
Calibración del Algoritmo de KMeans

A large, faint, light blue watermark of the Big Data Academy logo is centered in the background of the slide. It consists of a circular emblem with a stylized 'BDA' and the text 'BIG DATA ACADEMY' to its right.

Inercia de un clúster

Mide el grado de concentración de puntos en un clúster

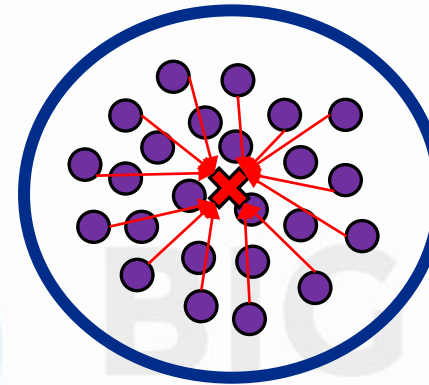
Fórmula

$$Inercia_{cluster} = \sum_{i=1}^n d(\text{centroide} \rightarrow \text{punto}_i)$$

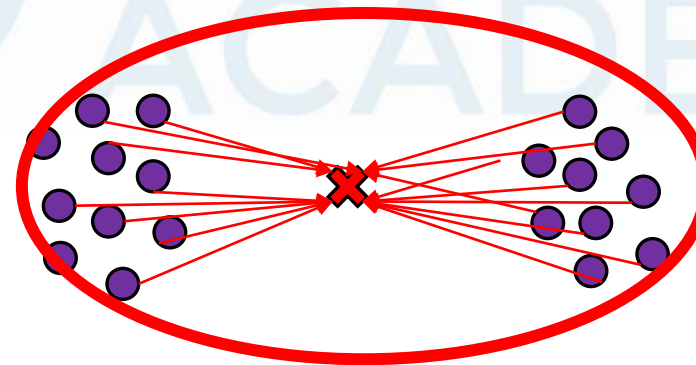
¿Qué valor de inercia es mejor?

Mientras más pequeña la inercia de un clúster, implica que los elementos se parecen más entre sí, **nos conviene inercia pequeñas**

Ejemplos:



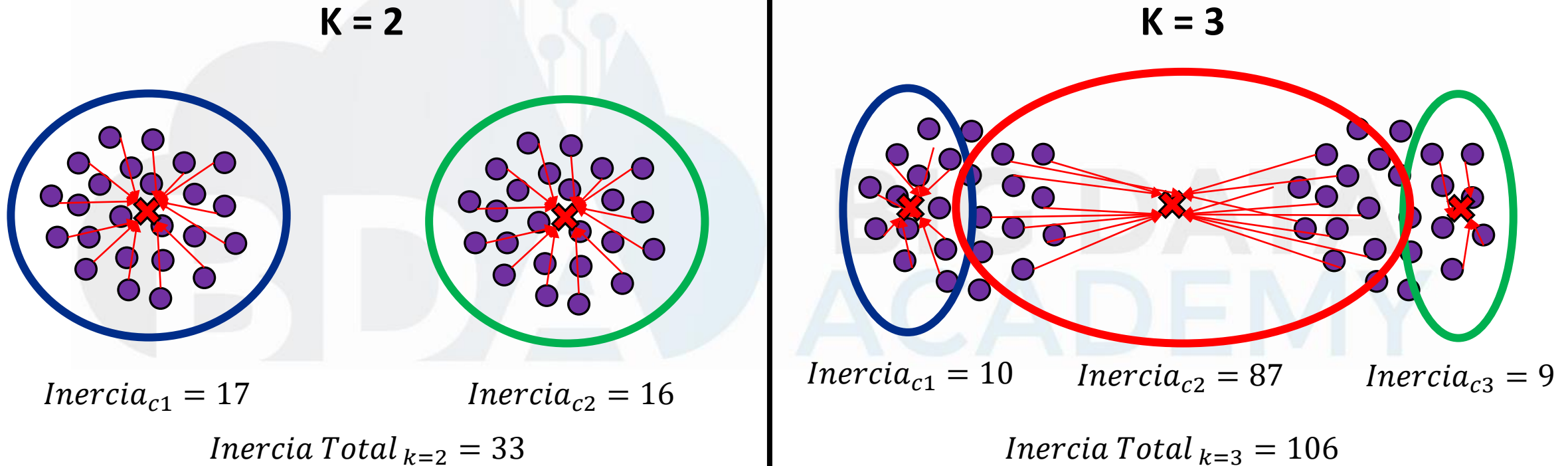
$$Inercia_{cluster\ 1} = 17$$



$$Inercia_{cluster\ 2} = 128$$

El “clúster 1” agrupa mejor los registros que el “clúster 2”

Inercia entre clases de una misma nube de puntos con dos valores de K diferentes



En esta nube de puntos, es mejor agrupar en K=2 que en K=3

Eligiendo el valor de “K” en función de la inercia

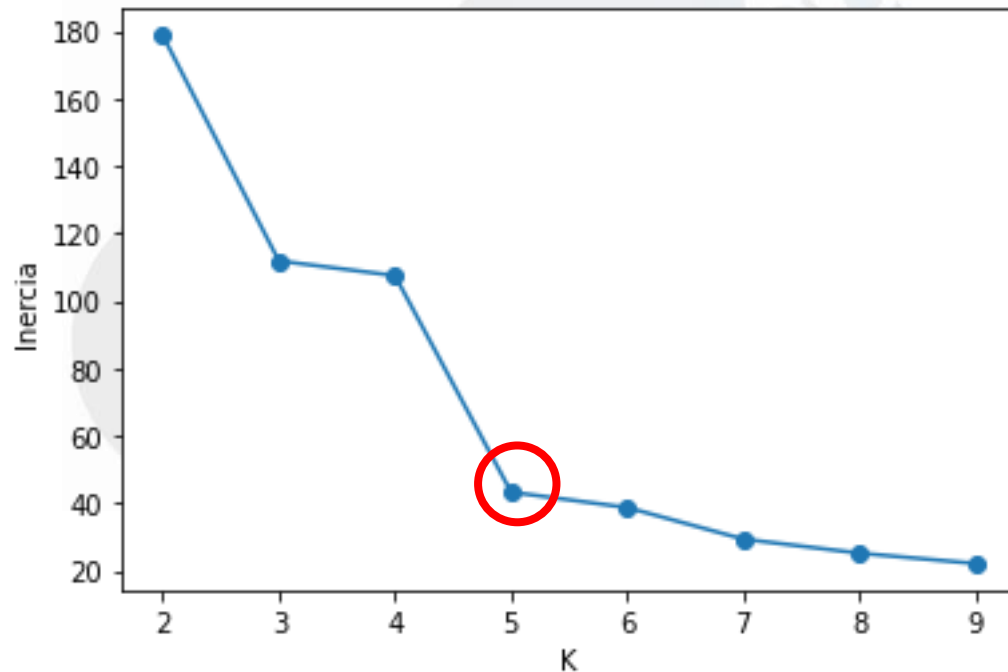


Gráfico de codo

Debemos de generar varios modelos de K-Means haciendo variar el valor de “K”, **nos quedaremos con el “K” en donde se encuentra la última disminución de inercia considerable.**