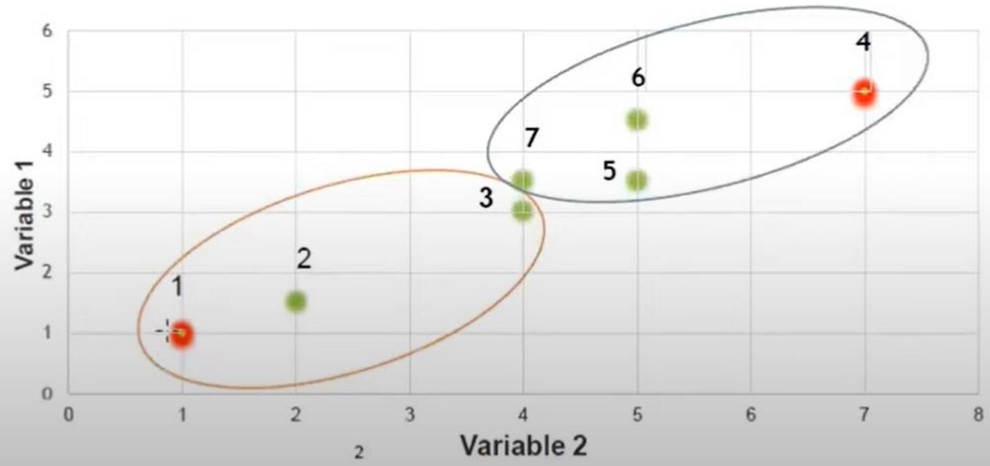
Example of K-means Clustering

A Simple example k-means (using K=2)

Individual	Variable 1	Variable 2
1	1	1
2	1.5	2
3	3	4
4	5	7
5	3.5	5
6	4.5	5
7	3.5	4.5

K= 2



Step 1:

Initialization: Randomly we choose following two centroids (k=2) for two clusters.
In this case the 2 centroid are: m1=(1.0,1.0) and m2=(5.0,7.0).

	Individual	Mean Vector
Group 1	1	(1.0, 1.0)
Group 2	4	(5.0, 7.0)

In iteration 1:

Step 2:

	Centroid 1	Centroid 2
1	$\sqrt{(1-1)^2+(1-1)^2}=0$	$\sqrt{(5-1)^2+(7-1)^2}=7.21$
2	$\sqrt{(1-1.5)^2+(1-2)^2}=1.12$	$\sqrt{(5-1.5)^2+(7-2)^2}=6.10$
3	$\sqrt{(1-3)^2+(1-4)^2}=3.61$	$\sqrt{(5-3)^2+(7-4)^2}=3.61$
4	$\sqrt{(1-5)^2+(1-7)^2}=7.21$	$\sqrt{(5-5)^2+(7-7)^2}=0$
5	$\sqrt{(1-3.5)^2+(1-5)^2}=4.72$	$\sqrt{(5-3.5)^2+(7-5)^2}=2.5$
6	$\sqrt{(1-4.5)^2+(1-5)^2}=5.31$	$\sqrt{(5-4.5)^2+(7-5)^2}=2.06$
7	$\sqrt{(1-3.5)^2+(1-4.5)^2}=4.30$	$\sqrt{(5-3.5)^2+(7-4.5)^2}=2.92$

Step 2:

- Thus, we obtain two clusters containing: {1,2,3} and {4,5,6,7}.
- Their new centroids are:

Group 1 $=(\frac{1+1.5+3}{3}), (\frac{1+2+4}{3}) = (1.83, 2.33)$

Group 2 $=(\frac{5+3.5+4.5+3.5}{4}), (\frac{7+5+5+4.5}{4}) = (4.12, 5.38)$

In iteration 2:

Step 3:

	Centroid 1	Centroid 2
1	$\sqrt{(1.83 - 1)^2 + (2.33 - 1)^2} = 1.57$	$\sqrt{(4.12 - 1)^2 + (5.38 - 1)^2} = 5.38$
2	$\sqrt{(1.83 - 1.5)^2 + (2.33 - 2)^2} = 0.47$	$\sqrt{(4.12 - 1.5)^2 + (5.38 - 2)^2} = 4.29$
3	$\sqrt{(1.83 - 3)^2 + (2.33 - 4)^2} = 2.04$	$\sqrt{(4.12 - 3)^2 + (5.38 - 4)^2} = 1.78$
4	$\sqrt{(1.83 - 5)^2 + (2.33 - 7)^2} = 5.64$	$\sqrt{(4.12 - 5)^2 + (5.38 - 7)^2} = 1.84$
5	$\sqrt{(1.83 - 3.5)^2 + (2.33 - 5)^2} = 3.15$	$\sqrt{(4.12 - 3.5)^2 + (5.38 - 5)^2} = 0.73$
6	$\sqrt{(1.83 - 4.5)^2 + (2.33 - 5)^2} = 3.78$	$\sqrt{(4.12 - 4.5)^2 + (5.38 - 5)^2} = 0.54$
7	$\sqrt{(1.83 - 3.5)^2 + (2.33 - 4.5)^2} = 2.74$	$\sqrt{(4.12 - 3.5)^2 + (5.38 - 4.5)^2} = 1.08$

Therefore, the new clusters are:

{1,2} and {3,4,5,6,7}

$$\text{Group 1} = \left(\frac{1+1.5}{2}\right), \left(\frac{1+2}{2}\right) \qquad = (1.25 , 1.5)$$

$$\text{Group 2} = \left(\frac{3+5+3.5+4.5+3.5}{5}\right), \left(\frac{4+7+5+5+4.5}{5}\right) \qquad = (3.9 , 5.1)$$

In iteration 3:

Step 3:

	Centroid 1	Centroid 2
1	$\sqrt{(1.83 - 1)^2 + (2.33 - 1)^2} = 1.57$	$\sqrt{(4.12 - 1)^2 + (5.38 - 1)^2} = 5.38$
2	$\sqrt{(1.83 - 1.5)^2 + (2.33 - 2)^2} = 0.47$	$\sqrt{(4.12 - 1.5)^2 + (5.38 - 2)^2} = 4.29$
3	$\sqrt{(1.83 - 3)^2 + (2.33 - 4)^2} = 2.04$	$\sqrt{(4.12 - 3)^2 + (5.38 - 4)^2} = 1.78$
4	$\sqrt{(1.83 - 5)^2 + (2.33 - 7)^2} = 5.64$	$\sqrt{(4.12 - 5)^2 + (5.38 - 7)^2} = 1.84$
5	$\sqrt{(1.83 - 3.5)^2 + (2.33 - 5)^2} = 3.15$	$\sqrt{(4.12 - 3.5)^2 + (5.38 - 5)^2} = 0.73$
6	$\sqrt{(1.83 - 4.5)^2 + (2.33 - 5)^2} = 3.78$	$\sqrt{(4.12 - 4.5)^2 + (5.38 - 5)^2} = 0.54$
7	$\sqrt{(1.83 - 3.5)^2 + (2.33 - 4.5)^2} = 2.74$	$\sqrt{(4.12 - 3.5)^2 + (5.38 - 4.5)^2} = 1.08$

- Therefore, there is no change in the cluster.
- Thus, the algorithm comes to a halt here and final result consist of 2 clusters {1,2} and {3,4,5,6,7}.