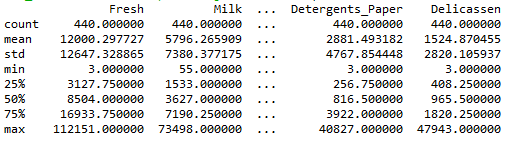
Laboratorio: Técnicas de agrupamiento

* + Análisis descriptivo de los datos.

El conjunto de datos se refiere a los clientes de un distribuidor de ventas al por mayor. Incluye el gasto anual en varias categorías de productos, esto se describe en el resultado de loa tabla creada de la función *describe* de pandas



* + Determinar el conjunto de modelización y el de validación.

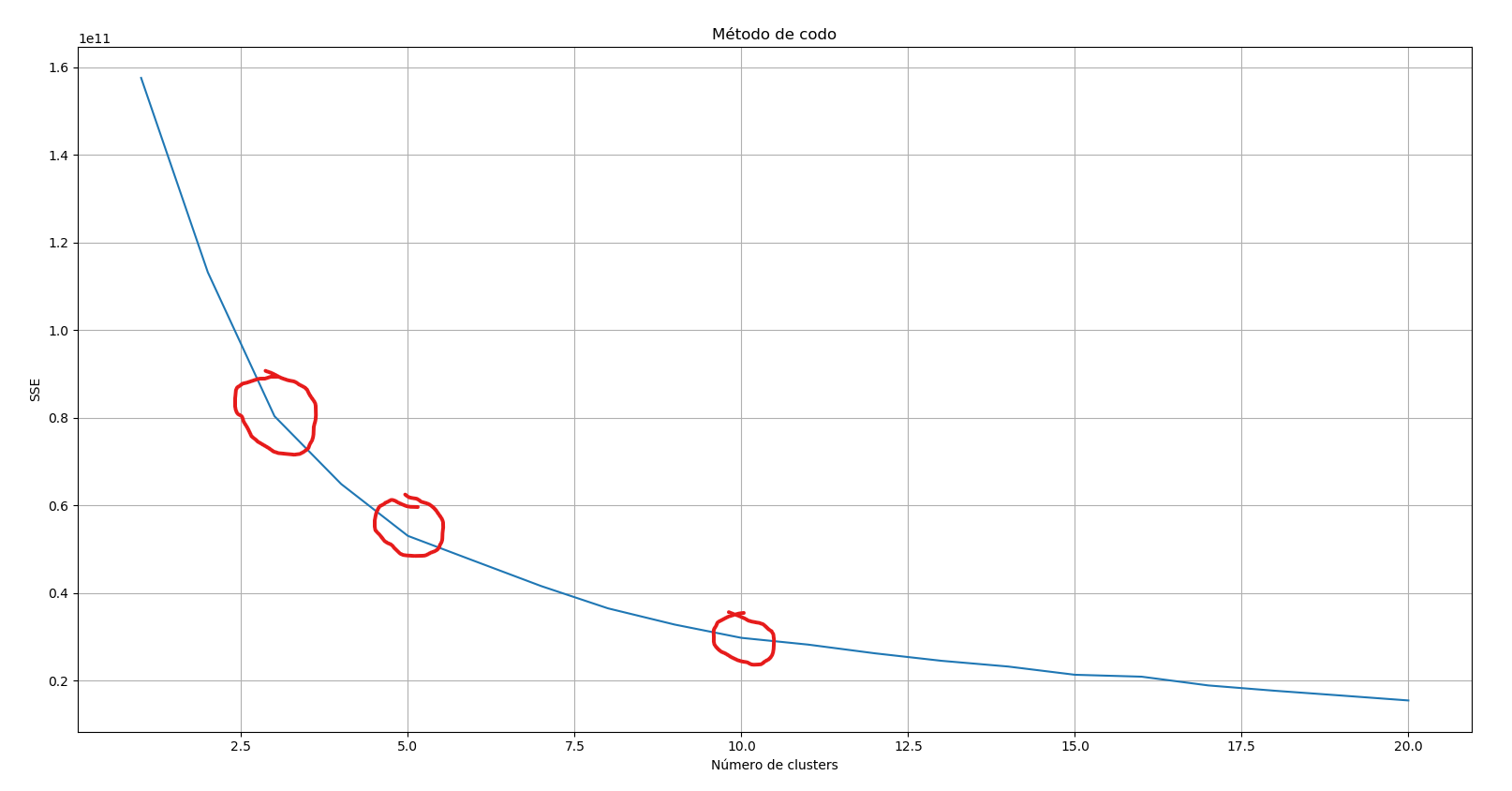
Se eliminaron las dos primeras columnas categóricas correspondientes a canal y región

* + Tratamiento de missing.

No se realiza tratamiento de missing, no se encuentran valores incompletos en el conjunto de datos

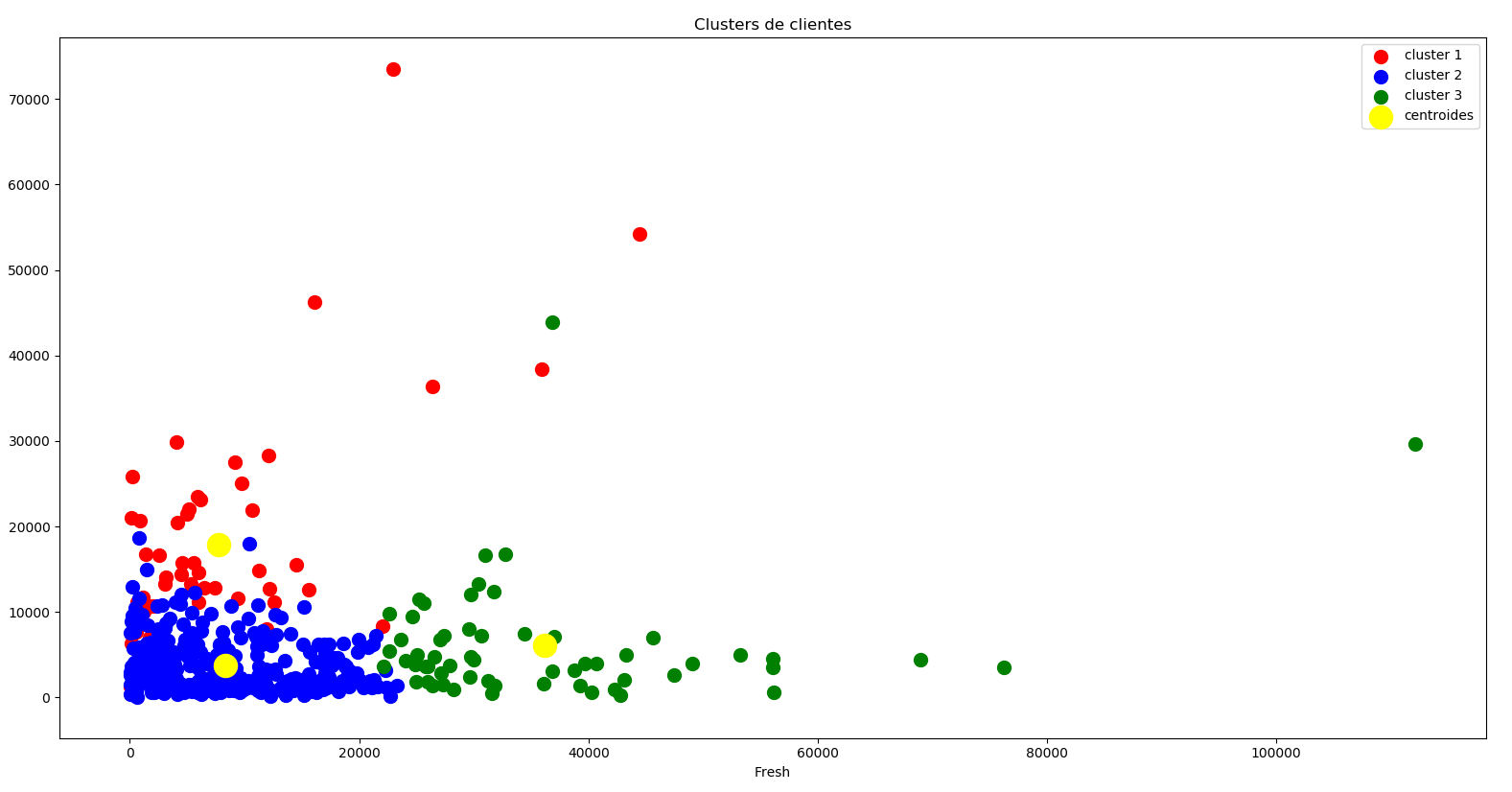
* + Calcular las métricas de evaluación de ajuste adecuadas.

Se evalúa con la gráfica de codo



Se eligen 3 clusters para una segmentación más sencilla de los clientes

* + Comentar los resultados obtenidos.



Se hizo una agrupación por solo dos dimensiones para demostrar el agrupamiento de los datos y se grafica los centroides, en este caso se puede utilizar el clustering para segmentar a los clientes por sus hábitos de consumo