K-means:- it is unsupervised learning Algorithm which divides data into K-groups (clusters).
The goal is to partition data into distinct groups (clusters) where data points in the same group are as similar as possible, and points in different groups are as dissimilar as possible.

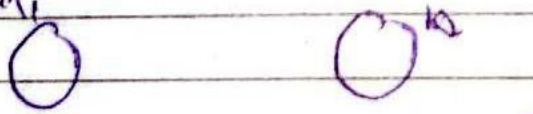
→ The "k" in K-Means represents the number of clusters you want to find. You have to tell the algorithm this number in advance.

→ exe- imagine you are a fruit seller, and you have a basket of fruits. They are only two types; Apples and Oranges, but they are all mixed together. Your task is to separate them into two groups based on their weight and size

→

ex-cs-no	age	amount.
c1	28	500
c2	40	1000
c3	30	800
c4	18	300
c5	28	1200
c6	35	1400
c7	45	1800

→ lets assume we have to make two clusters : step1:- define clusters so k=2.



so we have to decide which values lies on which clusters

Teacher's Signature

→ Step2 calculate the centroid value:
→ centroid value means on the basis of that we calculate the further values.

. so for each cluster we must have to initialize value, on the basis of that we will find other

→ so we can assign any value randomly from the given Table.

so → we assign

K_1 = (c1) 20,500

K_2 = (c2) 40,1000.

so for c_3 and further values we will find the distance, and we will check the nearest value.
if c_3 is nearest to c_1 then it will go on K_1
same for K2.

Distance formula:-

Euclidean Distance formula

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

(observed value, centroid value) → c_3 with c_1

$$c_3 = \sqrt{(30-20)^2 + (800-500)^2}$$

$c_3 = 300$

Now c_3 for c_2

so we will get

$c_3 = 200$

SOLO

Teacher's Signature

so less distance is 200.
so C3 will go to the C2 in K2 group

Note: so when you assign one value in K group. iso next step is we must have to update the centroid.

How?

so first we have value of C2 = 40 and the value of C3 = 30 so calculate the mean.

$$\frac{40+30}{2} = \frac{70}{2} = 35$$

$$\frac{800+100}{2} = 900.$$

so now exclude the prevous value and include the new one as a centroid.

(35, 900)

so 800 now C4 same with C1 and then C2.

$$C4 = \sqrt{(18-20)^2 + (300-500)^2}$$

$$C4 = \sqrt{(18-35)^2 + (300-900)^2}$$