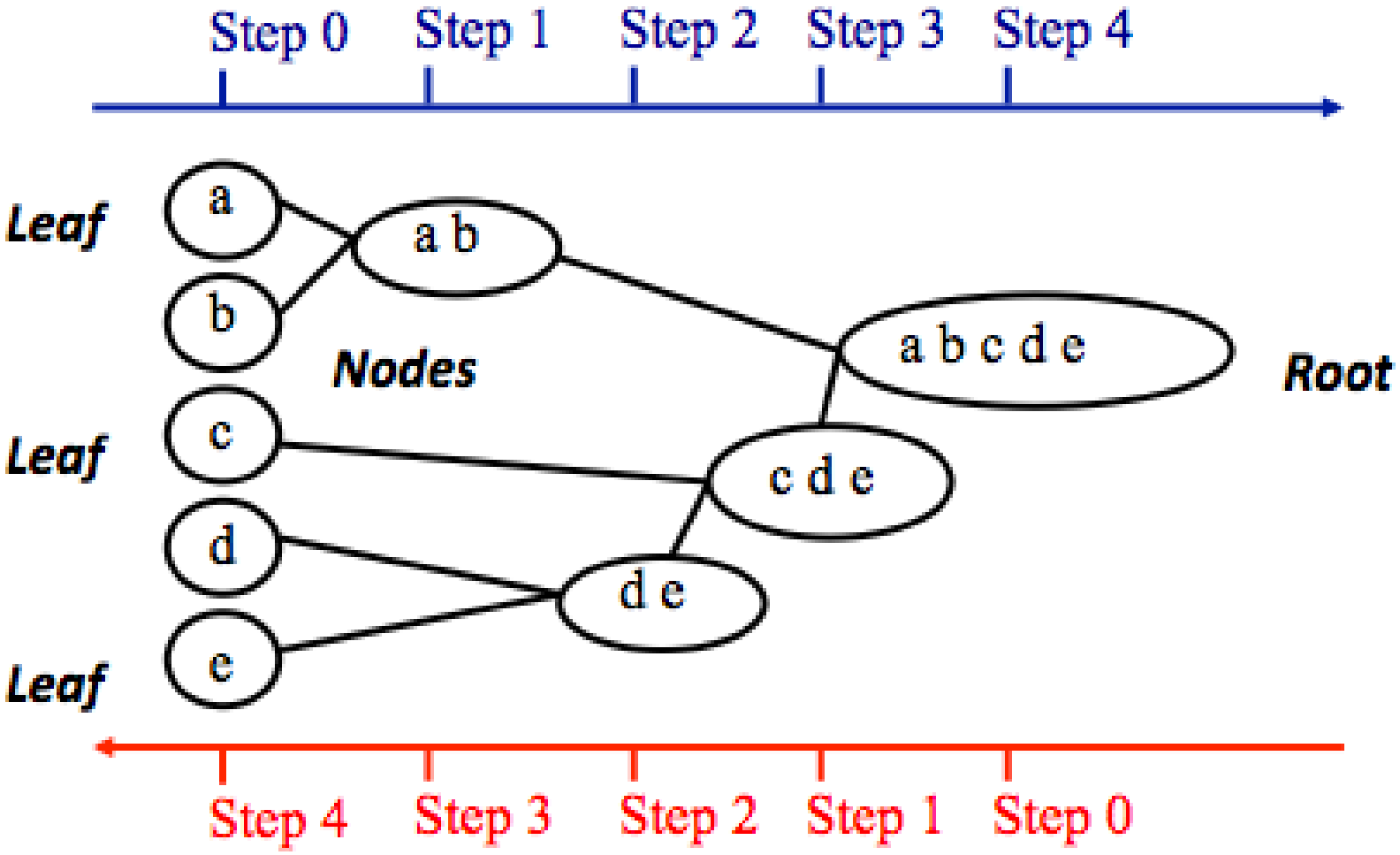
Hierarchical Clustering

BY MG ANALYTICS

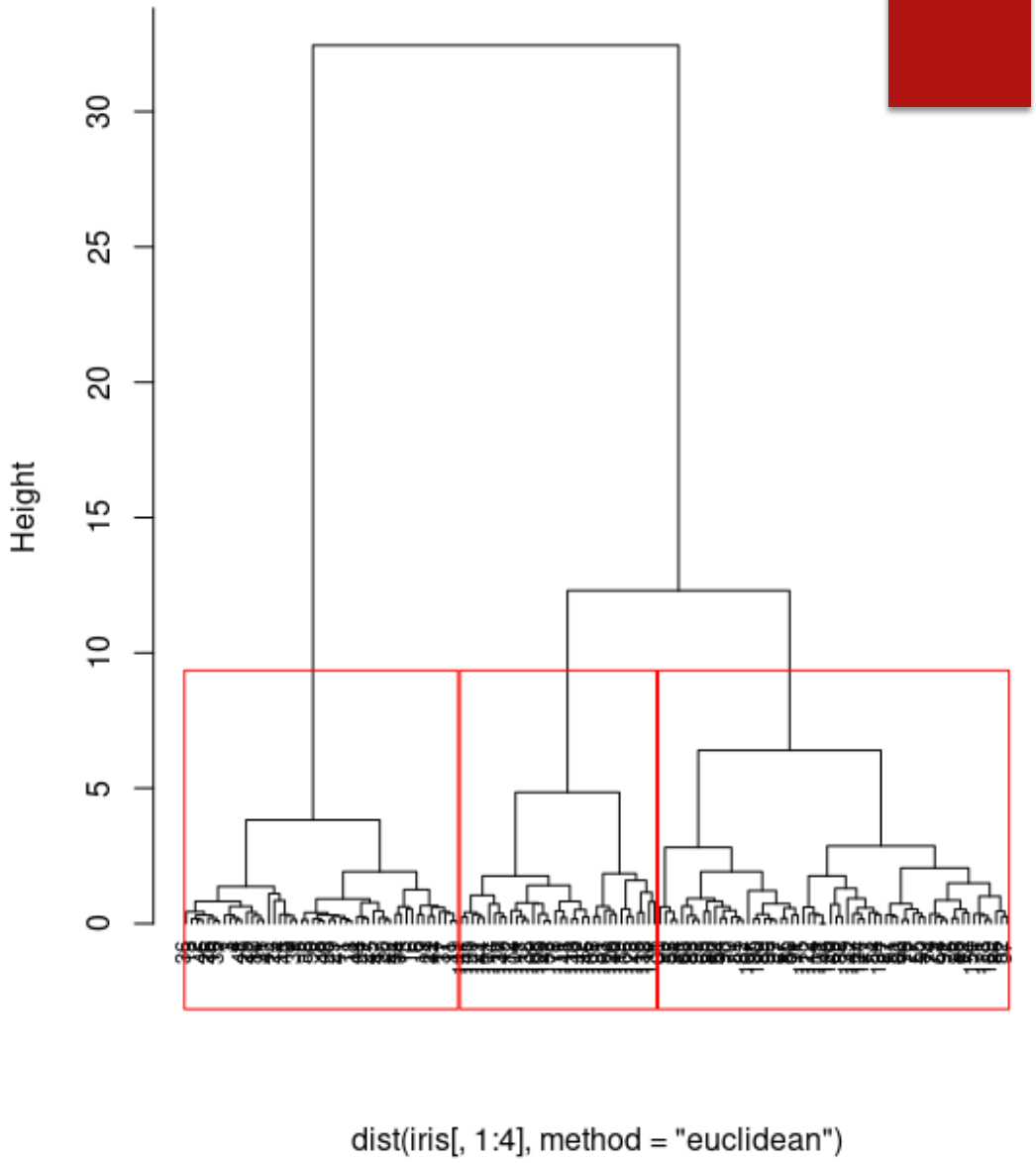
Agglomerative clustering

- It creates a hierarchy of clusters, and presents the hierarchy in a dendrogram.
- This method does not require the number of clusters to be specified at the beginning.
- Distance connectivity between observations is the measure.
- This is a bottom-up approach.
- Each observation starts in its own cluster.
- the similarity (or the distance) between each cluster is computed and then we merge the two most similar ones at each iteration until there is only one cluster left.

MG ANALYTICS
MG ANALYTICS


Step 0
Step 1
Step 2
Step 3
Step 4
Leaf
a
b
Leaf
c
d
Leaf
e
a b
Nodes
c d e
d e
a b c d e
Root
Step 4
Step 3
Step 2
Step 1
Step 0

- A tree that shows how clusters are merged/split hierarchically
- Each node on the tree is a cluster
- Each leaf node is a singleton cluster



- Computationally heavy.
- Suitable for smaller data set.
- Gives same clusters every time.
- No need to provide number of clusters at the beginning
- Difficulty handling different sized clusters and irregular shapes.