**ADDITIONAL NOTES FOR JULY 25th 2020 CLASS**

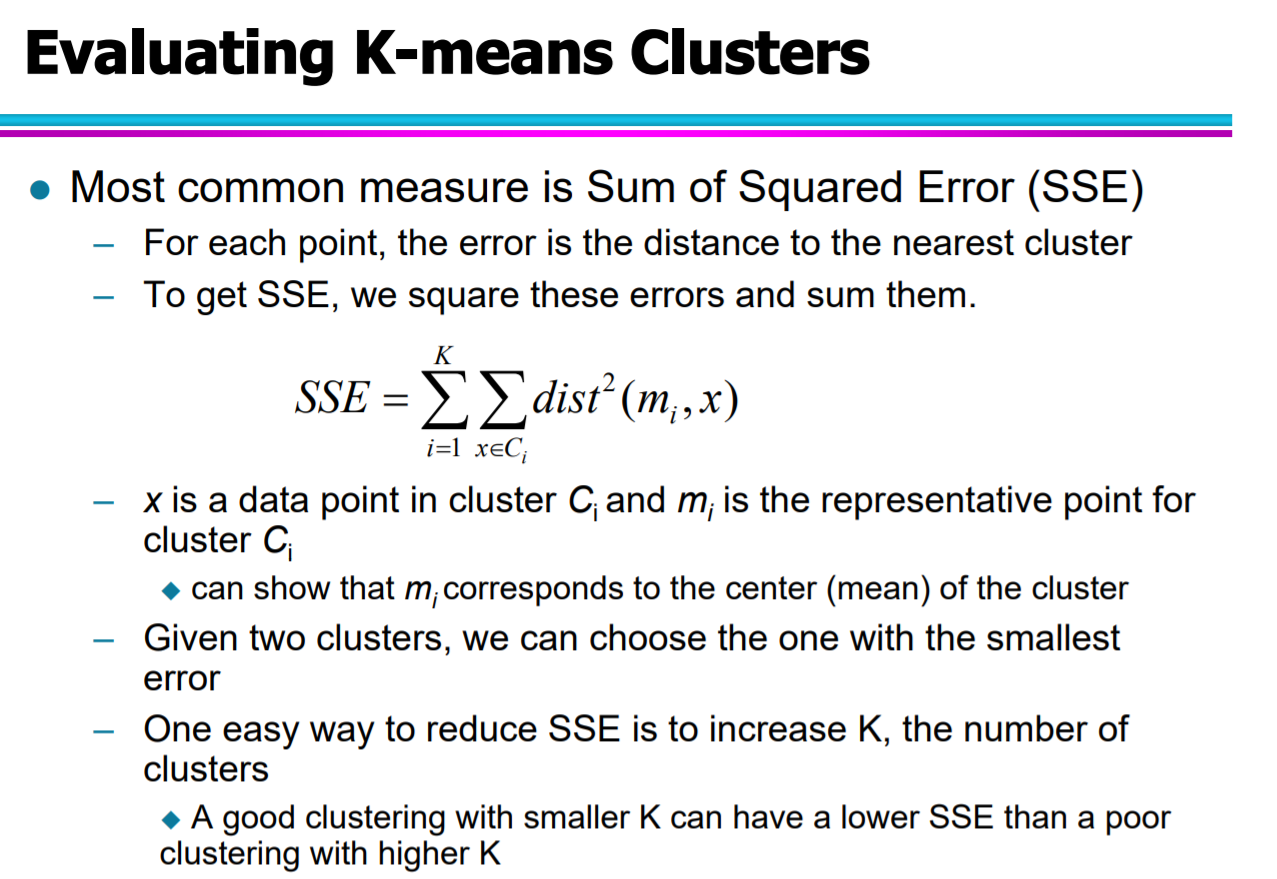
Partition Around Medoids (PAM)

PAM is just one kind of K-medoids algorithm. The difference is in new medoid selection (per iteration):

K-medoids selects object that is closest to the medoid as a next medoid\*\*\*

PAM tries out all of the objects in the cluster as a new medoid that will lead to lower SSE.

Error Sum of Squares (SSE) SSE is the sum of the squared differences between each observation and its group's mean. It can be used as a measure of variation within a cluster. If all cases within a cluster are identical the SSE would then be equal to 0.



**CLARA (Clustering Large Applications, (Kaufman and Rousseeuw 1990)) is an extension to k-medoids (PAM) methods to deal with data containing a large number of objects (more than several thousand observations) in order to reduce computing time and RAM storage problem.**