To perform agglomerative hierarchical cluster analysis on a data set :

1. **Find the similarity or dissimilarity between every pair of objects in the data set.** In this step, you calculate the distance between different nodes in the graph and store it in a proximity matrix
2. **Group the objects into a binary, hierarchical cluster tree:**  In this step, you link pairs of nodes that are in close proximity single linkage algorithm. As nodes are paired into binary clusters, the newly formed clusters are grouped into larger clusters until a hierarchical tree is formed.
3. **Determine where to cut the hierarchical tree into clusters:** Here we specify the number of clusters we want in the output , which is equivalent to specifying where the dendogram should be cut.

References:

1. <http://www.mathworks.com/help/stats/hierarchical-clustering.html#bq_679x-11>
2. Legendre, P. & Legendre, L. 1998. Numerical Ecology. Second English Edition. 853 pages.

3 R. Sibson (1973). ["SLINK: an optimally efficient algorithm for the single-link cluster method"](http://www.cs.gsu.edu/~wkim/index_files/papers/sibson.pdf). *The Computer Journal* (British Computer Society) **16** (1): 30–34.[doi](http://en.wikipedia.org/wiki/Digital_object_identifier):[10.1093/comjnl/16.1.30](http://dx.doi.org/10.1093%2Fcomjnl%2F16.1.30).

Algorithm for merging clusters in agglomerative clustering:

Single linkage clustering

The clusters are sequentially combined into larger clusters, until all elements end up being in the same cluster. At each step, the two clusters separated by the shortest distance are combined. We have taken the distance metric as the Euclidean distance.

Pictures: Single linkage clustering

In the Matlab example :

All the distances are < 5 . We have taken 10 as Inf for plotting the dendogram in Matlab.