## Unsupervised learning -- K means algorithm

?kmeans

mydata = mtcars

wss = 0

for (i in 1:15) {

wss[i] <- sum(kmeans(mydata, centers = i)$withinss)

}

## plotting

plot(1:15, wss, type = "b", xlab = "No. of clusters",

ylab = "Within groups sum of clusters")

## k means cluster analysis

fit = kmeans(mydata,5)

## each point which cluster it belong

fit$cluster

## to find center

fit$centers

## get cluster mean

?aggregate

aggregate(mydata, by = list(fit$cluster), FUN = mean)

## append cluster assignment

mydata = data.frame(mydata, fit$cluster)

head(mydata)

## cluster plot against first 2 principal components which vary

## parameter for most readable graph

## for forming cluster, we will use clusplot function which is in "cluster" package

install.packages("cluster")

library(cluster)

clusplot(mydata , fit$cluster,

color = T, shade = T,

labels = 2, lines = 0)