Clustering

3.1 Clusters

Clustering means picking a certain characteristic and using it to _____ _____.

Clustering is always subjective because _____.

While there are many different use cases of clustering, the key aspect is _____ _____.

In clustering, a vector of length 2 means that the data point represented by the vector has _____ _____.

3.2 Hierarchical Clustering

In every iteration of a hierarchical clustering algorithm, _____ _____.

Calculating a "centroid" is only possible in Euclidean spaces because _____ _____.

Using R as a programming language for hierarchical clustering gives you access to _____ _____.

A dendrogram is _____.

3.3 Point assignment

As the k-means clustering algorithm is dependent on centroids, _____.

To use the k-means algorithm, you first have to define _____.

When choosing k, the advantage of silhouettes over the elbow criterion is that _____ _____.

3.4   Modeling Decisions

Feature selection is the process of _____.

Applying normalization or standardization to variables is often necessary to _____
.

Scaling can be used to decrease and increase _____
.