DEFINITION

K-means clustering is a simple unsupervised learning algorithm that is used to solve clustering problems. It follows a simple procedure of classifying a given data set into a number of clusters, defined by the letter "k," which is fixed beforehand.

There are many types of clustering , k clustering is under exclusive clustering.

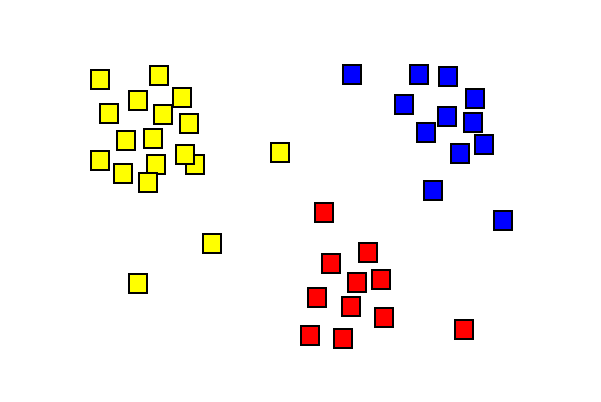
K here refers to the number clusters which is required.

K-means clustering is a method used for clustering analysis, especially in data mining and statistics. It aims to partition a set of observations into a number of clusters (k), resulting in the partitioning of the data into Voronoi cells. It can be considered a method of finding out which group a certain object really belongs to.

It is used mainly in statistics and can be applied to almost any branch of study. For example, in marketing, it can be used to group different demographics of people into simple groups that make it easier for marketers to target. Astronomers use it to sift through huge amounts of astronomical data; since they cannot analyze each object one by one, they need a way to statistically find points of interest for observation and investigation.

The algorithm:

1. K points are placed into the object data space representing the initial group of centroids.
2. Each object or data point is assigned into the closest k.
3. After all objects are assigned, the positions of the k centroids are recalculated.
4. Steps 2 and 3 are repeated until the positions of the centroids no longer move.



Bibliography:

* <https://www.youtube.com/watch?v=1XqG0kaJVHY>
* <https://www.techopedia.com/definition/32057/k-means-clustering>
* <https://brilliant.org/wiki/k-means-clustering/>