

Question 1:

Data = [1 2 5 6 8 9 10 11 13 15 17]

Initial means = [1 3 10]

Iteration 1:

Cluster 1: [1, 2], $\bar{X} = (1 + 2)/2 = 1.5$

Cluster 2: [5, 6], $\bar{X} = (5 + 6)/2 = 5.5$

Cluster 3: [8, 9, 10, 11, 13, 15, 17], $\bar{X} = (8 + 9 + 10 + 11 + 13 + 15 + 17)/7 = 11.857$

Iteration 2:

Cluster 1: [1, 2], $\bar{X} = (1 + 2)/2 = 1.5$

Cluster 2: [5, 6, 8], $\bar{X} = (5 + 6 + 8)/3 = 6.333$

Cluster 3: [9, 10, 11, 13, 15, 17], $\bar{X} = (9 + 10 + 11 + 13 + 15 + 17)/6 = 12.5$

Iteration 3:

Cluster 1: [1, 2], $\bar{X} = (1 + 2)/2 = 1.5$

Cluster 2: [5, 6, 8, 9], $\bar{X} = (5 + 6 + 8 + 9)/4 = 7$

Cluster 3: [10, 11, 13, 15, 17], $\bar{X} = (10 + 11 + 13 + 15 + 17)/5 = 13.2$

Iteration 4:

Cluster 1: [1, 2], $\bar{X} = (1 + 2)/2 = 1.5$

Cluster 2: [5, 6, 8, 9, 10], $\bar{X} = (5 + 6 + 8 + 9 + 10)/5 = 7.6$

Cluster 3: [11, 13, 15, 17], $\bar{X} = (11 + 13 + 15 + 17)/4 = 14$

Question 2:

1.

a. $\sqrt{(2 - 1)^2 + (3 - 2)^2} = \sqrt{2} = 1.41421$

b. $\sqrt{(2 - 1)^2 + (3 - 4)^2} = \sqrt{2} = 1.41421$

c. $\sqrt{(2 + 0)^2 + (3 - 3)^2} = 2$

d. $\sqrt{(2 - 2)^2 + (3 - 5)^2} = 2$

e. $\sqrt{(2 - 3)^2 + (3 - 5)^2} = \sqrt{5} = 2.23607$

f. $\sqrt{(2 - 3)^2 + (3 - 3)^2} = 1$

2.

a. (1,2), (1,4), (3,3)

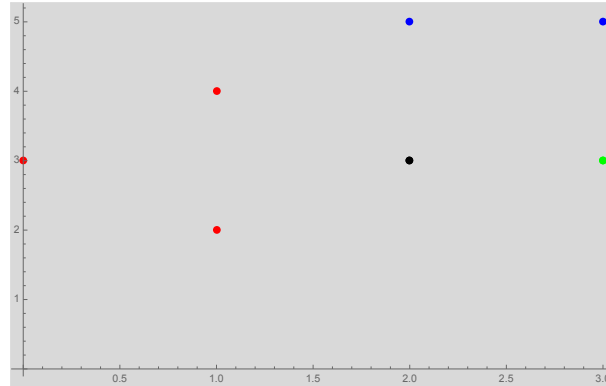
b. We should classify them as red.

3.

- a. For $K = 1$, we choose the closest data point (3, 3) which will classify the label as green.
- b. For $K = 5$, we will classify the label as red since there are three data points has been classified as red.

4.

- a. `ListPlot[{{{2, 3}}, {{1, 2}, {1, 4}, {0, 3}}, {{2, 5}, {3, 5}}, {{3, 3}}}, PlotStyle -> {Black, Red, Blue, Green}, Background -> LightGray, ImageSize -> Large]`



- b.
- c. My answers to the previous two questions agree with this plot, indicates that the plot is depends on the value of K. If $K = 1$, it will be classified as green and it will increase since other colors will become to the majority.