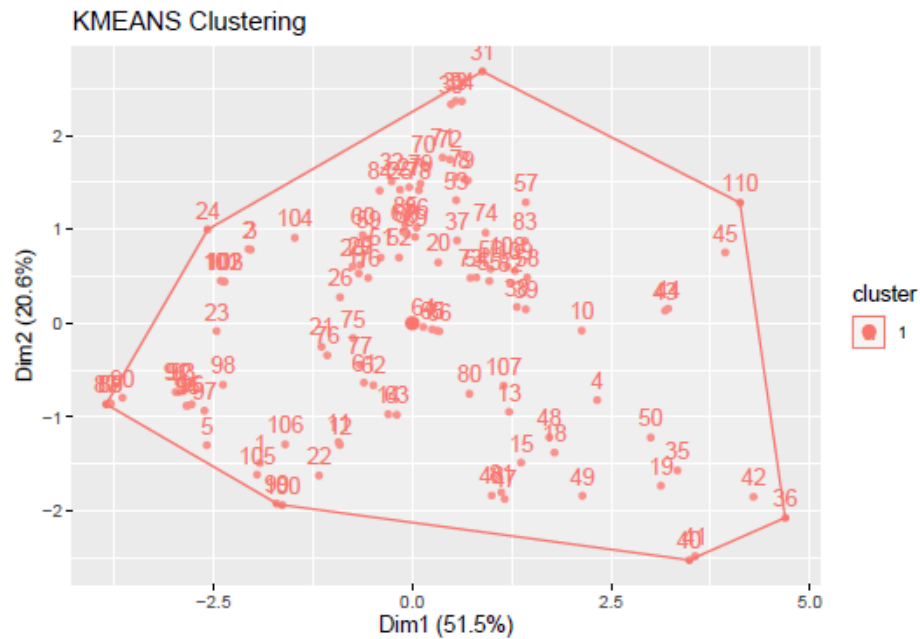
Dendrograma Coches Jefe



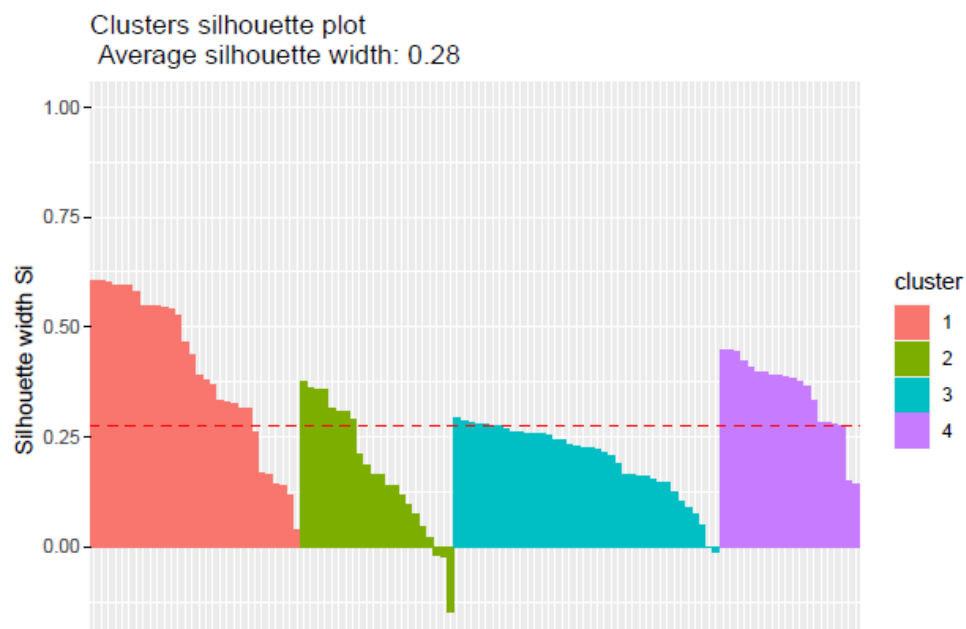
Este mapa de calor nos aporta más información, completando el dendrograma anterior.



Según el análisis clúster utilizando K-medias, aparecen los tres grupos claramente diferenciados, habiendo más información en el 2º. No obstante, también cabría la posibilidad de plantearse un 4º grupo, según este análisis, aunque ya no quedarían tan separados y habría solapamiento del 4º y el 3º.

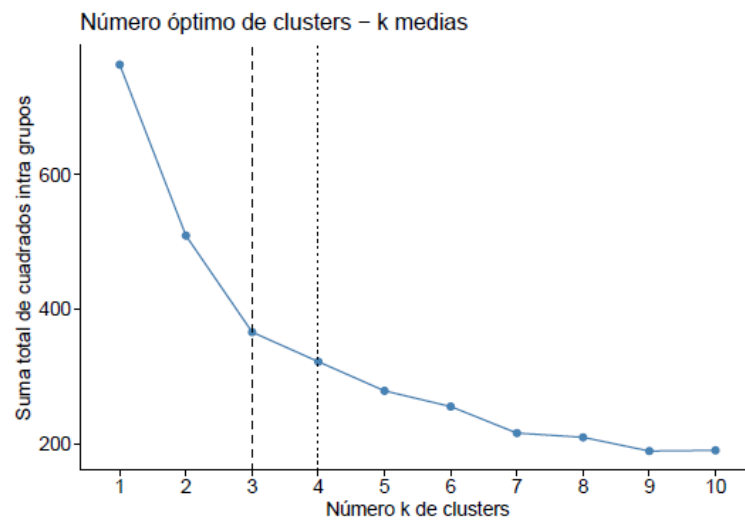


El gráfico de siluetas nos muestra 4 grupos representados, aunque la mayoría de información sería recogida por los 3 primeros grupos.



Número óptimo de clusters

El número óptimo de clusters, como se ha ido viendo a lo largo de los distintos gráficos es se encuentra comprendido entre tres y cuatro. Esto se puede ver representado en la siguiente gráfica, en el que el número óptimo de clústers o grupos viene definido según k-medias.



Como conclusión cabría decir, que se ha confirmado mediante numerosos análisis que el número elegido de grupos para distribuir los coches serían en grupos de tres. Todos los análisis y diferentes modelos afirman que 3 es el mejor número, como se puede observar en este histograma.

