

Q2) A)

- ① Cluster is a representation of the mean value of the objects in the cluster
- ② Input: set of objects (n), no of clusters (k)
Output: set of k clusters
- ③ Algorithm
 - i) Partition objects into k non-empty subsets and randomly assign data to each other
 - ii) Compute seed points as the centroids of the clusters of the current partition
(the centroid is the center i.e. mean point)
 - iii) Assign each object to the cluster with the nearest seed point
 - iv) go back to step 2. stop when no more new assignment.

Data set = {15, 15, 16, 19, 19, 20, 20, 21, 22, 28, 35, 40, 41, 42, 43, 44, 60, 61, 65}

no of clusters = 2

Randomly partition given data set
 $K_1 = 15, 16, 19, 20, 21, 22, 28$
mean = 20

Rebecca Dias
182027/19

TE CMPTN A

classmate

Date

Page

Rebecca

$x_2 = 35, 40, 41, 42, 43, 44, 60, 61, 65$

mean = 43

Reassign

n	cluster 1	cluster 2
15	5	28
15	5	28
16	4	27
19	10	24
19	10	24
20	0	23
20	0	23
21	1	22
22	2	21
28	8	15
35	15	18
40	20	3
41	4	2
42	22	1
43	23	0
44	24	1
60	40	17
61	41	18
65	45	22

mean (cluster 1) = 19.5

mean (cluster 2) = 47.88

Reassign

n	cluster 1	cluster 2
15	4.5	32.8
15	4.5	32.8
16	3.5	31.8
19	0.5	28.8
19	0.5	28.8
20	0.5	27.8
20	0.5	27.8
21	1.5	26.8
22	2.5	25.8
28	8.5	19.8
35	15.5	12.8
40	20.5	7.8
41	21.5	6.8
42	22.5	5.8
43	23.5	4.8
44	24.5	3.8
60	40.5	12.1
61	41.5	13.1
65	45.5	17.1

$$\text{mean (cluster 1)} = 19.5$$

$$\text{mean (cluster 2)} = 47.88$$

new mean is same as old mean stop
 \therefore final ans

$$K_1 = \{15, 15, 16, 19, 19, 20, 20, 21, 22, 28\}$$

$$K_2 = \{35, 40, 41, 42, 43, 44, 60, 61, 65\}$$