

WS #19 - hierarchical clustering

Wednesday, November 19, 2025

Your Name: _____

Names of people you worked with: _____

What's the best thing to eat or drink when it is raining cats and dogs?

Task:

Consider the distances between the following observations:

	A	B	C	D	E
A	0				
B	0.2	0			
C	0.6	0.5	0		
D	1	0.9	0.4	0	
E	0.9	0.8	0.5	0.3	0

Start with all objects in separate “clusters” (i.e., start with 5 clusters), by merging (complete linkage) one pair of clusters at a time, provide each clustering for $k = 5, 4, 3, 2, 1$.

Solution:

$k = 4$: Link A and B to get $-(AB), C, D, E$

$k = 3$: Link D and E to get $-(AB), C, (DE)$

$$d_{(AB)C} = \max(d_{AC}, d_{BC}) = 0.6$$

$$d_{(AB)D} = \max(d_{AD}, d_{BD}) = 1.0$$

$$d_{(AB)E} = \max(d_{AE}, d_{BE}) = 0.9$$

	AB	C	D	E
AB	0			
C	0.6	0		
D	1.0	0.4	0	
E	0.9	0.5	0.3	0

Link D and E!

$k = 2$: Link C with (DE) to get $-(AB), (CDE)$

$$d_{(AB)C} = 0.6$$

$$d_{(AB)(DE)} = \max(d_{AD}, d_{BD}, d_{AE}, d_{BE}) = 1.0$$

$$d_{(DE)C} = \max(d_{CD}, d_{CE}) = 0.5$$

	AB	C	DE
AB	0		
C	0.6	0	
DE	1.0	0.5	0

Link C with (DE) !

$k = 1$: Link all to get $-(ABCDE)$

$$d_{(AB)(CDE)} = d_{AD} = 1$$