K-means with K-means++ as the initialization

Data: Wholesale customers Data Set
https://www.kaggle.com/datasets/binovi/wholesale-customers-data-set

8 columns of 440 data
Values are very different
For example, channel (1, or 2) & regions (1, 2, or 3) have low magnitude,
but Fresh, Milk, Grocery, ect., have a higher magnitude.

Use **StandardScaler** //standardized the data from sklearn
https://ithelp.ithome.com.tw/m/articles/10265253

How many clusters?
Try cluster numbers 2 ~ 20, then sketch a plot (cluster heterogeneity vs cluster numbers) to decide the best cluster numbers.
Explain why you think it is the best cluster number. For example, you can use **.value_counts( )** to explore the clusters, or other smart method to support your answer.

There are two main codes: k-means++ initialization and k-means
You are strongly recommend to write your own. However, you are also allowed to only write one and used the function provided for the other. Please indicate which is your own code.
Due: 4/24 oral check: 5/01
(碩二同學可以都使用既有 functions)