

hierarchical-clustering

Data

Product	Sales
1	10
2	7
3	28
4	20
5	35

$$\sqrt{(10-7)^2} = \sqrt{9} = \mathbf{3}$$

$$\sqrt{(10-28)^2} = \sqrt{324} = \mathbf{18}$$

Data

Product	Sales
1	10
2	7
3	28
4	20
5	35

Proximity Matrix

ID	1	2	3	4	5
1	0	3	18	10	25
2	3	0	21	13	28
3	18	21	0	8	7
4	10	13	8	0	15
5	25	28	7	15	0

Step 1: Assign all the data points to an individual cluster

Data

Product	Sales
1	10
2	7
3	28
4	20
5	35

10

7

28

20

35

1

2

3






4

5

Step 1: Check the smallest distance proximity matrix and merge the data points

Proximity Matrix

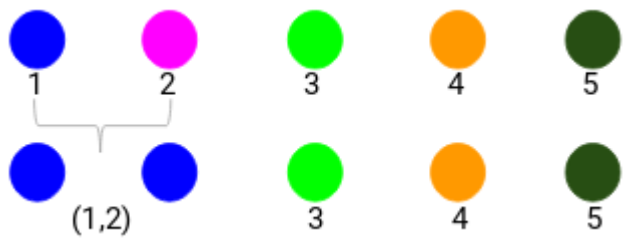
ID	1	2	3	4	5
1	0	3	18	10	25
2	3	0	21	13	28
3	18	21	0	8	7
4	10	13	8	0	15
5	25	28	7	15	0

Data	Product	Sales	10	7	28	20	35
	1	10					
	2	7	1	2	3	4	5
	3	28					
	4	20					
	5	35					

Step 2: Check the smallest distance proximity matrix and merge the data points.
 We have taken the **maximum** of the two marks (7, **10**) (1,2) to replace the marks for this cluster.
 You can also take **minimum, average**.

Proximity Matrix

ID	1	2	3	4	5
1	0	3	18	10	25
2	3	0	21	13	28
3	18	21	0	8	7
4	10	13	8	0	15
5	25	28	7	15	0



Data	Product	Sales
	(1,2)	10
	3	28
	4	20
	5	35

7,**10**

Data	Product	Sales
	1	10
	2	7
	3	28
	4	20
	5	35



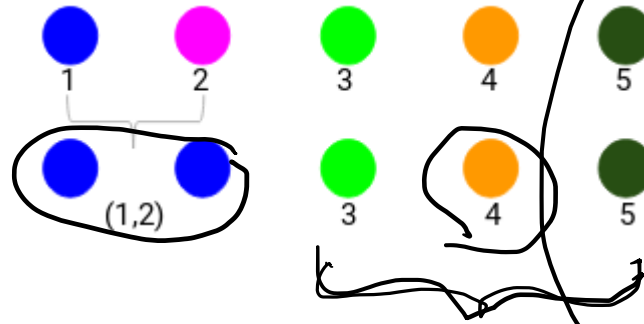
Step 2: Check the smallest distance proximity matrix and merge the data points.

We have taken the **maximum** of the two marks (7, **10**) (1,2) to replace the marks for this cluster.

You can also take **minimum, average**.

Proximity Matrix

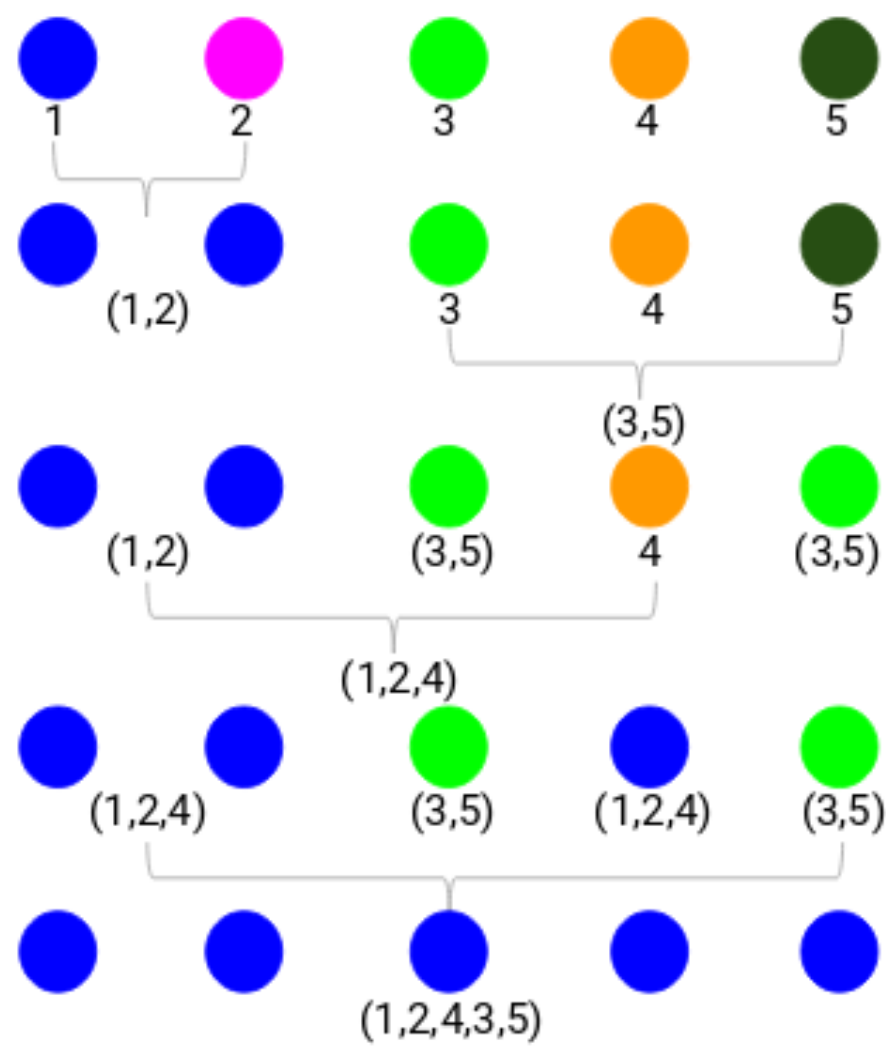
	1, 2	3, 5	4
1, 2	0	25	10
3, 5	25	0	15
4	10	15	0

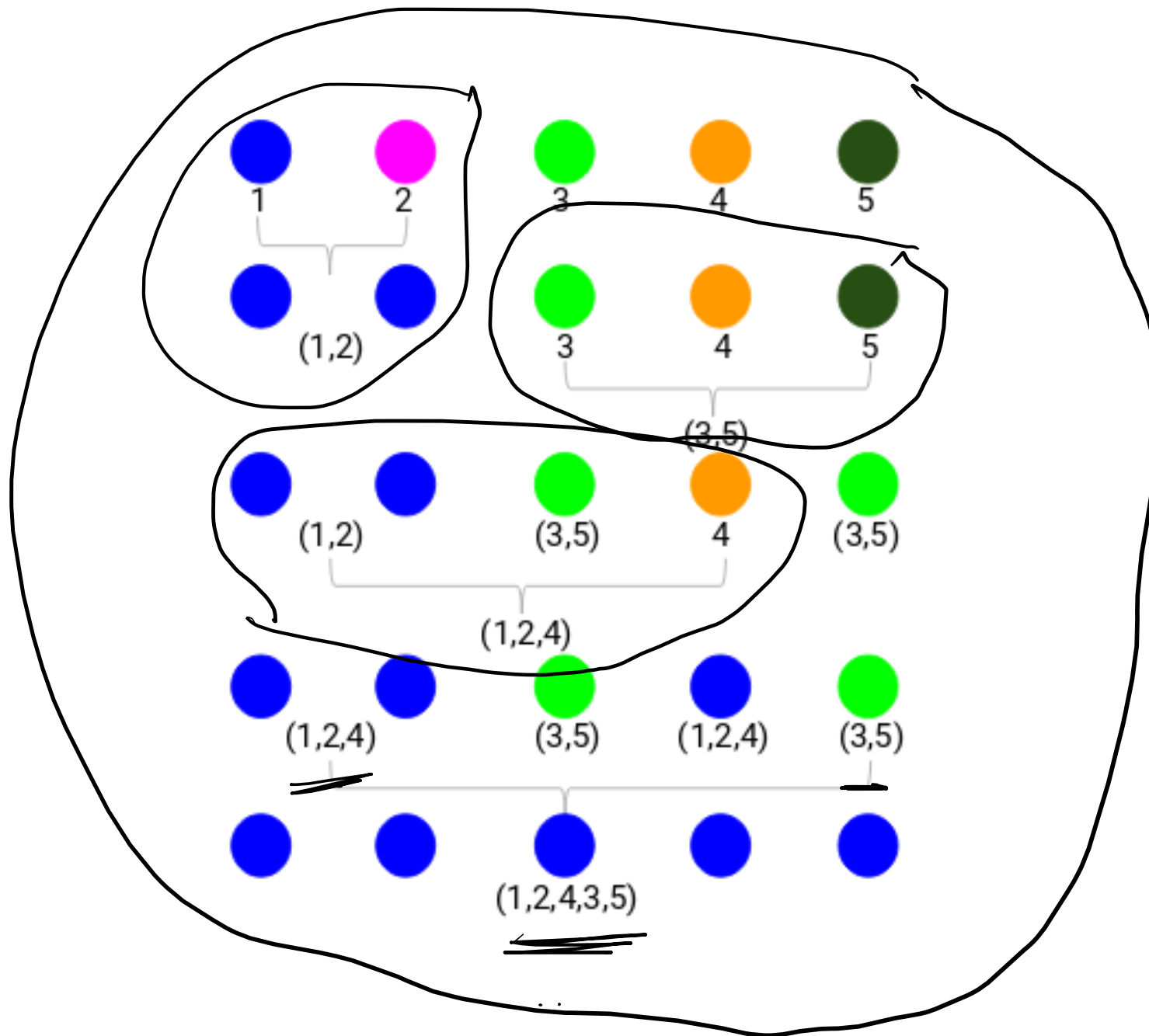


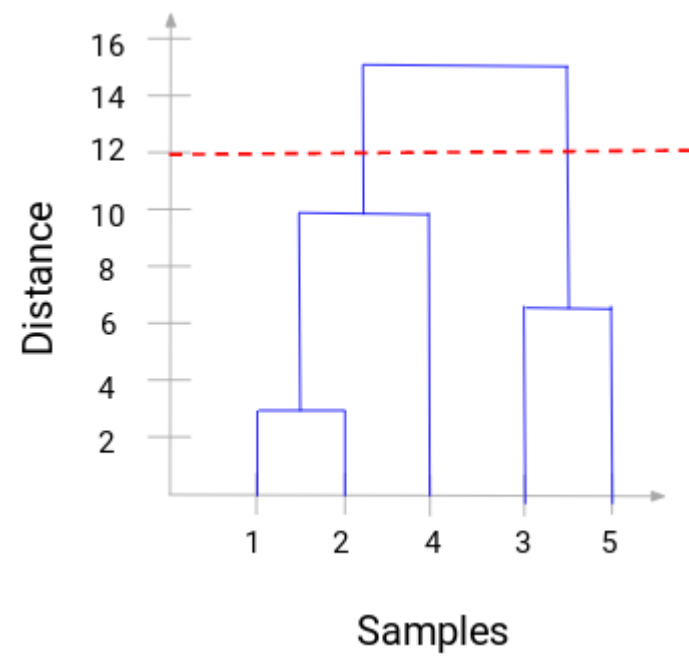
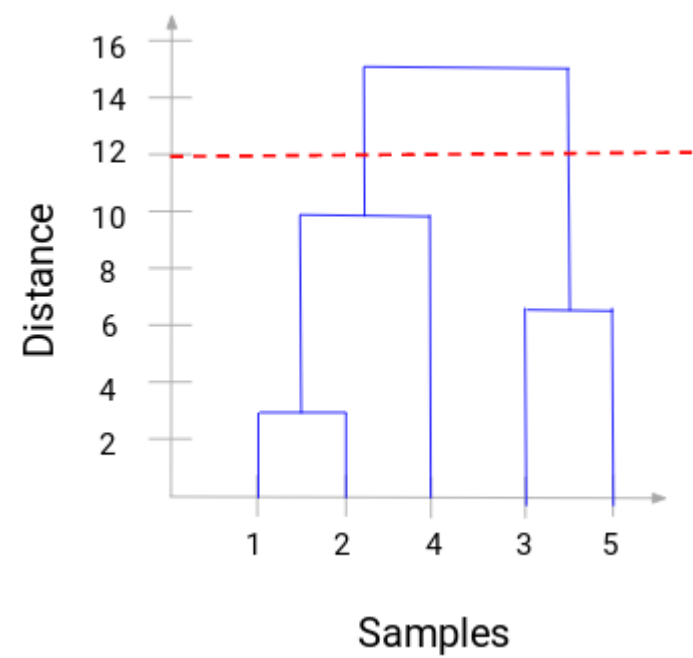
Data	Product	Sales
	(1,2)	10
	3	28
	4	20
	5	35

Data	Product	Sales
	1, 2, 4	20
	3, 5	35

Data	Product	Sales
	1, 2	10
	3, 5	35
	4	20



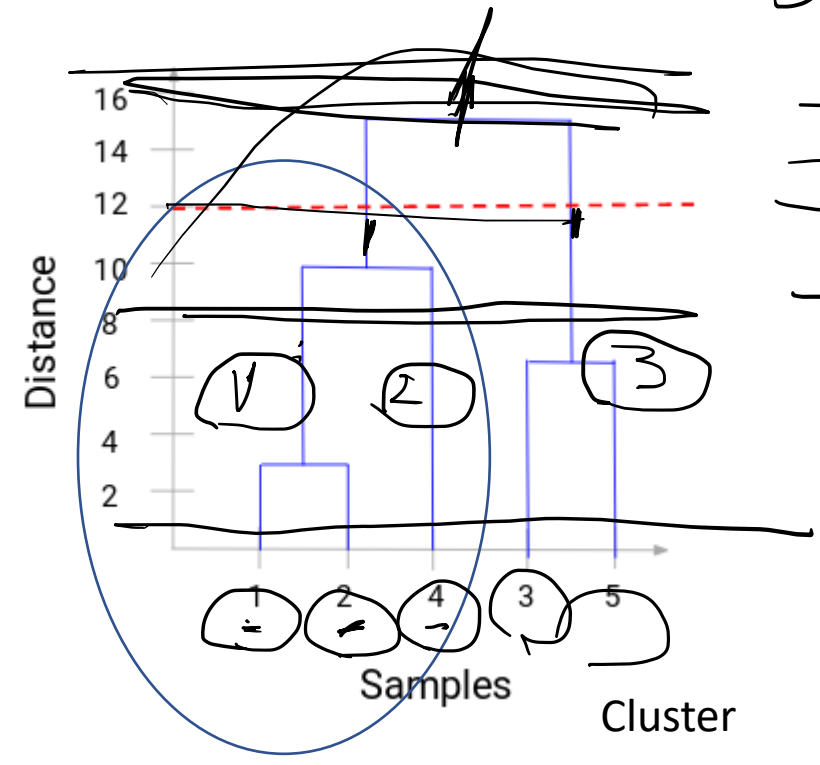
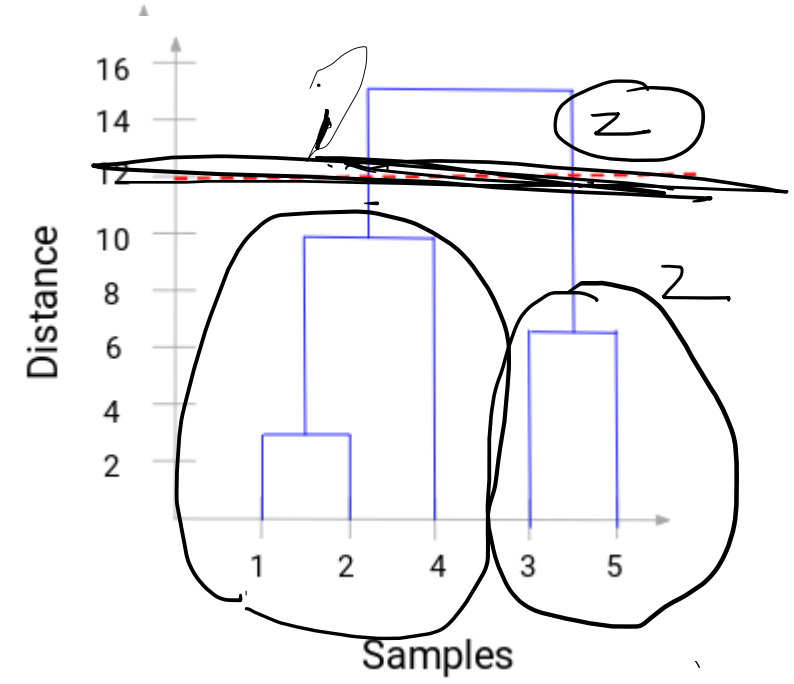




Dendrogram

data, proximity matrix, merge clusters

Dendrogram



1 Cluster -1,2,4