

```

#create .csv file with random dataset(Rating,Price,ALcohol,Sulphate,pH)

dt = read.csv(file.choose())
head(dt)
new_data <- dt[, -which(name(dt) == "pH")]
new_data

c1<-kmean(new_data,3)
c1

data <- new_data
wss<- sapply(1:15, function(k){kmeans(data ,k) $tot.withinss})
wss

#plot 3 cluster
plot(1:15,wss , type = "b" , pch = 19 , frame = FALSE,xlab="Number of cluster K",
     ylab="Total within cluster sum of square")

install.packages("cluster")
library(cluster)

clusplot(new_data,c1$cluster, color = TRUE ,
         shade=TRUE, labels=2,lines=0)
c1$cluster
c1$centers

clusters<-hclust(dist(dt[, 3:4]))
plot(clusters)

#ggplot and plot is missing
install.packages('ggplot2')
library(ggplot2)

ggplot(dt, aes(Alcohol,Sulphates,color = pH)) +
  geom_point(alpha = 0.4,size =3.5)+geom_point(col = clusterCut)
+ scale_color_manual(values = c('black','red','green'))

cluster <- hclust(dist(dt[,3:4]),method = 'average')
clusterCut1<-cutree(cluster , 3)
table(clusterCut1,dt$pH)
plot(clusters)

```