Find clusters using K-means, if we have several objects (4 types of medicines) and each object have two attributes or features as shown in the table below. The goal is to group these objects into K=2 group of medicine based on two features (PH & weight index)

Object	Attribute 1(X) Weight	Attribute 2(Y) PH
Medicine A	1	1
Medicine B	2	1
Medicine C	4	3
Medicine D	5	4

Solution

Each object represents one point with two attributes (X, Y) that can be represented as a coordinate in an attribute space as shown

$$D(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$ Euclidean Distance.
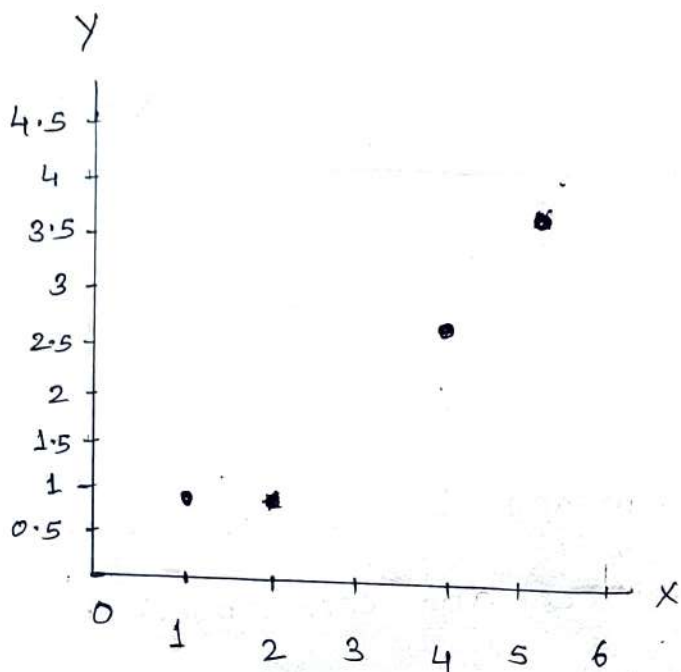
Take initial centroids: consider object A & object B
as the first centroids

consider initial centroids as

$$C_1 = (1, 1) - \qquad C_2 (2, 1)$$

DISTANCE MATRIX

$D_1 =$		$1, 1$	$2, 1$
	$1, 1$	0	1
	$2, 1$	1	0
	$4, 3$	3.6	2.83
	$5, 4$	5	4.24

Distance is calculated
using Euclidean
distance metric.

creating clusters based on distance.

$$C_1 = \{ (1, 1) \} \qquad C_2 = \{ (2, 1), (4, 3), (5, 4) \}$$

Recalculating mean values

$$C_1 = (1, 1)$$

$$C_2 = \frac{2+4+5}{3}, \frac{1+3+4}{3} = (3.66, 2.66)$$

creating distance matrix with newly calculated mean.

$D2 =$

	$(1,1)$	$(3.66, 2.66)$
$(1,1)$	0	3.14
$(2,1)$	1	2.36
$(4,3)$	3.61	0.47
$(5,4)$	5	1.89

The distance matrix is calculated using Euclidean distance Metric

New clusters as per new mean values & distance matrix.

$$C_1 = \{(1,1), (2,1)\}, \quad C_2 = \{(4,3), (5,4)\}$$

Recalculating the cluster means

$$C_1 = \frac{1+2}{2}, \frac{1+1}{2} = (1.5, 1)$$

$$C_2 = \frac{4+5}{2}, \frac{3+4}{2} = (9.5, 4.5)$$

creating distance matrix with newly computed means.

$D_3 =$

	C_1 $(1.5, 1)$	C_2 $(3.5, 4.5)$
$(1,1)$	0.5	4.30
$(2,1)$	0.5	3.54
$(4,3)$	3.20	0.71
$(5,4)$	4.61	0.71

$$C_1 = \{ (1,1), (2,1) \}$$

$$C_2 = \{ (4,3), (5,4) \}$$

etu We can stop here as clusters remain unchanged.