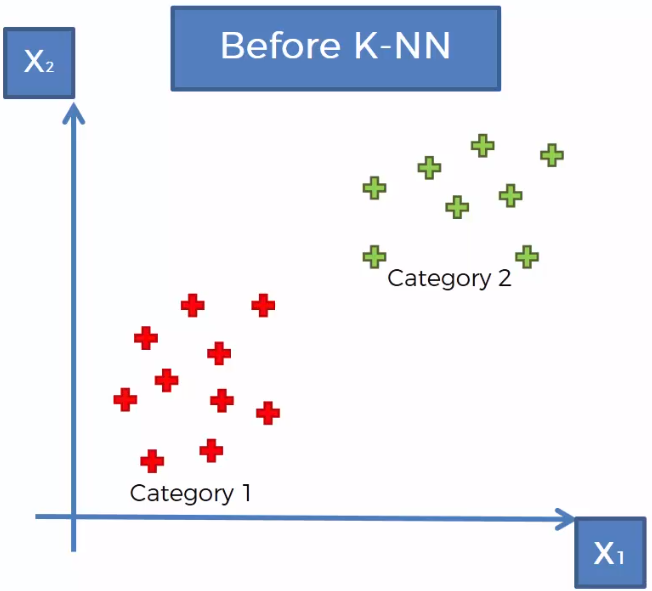
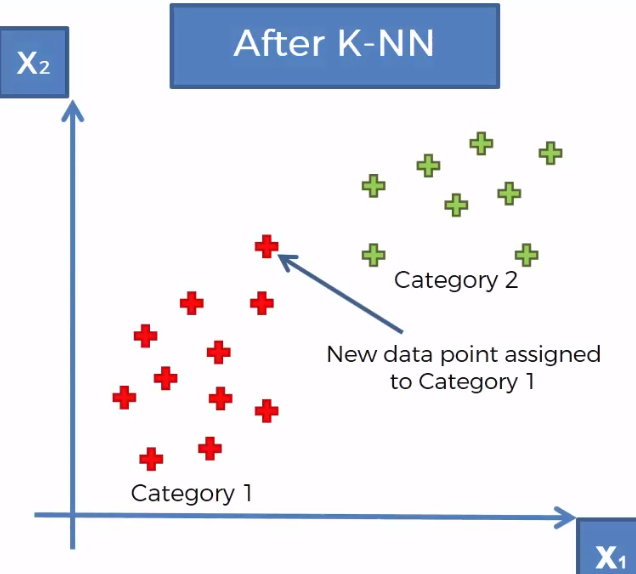
***K-Nearest Neighbors (KNN)***

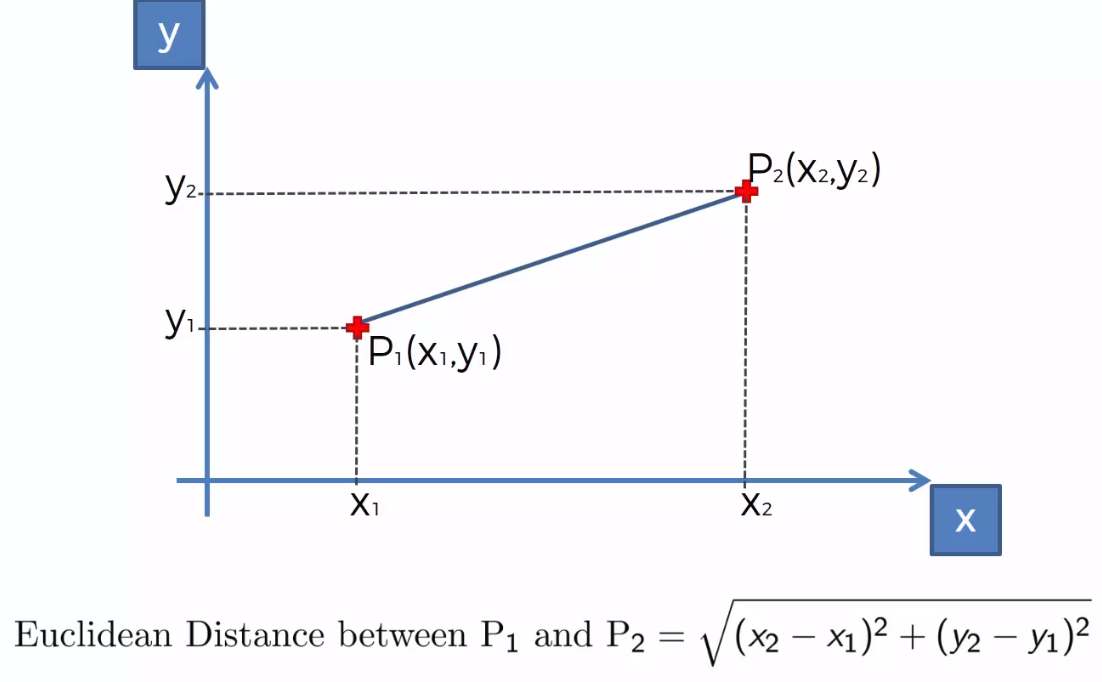
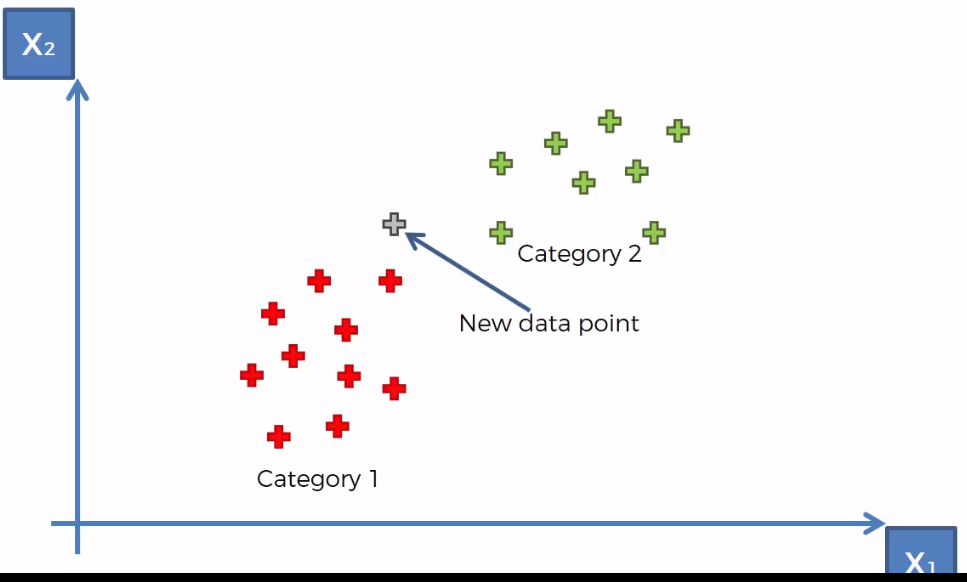
* We have 2 categories in the dataset:



* When we add a new DP, which category should it go in? KNN will help predict/decide this



* Step 1 = Choose the number of neighbors, K
* Common default = 5
* Step 2 = When we get a new DP, find the K nearest neighbors to that DP via the **Euclidean distance**
* If we don’t have the Euclidean distance, can use the **Manhattan distance,** or another
* Step 3 = Among the K nearest neighbors, count the DP’s in each category
* Step 4 = Assign new DP to the neighbor category w/ the highest number of DP’s counted





* So above, we assign the new DP to the red category (category 1)