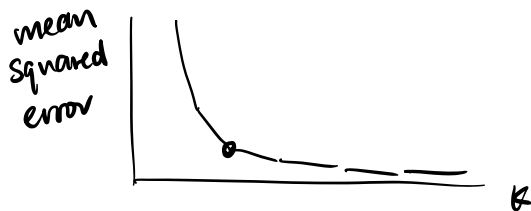


General Notes (K Means)

- Unsupervised learning for classification

- Choosing K using the elbow method:

Run K -means algo on $K \in [1, 10]$, graph the mean squared error. The "elbow" K value is the best choice



K - Means Algorithm

1. initialize centroids randomly for each K .
2. sort data based to the centroid closest to them (have the closest distance)
3. Update centroids to be the mean of all data pts sorted to that centroid
4. Repeat 2-3 until no more updates while re-sorting

General Notes (KNN)

- Both a classifier & a regressor!
- Intuition: look at the K nearest neighbor of an unknown data point & assign it the most common class of its neighbors $\rightarrow$ classification

KNN Regressor just takes the mean of the closest K neighbors