



ANALISIS LAS CASAS DEL JEFE

Máster en Data Science

Juan Manuel López Zafrá

Marta Divassón Carribero

Fecha: 23 noviembre 2018

ÍNDICE

- 1. Introducción**
- 2. Análisis exploratorio**
- 3. Visión de distribución**
- 4. Conclusión**

Introducción

En el siguiente informe analizaremos la distribución en las 10 residencias con un máximo de 15 coches por residencia de 76 coches que hemos organizado en clústeres. Es decir, el número máximo de plazas existente es de 150, de esta forma sobran 74 plazas, por tanto, podemos distribuir sin problema.

Primero ofreceremos una solución óptima al número de clústeres elegidos según los modelos, ya que, al principio, en un análisis exploratorio se llegó a la conclusión de que 6 clústeres era el número óptimo, pero según avanzamos en métodos han primado 3 clústeres. En la metodología se explicará esta elección de segmentación en grupos. La finalidad de esto es poder descubrir características y propiedades que producen dicho agrupamiento y así observar aquellos aspectos de forma más cercana para que no pasen inadvertidos.

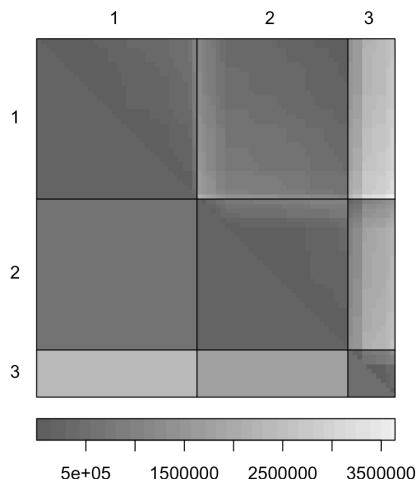
En cada grupo, existe la mayor homogeneidad posible en base a una serie de características, pero para poder tomar decisiones desde el punto de vista de negocio para optimizar la distribución por casas, vamos a diferenciar por variables para ver aquellos que pueden llegar antes a sitios más alejados, por ejemplo, porque tengan más potencia o más aceleración para llegar en un menor tiempo, o una mayor velocidad para llegar al objetivo.

Análisis exploratorio

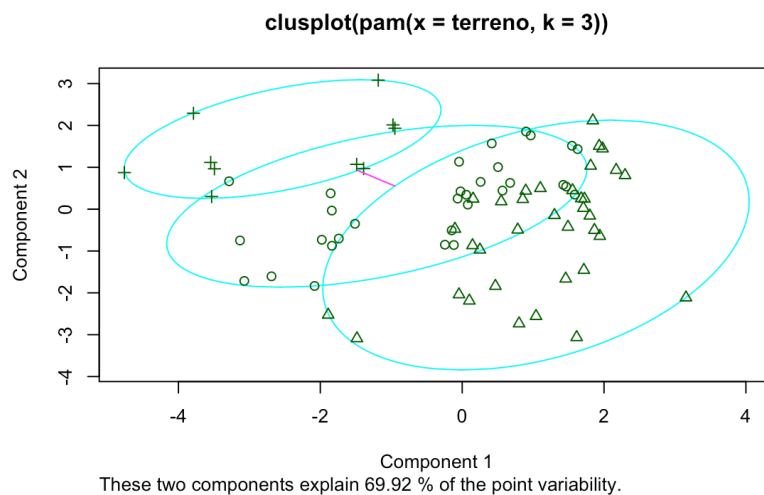
En la selección del número de clústeres empleamos varios métodos, el de "**Kmeans**", el de partición alrededor de los **medioides** o PAM y el **jerárquico**, que indican a ojo que el número de clústeres óptimo puede ser de 6. Pero hay que hacer un método más desarrollado y completo, en este caso el método "complete" que optimiza y distribuye en clústeres según las observaciones y sus características. Con la función "NbClust" se llega a la selección óptima de 3 clústeres, con un valor de índice de 0.4122 (es decir, de los 10 clústeres por los que se puede segmentar, el valor con mayor peso está en una distribución de 3 clústeres). No nos enfrentamos al problema de que haya un excesivo número de conglomerados, lo cual supondría un coste mayor ya sea por cuestiones de investigación o por una mala segmentación.

Finalmente, mediante el "**Hubert Index**" se selecciona el número óptimo de clústeres, y coincide con el analizado anteriormente.

Después de seleccionar el número de conglomerados, calculamos la matriz de distancia y lo visualizamos en el gráfico siguiente.



Finalmente, según los clústeres y las variables escogidas con el método de análisis compontes, tenemos casi un 70% de la varianza explicativa. Se hace una división en los 3 clústeres escogidos y se refleja en este grafico.



Visión de distribución

Mediante una validación interna de los clústeres obtenemos que en el primero se agrupan 44 observaciones, 21 en el segundo y 11 en el tercer clúster.

Hay que destacar que, para distribuir por **todas** las casas, vamos a calcular la distancia desde el punto original, siendo la universidad CUNEF hasta los puntos determinados. Si quisiéramos minimizar en costes, entonces tendríamos que distribuiré total de 76 coches, siempre teniendo en cuenta los clústeres formados, de 15 en 15 en los puntos más cercanos y evitar puntos a los que sea imposible llegar por carretera, como Córcega, cuyo coste será más elevado ya que hay que utilizar la vía marítima para su acceso. Vamos a asumir por tanto que al ser coches de coleccionista son todos de gasolina. El precio de la gasolina a día 22 de noviembre es de 1.414 €/l.

En el **primero** se agrupan las observaciones: 1,3,4,5,6,14,15,16,20,21,22,23,24,25,41,42,43,44, 46, 47, 48, 49, 50, 54, 55, 56, 57, 58, 59, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73 75 y 76.

La media según las observaciones de las medioides se encuentra en la 43 de marca "Mitsubishi" con un precio de 4747813, una potencia de 181, peso de 1925, 7 plazas a una velocidad km/h de 175 y aceleración de 12.1.

Las características de estos coches son de precio medio, son aquellos coches que se sitúan en el rango medio, con un nivel de potencia elevado (180), pero compuestos por coches con más plazas, por tanto, coches más familiares con un gran peso y una velocidad media alta. Pueden ser los coches tipo todoterreno.

Estos serán trasladados a los puntos 2,3,8 y 10.

Por tanto, la distancia hasta el punto 2 desde el punto de origen establecido es de 1273 kilómetros cuantificado en 12h y 10 minutos. Si asumimos para todos los coches que consumen una media de 10l/100 km, ya que son coches antiguos de gasolina (aunque por autopista el consumo se estabiliza), entonces el precio de recorrer 1273 km es de **1796,203 €** por coche.

A cada punto llevaremos 11 coches, ya que son 44 coches en total distribuidos en 4 puntos distintos.

La distancia hasta llegar al punto 3 es de 1254 km y un total de 11h y 55 minutos. Por tanto, el coste de llegar 11 coches es de **1773,156 €** por coche.

Para llegar al tercer punto o punto 8, la distancia es de 855 km y 8 horas 28 minutos. El coste por cada coche se reduce a **1208,97 €**.

Finalmente, la distancia al ultimo punto, es decir, hasta San Remo, es de 1327 km con un total de 12 horas y 44 minutos. El coste escala a **1876,378 €**, por coche.

Por tanto, al tener 44 observaciones, vamos a llevar el máximo de coches a los 2 puntos más baratos, esto es, hasta el Punto 3 situado en Niza y hasta el punto 8 situado en Yves, ambos en Francia.

El coste será total será de: 26.597,343 € al punto 3, 18134,55 € al punto 8, 17962,03 € al punto 2 y 7505,512 € al ultimo punto. Esto hace un total de **70.199,435 €**.

El **segundo** clúster compuesto de las siguientes observaciones: 7,26,27,28,29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 51, 52, 53 y 74.

La media de este clúster está en la observación 31, siendo la marca "Land Rover", con precio de 9560000, potencia de 225 y peso de 2220, 5 plazas, velocidad de 196 km/h y aceleración de 10.7.

Se almacenarán en los puntos: 1,4,6 y 7.

La distancia al primer punto es de 1272 kilómetros al primer punto, un total de 12 horas y 14 minutos lo cual es un coste de **1798, 608 €** por coche. La distancia al segundo punto (En la isla de Córcega) es de 18 horas y 36 minutos. Esta ruta incluye un ferry, que resta 5 horas de trayecto, por tanto, el viaje en coche total es de 13 horas y 36 minutos. El número de kilómetros recorrido es de 1506 kilómetros, lo cual supone un coste de **2129,484 €** cada vehículo.

Para llegar al tercer punto situado en Suiza, se tarda un total de 13 horas y 21 minutos siendo 1436 kilómetros totales. Asimismo, el coste de recorrido es de **2030,504 €** por coche. Finalmente, para llegar a Chemin de Chavatte en Suiza, se tardan 14 horas y 56 minutos lo cual suponen 1528 kilómetros, con un coste de **2160,592 €**.

Los puntos que suponen un menor coste son el punto 1 situado a los alrededores de París al que llevaríamos 15 coches (coste total de **26.979,12 €**), el tercer punto en Suiza, llevando 4 coches con un coste total de **8.122,016 €**. Para tener al menos un coche en cada casa llevaríamos 1 a la isla de Córcega, lo que supone **2129,484 €** y otro al último punto en Suiza que **son 2160,592 €**. Para concluir, el coste total es de **39.390,636 €**.

En el **tercer** y último conglomerado, se agrupan las observaciones: 2, 8, 9, 10, 11, 12, 13, 17, 18 y 19 y 60.

La media de este se sitúa en la observación 60, con marca "Nissan", con un precio de 349800, con potencia de 100, peso de 1750, 5 plazas, velocidad de 145 y aceleración de 19.0.

Estos irán a los puntos 9 y 5. Son 11 coches en total.

La primera mitad, aquellos con características mas pobres en velocidad y potencia y que sean de gasolina, irán al punto **9** situado en **Andorra**. Este es el punto más cercano, a una distancia de 8 horas y 9 minutos desde el punto de partida con un total de 723 km. El promedio de este grupo recorre 145km/h. Si consumen una media de 10l por cada 100 km (aproximadamente) entonces, la hora tiene un coste de 141,4€ y al ser 8 horas y 9 minutos, el coste de gasolina es de **1022,322€** por coche aproximadamente hasta Andorra.

La segunda mitad, irá al punto 5, situado en Saint-Räphael (Francia) a 12 horas y 30 minutos y 1.221 km. Asumamos que si consumen una media de 10L/h entonces el coste de la gasolina es de **1726, 494€**. Si llevamos 7 coches al punto más cercano, en Andorra, asumiríamos un coste de **7156,254 €**. Si llevamos solo 4 coches al segundo punto, el coste sería de **6905,976€**. Por tanto, el coste total es de **14.062,23 €**.

Conclusión

Asumiendo que todos gastan la misma cantidad de litros cada 100 kilómetros, el grupo que menor coste incurre es el tercer clúster con un coste total de **14.062,23 €** (11 coches totales). El segundo clúster constituido por los coches más completos en cuanto a calidad, velocidad y potencia incurren unos costes totales de **39.390,636 €** (21 coches totales). Finalmente, el primer clúster, con los

coches de gama media, pero de mayor peso generalmente, suponen unos costes de **70.199,435 €**.

Todos estos costes se asumen sin tener en cuenta el gasto de peajes y gastos adicionales por cada kilometro de la carrocería de los coches. Al ser coches antiguos hay que calcular el gasto de las ruedas, etc... La mejor forma de llevarlos a su destino sería a través de un camión de transporte, que a parte de abaratar en costes no incluiría el coste adicional de gasto físico del coche por llegar a su destino. He iría repartiendo por cada casa más cercana aquellos coches con características homogéneas (como hemos clasificado en coches).

GitHub **link:** <https://github.com/martadiva/REDUCCION-DIMENSION/blob/master/CLUSTER%20COCHES%20EXTENDED>