- In unsupervised learning, k = number of clusters. It comes from your domain understanding.
- We determine centroids, find Euclidian distance for every data point, and assign them to each centroid.
- Silhouette Score will give you a value between -1 and 1. A higher value is better.