



How Hierarchical clustering works?

Customer	<i>Average Bet per Visit</i> (x_1)	<i>Visit Per Month</i> (x_2)
A	15	2
B	30	2
C	80	8
D	120	9
E	90	12
F	50	5

A casino wants to segment its customers to improve marketing and loyalty programs. They track two key metrics for each customer:



How Hierarchical clustering works?

Customer	<i>Average Bet per Visit</i> (x_1)	<i>Visit Per Month</i> (x_2)
A	15	2
B	30	2
C	80	8
D	120	9
E	90	12
F	50	5

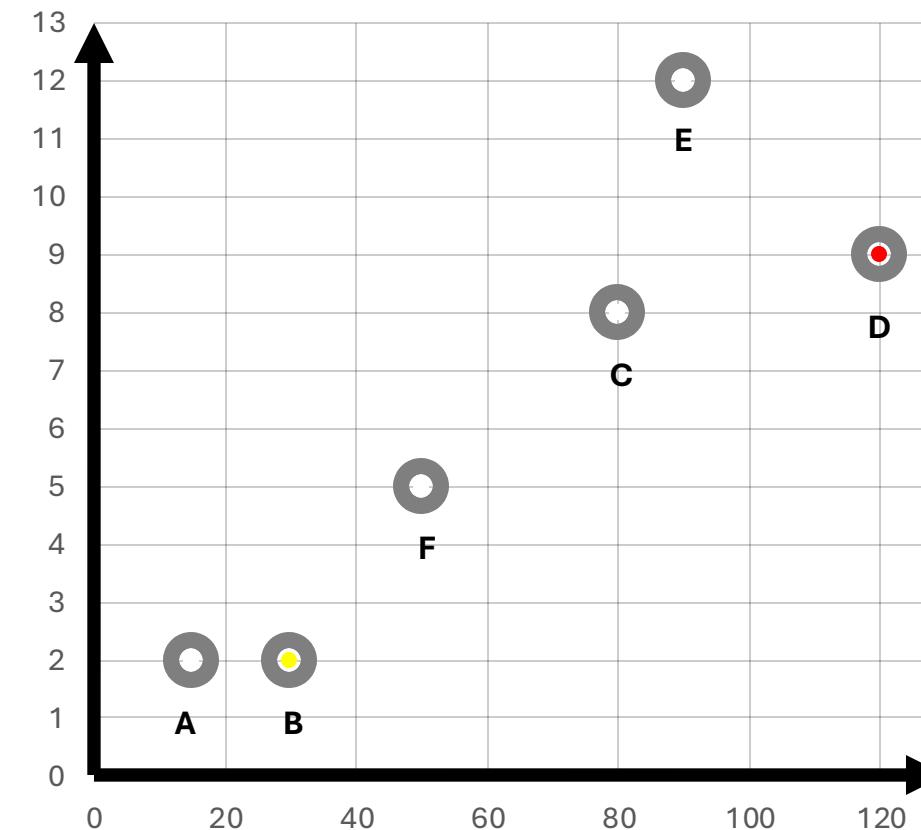
Average Bet Amount per Visit. Measures how much a customer typically wagers each visit in dollars.

How Hierarchical clustering works?

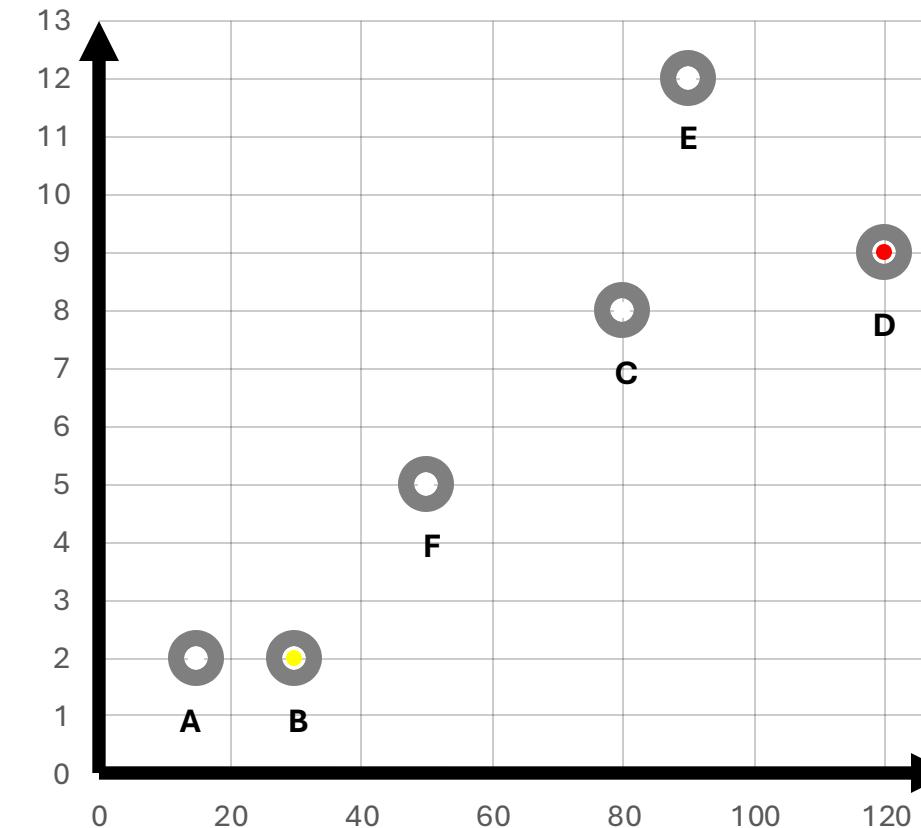
Customer	<i>Average Bet per Visit</i> (x_1)	<i>Visit Per Month</i> (x_2)
A	15	2
B	30	2
C	80	8
D	120	9
E	90	12
F	50	5

Visit Per Month. Measures how often a customer comes to the casino per month.

How Hierarchical clustering works?

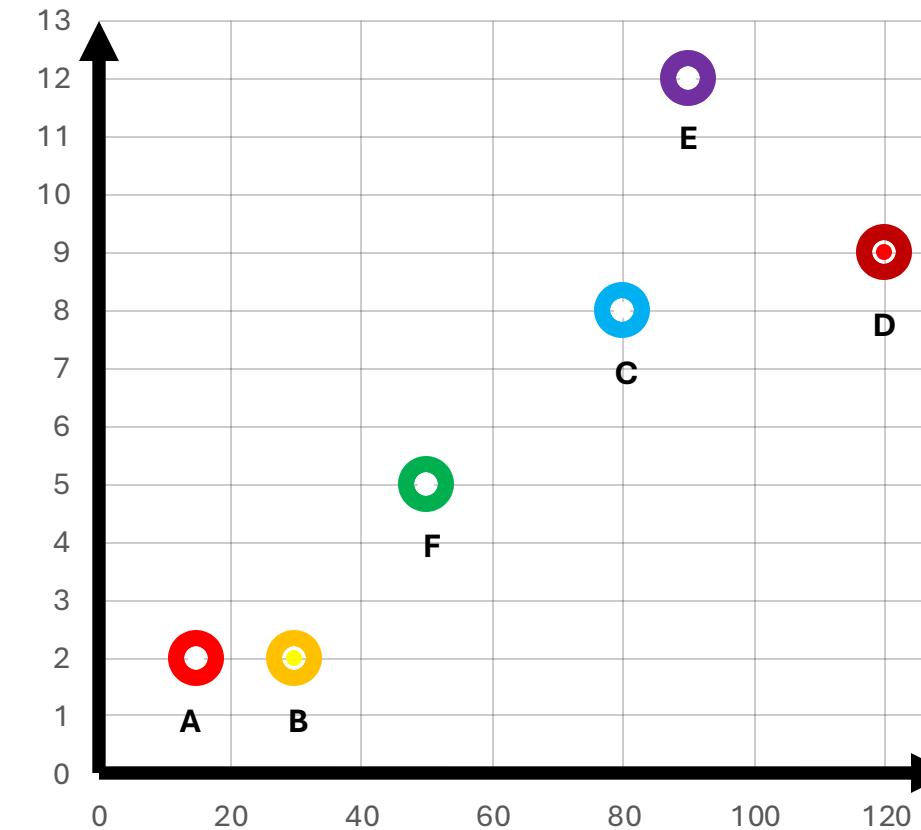


How Hierarchical clustering works?



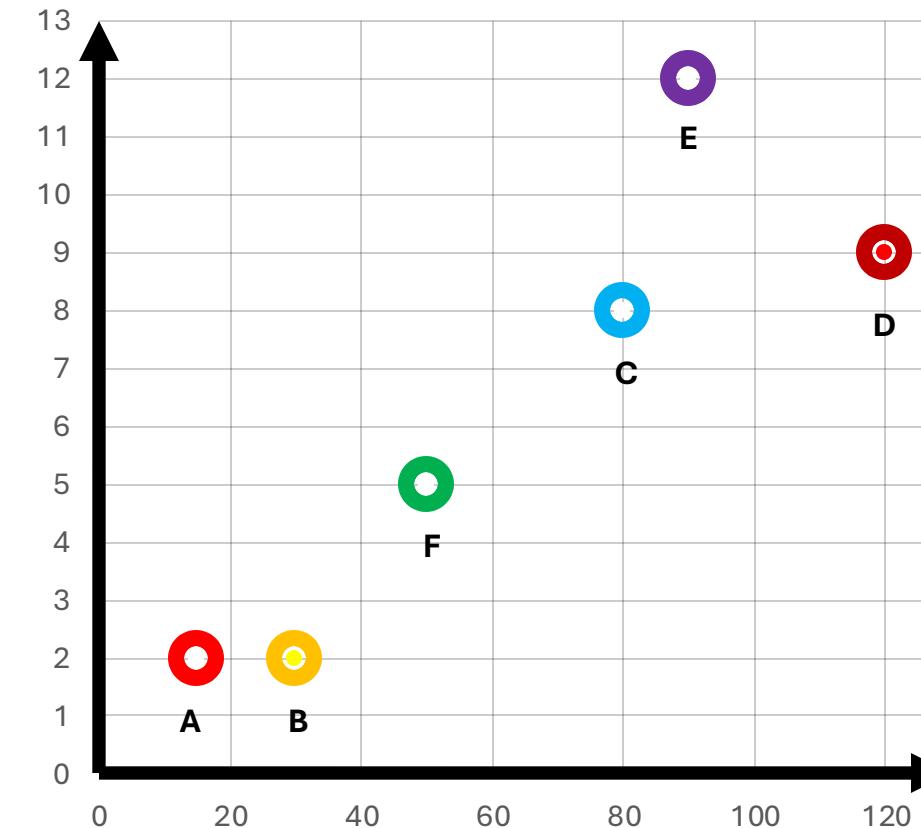
Step 1: Treat each data point as its own individual cluster

How Hierarchical clustering works?



Step 1: Treat each data point as its own individual cluster

How Hierarchical clustering works?



Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0					
B		0				
C			0			
D				0		
E					0	
F						0

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15				
B	15	0				
C			0			
D				0		
E					0	
F						0

$$d(A, B) = \sqrt{(A_{bet} - B_{bet})^2 + (A_{visit} - B_{visit})^2}$$

$$d(A, B) = \sqrt{(15 - 30)^2 + (2 - 2)^2}$$

$$d(A, B) = \sqrt{225}$$

$$d(A, B) = 15$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3			
B	15	0				
C	65.3		0			
D				0		
E					0	
F						0

$$d(A, C) = \sqrt{(A_{bet} - C_{bet})^2 + (A_{visit} - C_{visit})^2}$$

$$d(A, C) = \sqrt{(15 - 80)^2 + (2 - 8)^2}$$

$$d(A, C) = \sqrt{4261}$$

$$d(A, C) = 65.3$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3		
B	15	0				
C	65.3		0			
D	105.3			0		
E					0	
F						0

$$d(A, D) = \sqrt{(A_{bet} - D_{bet})^2 + (A_{visit} - D_{visit})^2}$$

$$d(A, D) = \sqrt{(15 - 120)^2 + (2 - 9)^2}$$

$$d(A, D) = \sqrt{11074}$$

$$d(A, D) = 105.3$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	
B	15	0				
C	65.3		0			
D	105.3			0		
E	75.7				0	
F						0

$$d(A, E) = \sqrt{(A_{bet} - E_{bet})^2 + (A_{visit} - E_{visit})^2}$$

$$d(A, E) = \sqrt{(15 - 90)^2 + (2 - 12)^2}$$

$$d(A, E) = \sqrt{5725}$$

$$d(A, E) = 75.7$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0				
C	65.3		0			
D	105.3			0		
E	75.7				0	
F	35.1					0

$$d(A, F) = \sqrt{(A_{bet} - F_{bet})^2 + (A_{visit} - F_{visit})^2}$$

$$d(A, F) = \sqrt{(15 - 50)^2 + (2 - 5)^2}$$

$$d(A, F) = \sqrt{1234}$$

$$d(A, F) = 35.1$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.



How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36			
C	65.3	50.36	0			
D	105.3			0		
E	75.7				0	
F	35.1					0

$$d(B, C) = \sqrt{(B_{bet} - C_{bet})^2 + (B_{visit} - C_{visit})^2}$$

$$d(B, C) = \sqrt{(30 - 80)^2 + (2 - 8)^2}$$

$$d(B, C) = \sqrt{2536}$$

$$d(B, C) = 50.36$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3		
C	65.3	50.36	0			
D	105.3	90.3		0		
E	75.7				0	
F	35.1					0

$$d(B, D) = \sqrt{(B_{bet} - D_{bet})^2 + (B_{visit} - D_{visit})^2}$$

$$d(B, D) = \sqrt{(30 - 120)^2 + (2 - 9)^2}$$

$$d(B, D) = \sqrt{8149}$$

$$d(B, D) = 90.3$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3	60.8	
C	65.3	50.36	0			
D	105.3	90.3		0		
E	75.7	60.8			0	
F	35.1					0

$$d(B, E) = \sqrt{(B_{bet} - E_{bet})^2 + (B_{visit} - E_{visit})^2}$$

$$d(B, E) = \sqrt{(30 - 90)^2 + (2 - 12)^2}$$

$$d(B, E) = \sqrt{3700}$$

$$d(B, E) = 60.8$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3	60.8	20.2
C	65.3	50.36	0			
D	105.3	90.3		0		
E	75.7	60.8			0	
F	35.1	20.2				0

$$d(B, F) = \sqrt{(B_{bet} - F_{bet})^2 + (B_{visit} - F_{visit})^2}$$

$$d(B, F) = \sqrt{(30 - 50)^2 + (2 - 5)^2}$$

$$d(B, F) = \sqrt{409}$$

$$d(B, F) = 20.2$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3	60.8	20.2
C	65.3	50.36	0			
D	105.3	90.3		0		
E	75.7	60.8			0	
F	35.1	20.2				0

$$d(B, F) = \sqrt{(B_{bet} - F_{bet})^2 + (B_{visit} - F_{visit})^2}$$

$$d(B, F) = \sqrt{(30 - 50)^2 + (2 - 5)^2}$$

$$d(B, F) = \sqrt{409}$$

$$d(B, F) = 20.2$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3	60.8	20.2
C	65.3	50.36	0	40		
D	105.3	90.3	40	0		
E	75.7	60.8			0	
F	35.1	20.2				0

$$d(C, D) = \sqrt{(C_{bet} - D_{bet})^2 + (C_{visit} - D_{visit})^2}$$

$$d(C, D) = \sqrt{(80 - 120)^2 + (8 - 9)^2}$$

$$d(C, D) = \sqrt{1601}$$

$$d(C, D) = 40$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3	60.8	20.2
C	65.3	50.36	0	40	10.77	
D	105.3	90.3	40	0		
E	75.7	60.8	10.77		0	
F	35.1	20.2				0

$$d(C, E) = \sqrt{(C_{bet} - E_{bet})^2 + (C_{visit} - E_{visit})^2}$$

$$d(C, E) = \sqrt{(80 - 90)^2 + (8 - 12)^2}$$

$$d(C, E) = \sqrt{116}$$

$$d(C, E) = 10.77$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3	60.8	20.2
C	65.3	50.36	0	40	10.77	30.15
D	105.3	90.3	40	0		
E	75.7	60.8	10.77		0	
F	35.1	20.2	30.15			0

$$d(C, F) = \sqrt{(C_{bet} - F_{bet})^2 + (C_{visit} - F_{visit})^2}$$

$$d(C, F) = \sqrt{(80 - 50)^2 + (8 - 5)^2}$$

$$d(C, F) = \sqrt{909}$$

$$d(C, F) = 30.15$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3	60.8	20.2
C	65.3	50.36	0	40	10.77	30.15
D	105.3	90.3	40	0	30.15	
E	75.7	60.8	10.77	30.15	0	
F	35.1	20.2	30.15			0

$$d(D, E) = \sqrt{(D_{bet} - E_{bet})^2 + (D_{visit} - E_{visit})^2}$$

$$d(D, E) = \sqrt{(120 - 90)^2 + (9 - 12)^2}$$

$$d(D, E) = \sqrt{909}$$

$$d(D, E) = 30.15$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3	60.8	20.2
C	65.3	50.36	0	40	10.77	30.15
D	105.3	90.3	40	0	30.15	70.1
E	75.7	60.8	10.77	30.15	0	
F	35.1	20.2	30.15	70.1		0

$$d(D, F) = \sqrt{(D_{bet} - F_{bet})^2 + (D_{visit} - F_{visit})^2}$$

$$d(D, F) = \sqrt{(120 - 50)^2 + (9 - 5)^2}$$

$$d(D, F) = \sqrt{4916}$$

$$d(D, F) = 70.1$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3	60.8	20.2
C	65.3	50.36	0	40	10.77	30.15
D	105.3	90.3	40	0	30.15	70.1
E	75.7	60.8	10.77	30.15	0	40.6
F	35.1	20.2	30.15	70.1	40.6	0

$$d(E, F) = \sqrt{(E_{bet} - F_{bet})^2 + (E_{visit} - F_{visit})^2}$$

$$d(E, F) = \sqrt{(90 - 50)^2 + (12 - 5)^2}$$

$$d(E, F) = \sqrt{1649}$$

$$d(E, F) = 40.6$$

Step 2: Calculate the distance between every pair of data points using a distance metric like Euclidean distance.

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3	60.8	20.2
C	65.3	50.36	0	40	10.77	30.15
D	105.3	90.3	40	0	30.15	70.1
E	75.7	60.8	10.77	30.15	0	40.6
F	35.1	20.2	30.15	70.1	40.6	0

Step 3: Use **single linkage** to merge the closest pair by finding the smallest distance in the matrix

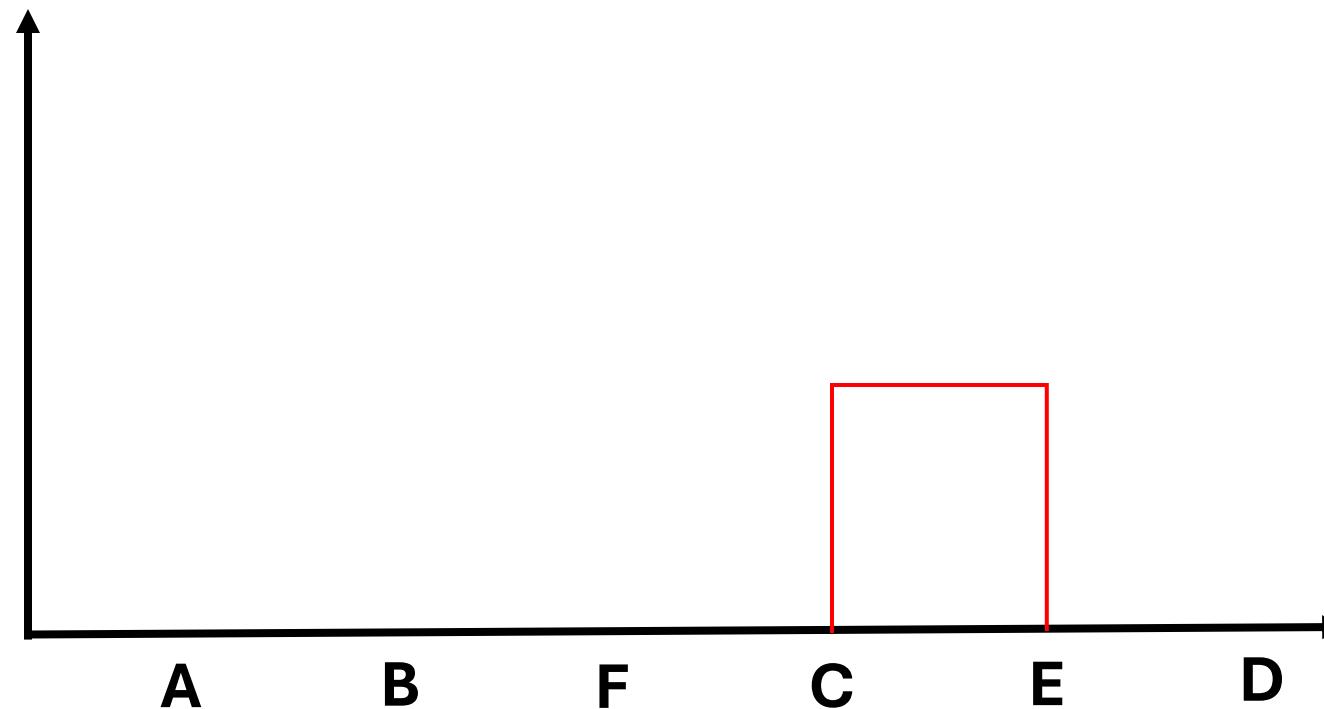


How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3	60.8	20.2
C	65.3	50.36	0	40	10.77	30.15
D	105.3	90.3	40	0	30.15	70.1
E	75.7	60.8	10.77	30.15	0	40.6
F	35.1	20.2	30.15	70.1	40.6	0

The smallest distance in the matrix is C-E

How Hierarchical clustering works?



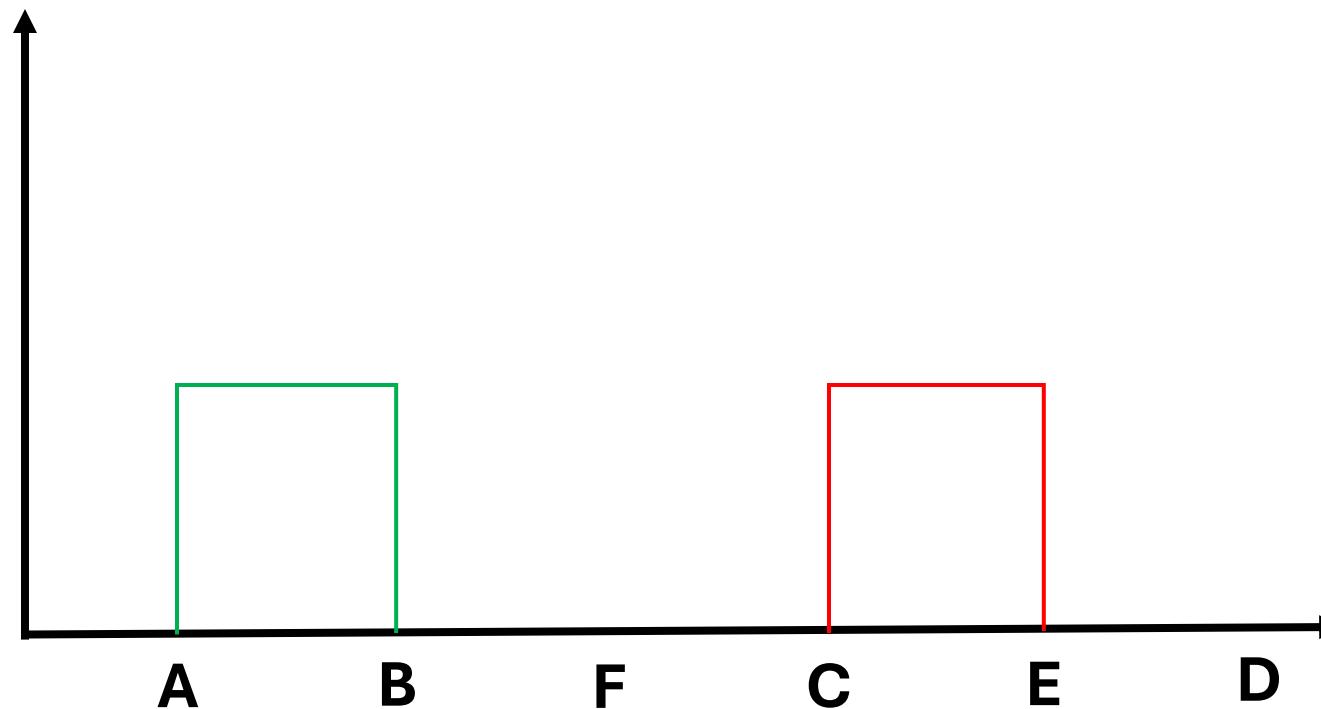
We will merge **C** and **E** creating the cluster **CE**

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3	60.8	20.2
C	65.3	50.36	0	40	10.77	30.15
D	105.3	90.3	40	0	30.15	70.1
E	75.7	60.8	10.77	30.15	0	40.6
F	35.1	20.2	30.15	70.1	40.6	0

The next smallest distance is A-B

How Hierarchical clustering works?



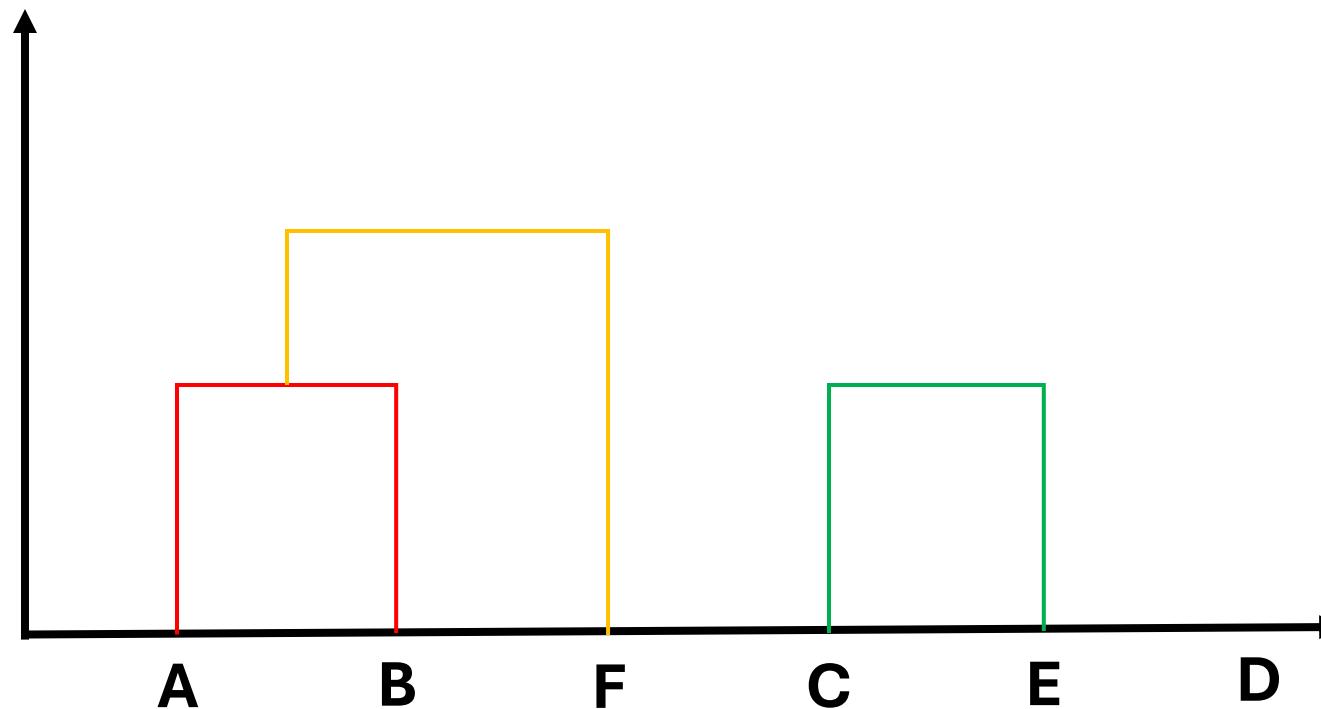
We will merge **A** and **B** creating the cluster **AB**

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3	60.8	20.2
C	65.3	50.36	0	40	10.77	30.15
D	105.3	90.3	40	0	30.15	70.1
E	75.7	60.8	10.77	30.15	0	40.6
F	35.1	20.2	30.15	70.1	40.6	0

The next smallest distance is B-F

How Hierarchical clustering works?



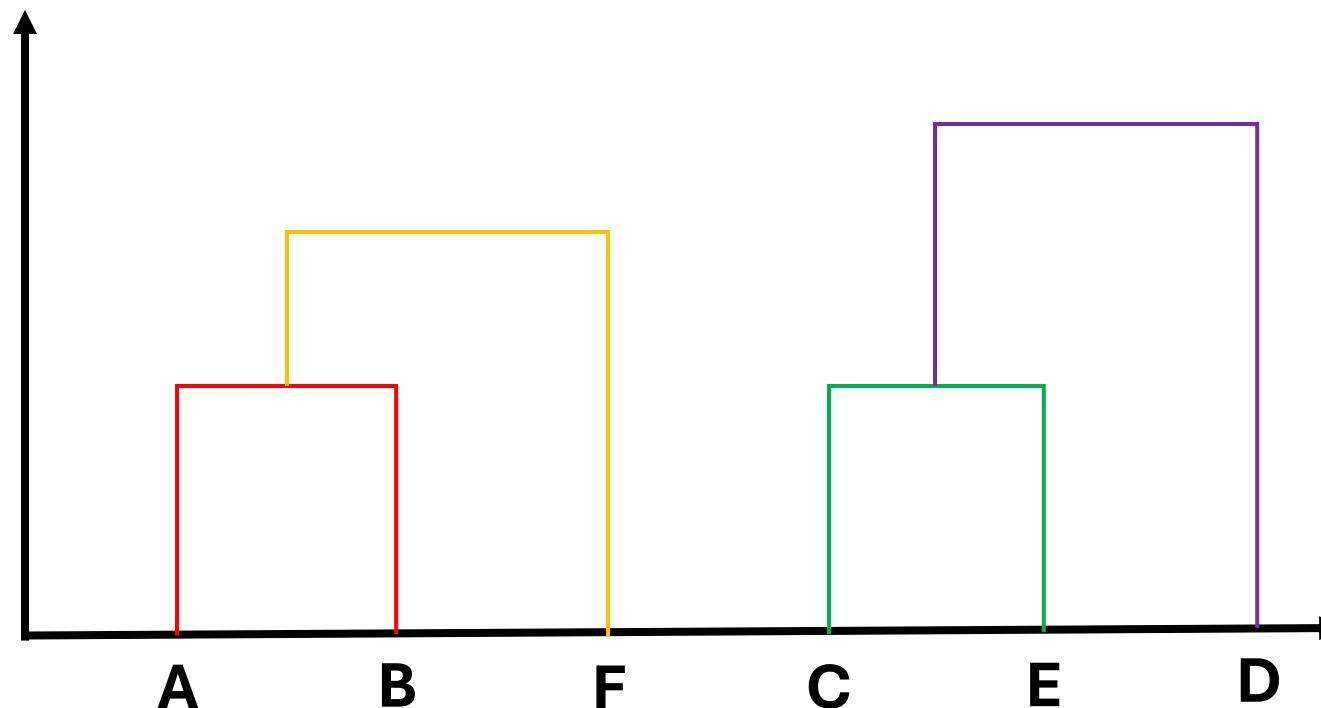
Since **B** is already in cluster **AB**, **F** will be merged, and we will create cluster **ABF**

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3	60.8	20.2
C	65.3	50.36	0	40	10.77	30.15
D	105.3	90.3	40	0	30.15	70.1
E	75.7	60.8	10.77	30.15	0	40.6
F	35.1	20.2	30.15	70.1	40.6	0

The next smallest distance is D-E

How Hierarchical clustering works?



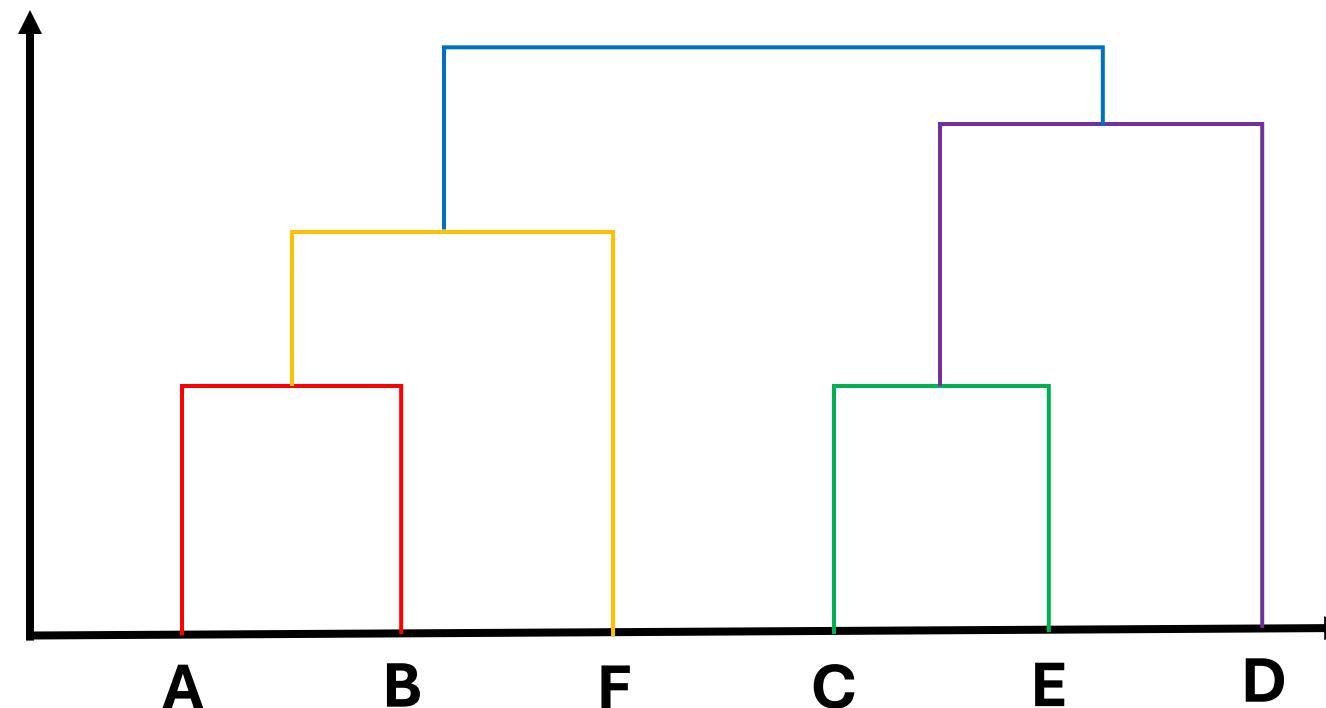
Since **E** is already in cluster **CE**, **D** will be merged, and we will create cluster **CED**

How Hierarchical clustering works?

	A	B	C	D	E	F
A	0	15	65.3	105.3	75.7	35.1
B	15	0	50.36	90.3	60.8	20.2
C	65.3	50.36	0	40	10.77	30.15
D	105.3	90.3	40	0	30.15	70.1
E	75.7	60.8	10.77	30.15	0	40.6
F	35.1	20.2	30.15	70.1	40.6	0

The next smallest distance is C-F

How Hierarchical clustering works?



Since **C** is already in cluster **CED** and **F** is already in cluster **ABF**, **ABF** and **CED** will be merged creating cluster **ABFCED**