CPSC 483 PROJECT 2

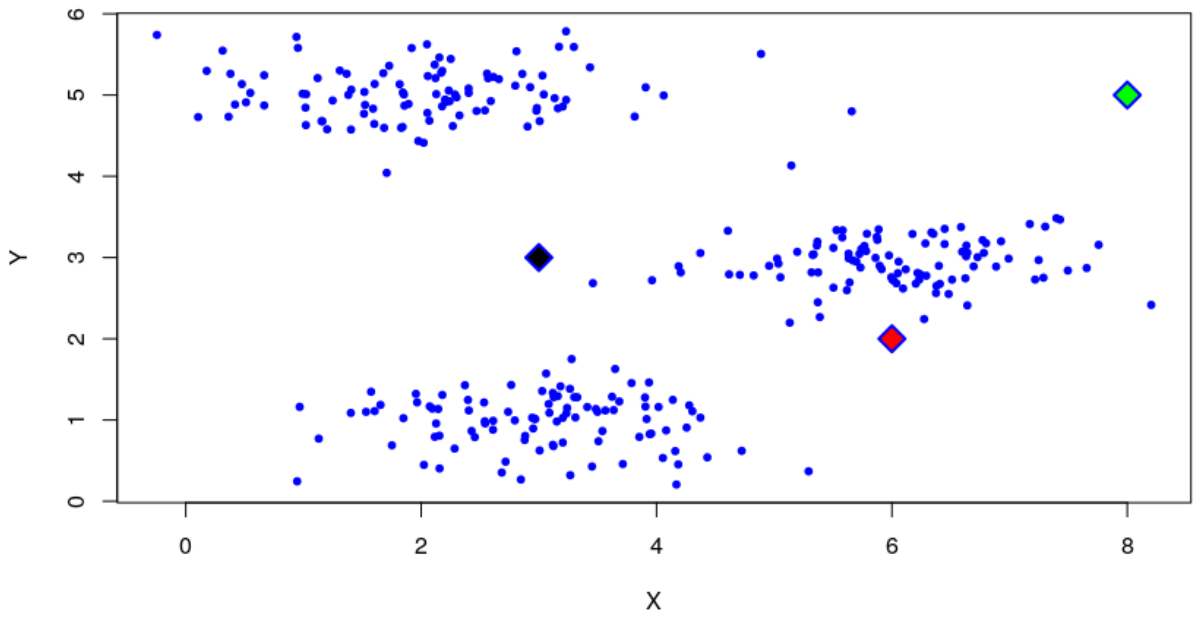
Yong Kim

- Assignment
To implement **Kmeans** algorithm without any ML libraries and packages

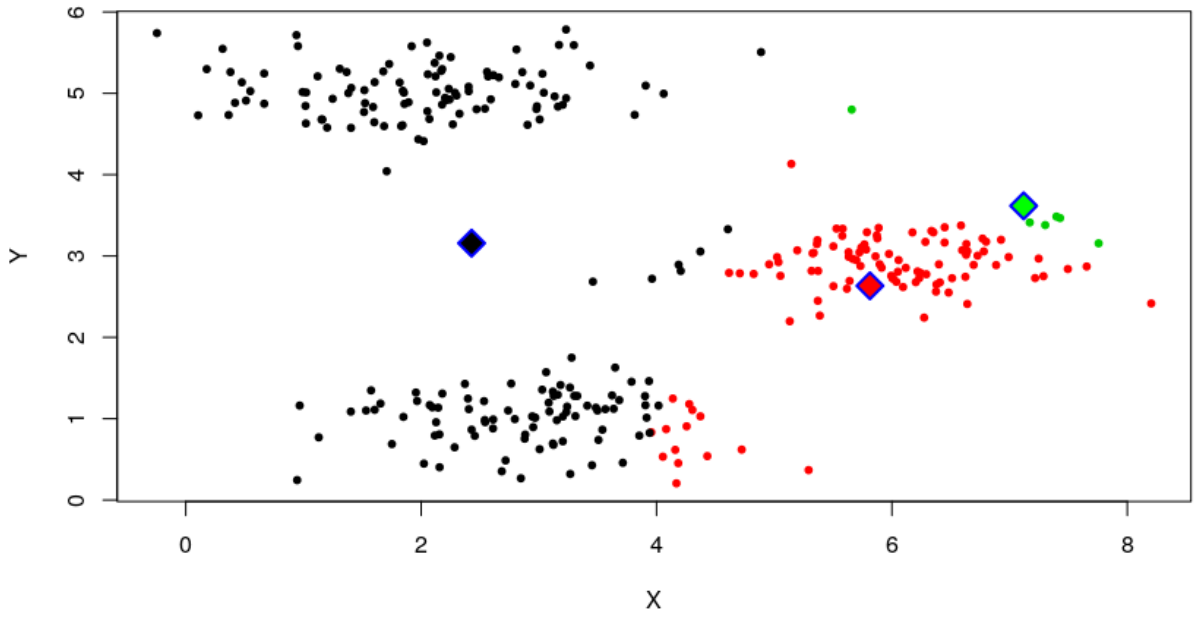
1) **Environment**
   R / R-studio

2) **Usage**
   1. Execute R-studio
   2. place the input file, "kmeansdata.mat" in the R-studio working directory.
   3. Execute **kmeans_YKim.R**
   or copy the source code, then enter
   4. Execute **K_means(10, df, centroid)**,
   10 is the number of iteration,
   df is the imported dataset,
   and centroid is the data frame of 3 centroids.
   "df" and "centroid" are hard-coded. Therefore, if arguments of other names will cause an error.

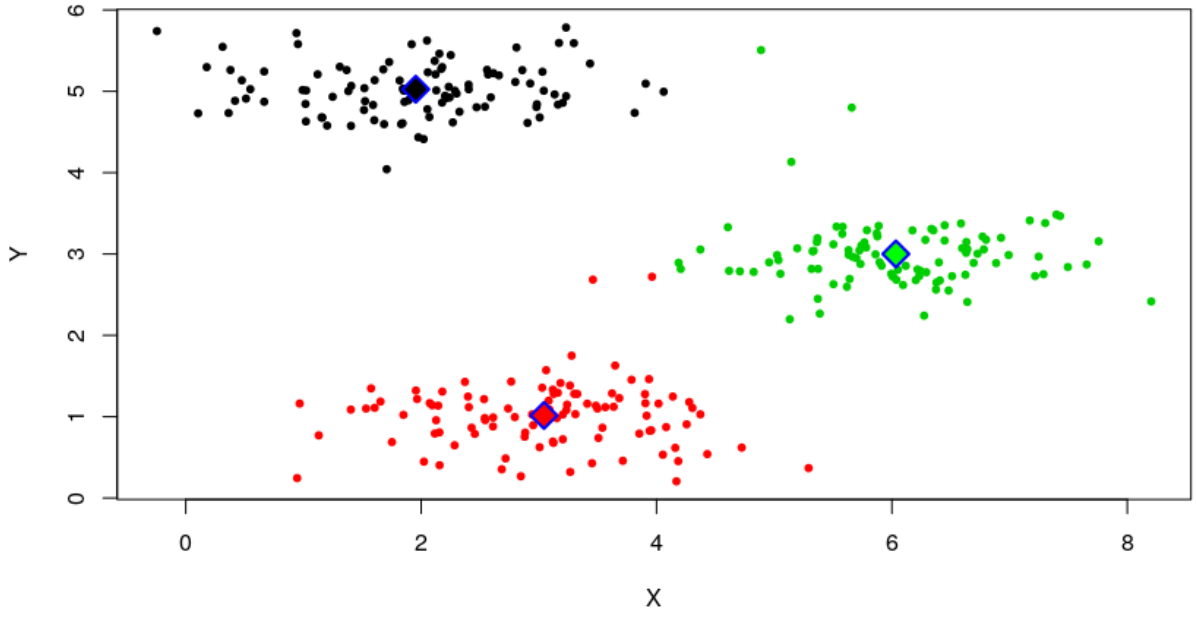
3) **3 plots**



Initial point distribution



After 1st iteration



Final point distribution

4) Changing centroids

```
> K_means(10, df, centroid)
[1] 1
       X       Y
1 2.428301 3.157924
2 5.813503 2.633656
3 7.119387 3.616684
[1] 2
       X       Y
1 2.313255 3.228306
2 5.332738 2.431596
3 6.865362 3.232940
[1] 3
       X       Y
1 2.196925 3.421367
2 4.835554 2.129767
3 6.656005 3.075135
[1] 4
       X       Y
1 1.982412 4.025078
2 3.911508 1.470605
3 6.340086 3.053666
[1] 5
       X       Y
1 1.953995 5.025570
2 3.126637 1.112171
3 6.129195 3.016063
[1] 6
       X       Y
1 1.953995 5.025570
2 3.043671 1.015410
3 6.033667 3.000525
[1] 7
       X       Y
1 1.953995 5.025570
2 3.043671 1.015410
3 6.033667 3.000525
[1] 8
       X       Y
1 1.953995 5.025570
2 3.043671 1.015410
3 6.033667 3.000525
[1] 9
       X       Y
1 1.953995 5.025570
2 3.043671 1.015410
3 6.033667 3.000525
[1] 10
       X       Y
1 1.953995 5.025570
2 3.043671 1.015410
3 6.033667 3.000525
```