



# DATA SCIENCE

*Estudio de la felicidad con K-Means*



Escuela Técnica Superior de  
**Ingeniería Informática**

REALIZADO POR:

*Yassin Lalj El Assri*

# ¿CÓMO INFLUYE LA CONFIANZA EN EL GOBIERNO, LA ECONOMÍA Y LA SALUD EN LA FELICIDAD DE LA GENTE?

## INTRODUCCIÓN

La pregunta que me hago es en realidad una triple pregunta, pero realmente me interesaba analizar cómo se desarrollaba el comportamiento de la felicidad con estas tres variables, las cuales considero las más importantes del dataset.

Las personas vivimos en sociedad, por tanto, la política y por ende nuestro gobierno afecta directamente a nuestra vida diaria, pero... ¿llega a ser un factor determinante en nuestra felicidad?, por otro lado, ¿el dinero nos da la felicidad?, y por último, si no cuidamos nuestra salud, ¿impactará en nuestra felicidad?

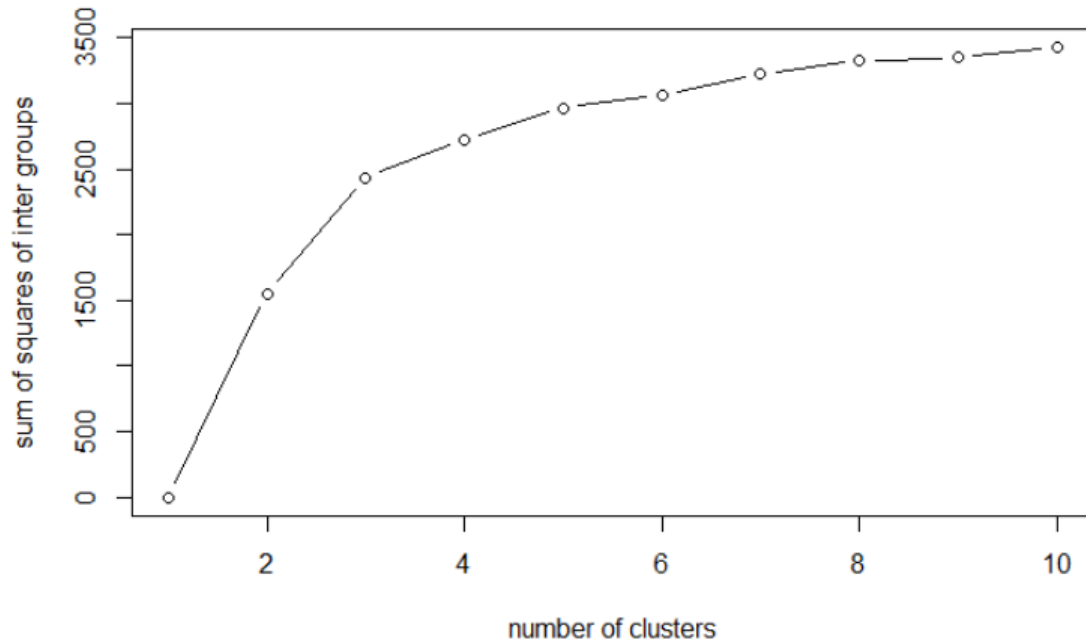
A priori, casi cualquier persona podría responder estas preguntas con facilidad, pero ¿qué nos responde la ciencia? A continuación, en esta memoria haré un estudio de cómo afectan estas tres variables en nuestra felicidad a partir de una técnica de Data Science reconocida como K-Means.

## ¿POR QUÉ K-MEANS?

K-Means es un algoritmo de clasificación no supervisada (clusterización) que tiene como función segmentar información de un dataset en  $k$  grupos basándose en las características o peculiaridades de los datos.

Como vemos, este algoritmo nos permite crear distintas agrupaciones o poblaciones según cómo se comporten los datos. De este modo, creo que resulta una técnica de interés para estudiar cómo se divide la gente cuando hablamos de su nivel de felicidad respecto a las variables que comentamos anteriormente. En resumen, con esta técnica podremos conocer si hay más gente que considere que el dinero les da la felicidad o si la población predominante es aquella que no, y lo mismo para las dos otras variables.

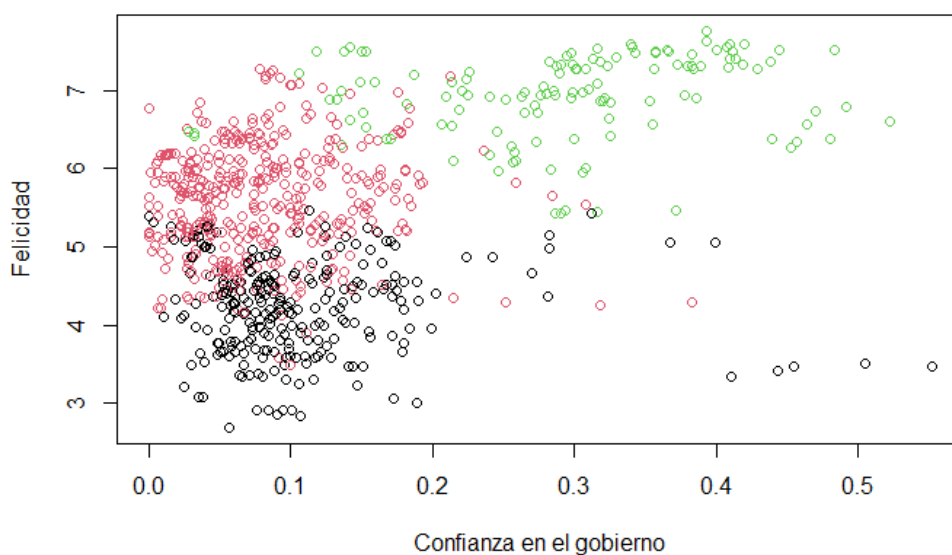
Antes de pasar a las preguntas, realicemos un análisis previo de cómo hemos aplicado K-Means para nuestro objeto de estudio. Primeramente, el primer paso sería el de obtener el número de clústers óptimos para realizar la clasificación. Como no hay forma de conocer el número óptimo de forma exacta, lo que se ha realizado es una función que nos devuelve una gráfica que nos da una pista de cuántos clústers integrupales pueden existir en nuestro dataset, siendo la gráfica resultante la siguiente:



Según he leído en diferentes fuentes de investigación, el número de clústers óptimo se encuentra donde la gráfica forma como el “pico de un codo”. Vemos que para nuestro caso ese pico se encuentra en el 3, luego establecemos ese número como el número de clústers que tendrá nuestro problema. Sabiendo esto, podemos dar paso a responder a las preguntas.

## ¿CÓMO INFLUYE LA CONFIANZA EN EL GOBIERNO EN LA FELICIDAD?

La gráfica resultante tras realizar el análisis de K-Means es el siguiente:



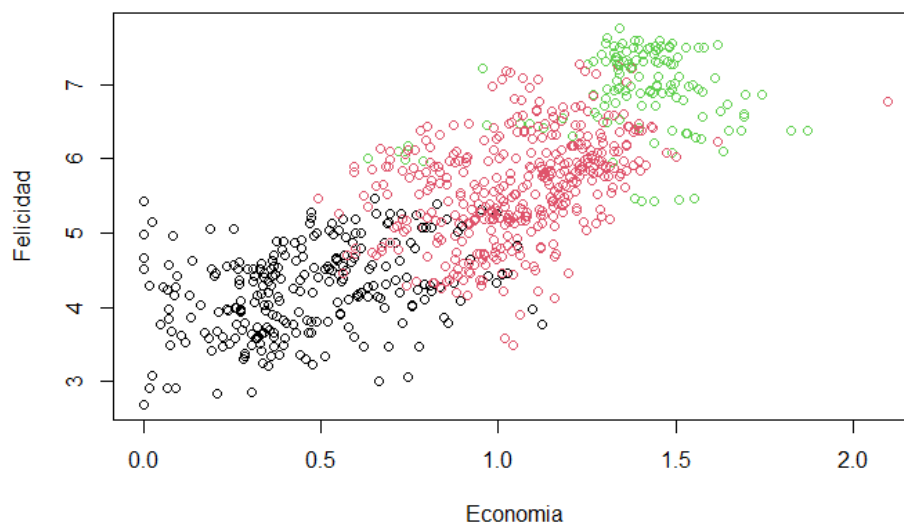
A primera vista podemos observar que el grupo predominante es el rojo, en el cual se identifica que la confianza en el gobierno no es un factor clave para su felicidad. Es decir, para la mayoría de personas de este dataset, el gobierno no determina su felicidad.

Por otro lado, el segundo grupo predominante, el cual es el negro, nos muestra que a menor confianza en el gobierno (o mayor corrupción) menos felices son. Y el último grupo, el cual es el verde, nos muestra que su alta felicidad no viene determinada por la política.

Por tanto, podemos observar que, para dos grupos de la clasificación, la confianza en el gobierno no es factor clave para su felicidad. De este modo, la conclusión que extraigo es que para la mayoría la política no determina su felicidad, al haber dos grupos que lo prueban.

## ¿CÓMO INFLUYE LA ECONOMÍA EN LA FELICIDAD?

La gráfica resultante tras realizar el análisis de K-Means es el siguiente:



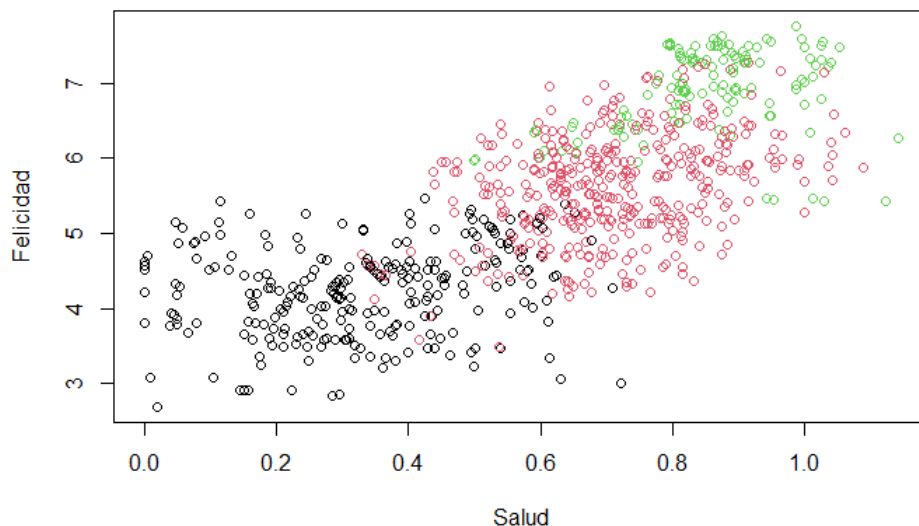
En este caso vemos un crecimiento lineal en la gráfica. Por un lado, en el clúster negro se refleja que a menor producción económica, menor es el rango de felicidad. Luego, el clúster rojo es el predominante, y muestra que la economía no impacta tanto en la felicidad de las personas de este grupo, simplemente existe una conformidad con una producción económica estable (ni riqueza, ni pobreza). Por último, vemos que el clúster más pequeño es el verde, y es aquel donde el nivel de economía sí determina en gran medida la felicidad de estas personas.

Por ende, extraigo como conclusión del análisis que el dinero no da tanto la felicidad, más bien, es la ausencia del dinero la que hace desvanecer cualquier atisbo de la misma, y esto es evidentemente, pues sin dinero no se pueden cubrir necesidades

básicas, y esto es lo que produce la pobreza. También observo que son pocas las personas que necesitan lujos para ser felices.

## ¿CÓMO INFLUYE LA SALUD EN LA FELICIDAD?

La gráfica resultante tras realizar el análisis de K-Means es el siguiente:



Para este caso vemos que se comporta de forma parecida al de la economía. Observamos que el clúster negro es menos feliz cuanto menos salud posee. Por otro lado, el clúster rojo demuestra que si se tiene una salud buena, la felicidad aumenta considerablemente. Por último, el clúster verde refleja que si se tiene una muy buena salud, la felicidad es bastante alta.

Claramente se extrae de este análisis que a mayor salud, mayor es la felicidad de las personas, y esto es obvio ya que normalmente muchas enfermedades y patologías vienen dadas por una mala salud tanto mental como física.

## CONCLUSIÓN

Ante las preguntas que me he realizado, creo que la ciencia ha respondido coherentemente, o al menos como yo me esperaba. Por un lado, hemos visto que la política no es determinante para ser feliz. Luego se ha comprobado que el dinero no da la felicidad, pero sí la quita cuando no lo hay. Y por último también vemos que la salud afecta a la felicidad por completo. Me parece que las respuestas se ajustan a lo que podría responder la mayoría de personas, por tanto, podemos observar la potencia que nos ofrece el Data Science y lo atractivo que resulta para estudiar comportamientos psicológicos como es el caso de la felicidad en las personas.