Spherical clusters
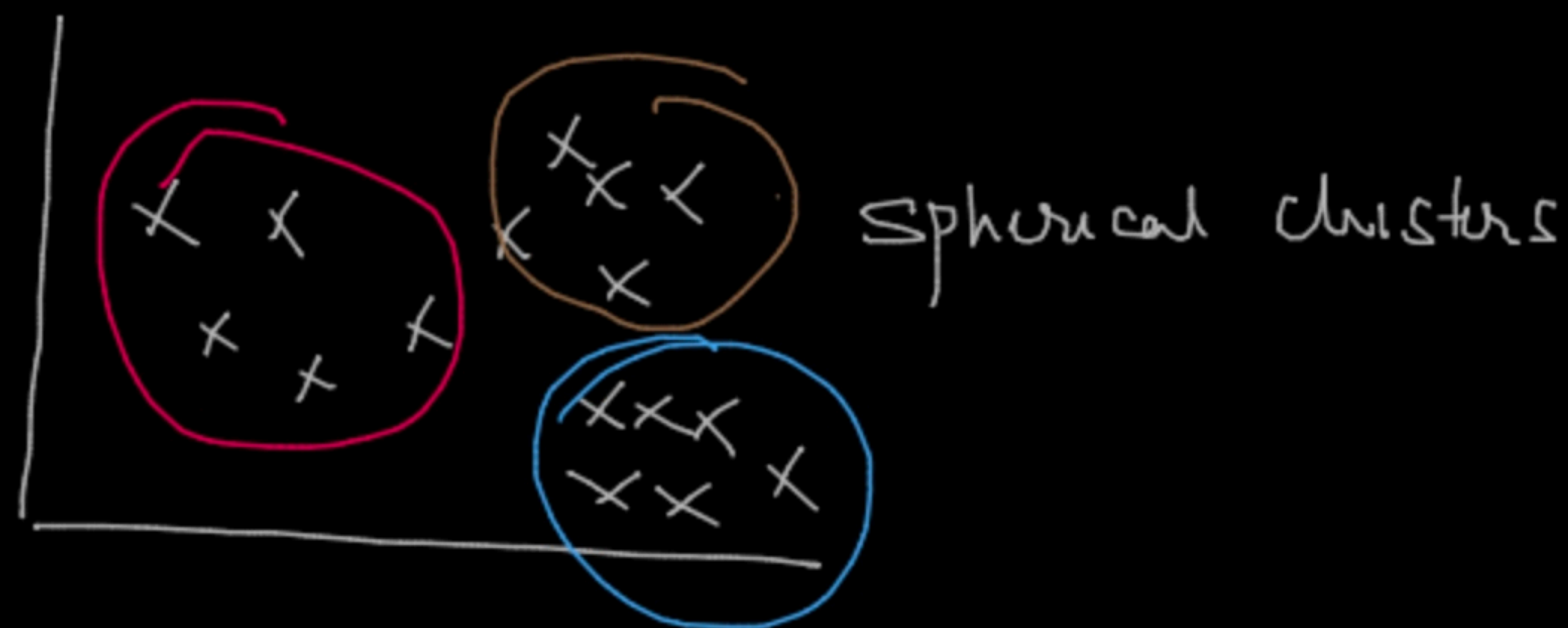
Non Spherical clusters — Hierarchical clustering.
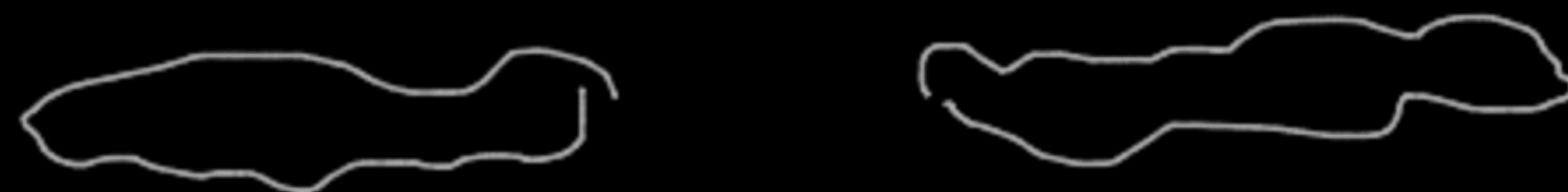
Hierarchical clustering
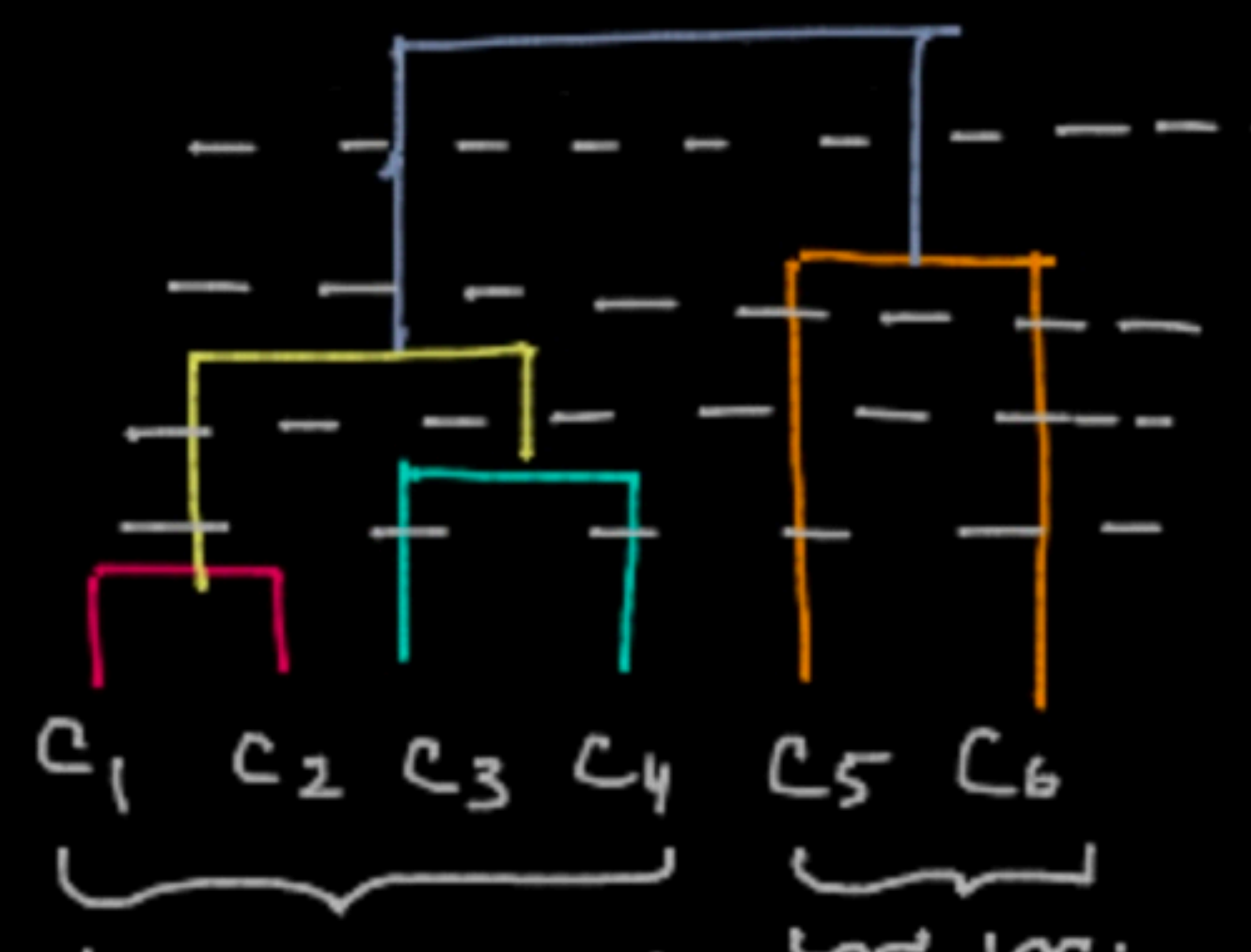
Agglomerative   Divisive.

→ Feature
  space

C1
C2
C6
C5
C3 C4
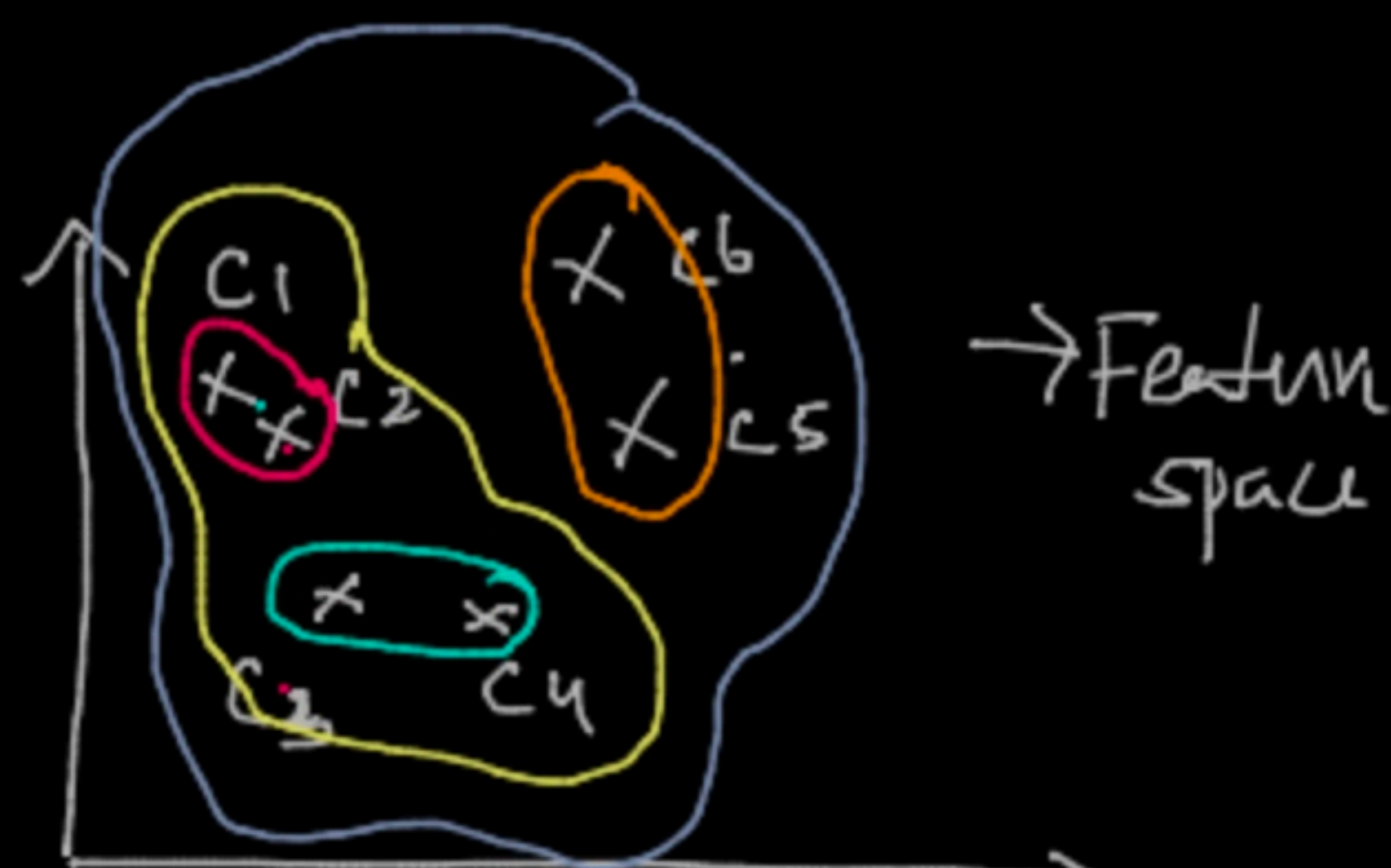
Start by assuming each point is
a cluster

Keep merging clusters which are
close to each other, until you are
left with a single cluster

C_1 C_2 C_3 C_4 C_5 C_6

Start with 1 cluster
Divide the clusters
until each cluster has
only one point in it

P_1 P_2

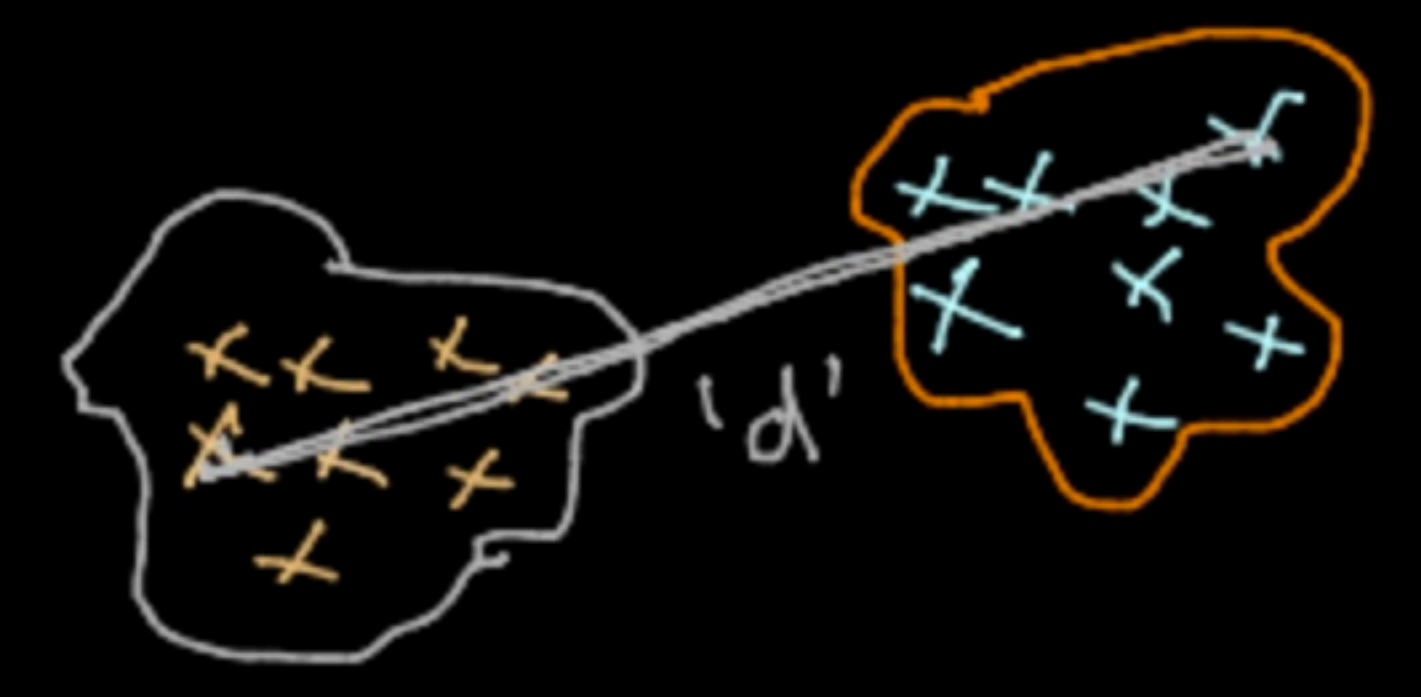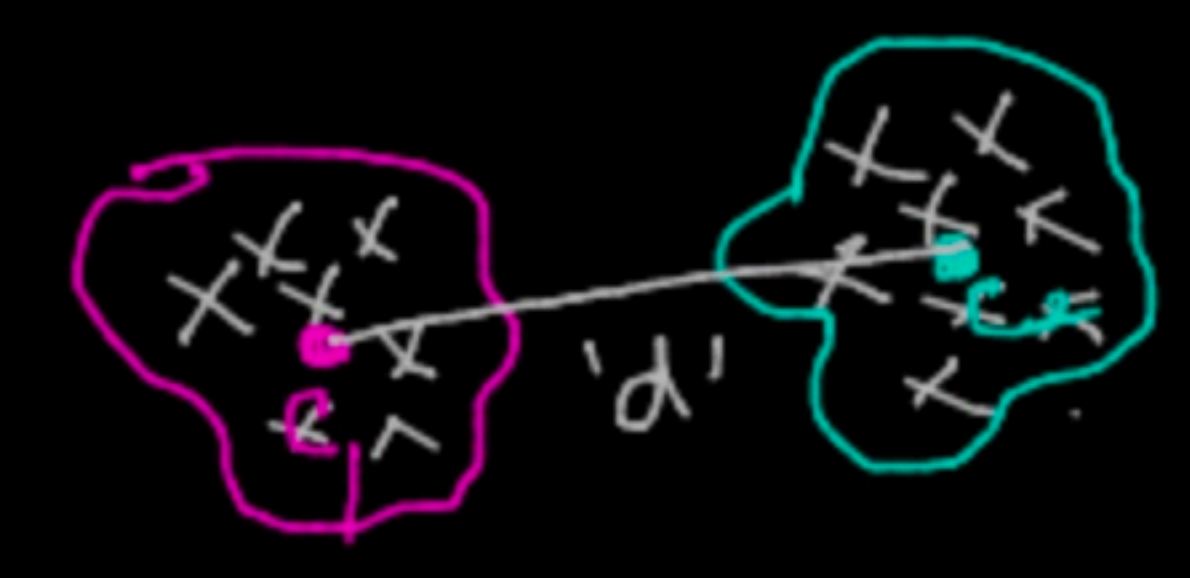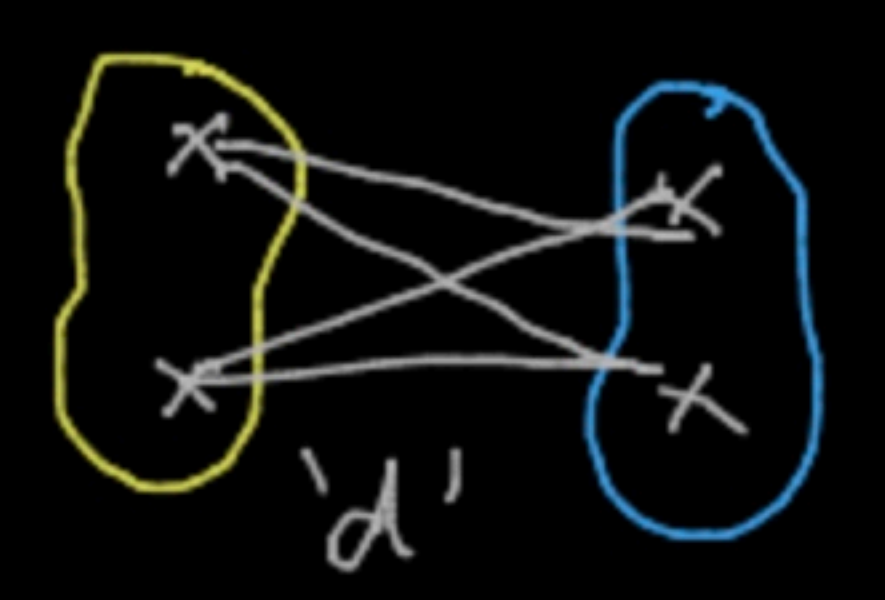Linkage — How we measure the distance between clusters.

Single linkage

Complete linkage

Centroid linkage -

'd'

'd'

'd'

Average linkage

'd'

$\begin{cases} 1. \text{ linkage } \checkmark \\ 2. \text{ Euclidean } \checkmark \rightarrow 3 \end{cases}$

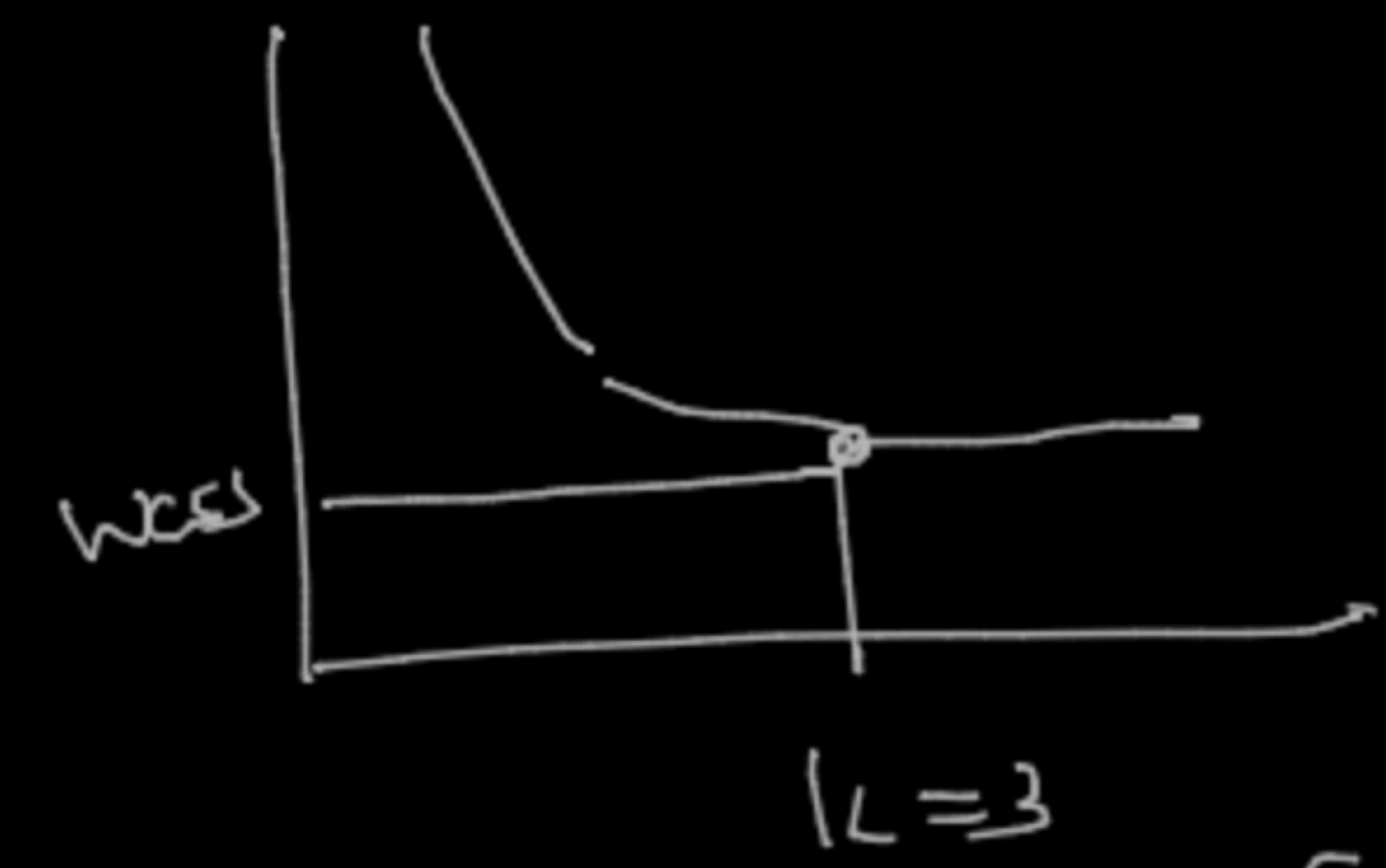3. How many clusters?

Standardise-

Scaling.
_____
Age Income

$$Z_k \Rightarrow \frac{X - \mu}{\sigma} \quad -3 \text{ to } +3. \checkmark$$

Normalize

$$N_k = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad 0 \text{ to } 1. \checkmark$$

EDA →

✓ 1. Max inter cluster — WCSS
✓ 2. Min intracluster dist

$IL = 3$

WCSS

✓ Durin Index } — $\frac{max}{min}$
✓ Silhoulette }

{
KMeans ✓ —
HAC — ✓
DISSCAN —