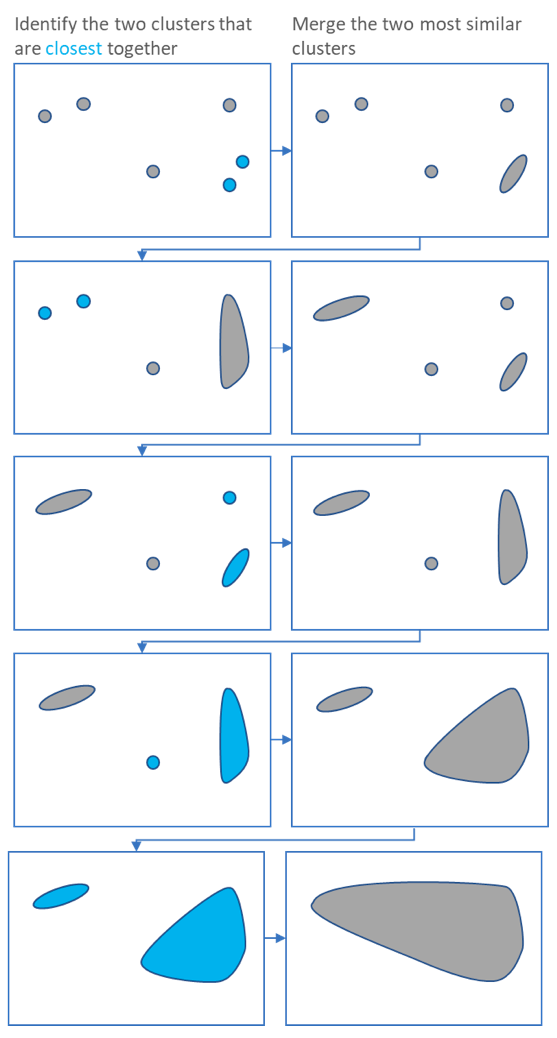
***Hierarchical clustering***, also known as hierarchical cluster analysis, is an algorithm that groups similar objects into groups called clusters. The endpoint is a set of clusters, where each cluster is distinct from each other cluster, and the objects within each cluster are broadly similar to each other.

Required data

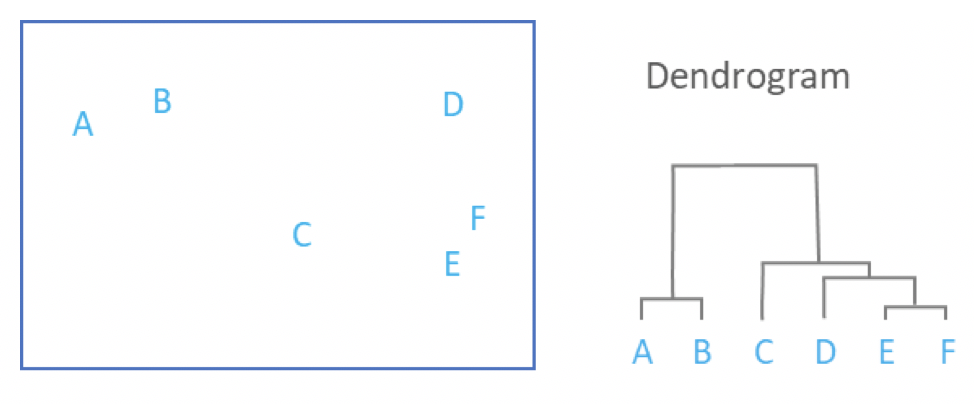
Hierarchical clustering can be performed with either a *distance matrix*or *raw data.* When raw data is provided, the software will automatically compute a distance matrix in the background. The distance matrix below shows the distance between six objects.

## How hierarchical clustering works

Hierarchical clustering starts by treating each observation as a separate cluster. Then, it repeatedly executes the following two steps: (1) identify the two clusters that are closest together, and (2) merge the two most similar clusters. This iterative process continues until all the clusters are merged together. This is illustrated in the diagrams below.



The main output of Hierarchical Clustering is a *[dendrogram](https://www.displayr.com/what-is-dendrogram/" \t "_blank)*, which shows the hierarchical relationship between the clusters:



## Measures of distance (similarity)

In the example above, the distance between two clusters has been computed based on the length of the straight line drawn from one cluster to another. This is commonly referred to as the Euclidean distance. Many other distance metrics have been developed.

The choice of distance metric should be made based on theoretical concerns from the domain of study. That is, a distance metric needs to define similarity in a way that is sensible for the field of study. For example, if clustering crime sites in a city, city block distance may be appropriate. Or, better yet, the time taken to travel between each location. Where there is no theoretical justification for an alternative, the Euclidean should generally be preferred, as it is usually the appropriate measure of distance in the physical world.