Problem 1.

Cluster1: (2, 4) (3,3) (3,5) (4,8) (5,5)

Cluster2: (2,6) (2,8) (5,7)

**New Centroid:**

C1.X = 2+3+3+4+5/5 = 17/5 = 3.4

C1.Y = 4+3+5+8+5/5 = 25/5 = 5

C2.X = 2+2+5/3 = 3

C2.Y = 6+8+7/3 = 7

Calculate distances between points and new centroids

D((2,4),c1) = (2,4)-(3.4,5) = 2.4

D((2,4),c2) = (2,4)-(3,7) = 4

(3,3)-(3.4,5) = 2.4

(3,3)-(3,7) = 4

(3,5)-(3.4,5) = 0.4

(3,5)-(3,7) = 2

(4,8)-(3.4,5) = 3.6

(4,8)-(3,7) = 2

(5,5)-(3.4,5) = 1.6

(5,5)-(3,7) = 4

(2,6)-(3.4,5) = 2.4

(2,6)-(3,7) = 2

(2,8)-(3.4,5) = 4.4

(2,8)-(3,7) = 2

(5,7)-(3.4,5) = 3.6

(5,7)-(3,7) = 2

Cluster1: (2, 4) (3,3) (3,5) (5,5)

Cluster2: (2,6) (2,8) (5,7) (4,8)

**New Centroid:**

C1.X = 2+3+3+5/4 = 13/4 = 3.25

C1.Y = 4+3+5+5/4 = 17/4 = 4.25

C2.X = 2+2+5+4/4 = 3.25

C2.Y = 6+8+7+8/4 = 7.25

Calculate distances between points and new centroids

(2,4)-(3.25,4.25) = 1.5

(2,4)-(3.25,7.25) = 4.5

(3,3)-( 3.25,4.25) = 1.5

(3,3)-( 3.25,7.25) = 4.5

(3,5)-( 3.25,4.25) = 1

(3,5)-( 3.25,7.25) = 2.5

(5,5)-( 3.25,4.25) = 2.5

(5,5)-( 3.25,7.25) = 4

(2,6)-( 3.25,4.25) = 3

(2,6)-( 3.25,7.25) = 2.5

(2,8)-( 3.25,4.25) = 5

(2,8)-( 3.25,7.25) = 2

(5,7)-( 3.25,4.25) = 4.5

(5,7)-( 3.25,7.25) = 2

(4,8)-( 3.25,4.25) = 4.5

(4,8)-( 3.25,7.25) = 1.5

Cluster1: (2, 4) (3,3) (3,5) (5,5)

Cluster2: (2,6) (2,8) (5,7) (4,8)

Problem 2.

Cluster 1 (2,5), (3,3), (4,6)

Cluster 2 (7,3), (7,6), (8,4), (9,5)

1. Min distance

|(2,5), (7,3)| = 7

|(2,5), (7,6)| = 6

|(2,5), (8,4)| = 7

|(2,5), (9,5)| = 7

|(3,3), (7,3)| = 4

|(3,3), (7,6)| = 7

|(3,3), (8,4)| = 6

|(3,3), (9,5)| = 8

|(4,6), (7,3)| = 6

**|(4,6), (7,6)| = 3**

|(4,6), (8,4)| = 6

|(4,6), (9,5)| = 6

So minimum distance = **3** between **(4,6), (7,6)**

1. **Mean distance**
2. Cluster 1 (2,5), (3,3), (4,6)
3. Cluster 2 (7,3), (7,6), (8,4), (9,5)

C1.X centroid = 2+3+4/3 = 3, C1.Y centroid = 5+3+6/3 = 4.67

C2.X centroid = 7+7+8+9/4 = 7.75, C2.Y = 3+6+4+5/4 = 4.5

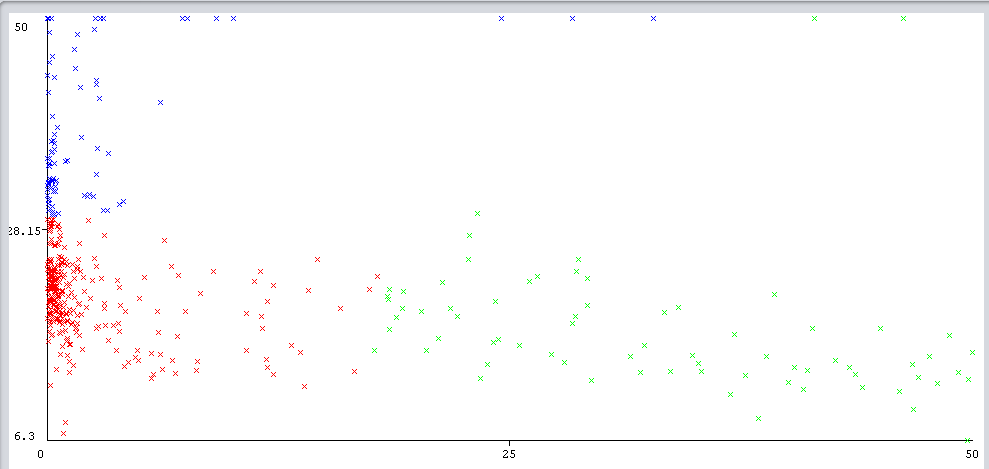
Mean distance = |c1 – c2| = |3-7.75| + |4.67-4.5| = 4.75 + 0.17 = 4.92

Problem 3.

3.1

Turning point is 3.3, so it seems K will lay between 3 & 4

3.2



Most of A1 with values below 10 and value of A2 greater than 28 belongs to cluster #0 (blue one)

If A1 value is greater than 20 then regardless of A2 value, points will be in cluster #2 (green one)