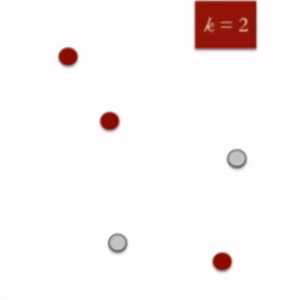
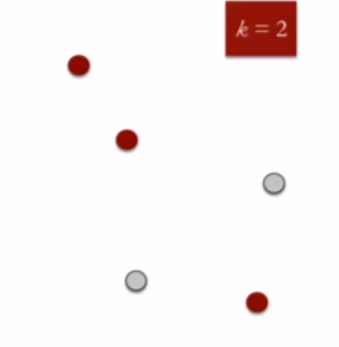
**K Means Clustering**

K means is an iterative clustering algorithm that aims to find local maxima in each iteration. This algorithm works in these 5 steps :

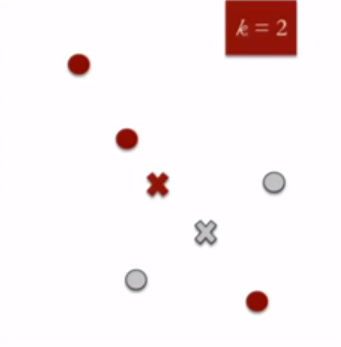
1. Specify the desired number of clusters K : Let us choose k=2 for these 5 data points in 2-D space.

[](https://www.analyticsvidhya.com/wp-content/uploads/2016/11/clustering-2.png)

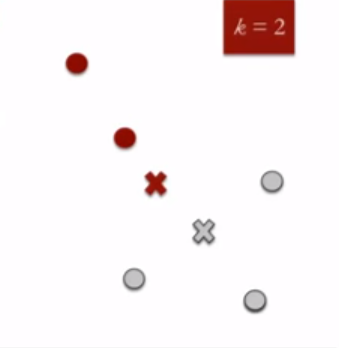
1. Randomly assign each data point to a cluster : Let’s assign three points in cluster 1 shown using red color and two points in cluster 2 shown using grey color.

[](https://www.analyticsvidhya.com/wp-content/uploads/2016/11/clustering-2-1.png)

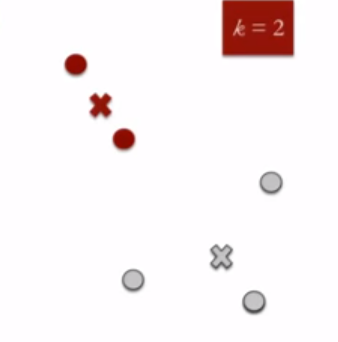
1. Compute cluster centroids : The centroid of data points in the red cluster is shown using red cross and those in grey cluster using grey cross.

[](https://www.analyticsvidhya.com/wp-content/uploads/2016/11/clustering-3.png)

1. Re-assign each point to the closest cluster centroid : Note that only the data point at the bottom is assigned to the red cluster even though its closer to the centroid of grey cluster. Thus, we assign that data point into grey cluster

[](https://www.analyticsvidhya.com/wp-content/uploads/2016/11/clustering-4.png)

1. Re-compute cluster centroids : Now, re-computing the centroids for both the clusters.

[](https://www.analyticsvidhya.com/wp-content/uploads/2016/11/clustering-5.png)

1. Repeat steps 4 and 5 until no improvements are possible : Similarly, we’ll repeat the 4th and 5th steps until we’ll reach global optima. When there will be no further switching of data points between two clusters for two successive repeats. It will mark the termination of the algorithm if not explicitly mentioned.