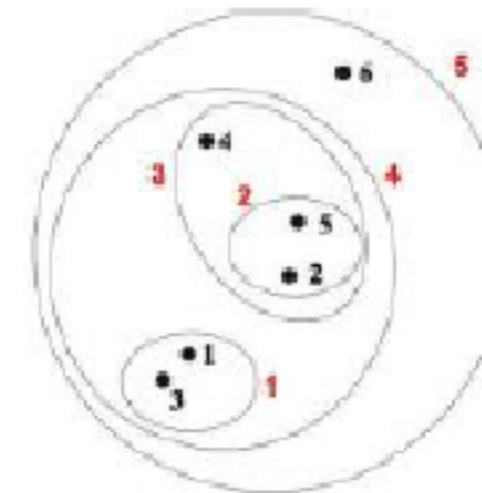
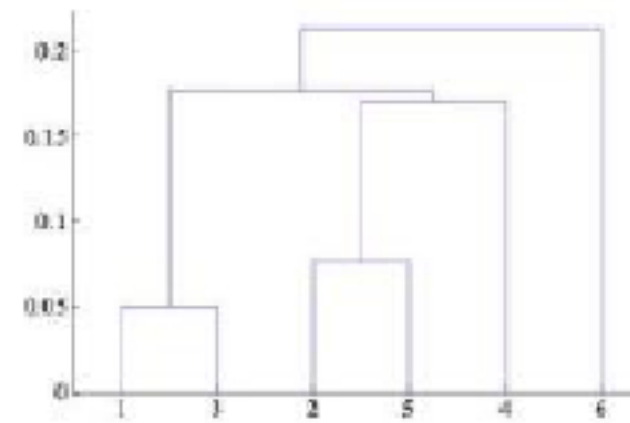
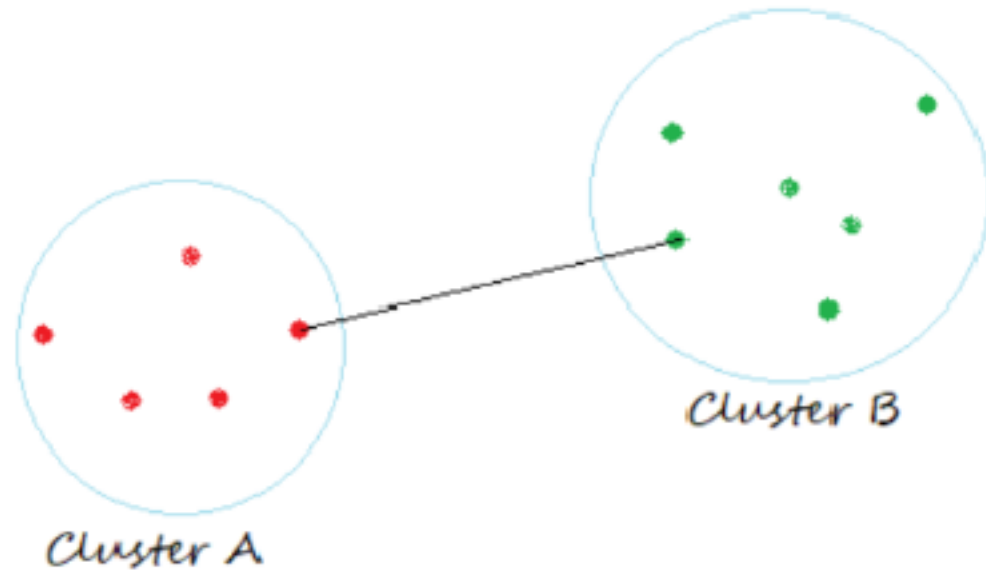


- 2 Hierarchical clustering (e.g., agglomerative clustering, divisive clustering)
- Partitions can be visualized using a tree structure (a dendrogram)
 - Does not need the number of clusters as input
 - Possible to view partitions at different levels of granularities using different K

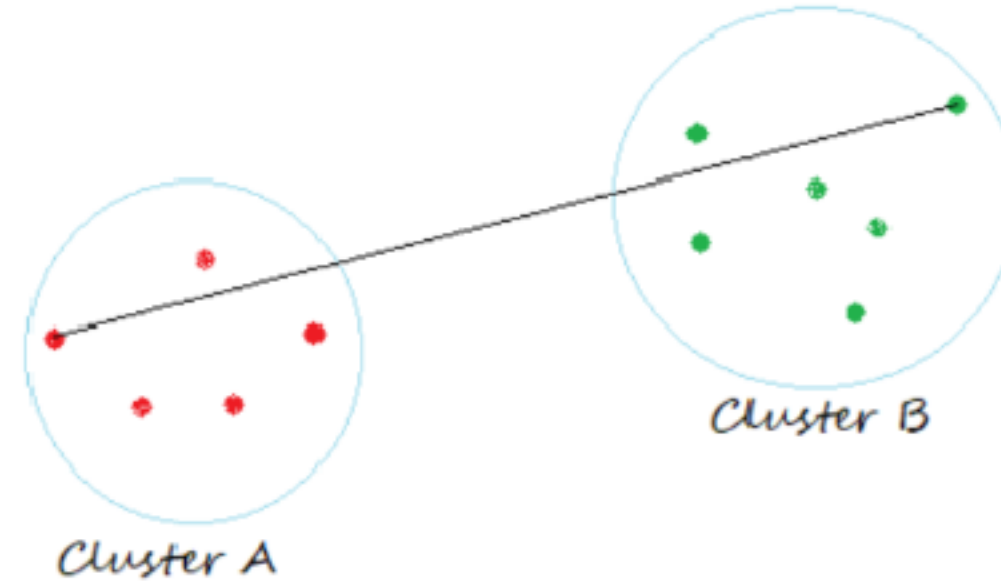


Different types agglomerative clustering

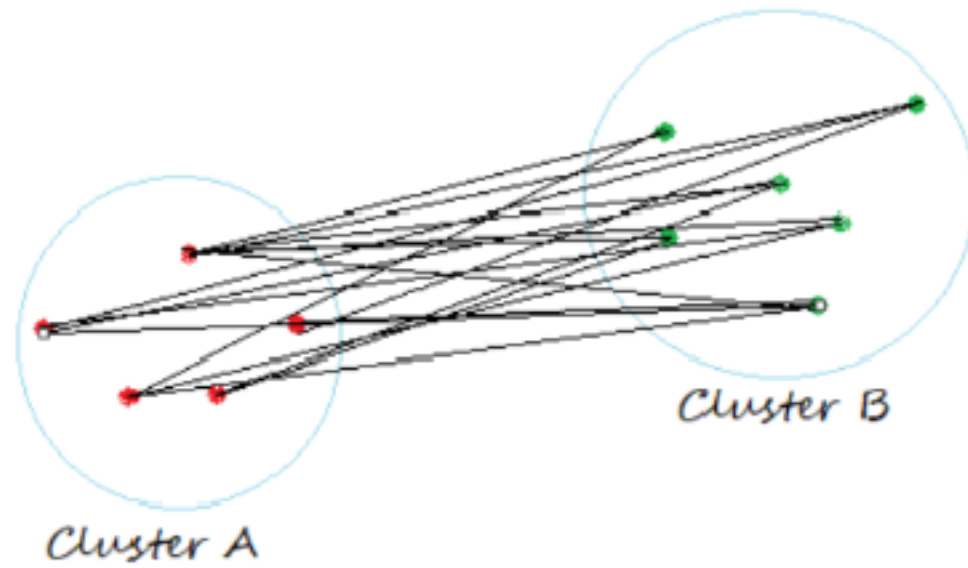
Single Linkage



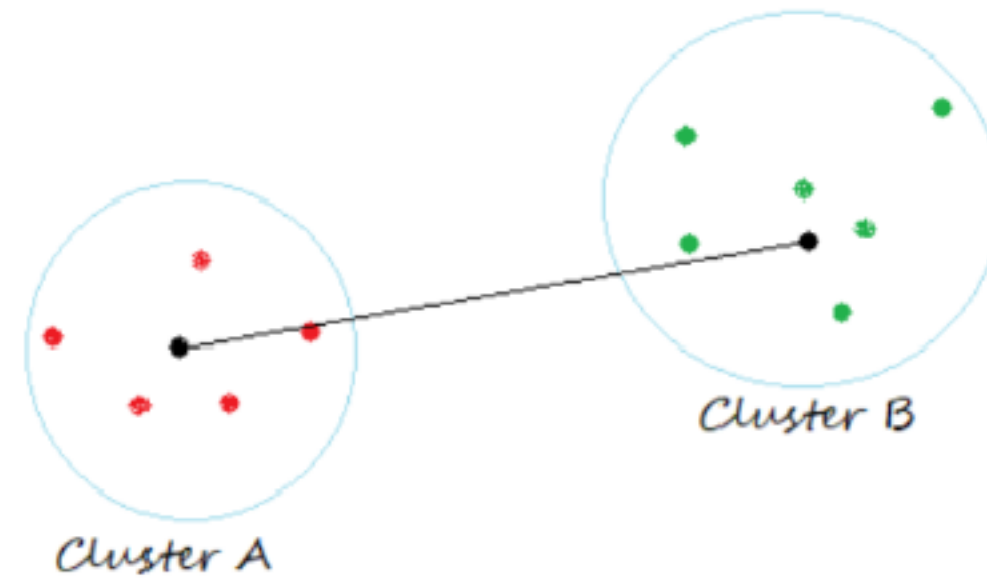
Complete Linkage



Average Linkage



Centroid Linkage



Single Linkage Hierarchical Clustering

Algorithm:

1. Calculate how far each datapoint is from each other
2. Find the closest datapoints and cluster them
3. Recalculate how far each datapoint is from each other and from other clusters
4. Repeat step 2 until convergence.

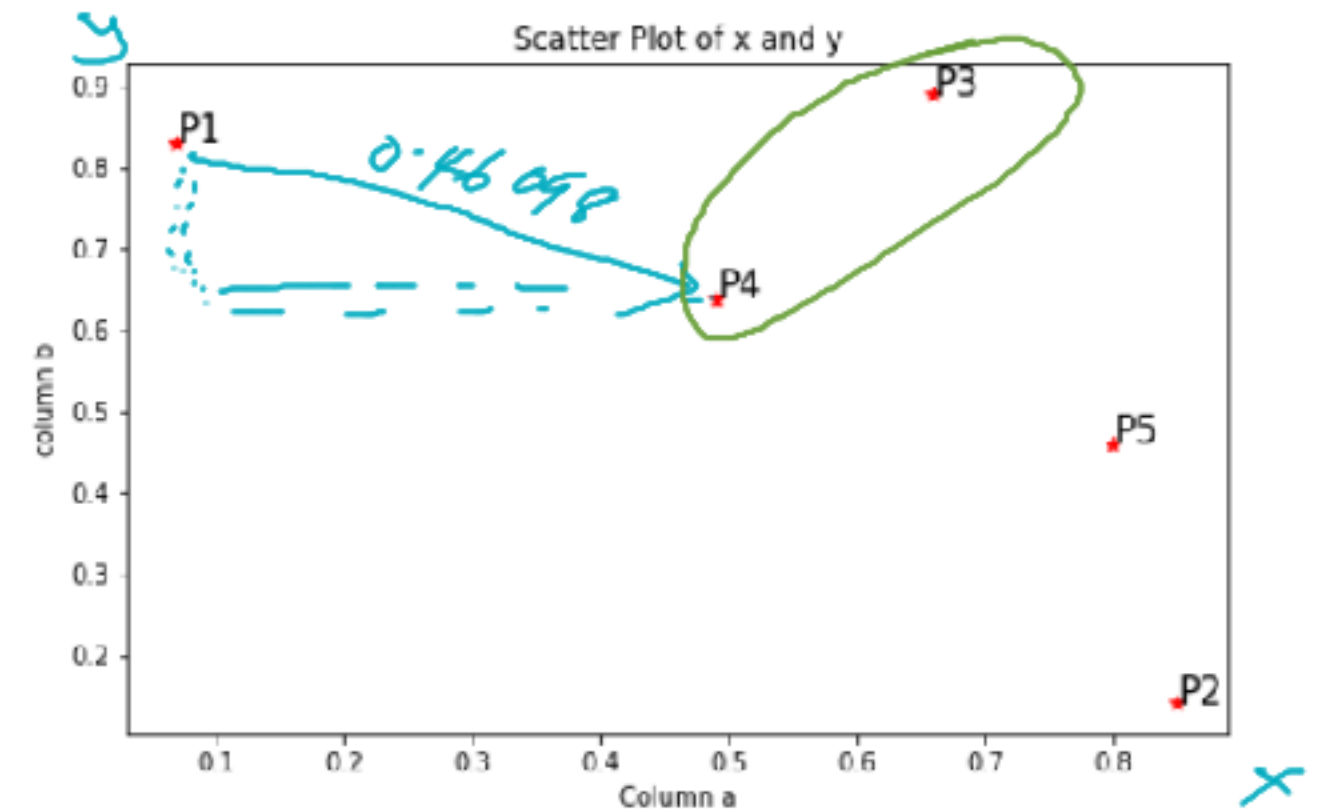
Source:

<https://www.analyticsvidhya.com/blog/2021/06/single-link-hierarchical-clustering-clearly-explained/>

Single Linkage Hierarchical Clustering-Example

	P ₁	P ₂	(P ₃ , P ₄)	P ₅
P ₁	0			
P ₂	1.04	0		
(P ₃ , P ₄)	0.46098		0	
P ₅	0.81841	0.32		0

Point	a	b
P1	0.07	0.83
P2	0.85	0.14
P3	0.66	0.89
P4	0.49	0.64
P5	0.80	0.46



Calculate the euclidean distance

For our convenience, will be considering only the lower bound values of the matrix as shown below.

①

dist (P₃, P₄) and P₁

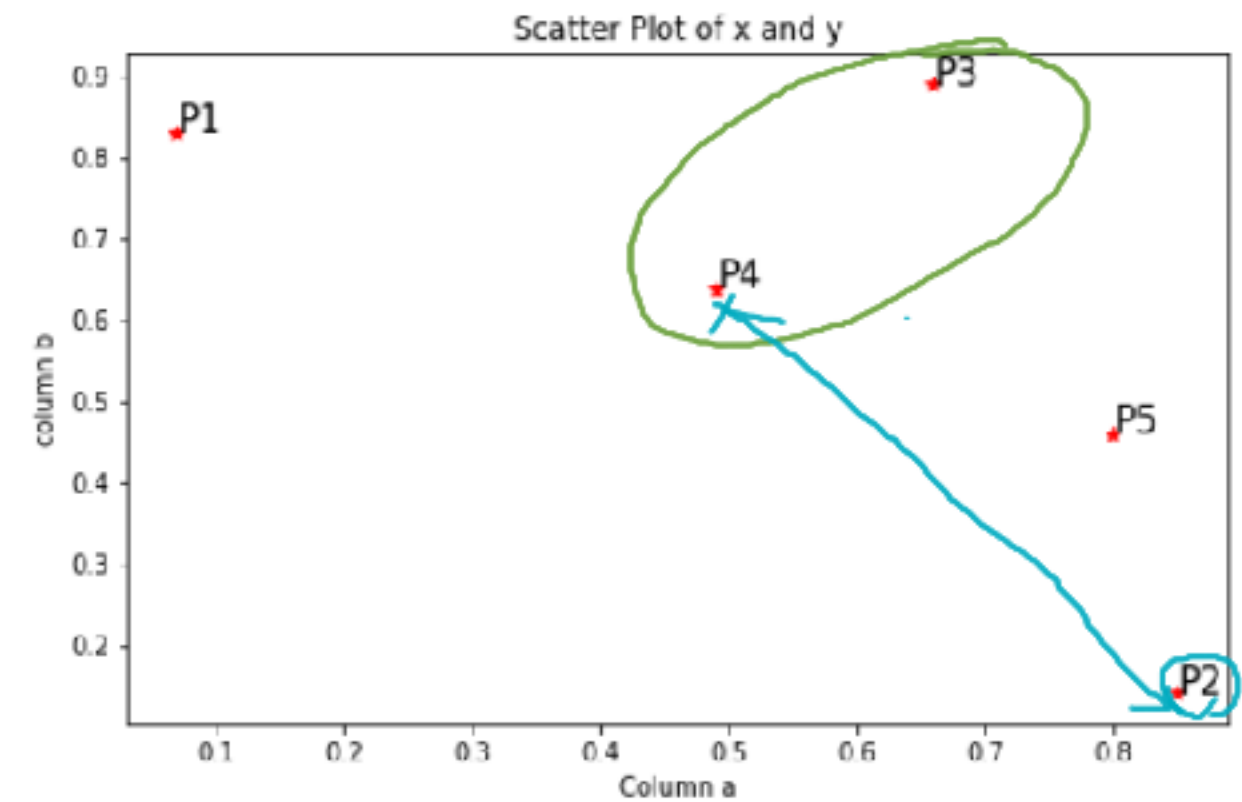
$$\begin{aligned}
 & \min (\text{dist}(P_3, P_1), \text{dist}(P_4, P_1)) \\
 &= \min (0.59304, 0.46098) \\
 &= \underline{0.46098}
 \end{aligned}$$

	P1	P2	P3	P4	P5
P1	0				
P2	1.04139	0			
P3	0.59304	0.77369	0		
P4	0.46098	0.61612	0.30232	0	
P5	0.81841	0.32388	0.45222	0.35847	0

Single Linkage Hierarchical Clustering-Example

	P ₁	P ₂	(P ₃ , P ₄)	P ₅
P ₁	0			
P ₂	0.59304	0		
(P ₃ , P ₄)	0.46098	0.61612	0	
P ₅	0.81841	0.32388	0.45222	0

	a	b
Point		
P1	0.07	0.83
P2	0.85	0.14
P3	0.66	0.89
P4	0.49	0.64
P5	0.80	0.46



For our convenience, will be considering only the lower bound values of the matrix as shown below.

② dist (P₃, P₄) and P₂

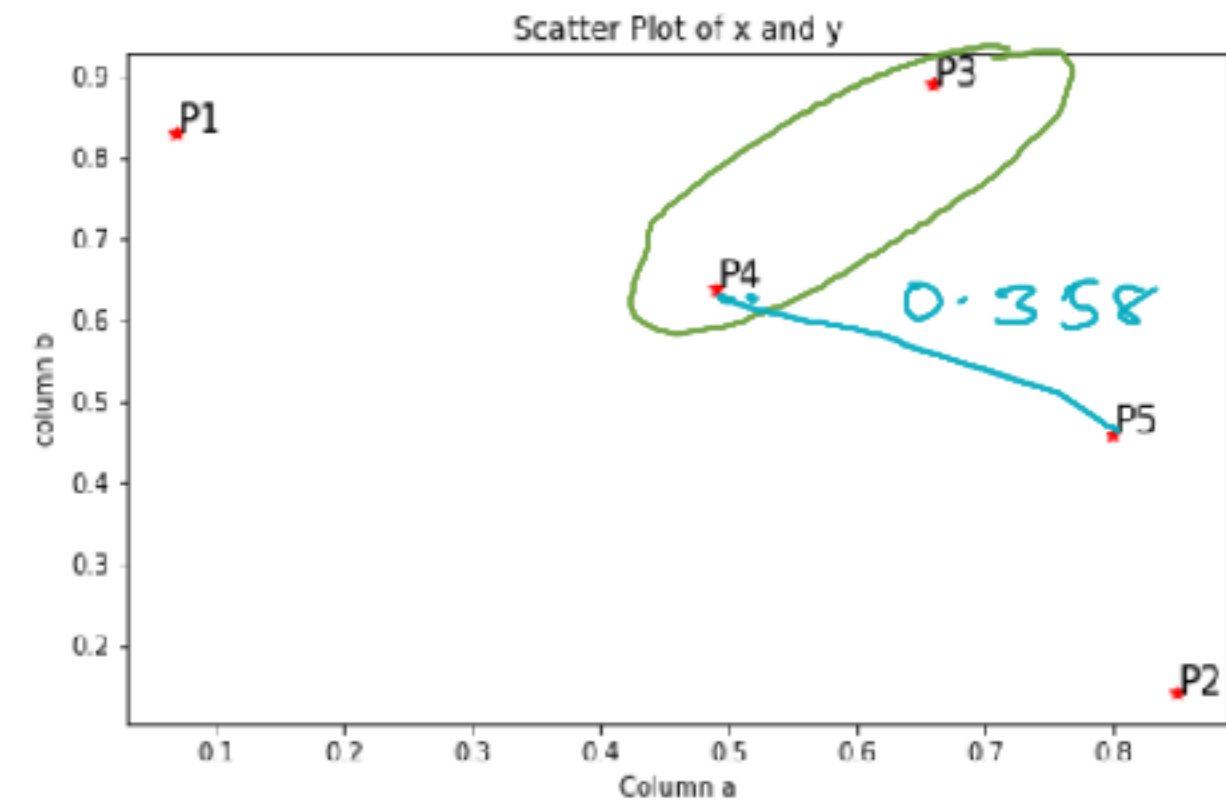
$$\begin{aligned}
 & \min (\text{dist}(P_3, P_2), \text{dist}(P_4, P_2)) \\
 &= \min (0.77369, 0.61612) \\
 &= \underline{0.61612}
 \end{aligned}$$

	P1	P2	P3	P4	P5
P1	0				
P2	1.04139	0			
P3	0.59304	0.77369	0		
P4	0.46098	0.61612	0.30232	0	
P5	0.81841	0.32388	0.45222	0.35847	0

Single Linkage Hierarchical Clustering-Example

	P ₁	P ₂	(P ₃ , P ₄)	P ₅
P ₁	0			
P ₂	1.04139	0		
(P ₃ , P ₄)	0.46098	0.61612	0	
P ₅	0.81841	0.32388	0.35847	0

	a	b
Point		
P1	0.07	0.83
P2	0.85	0.14
P3	0.66	0.89
P4	0.49	0.64
P5	0.80	0.46



③ dist (P₃, P₄) and P₅

$$\begin{aligned}
 & \min (\text{dist}(P_3, P_5), \text{dist}(P_4, P_5)) \\
 &= \min (0.45222, 0.35847) \\
 &= \underline{0.35847}
 \end{aligned}$$

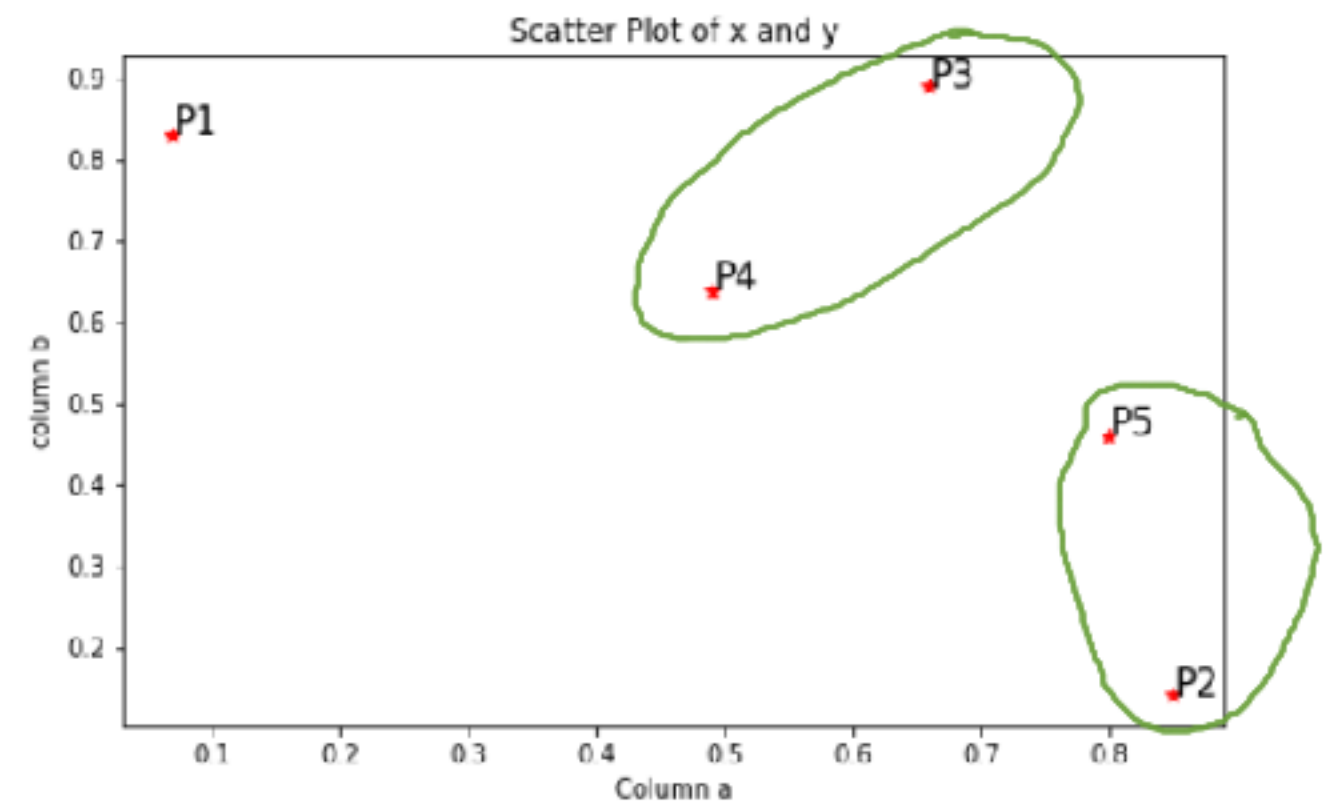
For our convenience, will be considering only the lower bound values of the matrix as shown below.

	P1	P2	P3	P4	P5
P1	0				
P2	1.04139	0			
P3	0.59304	0.77369	0		
P4	0.46098	0.61612	0.30232	0	
P5	0.81841	0.32388	0.45222	0.35847	0

Single Linkage Hierarchical Clustering-Example

	P ₁	P ₂	(P ₃ , P ₄)	P ₅
P ₁	0			
P ₂	1.04139	0		
(P ₃ , P ₄)	0.46098	0.61612	0	
P ₅	0.81841	0.32388	0.35847	0

	a	b
Point		
P1	0.07	0.83
P2	0.85	0.14
P3	0.66	0.89
P4	0.49	0.64
P5	0.80	0.46



For our convenience, will be considering only the lower bound values of the matrix as shown below.

	P1	P2	P3	P4	P5
P1	0				
P2	1.04139	0			
P3	0.59304	0.77369	0		
P4	0.46098	0.61612	0.30232	0	
P5	0.81841	0.32388	0.45222	0.35847	0

To cluster (P₂, P₅)

Calculate new distance

	P ₁	(P ₂ , P ₅)	(P ₃ , P ₄)
P ₁			
(P ₂ , P ₅)			
(P ₃ , P ₄)			

	P ₁	(P ₂ , P ₅)	(P ₃ , P ₄)
P ₁	0		
(P ₂ , P ₅)	0.81	0	
(P ₃ , P ₄)	0.46		0

Distance (P₂, P₅) and P₁

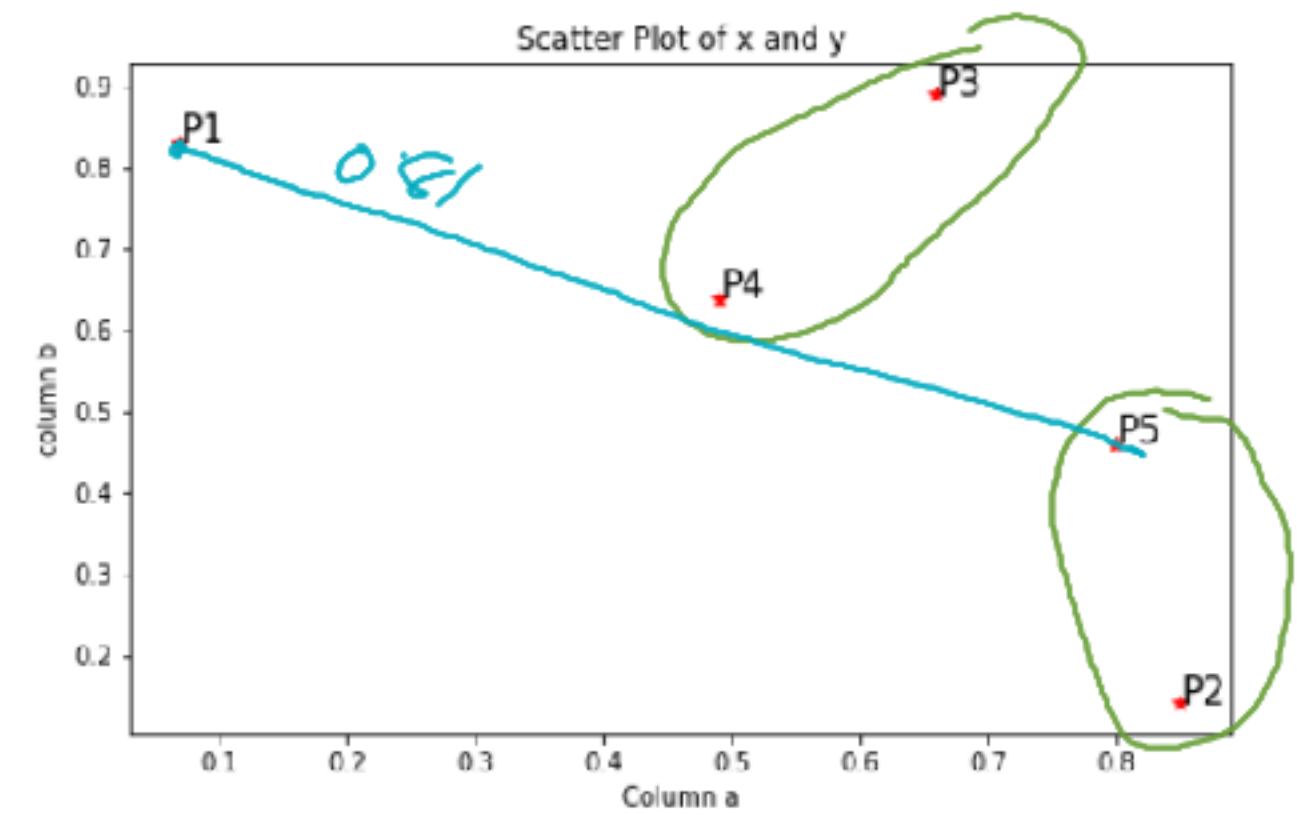
$$= \min(\text{dist}(P_2, P_1), \text{dist}(P_5, P_1))$$

$$= \min(1.04, 0.81)$$

$$= \underline{0.81}$$

Single Linkage Hierarchical Clustering-Example

	a	b
Point		
P1	0.07	0.83
P2	0.85	0.14
P3	0.66	0.89
P4	0.49	0.64
P5	0.80	0.46



For our convenience, will be considering only the lower bound values of the matrix as shown below.

	P1	P2	P3	P4	P5
P1	0				
P2	1.04139	0			
P3	0.59304	0.77369	0		
P4	0.46098	0.61612	0.30232	0	
P5	0.81841	0.32388	0.45222	0.35847	0

	P ₁	(P ₂ , P ₅)	(P ₃ , P ₄)
P ₁	0		
(P ₂ , P ₅)	0.81	0	
(P ₃ , P ₄)	0.46	0.35	0

Distance (P₂, P₅) and (P₃, P₄)

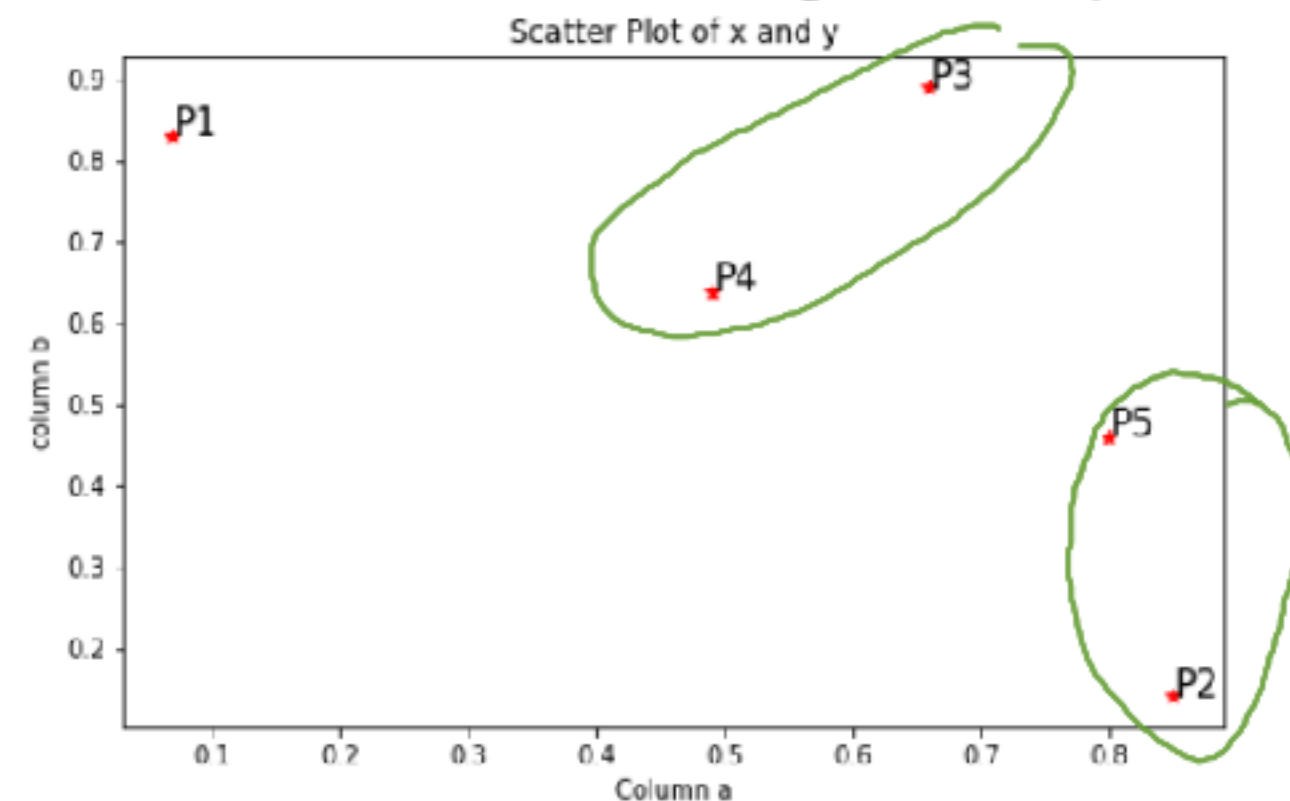
$$= \min (\text{dist}(P_2, P_3), \text{dist}(P_5, P_3), \text{dist}(P_2, P_4), \text{dist}(P_5, P_4))$$

$$= \min (0.77, 0.45, 0.61, 0.35)$$

$$= \underline{0.35}$$

Single Linkage Hierarchical Clustering-Example

	a	b
Point		
P1	0.07	0.83
P2	0.85	0.14
P3	0.66	0.89
P4	0.49	0.64
P5	0.80	0.46



For our convenience, will be considering only the lower bound values of the matrix as shown below.

	P1	P2	P3	P4	P5
P1	0				
P2	1.04139	0			
P3	0.59304	0.77369	0		
P4	0.46098	0.61612	0.30232	0	
P5	0.81841	0.32388	0.45222	0.35847	0

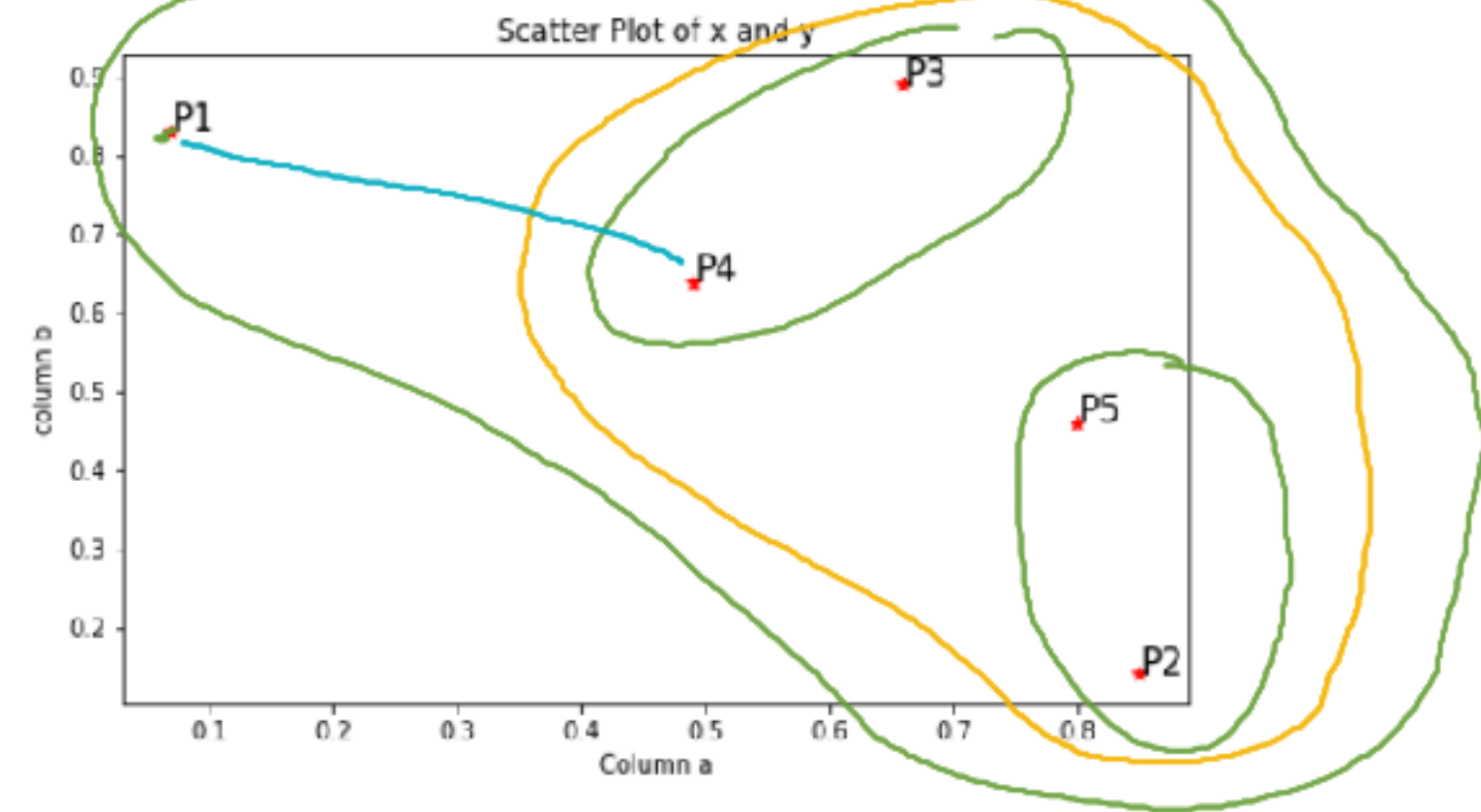
	P_1	(P_2, P_5)	(P_3, P_4)
P_1	0		
(P_2, P_5)	0.81	0	
(P_3, P_4)	0.46	0.35	0



	P_1	(P_2, P_5, P_3, P_4)
P_1	0	
(P_2, P_5, P_3, P_4)	0.46098	0

Single Linkage Hierarchical Clustering-Example

	a	b
Point		
P1	0.07	0.83
P2	0.85	0.14
P3	0.66	0.89
P4	0.49	0.64
P5	0.80	0.46



For our convenience, will be considering only the lower bound values of the matrix as shown below.

	P1	P2	P3	P4	P5
P1	0				
P2	1.04139	0			
P3	0.59304	0.77369	0		
P4	0.46098	0.61612	0.30232	0	
P5	0.81841	0.32388	0.45222	0.35847	0

