Javier Díaz Machado

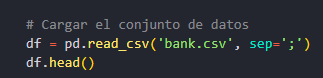
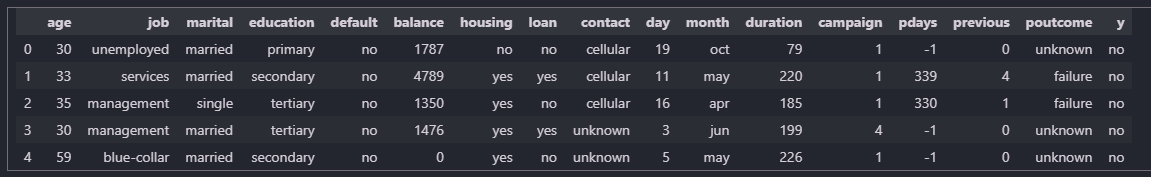
Segmentación de clientes según datos bancarios

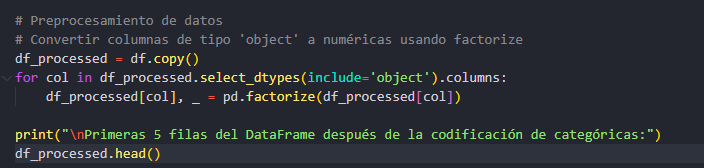
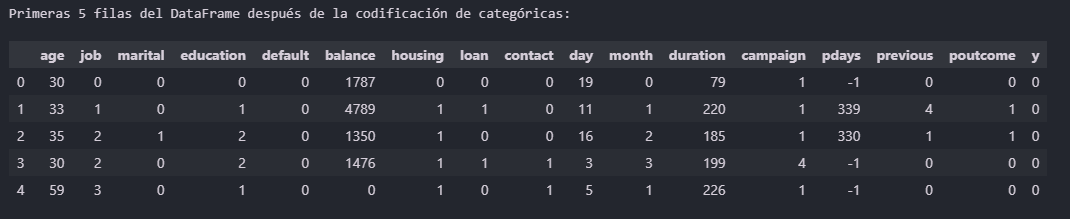
Link: [github](https://github.com/Javiert54/bigData/blob/main/CEIABD_SNS/UT4%20-%20Algoritmos%20y%20herramientas%20para%20el%20aprendizaje%20no%20supervisado/Actividad%204.1%20-%20Segmentación%20de%20clientes%20según%20datos%20bancarios/SNS_ACT4_1JavierDíazMachado.ipynb)

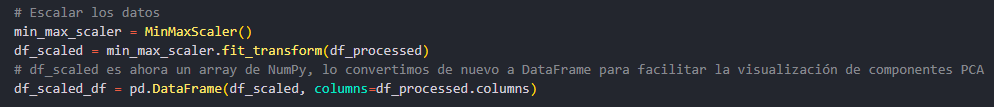
Carga del dataset

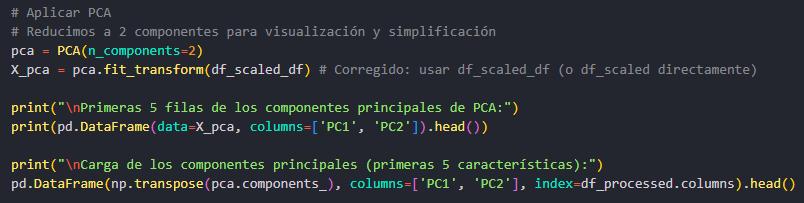
Decidí utilizar el dataset Bank Marketing para realizar esta tarea porque era el que me parecía más cómodo.

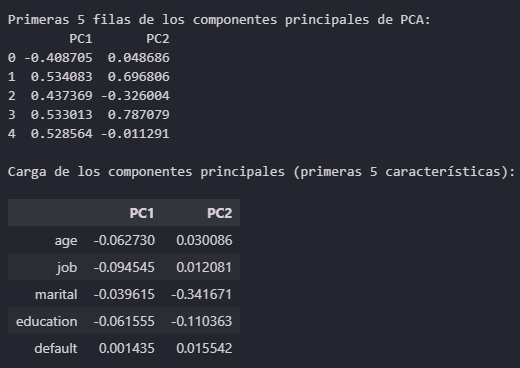
Se carga el dataset:

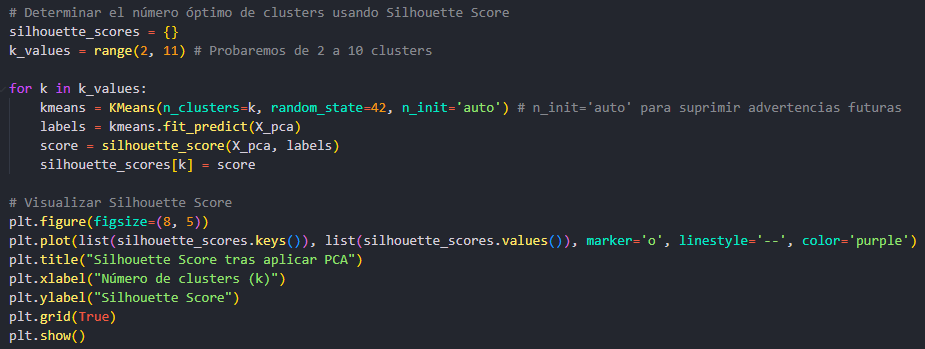


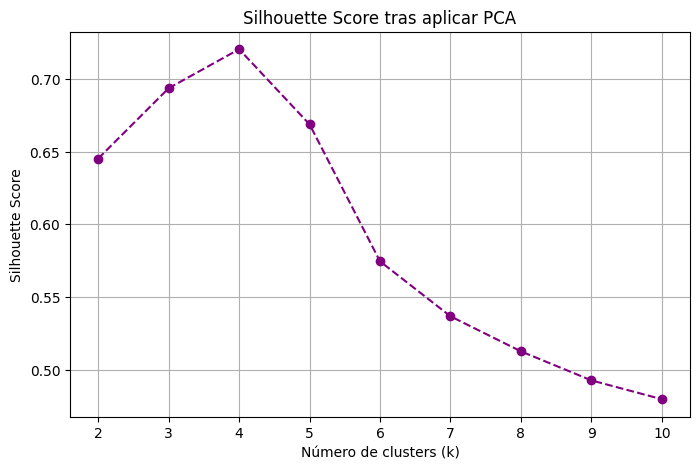
Convertimos todos los valores object a números:

Escalamos los datos:

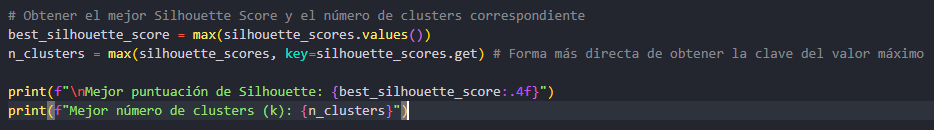
Aplicamos PCA:



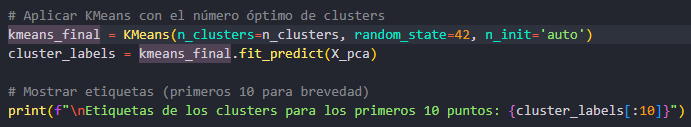
Sacamos el número óptimo de clusters usando SIlhoutte Score y mostramos un gráfico con este:



Obtener la mejor puntuación obtenida en Silhoutte score y el número de clusters correspondiente a la puntuación

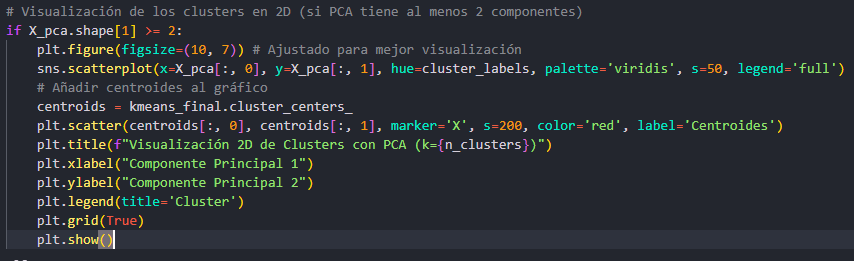


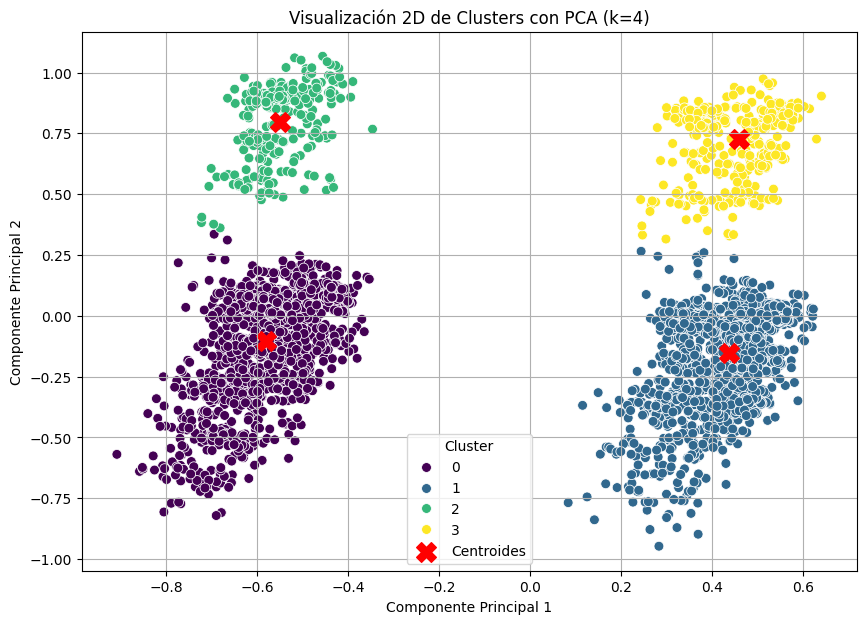
Agrupamos los datos en clusters utilizando **K-Means**, aplicando el número óptimo de grupos





Ponemos un gráfico para poder ver los datos agrupados por clusters:





Finalmente, añadimos los resultados y exportamos en un csv:

