DBSCAN.
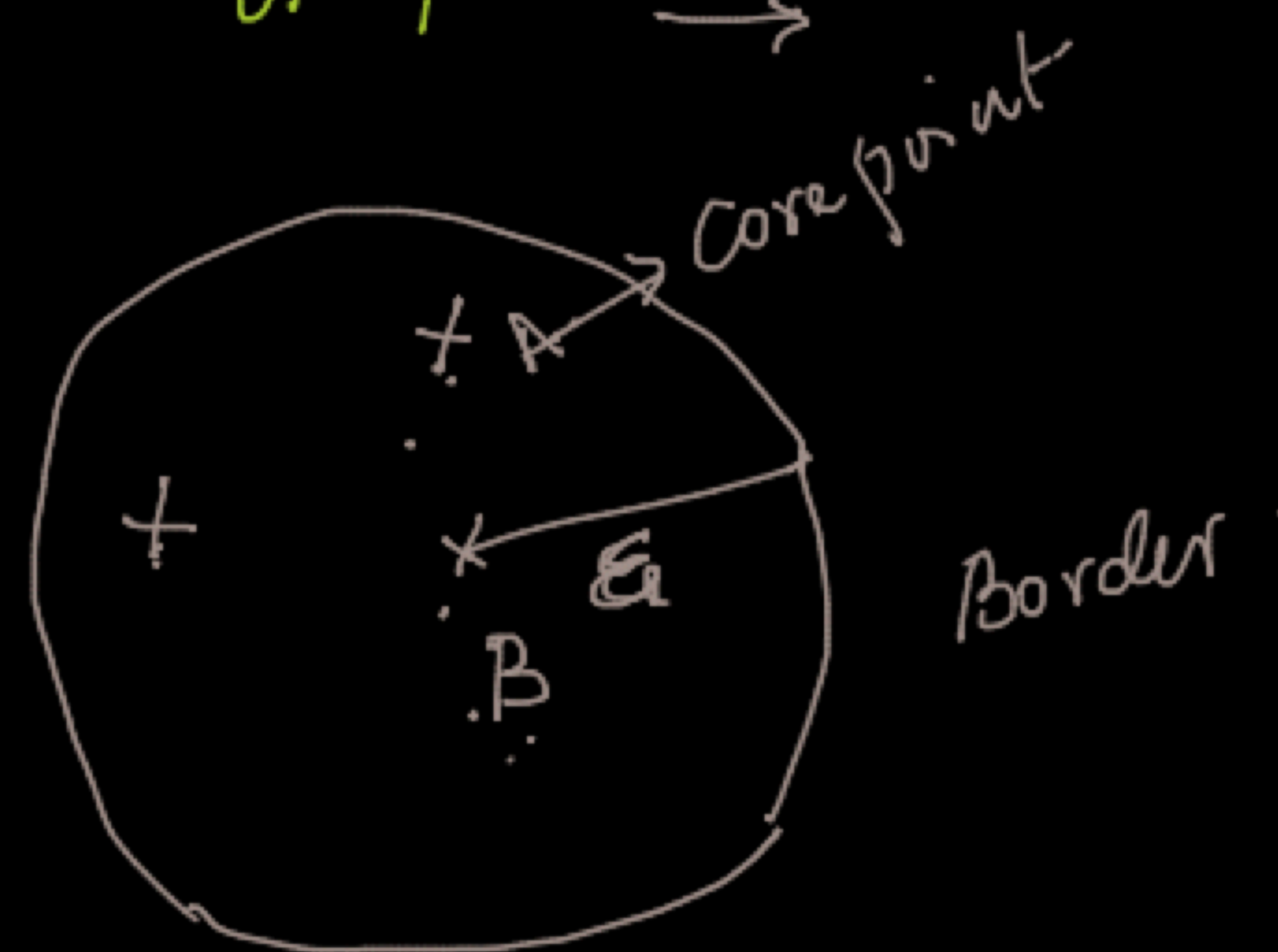
— Density Based spatial Clustering of Applications with Noise
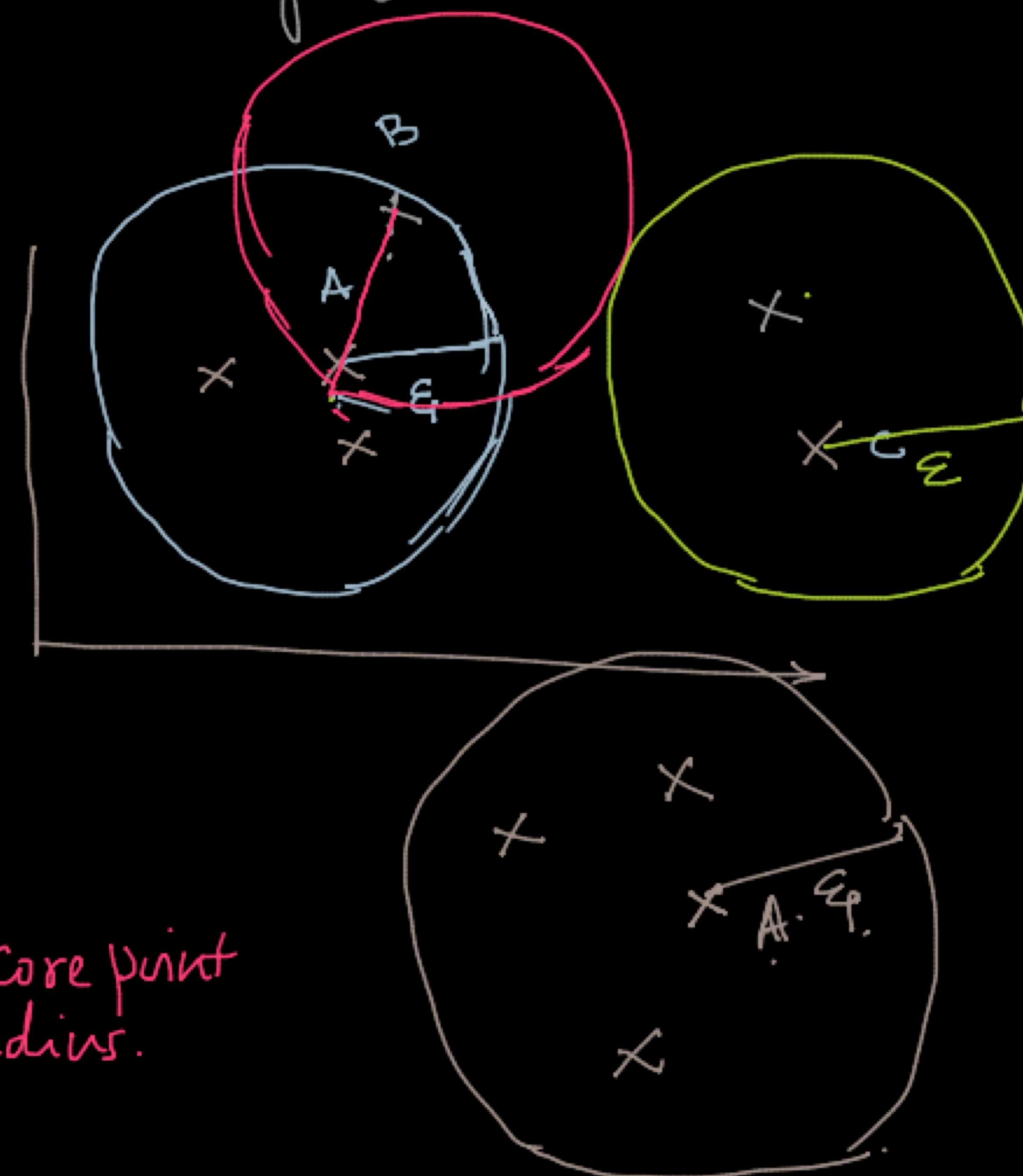
Min points = 3 ✓

Epsilon distance (ε) ✓

— Hyperparameters →

Core point

Border Point

Noise point —

≥ Minpoints

A → Core point.

B → Border points

< Minpoints

— Includes a Core point within ε-radius.

Noise → < Minpoints
→ no core points within ε Radius.

→ Core point
+A
+ B ε

Border.

Noise
+ C ε

< Minpoints.
→ No core points within the ε radius

Noise point

+ + A ε.
+

	x_1	x_2	x_3		
A	—	—	—	Core point	✓
B	—	—	—	Core point	✓
C	—	—	— .	Border point	✓
D	—	—	—	Core point	✓
~~E~~	—	—	—	~~Noise point~~	✓ ✗
F	—	—	—	Border point	✓

Reachable

Min Points.

ε - distance.
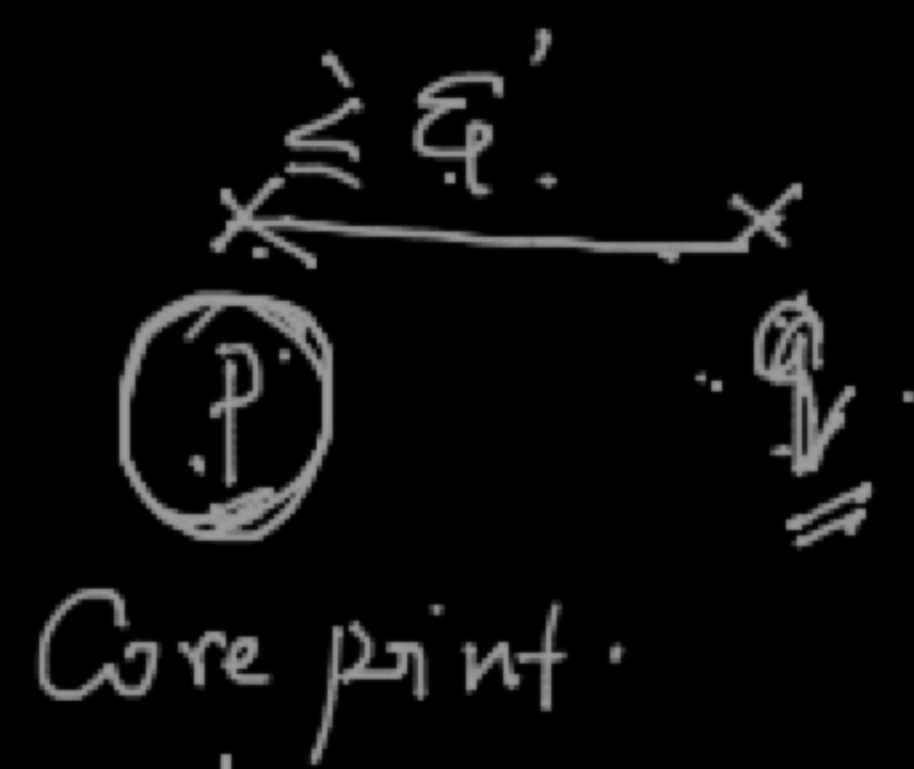
  Hyper param

Core point
Border point
Noise point.

Directly Density Reachable points
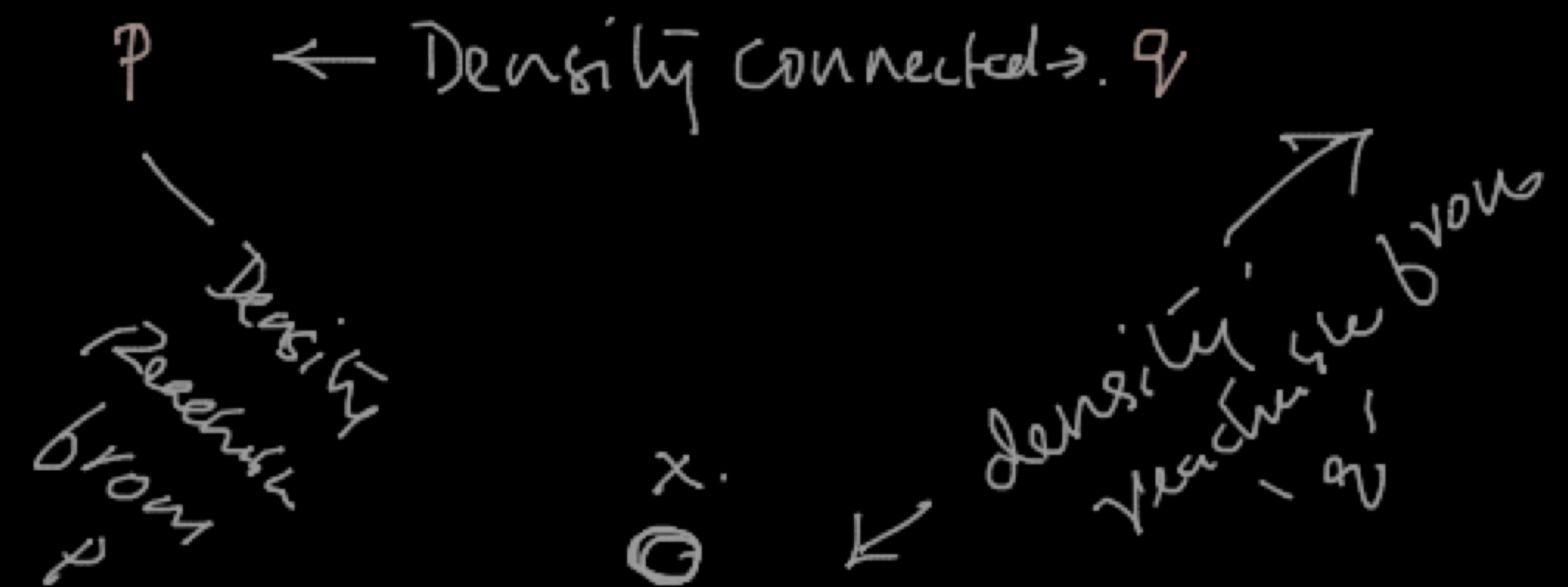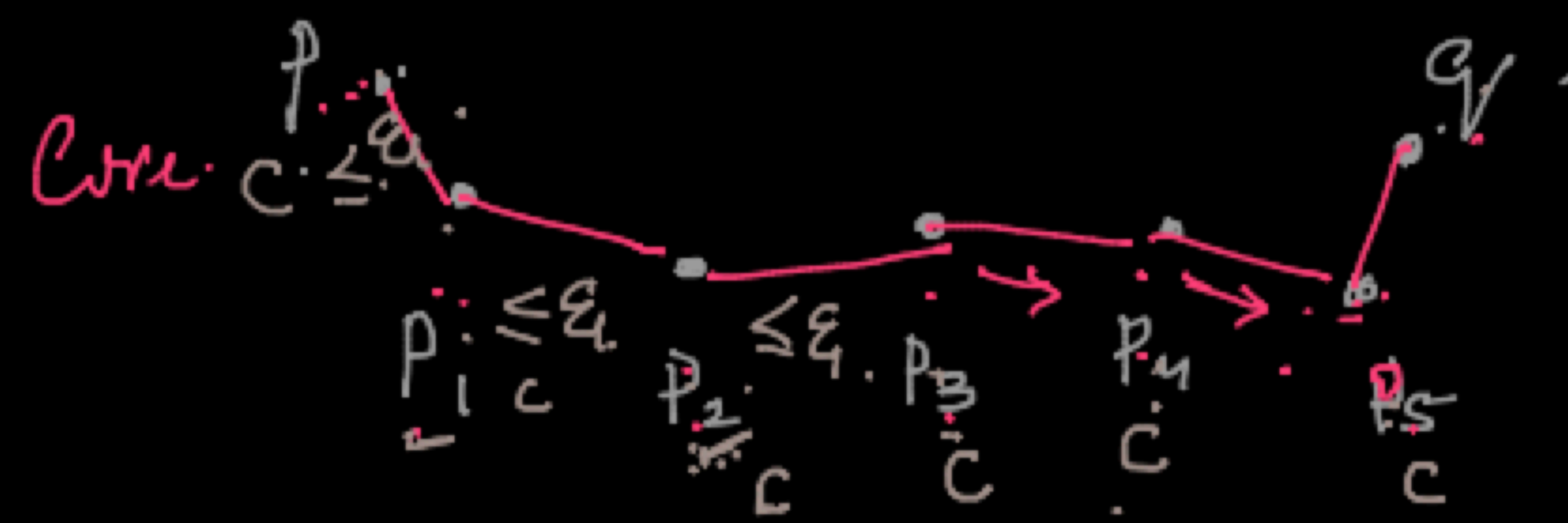Density Reachable. Points.
Density Connected. points.

Reachability.

1. Directly density reachable
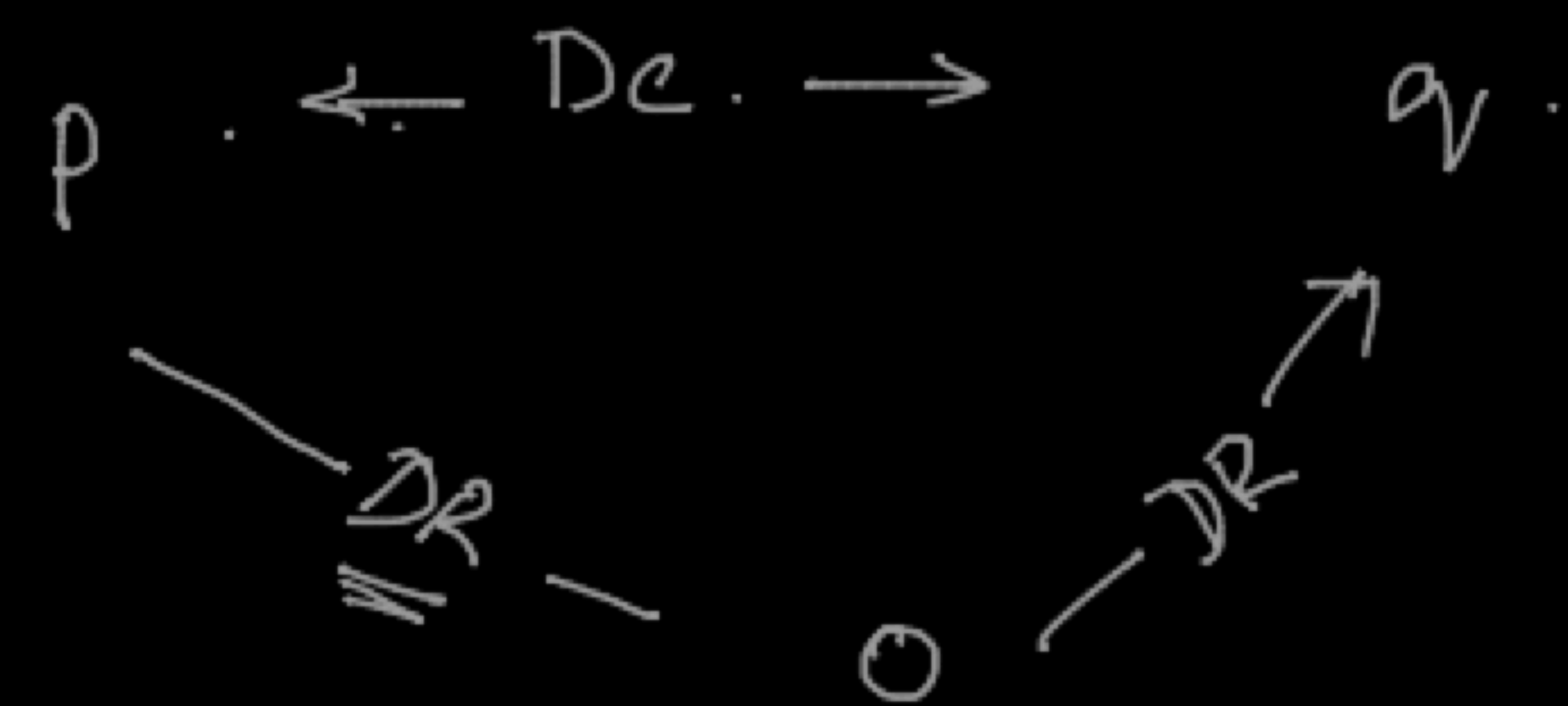
$\to$ 'q' is directly density reachable from 'p'

Core point.
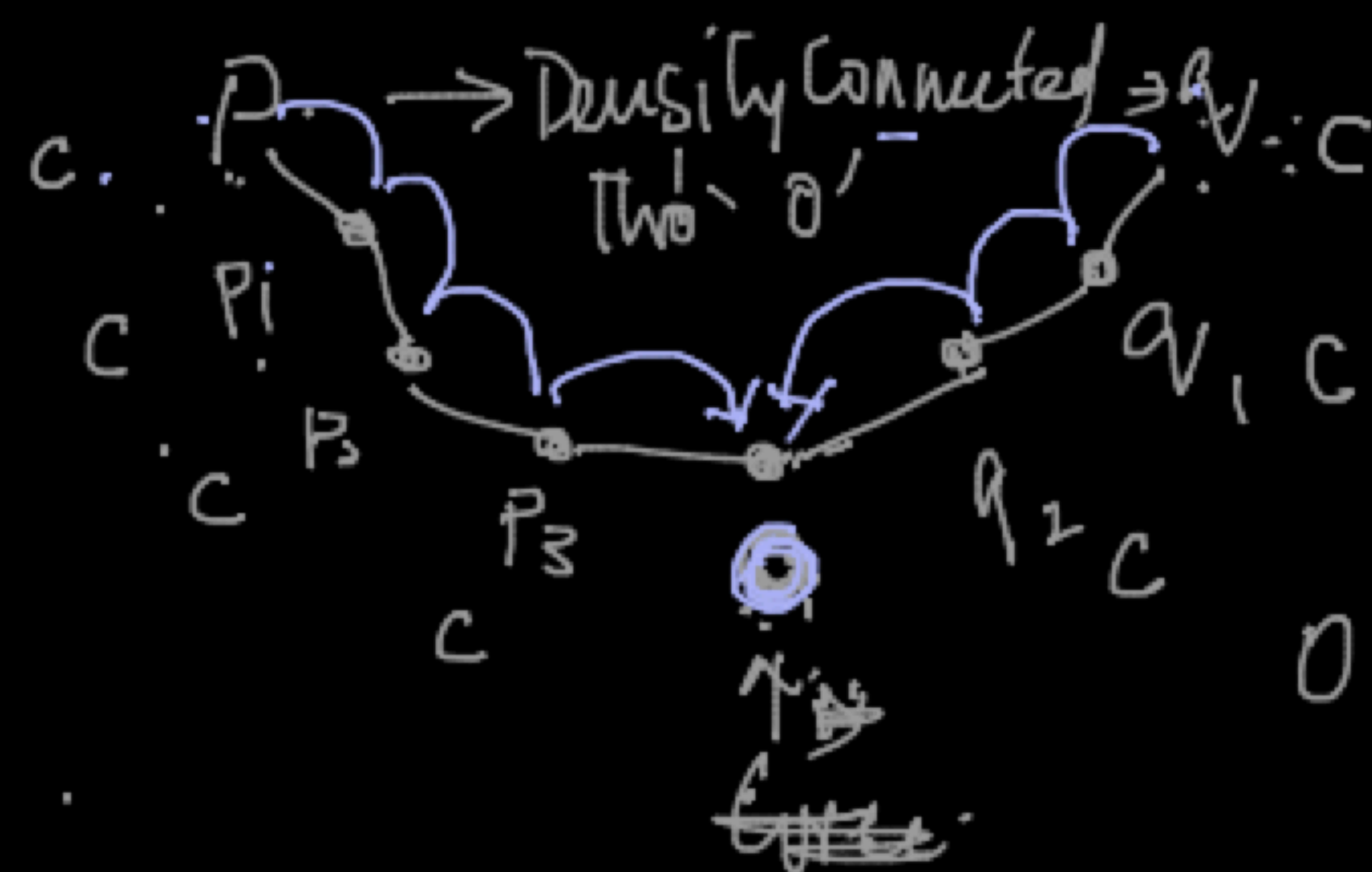
2. Density Reachable ✓✓

$P, P_1, P_2, P_3, P_4, P_5 \to$ Core points.

ε

Density Connected $\ni q$
Tho' O'

3. Density Connected Points.

$P \quad \leftarrow$ Density connected $\to$ q

Density Reachable from p

density reachable from q'

$x.$

O

$P \quad \leftarrow DC \longrightarrow \quad q.$

$\searrow DR \qquad \nearrow DR$

O

O is density reachable from p.

$x_1 \ x_2 \ x_5 \ \cdot \quad \cdot \quad \cdot \quad \cdot \quad x_{100} \rightharpoonup y$

$\rightarrow$ M

B

D_1	D_2
—	—

M.

B

D_2

$x_1 \quad x_2$

y

x

512

512 | Pixels.

$\rightarrow$

$\rightarrow D_1$

$\rightarrow D$
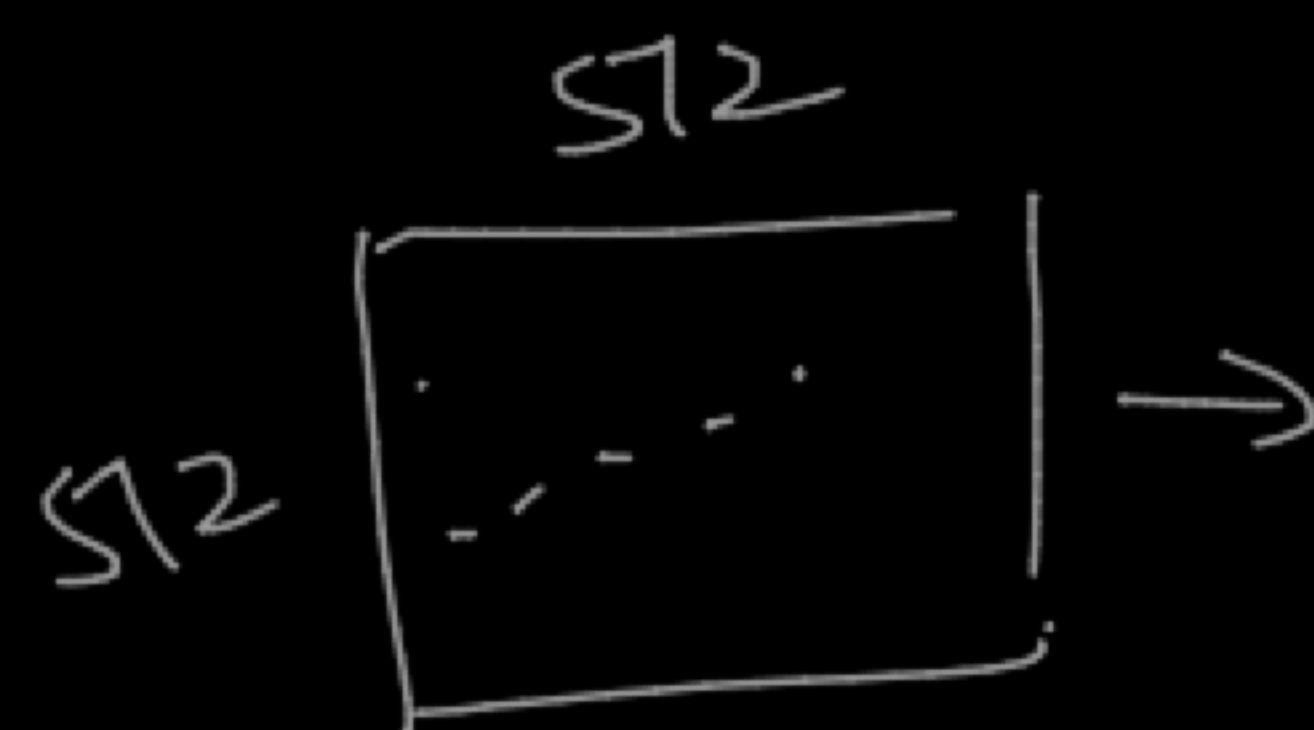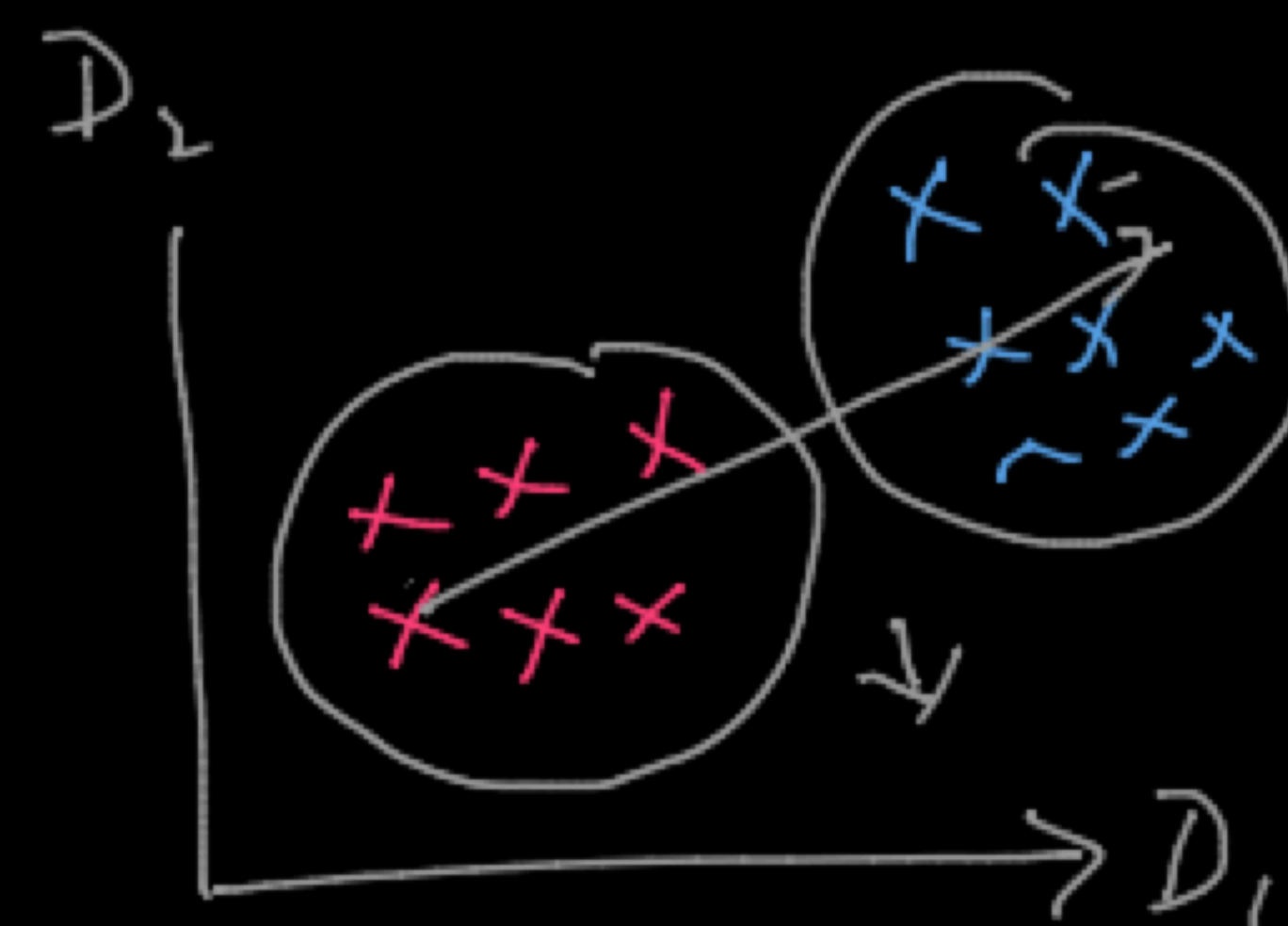
Minpoints $(4) \rightarrow \varepsilon$.

Minpoints $\rightarrow$ 2d or $(3d)$

$x_1 \quad x_2 \quad x_3 \quad \ldots \quad x_d \quad \rightarrow$
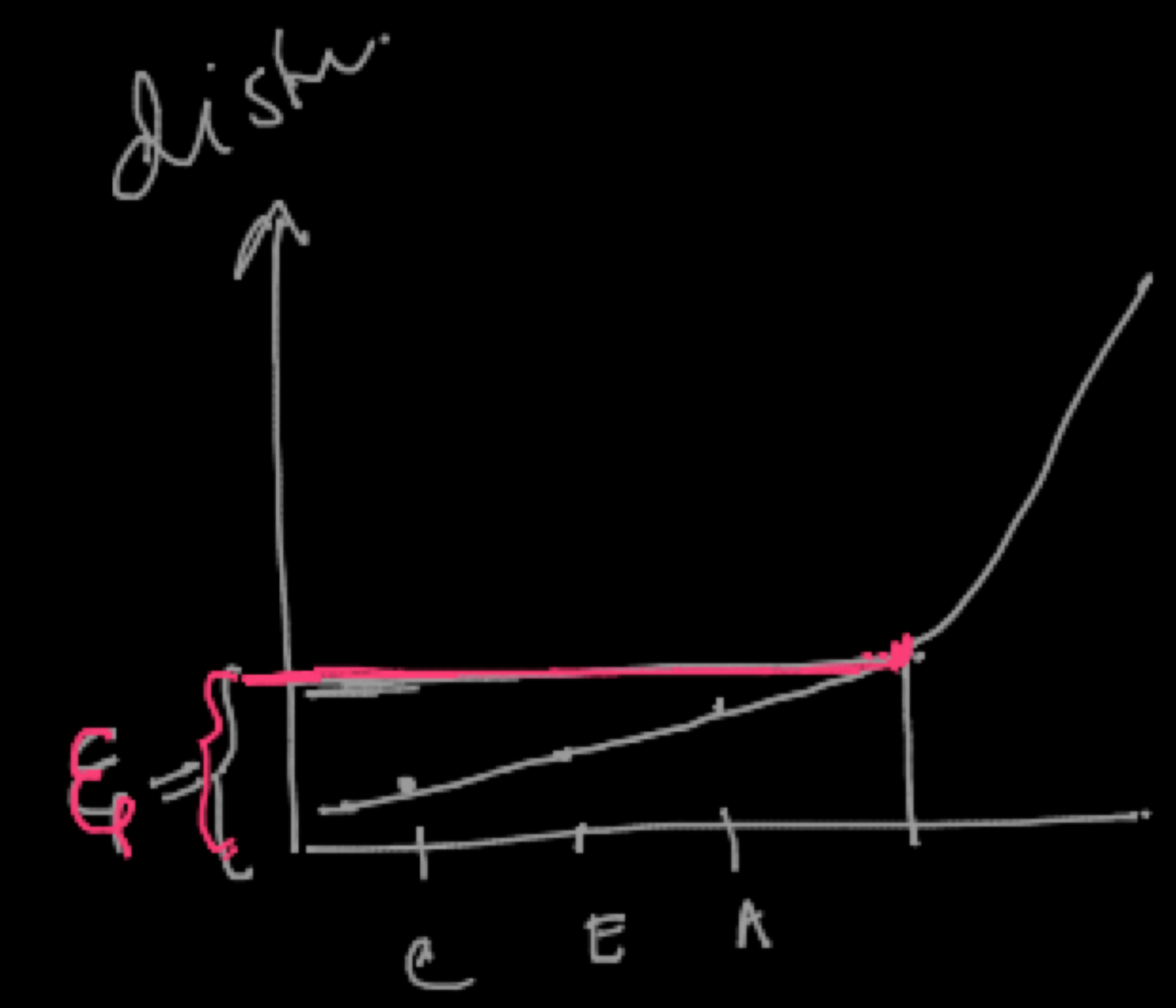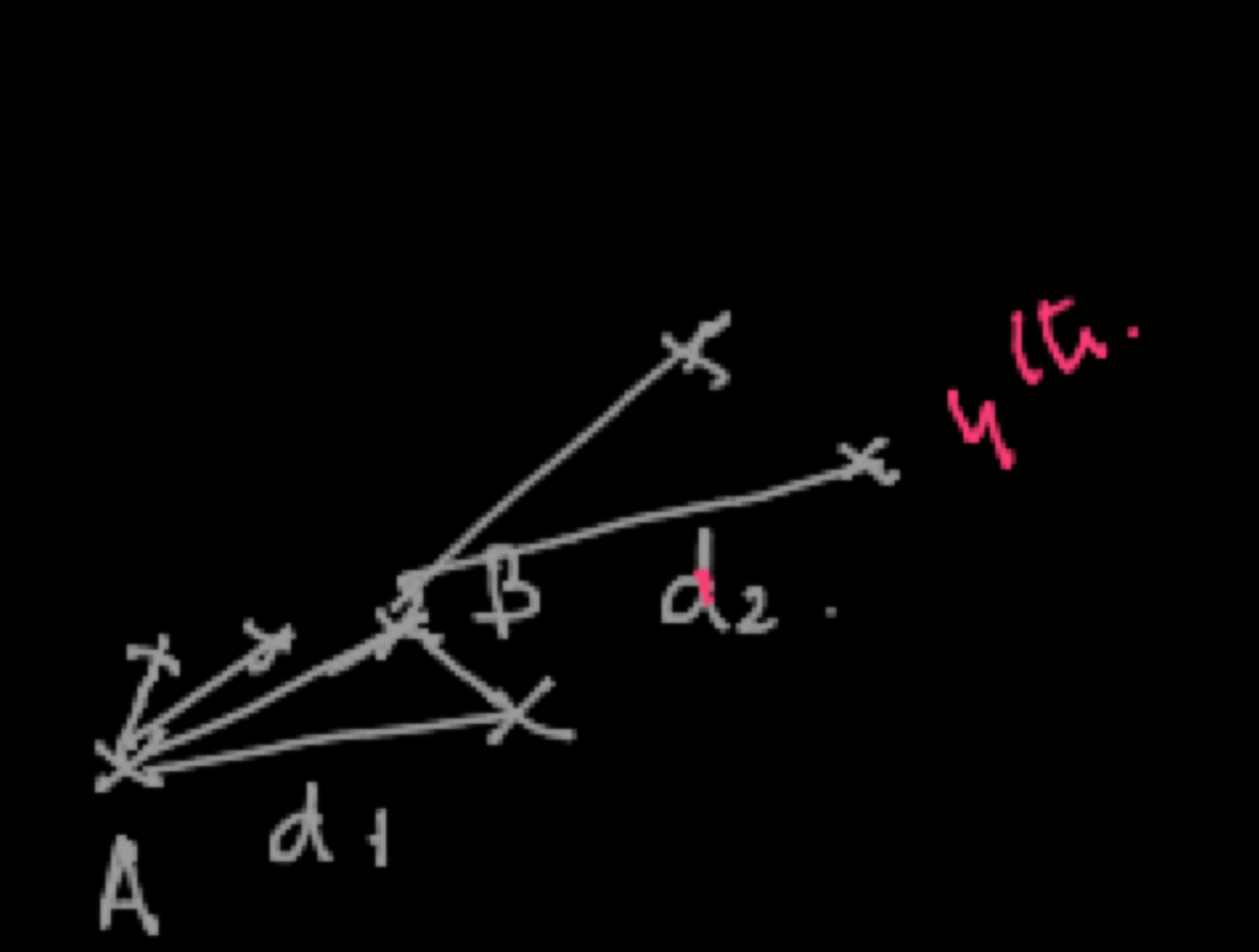
A	d_1
B.	d_2
C	d_3
D.	d_4
E	d_5
F	d_6

ascending Order

dist.

ε {

e   E   A

C	d_3
E	d_5
A	d_1
B	d_2
D	d_4
F	d_6
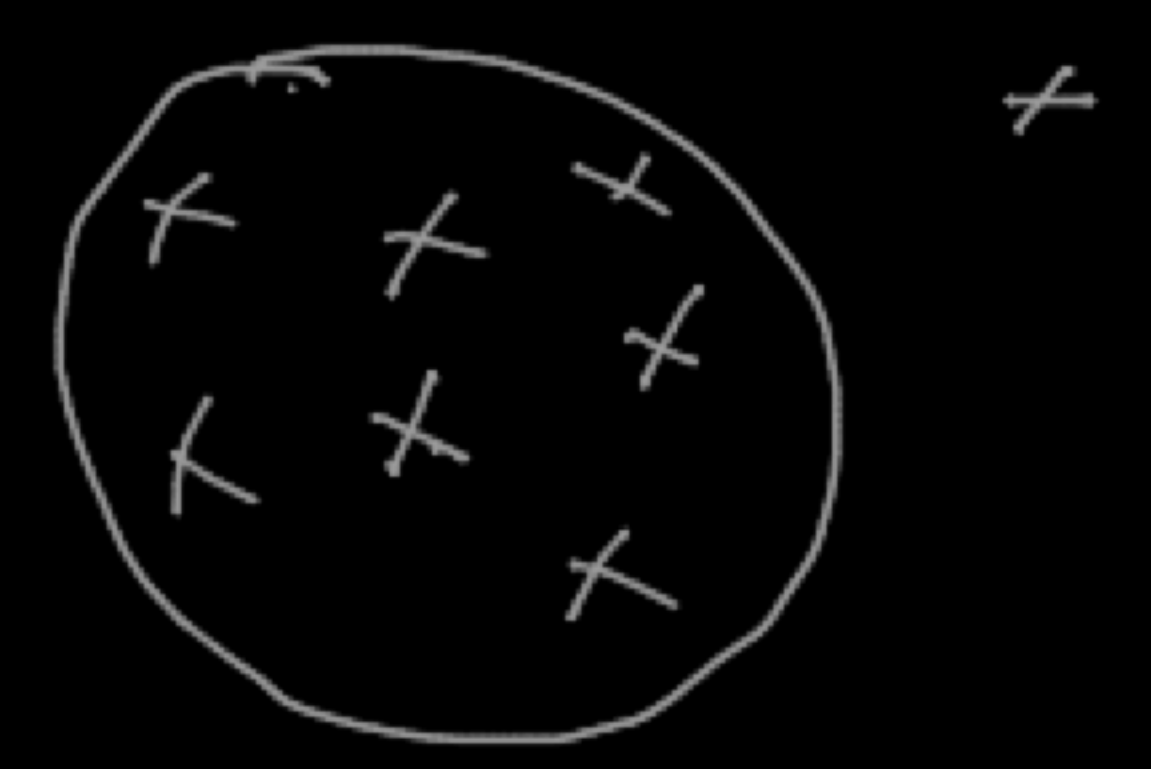
$$\text{Dunn Index} = \frac{\text{Smallest distance between points not in a cluster. } \uparrow \checkmark}{\text{largest distance between points within a cluster. } \downarrow}$$
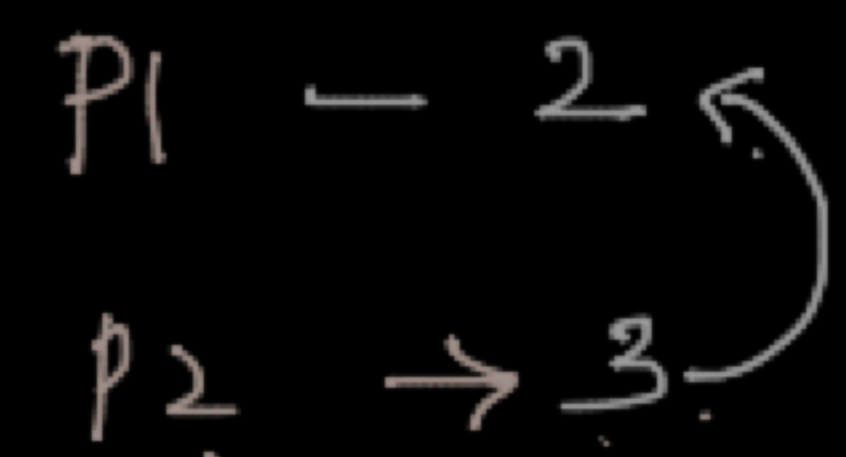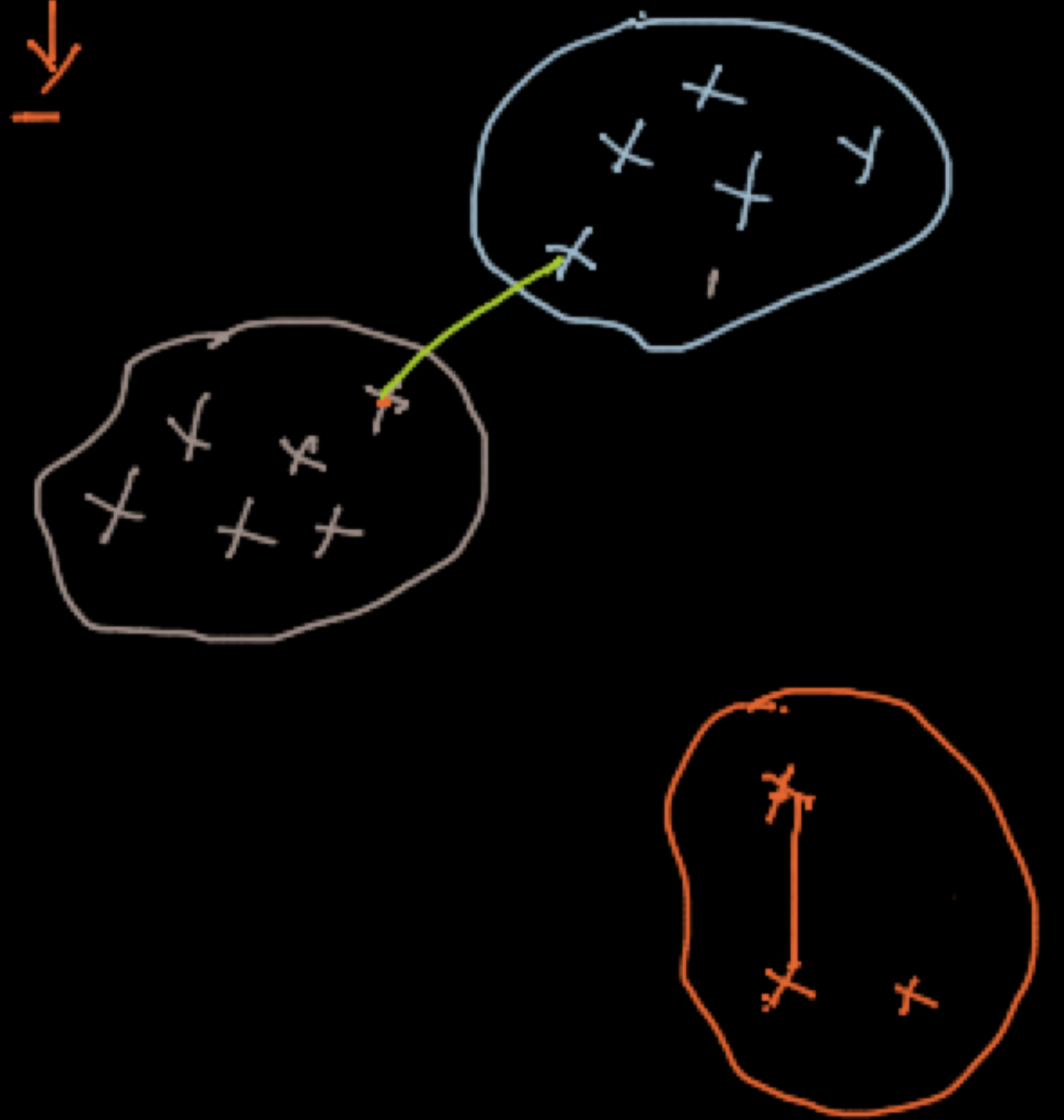
$\Rightarrow$ As large as possible.

$$\text{Silhoutte Score} = \frac{(b-a) \uparrow}{\max(b,a) \rightarrow}$$

$\approx \frac{1}{}, \; 0 \Rightarrow$
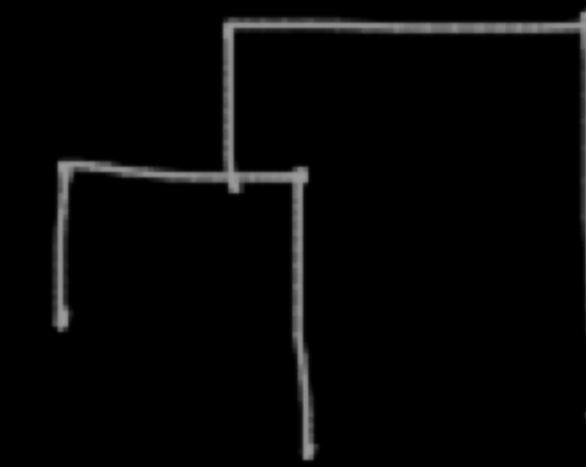
$-$Ve values $\rightarrow$ Wrong grouping of datapoints.

a - average intracluster distance. $\checkmark$ small

b - average intercluster distance. $--$ larg.
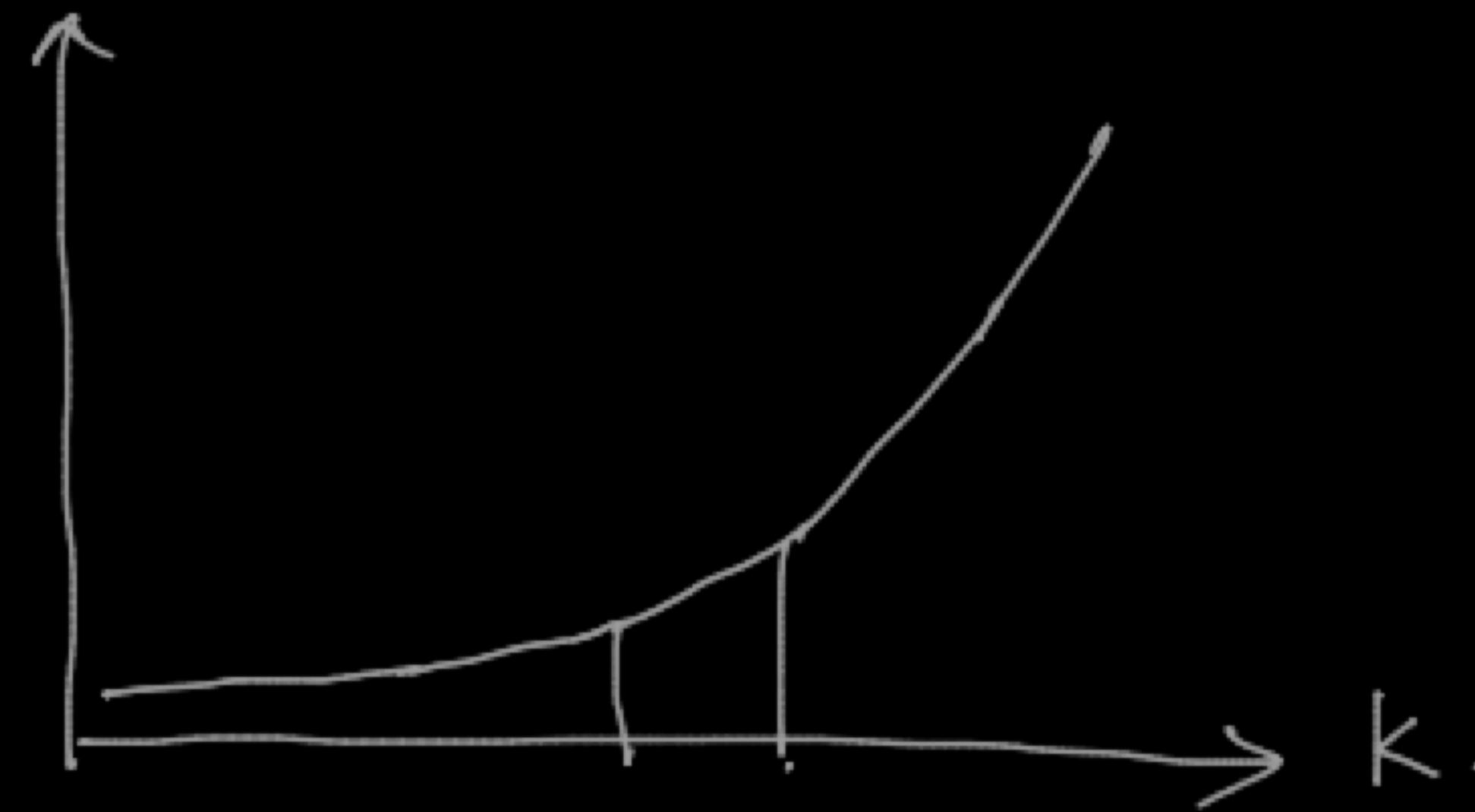
1. Small intracluster distance.

2. Inter Cluster distances

$$P1 \; - \; 2 \; \curvearrowright$$

$$P2 \; \rightarrow \; 3.$$

How many
groups.

$\begin{cases} \overset{\nearrow}{\textcircled{k}} \to & \text{K-means.} \\ & \to \text{How many groups.} \\ & \to \text{HAC.} \\ & \to \end{cases}$

DBSCAN $\to$

$\textcircled{3} \to \bigcirc \quad \bigcirc \quad \bigcirc$ WCSS.

i Dun Ind. $\Rightarrow$ — ✓

, Silh. $\Rightarrow$ — ✓

$\to \textcircled{4} \to \bigcirc \quad \bigcirc \quad \bigcirc$

— ✓ $= 3$

— ✓

K.

K $\to$ No. of groups.

$\begin{cases} \text{K-Means} \\ \text{HAC.} \end{cases}$

DBSCAN · X

X   K.

$\to$ Density