Assignment-6

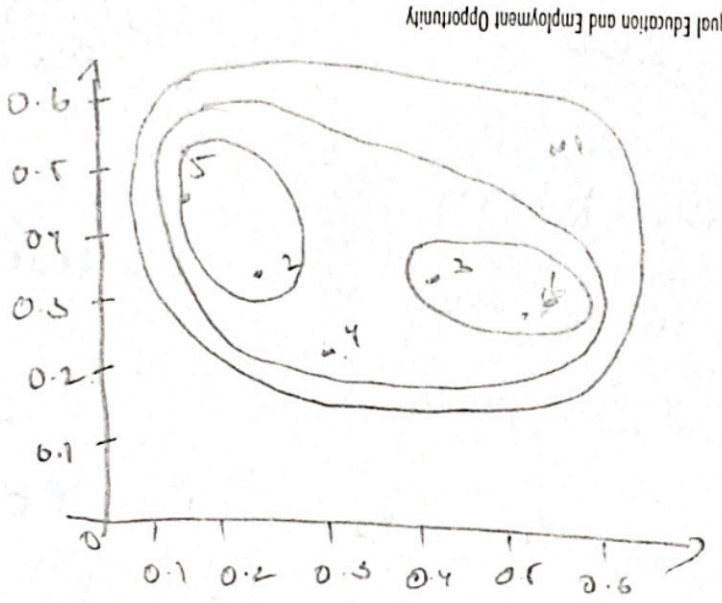
G. Ramya Sudha
2007-33458

(1) Single link proximity function

Given below in the table with euclidean distance b/w each individual point

	P_1	P_2	P_3	P_4	P_5	P_6
P_1	0					
P_2	0.2357	0				
P_3	0.2218	0.1485	0			
P_4	0.3688	0.2042	0.1513	0		
P_5	0.3421	0.1381	0.2843	0.2932	0	
P_6	0.2347	0.2540	0.1100	0.2216	0.3921	0

Considering lower bond value, as upper bond values are equal to lower bond & finding clusters



Graphical representation of the given six points

→ merging two cluster points & updating the distance as seen in above table min val is blw P_3 & P_6 after merging two number we need to find the distance with other member using ecuildian distance

re-calculating the distance after first cluster

⇒ min [dist (P_3, P_6), (P_1)]

⇒ min [(0.2218), (0.2342)]

⇒ 0.2218

→ distance min ((distance (P3, P6) P2))

⇒ min ((0.1483), (0.2540))

= 0.1483

→ min (dist (P5, P6) P4) ⇒ min (dist (P3, P4) (P6, P4))

⇒ 0.1513

→ min (dist (P3, P6), P5))

min (dist (P3, P5), (P6, P5))

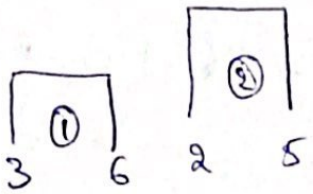
⇒ 0.2843

now updating matrix

	P1	P2	P3, P6	P4	P5
P1	0	0.2357			
P2	0.2357	0			
P3, P6	0.2217	0.1483	0		
P4	0.3488	0.2042	0.1513	0	
P5	0.3421	0.1388	0.2843	0.2	0

(4)

we find min ~~value~~ as 0.1888 b/w P2 & P5



Updating distance matrix

⇒ min [dist (P2, P5), P1]

⇒ min ((0.2357), (0.3421))

⇒ 0.2357

⇒ min (dist(P2, P5) (P3, P6)

⇒ min (dist (P2, (P3, P6), (P5, (P3, P6))

⇒ 0.1483

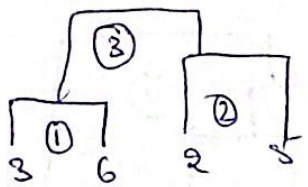
⇒ min [dist (P2, P5), P4)]

min [dist (0.2042), (0.2952)

⇒ 0.2042

	P1	P2,P5	P3,P6	P4
P1	0			
P2,P5	0.2357	0		
P3,P6	0.2218	0.1483	0	
P4	0.3688	0.2042	0.1513	0

min value ai 0.1483 b/w P2,P5 & P3,P6



⇒ updating distance matrix $\min[dist((P_2,P_5),(P_3,P_6)),P_1)]$

$\min[dist((P_2,P_5);P_1)((P_3,P_6),P_1)]$

⇒ 0.2218

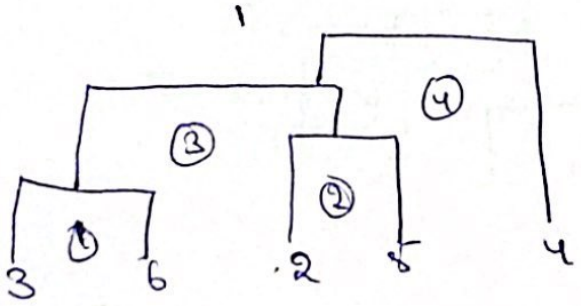
⇒ $\min[dist((P_2,P_5),P_4)),((P_3,P_6),P_4))$

⇒ 0.1513

Equal Education and Employment Opportunity

	P_1	P_2, P_5, P_3, P_6	P_4
P_1	0		
P_2, P_5, P_3, P_6	0.2218	0	
P_4	0.3588	0.1513	0

the minimum value is 0.1513 b/w P_4 & P_2, P_5, P_3, P_6

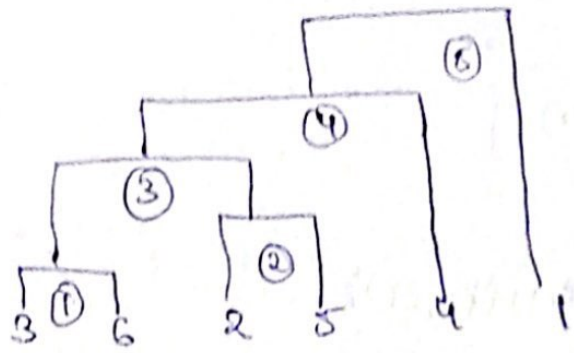


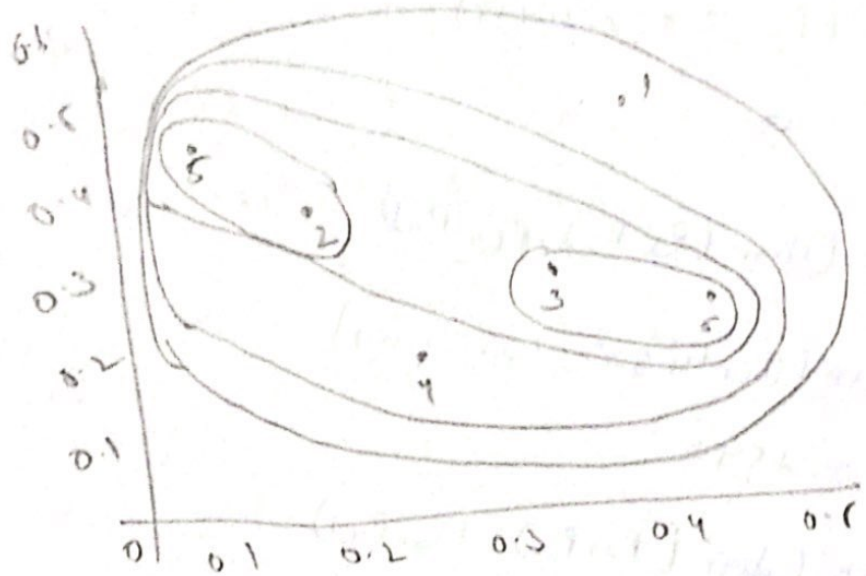
Updating distance matrix $\min\{dist(P_2, P_5, P_3, P_6), P_4\}$

$\Rightarrow 0.2214$

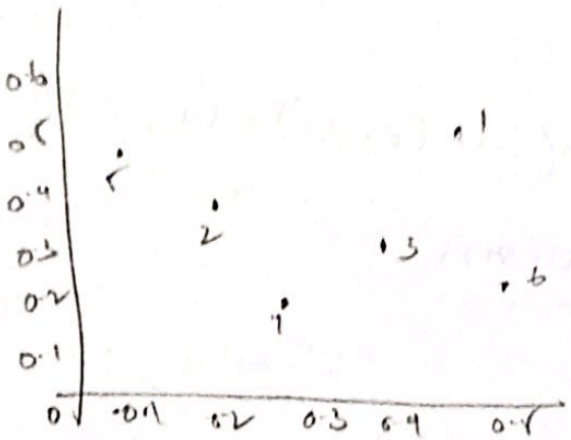
	P_1	P_2, P_5, P_3, P_6, P_4
P_1	0	
P_2, P_5, P_3, P_6, P_4	0.2218	0

find clusters blw P1 & P2, P5, P3, P6, P4





Complete-link :- ploting x, & y values with given matrix



from data we can say min value is 0.1100 blw P3 & P6

Cluster

→ max (dist (P3, P1), (P6, P1))

⇒ max ([0.2218, 0.2347])

⇒ 0.2347

→ max (dist (P3, P2), (P6, P2))

⇒ max ([0.1485], [0.2540])

→ 0.2540

→ max (dist [P3, P4), (P6, P4))

⇒ max [0.1617, 0.2216)

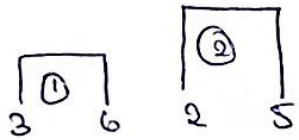
⇒ 0.2216

→ ~~optdrums~~ max (dist (P3, P5), (P6, P5))

⇒ 0.3921

	P_1	P_2	P_3, P_6	P_4	P_5
P_1	0				
P_2	0.2357	0			
P_3, P_6	0.2347	0.2590	0		
P_4	0.3680	0.2042	0.1426	0	
P_5	0.3421	0.1388	0.3921	0.2932	0

⇒ min val of lower bond is 0.1388 b/w P_2 & P_5



→ max (dist $(P_2, P_5), (P_5, P_1)$)

⇒ max [0.2357, 0.3421)

= 0.3421

→ max [dist $([P_2, (P_3, P_6)], [P_5, (P_3, P_6)])$]

⇒ 0.3921

→ max [dist $[P_2, P_4), [P_5, P_4)]$

⇒ 0.2932

(10)

	P1	P2, P5	P3, P6	P4
P1	0			
P2, P5	0.3421	0		
P3, P6	0.2347	0.3921	0	
P4	0.3688	0.2132	0.2216	0

min val is 0.2216 blw P3, P6 & P4

→ Updating distance matrix

→ max [dist ((P3, P6), P1), (P4, P1))]

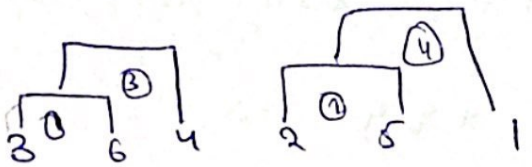
⇒ 0.3688

→ max [dist ((P3, P6), (P2, P5)), (P4 (P2, P5))]

⇒ 0.3921

Updating matrix for cluster P_3, P_6, P_4

	P_1	P_2, P_5	P_3, P_6, P_4
P_1	0		
P_2, P_5	0.3421	0	
P_3, P_6, P_4	0.3683	0.3921	0



updating distance matrix $\max[dist[P_2, P_5), (P_3, P_6, P_1]]$

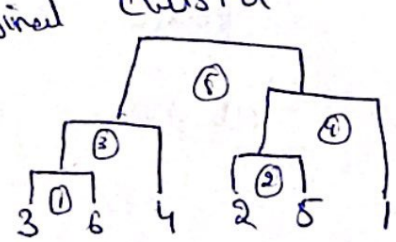
$(P_1, (P_3, P_6, P_4)])$

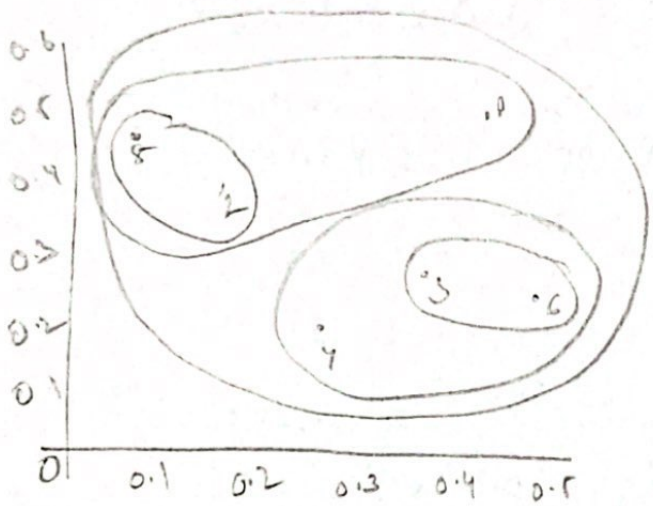
$\Rightarrow 0.3921$

updating distance matrix.

	P_2, P_5, P_1	P_3, P_6, P_4
P_2, P_5, P_1	0	
P_3, P_6, P_4	0.3921	0

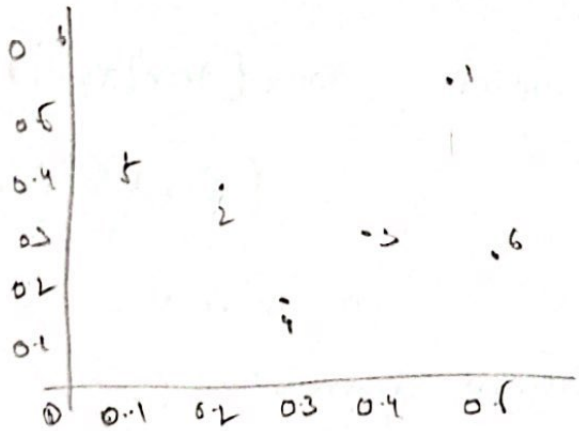
$\Rightarrow$ final cluster





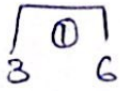
Average link :- looks at distance b/w all pairs & avg of the distances

from Given data plotting x & y axis



found min value of 0.1100 b/w P3 & P6

Cluster



→ AVG[dist(P3, P6), P1]

⇒ 1/2 [dist [P3, P1) + (P6, P1))

⇒ 0.2282

→ Avg [dist (P3, P6), P2] ⇒ 1/2 [dist [P3, P2), dist [P6, P2)]

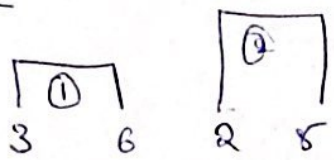
⇒ 0.2011

→ Avg [dist(P3, P6) P4] ⇒ 0.1864

→ Avg [dist[P3, P6), P5] = 0.3382

	P1	P2	P3, P6	P4	P5
P1	0				
P2	0.3354	0			
P3, P6	0.2282	0.2011	0		
P4	0.3688	0.2042	0.1864	0	
P5	0.3421	0.1588	0.3382	0.2932	0

2nd Cluster



→ Avg [dist [P2, P5), P1)]

dist [(P2, P5), P1] = 1/2 [dist (P2, P1) + dist (P5, P1))

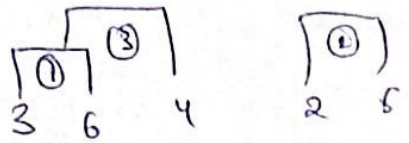
⇒ 0.2889

→ dist [(P2, P5)(P3, P6)] = 0.2696

→ dist ((P2, P5), P4)) = 0.2487

	P1	P2, P5	P3, P6	P4
P1	0			
P2, P5	2.889	0		
P3, P6	0.2282	0.2696	0	
P4	0.3688	0.2487	0.1864	0

→ min value is at P4 & (P3, P6) which is 0.1864



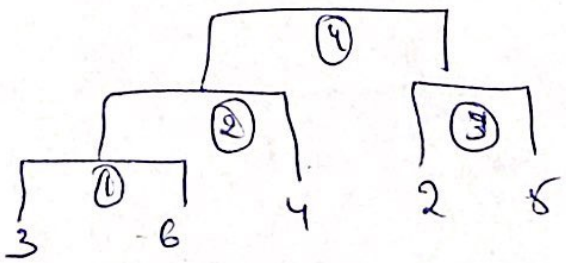
→ dist $[((P_3, P_6, P_4), P_5)] = \frac{1}{2}[dist((P_3, P_6), P_4) + (P_4, P_5)]$

$= 0.2986$

→ dist $[((P_3, P_6, P_4), (P_2, P_5)] = 0.2591$

	P_1	$P_2 P_5$	P_3, P_6, P_4
P_1	0		
$P_2 P_5$	0.2887	0	
$P_3, P_6 P_4$	0.2965	0.2591	0

Cluster



→ ~~dist~~ dist

dist $[(P_3, P_6, P_4), (P_2, P_5)] \Rightarrow 0.2937$