



## With Two Clusters:

We quickly discovered there was no linearity between the number of clusters we used and the subsequent silhouette score – we initially tried gradually increasing the number of clusters until we discovered the optimal amount only to discover that each increase gave an unpredictable result. One increase gave a higher silhouette score, then another increase gave a lower, then higher, lower, etc. Based on this, we concluded that splitting the data points evenly in half with 2 clusters would give the highest score:

