**Data Visualization Lab(L13+L14)**

**Name: Alpanshu Kataria**

**Reg. No.: 18BCE1267**

## Faculty: Dr.Parvathi.R

Lab 2

**Task:**

## KMeans\_Animation on wheat seeds dataset.

## 

## Code:

## #require("datasets")

## #data("seeds")

## seeds <- read.csv("D:/ViT/Sem 6/Data viz/Lab/Lab2/seeds.csv")

## str(seeds)

## summary(seeds)

## head(seeds)

## seeds.new<- seeds[,c(1,2,3,4,5,6,7)]

## seeds.class<- seeds[,"Type"]

## head(seeds.new)

## head(seeds.class)

## normalize <- function(x){

## return ((x-min(x))/(max(x)-min(x)))

## }

## seeds.new$Area<- normalize(seeds.new$Area)

## seeds.new$Perimeter<- normalize(seeds.new$Perimeter)

## seeds.new$Compactness<- normalize(seeds.new$Compactness)

## seeds.new$Kernal.Length<- normalize(seeds.new$Kernel.Length)

## seeds.new$Kernal.Width<- normalize(seeds.new$Kernel.Width)

## seeds.new$Asymmetry.Coeff<- normalize(seeds.new$Asymmetry.Coeff)

## seeds.new$Kernal.Groove<- normalize(seeds.new$Kernel.Groove)

## seeds.new<- seeds.new[,c(1,2,3,6,8,9,10)]

## head(seeds.new)

## head(seeds.class)

## result<- kmeans(seeds.new,3)

## result$size # gives no. of records in each cluster

## result$centers # gives value of cluster center datapoint value(3 centers for k=3)

## result$cluster

## par(mfrow=c(2,2), mar=c(5,4,2,2))

## plot(seeds.new[c(1,2)], col=result$cluster)

## plot(seeds.new[c(1,2)], col=seeds.class)

## plot(seeds.new[c(3,4)], col=result$cluster)

## plot(seeds.new[c(3,4)], col=seeds.class)

## plot(seeds.new[c(5,6)], col=result$cluster)

## plot(seeds.new[c(5,6)], col=seeds.class)

## result$cluster <- as.factor(result$cluster)

## library(ggplot2)

## ggplot(seeds.new, aes(Kernal.Length, Kernal.Width, color = result$cluster)) + geom\_point()

## plot(seeds.new[,], col=result$cluster)

## table(result$cluster,seeds.class)

## library(animation)

## km1<-kmeans.ani(seeds.new,3)

## library(factoextra)

## fviz\_cluster(result, data = seeds.new)

## k2 <- kmeans(seeds.new, centers = 2, nstart = 25)

## k3 <- kmeans(seeds.new, centers = 3, nstart = 25)

## k4 <- kmeans(seeds.new, centers = 4, nstart = 25)

## k5 <- kmeans(seeds.new, centers = 5, nstart = 25)

## p1 <- fviz\_cluster(k2, geom = "point", data = seeds.new) + ggtitle("k = 2")

## p2 <- fviz\_cluster(k3, geom = "point", data = seeds.new) + ggtitle("k = 3")

## p3 <- fviz\_cluster(k4, geom = "point", data = seeds.new) + ggtitle("k = 4")

## p4 <- fviz\_cluster(k5, geom = "point", data = seeds.new) + ggtitle("k = 5")

## library(gridExtra)

## grid.arrange(p1, p2, p3, p4, nrow = 2)

## Console Output:

## 

## 

## 

## 

## Plot Output:

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 