Roll No.: 40023

Date:

**Aim:** To perform practical of Clustering.

**Program Code:**

install.packages("ggplot2")

library(ggplot2)

scatter <- ggplot(data=iris,aes(x=Sepal.Length,y=Sepal.Width))

scatter + geom\_point(aes(color=Species,shape=Species))+

theme\_bw()+

xlab("Sepal Length")+ylab("Sepal Width")+

ggtitle("Sepal Length-Width")

ggplot(data=iris,aes(Sepal.Length,fill=Species))+

theme\_bw()+

geom\_density(alpha=0.25)+

labs(x="Sepal.Length",title="Species vs Sepal Length")

vol <- ggplot(data=iris,aes(x=Sepal.Length))

vol + stat\_density(aes(ymax=..density..,ymin=-..density..,fill=Species,color=Species),geom="ribbon",position="identity")+

facet\_grid(.~Species)+coord\_flip()+theme\_bw()+labs(x="Sepal Length",title="Species vs Sepal Length")

vol <- ggplot(data=iris,aes(x=Sepal.Width))

vol + stat\_density(aes(ymax=..density..,ymin=-..density..,fill=Species,color=Species),geom="ribbon",position="identity")+

facet\_grid(.~Species)+coord\_flip()+theme\_bw()+labs(x="Sepal Width",title="Species vs Sepal Width")

irisData <- iris[,1:4]

totalwSS<-c()

for(i in 1:15)

{clusterIRIS<- kmeans(irisData,centers = i)

totalwSS[i] <-clusterIRIS$tot.withinss}

plot(x=1:15,y=totalwSS,type="b",xlab="Number of Clusters",ylab="Within groups sum-of-squares")

install.packages("NbClust")

library(NbClust)

par(mar=c(2,2,2,2))

nb<-NbClust(irisData,method="kmeans")

hist(nb$Best.nc[1,],breaks=15,main="Histogram for Number of Clusters")

install.packages("vegan")

library(vegan)

modelData<-cascadeKM(irisData,1,10,iter=100)

plot(modelData,sortg=TRUE)

modelData$results[2,]

which.max(modelData$results[2,])

library(cluster)

cl<-kmeans(iris[,-5],2)

dis<-dist(iris[,-5])^2

sil=silhouette(cl$cluster,dis)

plot(sil,main="Clustering Data with silhoutte plot using 2 Clusters",col=c("cyan","blue"))

library(cluster)

cl<-kmeans(iris[,-5],8)

dis<-dist(iris[,-5])^2

sil=silhouette(cl$cluster,dis)

plot(sil,main="Clustering Data with silhoutte plot using 8 Clusters",col=c("cyan","blue","orange","yellow","red","gray","green","maroon"))

install.packages("factoextra")

library(factoextra)

install.packages("clustertend")

library(clustertend)

genx<-function(x){

runif(length(x),min(x),(max(x)))}

random\_df<-apply(iris[,-5],2,genx)

random\_df<-as.data.frame(random\_df)

iris[,-5]<-scale(iris[,-5])

random\_df<-scale(random\_df)

res<-get\_clust\_tendency(iris[,-5],n=nrow(iris)-1,graph=FALSE)

res$hopkins\_stat

hopkins(iris[,-5],n=nrow(iris)-1)

res<-get\_clust\_tendency(random\_df,n=nrow(random\_df)-1,graph=FALSE)

res$hopkins\_stat

**Conclusion:** Practical of Clustering has been executed successfully.