DATA ANALYTICS – 4027

LAB-11

Name: Hari Krishna P

Reg No: 19BCE7675

DATE: 16/11/2021

Contents:

➢ **Clustering Algorithm**

Submitted to:

Prof . Hari Seetha

Clustering Algorithms implementation using R

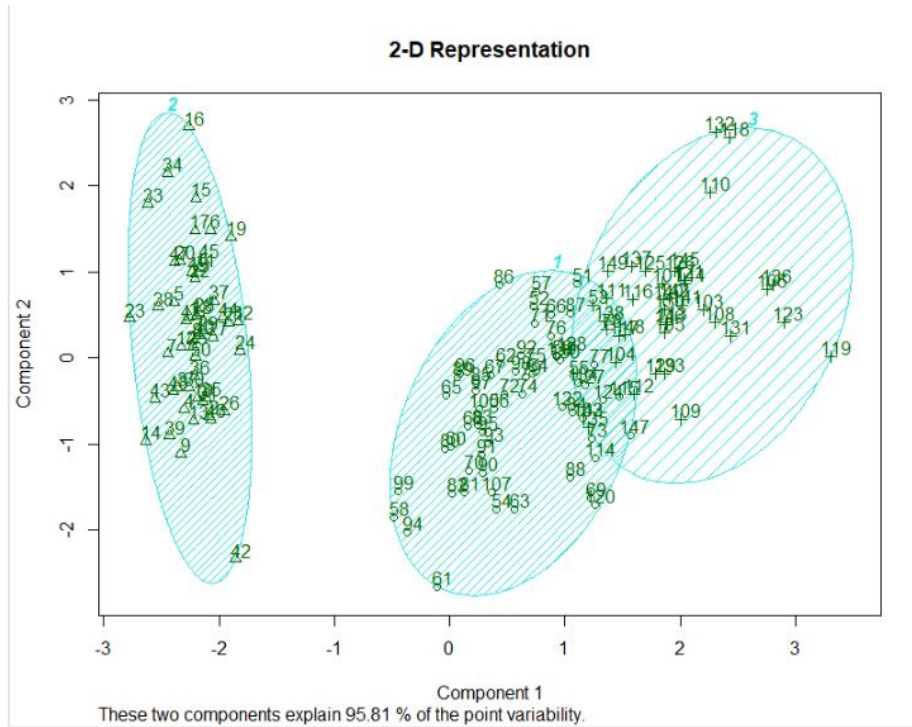**Kmeans clustering for iris dataset**

iris

dim(iris)

irisnew<-iris[c(1:4)]

kmresult<-kmeans(irisnew,3)

kmresult

library(cluster)

clusplot(irisnew, kmresult$cluster,main = "2-D representation of cluster data",shade = TRUE, labels=2, lines=0)

Code:

install.packages("party")

install.packages("e1071")

library(party)

dim(iris)

iris_new <- iris[c(1:4)]

iris_new

kmresult <- kmeans(iris_new,3)

kmresult

library(cluster)

clusplot(iris_new,kmresult$cluster, main="2-D Representation",shade = TRUE,labels = 2,lines = 0)

2-D Representation

Component 2 (vertical axis) vs Component 1 (horizontal axis)

These two components explain 95.81 % of the point variability.

IRIS USING CLUSTER:

CODE:

```
x=iris[,3:4] #using only petal length and width columns

head(x)

model=kmeans(x,3)

library(cluster)

clusplot(x,model$cluster)

clusplot(x,model$cluster,color=T,shade=T)
```
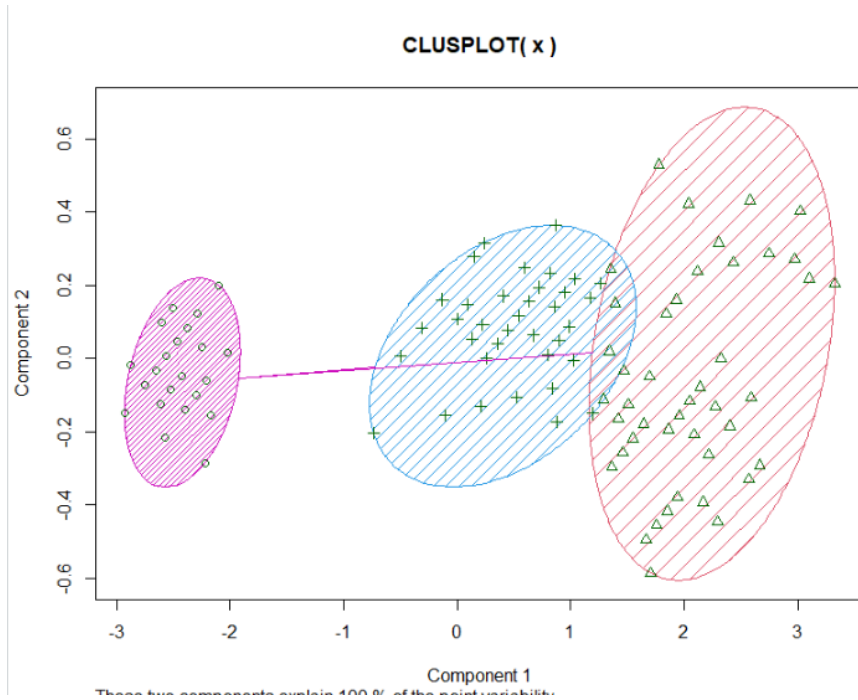
Refer

Implement clustering algorithm for clustering dataset available in UCI ML repository.

Code:

```
x= Amzn_anon_access_samples_history_2_0[,2:3]

head(x)


model=kmeans(x,2)

library(cluster)

clusplot(x,model$cluster)
```

OUTPUT:

CLUSPLOT(x)

Component 2

-20000   0   20000   40000   60000

Component 1
These two components explain 100 % of the point variability.