Sanjana Kothari

Q. Agglomerative clustering using single link.

(1) A = (2,6) G = (5,2)
B = (3,7) H = (7,3)
C = (5,8) I = (8,4)
D = (6,6) J = (10,6)
E = (5,5) K = (12,8)
F = (2,2)

Eyeballing gives approximately 3 clusters.

Distance matrix: (Using Euclidean distance)

	A	B	C	D	E	F	G	H	I	J	K
A	0										
B	$\sqrt{2}$	0									
C	$\sqrt{13}$	$\sqrt{5}$	0								
D	$\sqrt{16}$	$\sqrt{10}$	$\sqrt{5}$	0							
E	$\sqrt{10}$	$\sqrt{8}$	$\sqrt{9}$	$\sqrt{2}$	0						
F	$\sqrt{16}$	$\sqrt{26}$	$\sqrt{45}$	$\sqrt{32}$	$\sqrt{18}$	0					
G	$\sqrt{25}$	$\sqrt{29}$	$\sqrt{36}$	$\sqrt{17}$	$\sqrt{9}$	$\sqrt{9}$	0				
H	$\sqrt{34}$	$\sqrt{32}$	$\sqrt{29}$	$\sqrt{10}$	$\sqrt{8}$	$\sqrt{26}$	$\sqrt{5}$	0			
I	$\sqrt{40}$	$\sqrt{34}$	$\sqrt{25}$	$\sqrt{8}$	$\sqrt{10}$	$\sqrt{40}$	$\sqrt{13}$	$\sqrt{2}$	0		
J	$\sqrt{64}$	$\sqrt{50}$	$\sqrt{29}$	$\sqrt{16}$	$\sqrt{26}$	$\sqrt{68}$	$\sqrt{41}$	$\sqrt{18}$	$\sqrt{8}$	0	
K	$\sqrt{104}$	$\sqrt{82}$	$\sqrt{49}$	$\sqrt{40}$	$\sqrt{58}$	$\sqrt{136}$	$\sqrt{85}$	$\sqrt{50}$	$\sqrt{32}$	$\sqrt{8}$	0

Taking min-dist = $\sqrt{2}$ between A and B and clustering those points together.

	AB	C	D	E	F	G	H	I	J	K
AB	0									
C	$\sqrt{5}$	0								
D	$\sqrt{10}$	$\sqrt{5}$	0							
E	$\sqrt{8}$	$\sqrt{9}$	$\sqrt{2}$	0						
F	$\sqrt{16}$	$\sqrt{45}$	$\sqrt{32}$	$\sqrt{18}$	0					
G	$\sqrt{25}$	$\sqrt{36}$	$\sqrt{17}$	$\sqrt{9}$	$\sqrt{9}$	0				
H	$\sqrt{32}$	$\sqrt{29}$	$\sqrt{10}$	$\sqrt{8}$	$\sqrt{26}$	$\sqrt{5}$	0			
I	$\sqrt{34}$	$\sqrt{25}$	$\sqrt{8}$	$\sqrt{10}$	$\sqrt{40}$	$\sqrt{13}$	$\sqrt{2}$	0		
J	$\sqrt{50}$	$\sqrt{29}$	$\sqrt{16}$	$\sqrt{26}$	$\sqrt{68}$	$\sqrt{41}$	$\sqrt{18}$	$\sqrt{8}$	0	
K	$\sqrt{82}$	$\sqrt{49}$	$\sqrt{40}$	$\sqrt{58}$	$\sqrt{136}$	$\sqrt{85}$	$\sqrt{50}$	$\sqrt{32}$	$\sqrt{8}$	0

Taking min-distance = $\sqrt{2}$ between pts D and E.

	AB	C	DE	F	G	H	I	J	K
AB	0								
C	$\sqrt{5}$	0							
DE	$\sqrt{8}$	$\sqrt{5}$	0						
F	$\sqrt{16}$	$\sqrt{45}$	$\sqrt{18}$	0					
G	$\sqrt{25}$	$\sqrt{36}$	$\sqrt{9}$	$\sqrt{9}$	0				
H	$\sqrt{32}$	$\sqrt{29}$	$\sqrt{8}$	$\sqrt{26}$	$\sqrt{5}$	0			
I	$\sqrt{34}$	$\sqrt{25}$	$\sqrt{8}$	$\sqrt{40}$	$\sqrt{13}$	$\sqrt{2}$	0		
J	$\sqrt{50}$	$\sqrt{29}$	$\sqrt{16}$	$\sqrt{68}$	$\sqrt{41}$	$\sqrt{18}$	$\sqrt{8}$	0	
K	$\sqrt{82}$	$\sqrt{49}$	$\sqrt{40}$	$\sqrt{136}$	$\sqrt{85}$	$\sqrt{50}$	$\sqrt{32}$	$\sqrt{8}$	0

Taking min-distance = $\sqrt{2}$ b/w pts H and I.

	AB	C	DE	F	G	HI	J	K
AB	0							
C	$\sqrt{5}$	0						
DE	$\sqrt{8}$	$\sqrt{5}$	0					
F	$\sqrt{16}$	$\sqrt{45}$	$\sqrt{18}$	0				
G	$\sqrt{25}$	$\sqrt{36}$	$\sqrt{9}$	$\sqrt{9}$	0			
HI	$\sqrt{32}$	$\sqrt{25}$	$\sqrt{8}$	$\sqrt{26}$	$\sqrt{5}$	0		
J	$\sqrt{50}$	$\sqrt{29}$	$\sqrt{16}$	$\sqrt{68}$	$\sqrt{41}$	$\sqrt{8}$	0	
K	$\sqrt{82}$	$\sqrt{49}$	$\sqrt{40}$	$\sqrt{136}$	$\sqrt{85}$	$\sqrt{32}$	$\sqrt{8}$	0

Taking min-distance = $\sqrt{5}$ b/w pt C and cluster DE

	AB	CDE	F	G	HI	J	K
AB	0						
CDE	$\sqrt{5}$	0					
F	$\sqrt{16}$	$\sqrt{18}$	0				
G	$\sqrt{25}$	$\sqrt{9}$	$\sqrt{9}$	0			
HI	$\sqrt{32}$	$\sqrt{8}$	$\sqrt{26}$	$\sqrt{5}$	0		
J	$\sqrt{50}$	$\sqrt{16}$	$\sqrt{68}$	$\sqrt{41}$	$\sqrt{8}$	0	
K	$\sqrt{82}$	$\sqrt{40}$	$\sqrt{136}$	$\sqrt{85}$	$\sqrt{32}$	$\sqrt{8}$	0

Taking min-distance = $\sqrt{5}$ b/w pt G and cluster HI

	AB	CDE	F	GHI	J	K
AB	0					
CDE	$\sqrt{5}$	0				
F	$\sqrt{16}$	$\sqrt{18}$	0			
GHI	$\sqrt{25}$	$\sqrt{8}$	$\sqrt{9}$	0		
J	$\sqrt{50}$	$\sqrt{16}$	$\sqrt{68}$	$\sqrt{8}$	0	
K	$\sqrt{82}$	$\sqrt{40}$	$\sqrt{136}$	$\sqrt{32}$	$\sqrt{8}$	0

Taking min_dist = $\sqrt{5}$ b/w clusters AB and CDE

	ABCDE	F	GHI	J	K
ABCDE	0				
F	$\sqrt{16}$	0			
GHI	$\sqrt{8}$	$\sqrt{9}$	0		
J	$\sqrt{16}$	$\sqrt{68}$	$\sqrt{8}$	0	
K	$\sqrt{40}$	$\sqrt{136}$	$\sqrt{32}$	$\sqrt{8}$	0

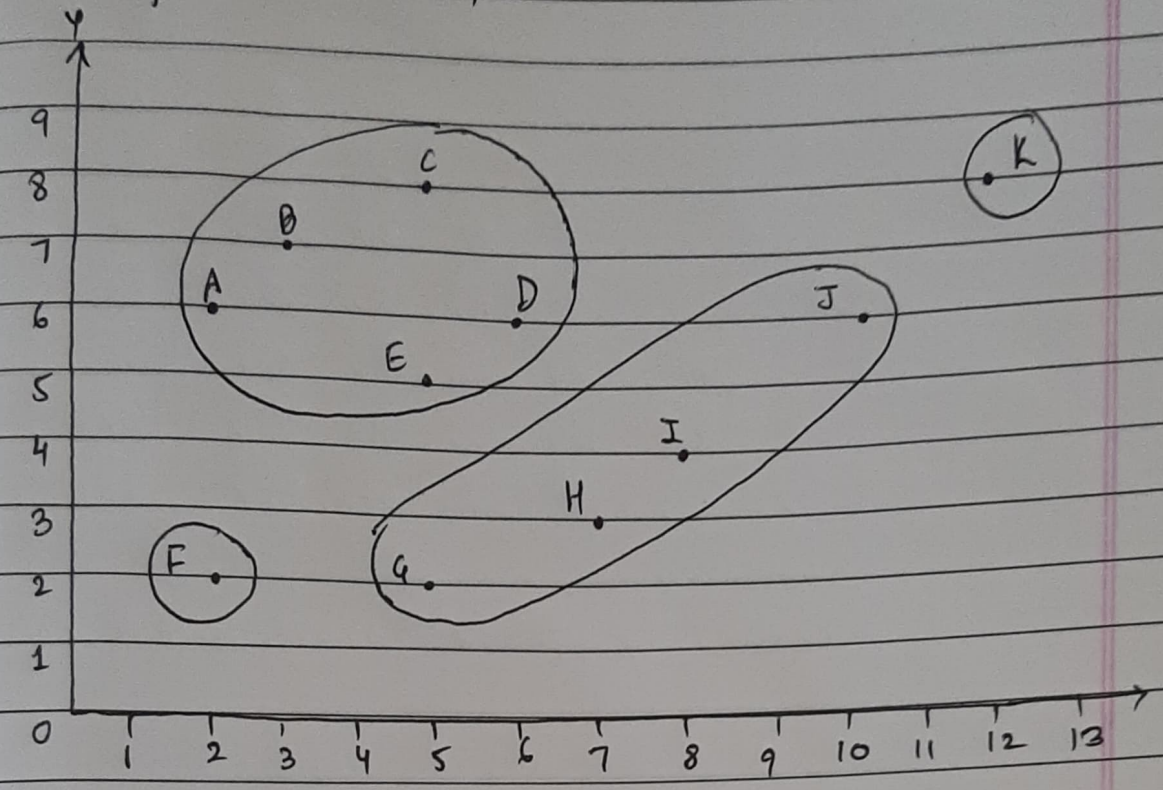
Taking min_dist = $\sqrt{8}$ b/w pt J and GHI

	ABCDE	F	GHIJ	K
ABCDE	0			
F	$\sqrt{16}$	0		
GHIJ	$\sqrt{8}$	$\sqrt{9}$	0	
K	$\sqrt{40}$	$\sqrt{136}$	$\sqrt{8}$	0

Taking min_dist = $\sqrt{8}$ b/w pt K and cluster GHIJ

	ABCDE	F	GHIJK
ABCDE	0		
F	$\sqrt{16}$	0	
GHIJK	$\sqrt{8}$	$\sqrt{9}$	0

If we stop at 4 clusters,



If we go down to 3 clusters,