HIERARCHICAL CLUSTERING

Hierarchical clustering is a type of unsupervised learning method used in data analysis to group similar objects into clusters based on their similarities. Unlike partitioning methods like k-means, hierarchical clustering does not require predefining the number of clusters. Instead, it creates a hierarchy of clusters that can be visualized as a tree-like diagram called a dendrogram.

TYPES OF HIERARCHICAL CLUSTERING

1. **Agglomerative (Bottom-Up):**

- Starts with each data point as its own cluster.
- Gradually merges clusters that are closest together based on a distance metric.
- Continues until all points are merged into a single cluster.

2. **Divisive (Top-Down):**

- Starts with all data points in a single cluster.
- Recursively splits the clusters into smaller ones based on differences.
- Continues until each data point is in its own cluster.

Ques. Consider data points given. Apply single linkage algorithm (agglomerative method) construct hierarchical clustering.

OBJECTS	X	Y
0	1	4
1	2	8
2	5	10
3	12	18
4	14	28

Sol :-

STEP 1: Find all the distances between all the points

Distance between 0 and 1

Distance $=\sqrt{((x2-x1)^2 + (y2-y1)^2)}$

$=\sqrt{((2-1)^2 +(8-4)^2)}$

$=\sqrt{17} = 4.12$

Cluster	0	1	2	3	4
0	--	4.12	7.21	17.80	27.29
1	--	--	3.60	14.14	23.32
2	--	--	--	10.63	20.12
3	--	--	--	--	10.19
4	--	--	--	--	--

STEP 2 : In the agglomerative table , the minimum distance is 3.60. Therefore, the cluster 1 and 2 are merged together.

Cluster	{1,2}	0	3	4
{1,2}	--	4.12	10.63	20.12
0	--	--	17.80	27.29
3	--	--	--	10.19
4	--	--	--	--

Find distance from {1,2} to difference points from {1,2} to 0

- Minimum of { (1,0) ,(2,0) }
- Minimum of { 4.12 , 7.21 } = 4.12

STEP 3 : The minimum distance in the agglomerative table is 4.12. Therefore, cluster {1,2} and 0 are merged together.

Cluster	{0,1,2}	3	4
{0,1,2}	--	10.63	20.12
3	--	--	10.19
4	--	--	--

STEP 4 : In the agglomerative table the minimum distance is 10.19. Therefore cluster 3 and 4 are merged together.

Cluster	{0,1,2}	{3,4}
{0,1,2}	--	10.63
{3,4}	--	--

Dendrogram :-