**TALLER DE CLUSTER**

**PRESENTADO POR:**

**JESUS DAVID SUAREZ PEÑA**

**PRESENTADO AL DOCENTE:**

**ALVARO AGUSTIN OÑATE BOWEN**

**UNIVERSIDAD POPULAR DEL CESAR**

**FACULTAD DE INGENIERÍA Y TECNOLOGÍAS**

**INGENIERIA DE SISTEMAS**

**VALLEDUPAR - CESAR**

**2020**

**AGRUPACION DE DATOS**

Del conjunto de datos german credit se agrupo los siguientes atributos para desarrollar el algoritmo k-means los cueles son:

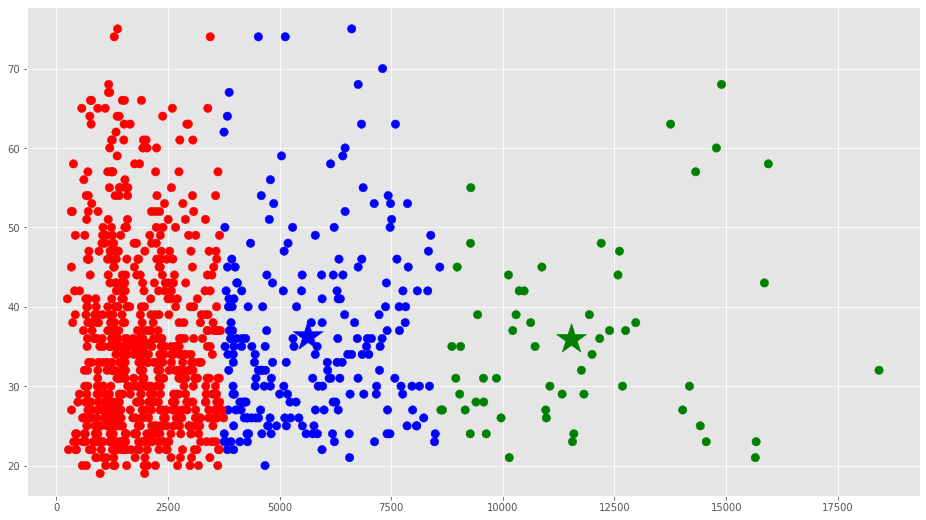
**MONTO DE CREDITO, EDAD, DURACION DE MES, TRABAJO, PROPOSITO, PLANES\_CUOTAS**

PARA UN K=3

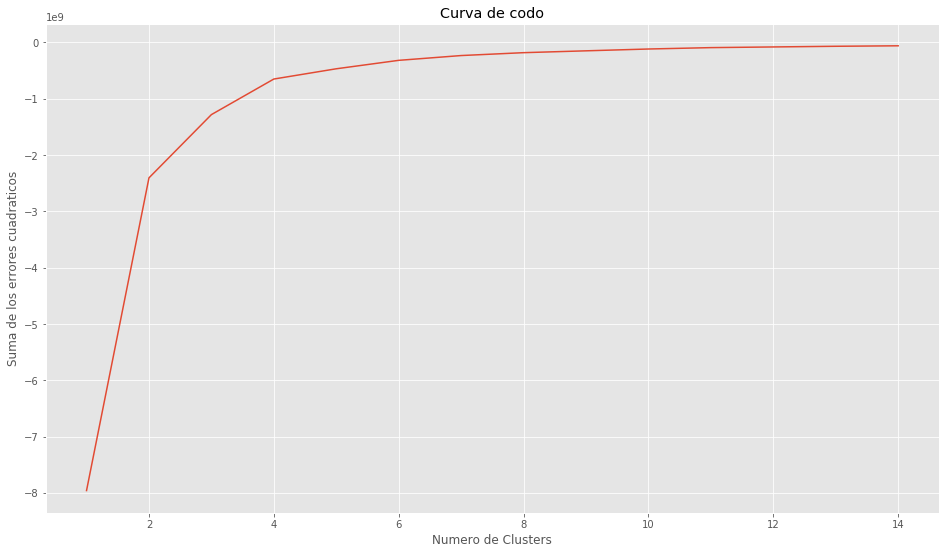
[[ 1871.74618585 35.30235784 16.65603329]

[11539.01694915 35.89830508 39.16949153]

[ 5640.57727273 36.25 29.92272727]]



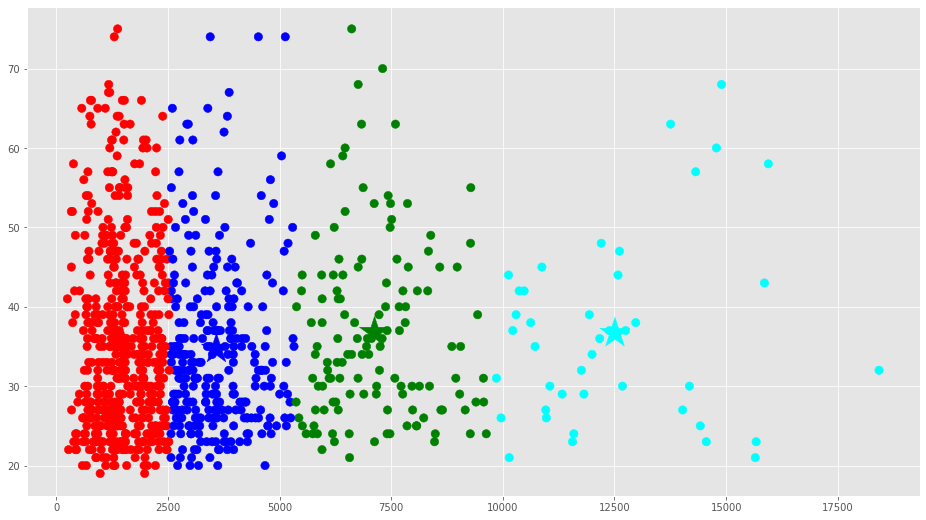
**Buscamos a K**



Como puede ver, el valor óptimo de k está entre 2 y 4, ya que la forma de codo se forma en k = 2 en el gráfico anterior

**Grafica cuando K=4**

visualizar los cuatro grupos que se formaron con el valor k óptimo. Puede ver claramente cuatro grupos en la imagen a continuación, con cada grupo representado por un color diferente.



[[ 1471.31066176 35.61029412 15.08272059]

[ 7127.52307692 36.8 33.34615385]

[ 3587.30633803 34.68309859 23.49295775]

[12511.71428571 36.66666667 40.26190476]]

Se considera K=4 como el k optimo por que se observa una mejor agrupación con sus centroides