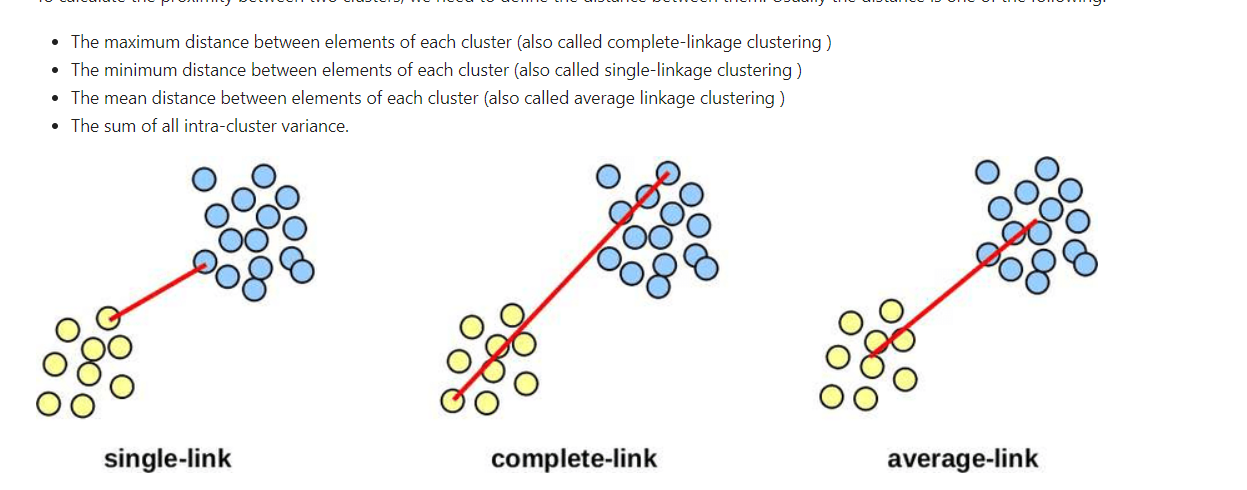
Hierarchical clustering is a method that seeks to build a hierarchy of clusters. Strategies for hierarchical clustering generally fall into two types:

* Agglomerative: This is a "bottom-up" approach: each observation starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy.
* Divisive: This is a "top-down" approach: all observations start in one cluster, and splits are performed recursively as one moves down the hierarchy.

In general, the merges and splits are determined in a greedy manner, and the results of hierarchical clustering are usually presented in a dendrogram.



Note: Euclidean, Manhattan, Mahalanobis, etc. distance formulas can be used when calculating distances for each of the above.

To calculate the proximity between two clusters, we need to define the distance between them. Usually, the distance is one of the following:

* The maximum distance between elements of each cluster (also called complete-linkage clustering )
* The minimum distance between elements of each cluster (also called single-linkage clustering )
* The mean distance between elements of each cluster (also called average linkage clustering
* The sum of all intra-cluster variance.

Checking to see that there are no more missing values:

DATA. isnull ( ) . sum( )

PCA in LAB 11