DBSCAN (Density Based Spatial Clustering of Applications with Noise).

Apply DBSCAN algorithm to given data points and create clusters with min_pts = 4 and $\varepsilon = 1.9$.

Datapoints :-

$P_1(3,7)$ $P_2(4,6)$ $P_3(5,5)$ $P_4(6,4)$ $P_5(7,3)$

$P_6(6,2)$ $P_7(7,2)$ $P_8(8,4)$ $P_9(3,3)$ $P_{10}(2,6)$

$P_{11}(3,5)$ $P_{12}(2,4)$

STEP-1 :-

Calculate the distance of every datapoints with other datapoints using the eudlidean distance formula and tabulate it.

$$d(A,B) = \sqrt{(x_2-x_1)^2+(y_2-y_1)^2}$$

(x_1, y_1) (x_2, y_2)

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}
P_1	0	1.4	2.8	4.2	5.66	5.83	6.4	5.83	4	1.41	2	3.16
P_2		0	1.41	2.82	4.2	4.47	5	4.47	3.16	2	1.41	2.82
P_3			0	1.41	2.82	3.16	3.61	3.16	2.82	3.16	2	3.16
P_4				0	1.41	2	2.23	2	3.16	4.47	3.16	4
P_5					0	1.41	1	1.41	4	5.83	4.24	5.08
P_6						0	1	2.82	3.16	5.65	4.24	4.47
P_7							0	2.23	4.12	6.4	5.8	5.38
P_8								0	5.09	6.32	5.09	6
P_9									0	3.16	2	1.41
P_{10}										0	1.41	2
P_{11}											0	1.41
P_{12}												0

STEP – 2 :-

Determine the neighbours for each data point. Considering the given $\varepsilon = 1.9$.

$N(P_1)$: P_2, P_{10}

$N(P_2)$: P_1, P_3, P_{11}

$N(P_3)$: P_2, P_4

$N(P_4)$: P_3, P_5

$N(P_5)$: P_4, P_6, P_7, P_8

$N(P_6)$: P_5, P_7

$N(P) < 1.9$

P_2	P_{10}
1.4	1.41

Check both Column and Row.

$N(P_7)$: P_5, P_6

$N(P_8)$: P_5

$N(P_9)$: P_{12}

$N(P_{10})$: P_1, P_{11}

$N(P_{11})$: P_2, P_{10}, P_{12}

$N(P_{12})$: P_9, P_{11}

P_1 – Border
Present in Core
P_2

Core

$P_2 \rightarrow P_1, P_3, P_{11}$

Consider P_2 also so, =4.

STEP-3 :-

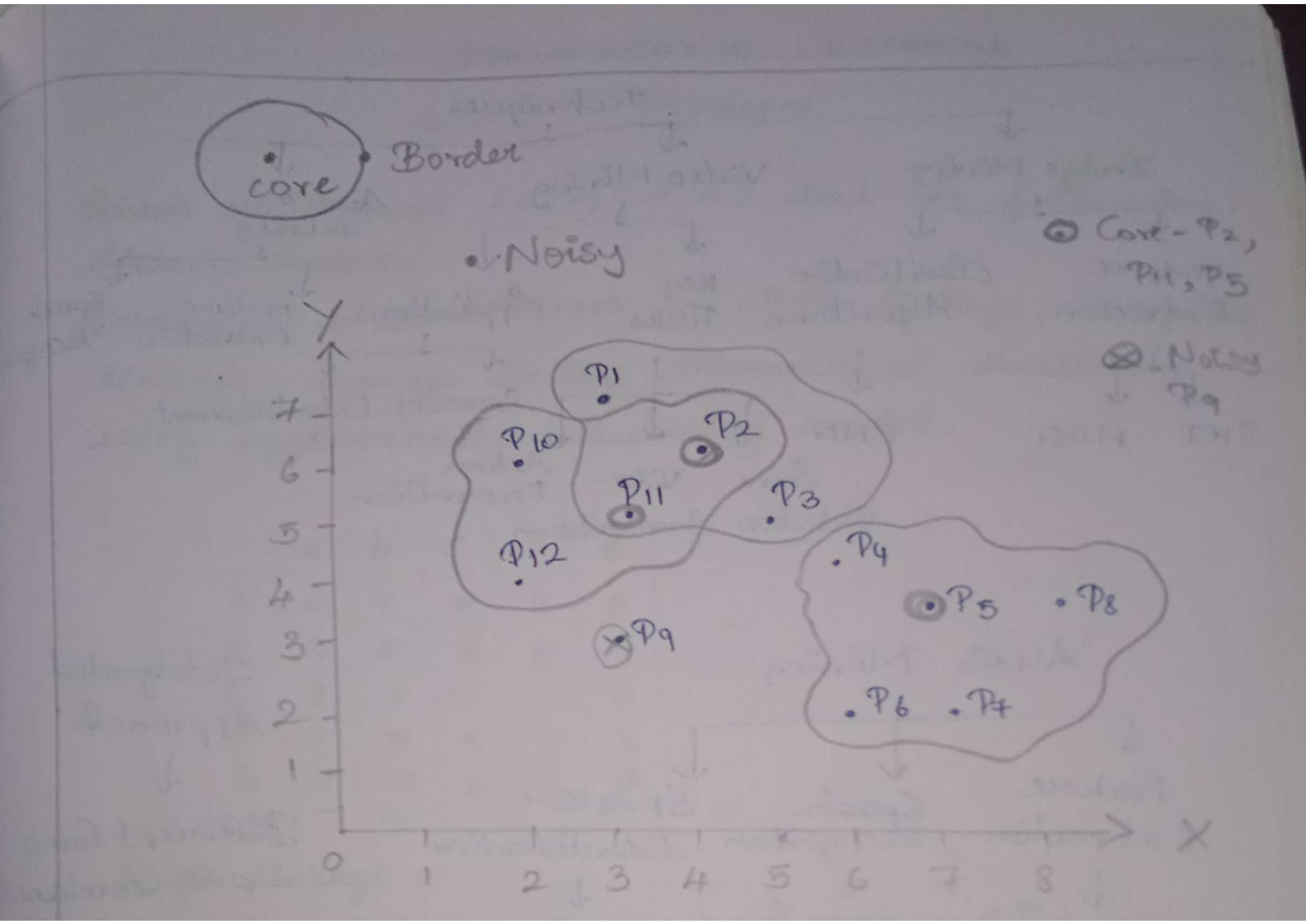
Determine the core datapoint, Noisy datapoint, and border datapoint using min_pts = 4.

Datapoints	Status	
P_1	Noisy	Border
P_2	Core *	
P_3	Noisy	Border
P_4	Noisy	Border
P_5	Core *	
P_6	Noisy	Border
P_7	Noisy	Border
P_8	Noisy	Border
P_9	Noisy	Noisy
P_{10}	Noisy	Border
P_{11}	Core *	
P_{12}	Noisy	Border

$< 4 \rightarrow$ Noisy
↳ Border
$>= 4 \rightarrow$ Core

minimum it should have 4 less than 4 are Noisy datapoint we cannot form cluster greater than 4 are core

Noisy → Not present in any core


core
Border
Noisy
Core - P2, P11, P5
Noisy P9
Y
X
P1
P2
P3
P4
P5
P6
P7
P8
P9
P10
P11
P12
0
1
2
3
4
5
6
7
8