```
title: "Untitled"
output:
 pdf_document: default
 word_document: default
#UNIVERSITY: STEVENS INSTITUE OF TECHNOLOGY
#Project : HW_08_clusters
#Purpose : Use Kmeans & hcLust methodology to develop a classification
model
#First Name : Sarthak
#Last Name : Ahir
#CWID: 10479028
#Date: 11/22/2021
rm(list=ls())
#Load the "wisc_bc_ContinuousVar" from canvas into R and perform the
analysis
newDataSet<-read.csv("~/Downloads/wisc_bc_ContinuousVar.csv")</pre>
#Summarizing each column
summary(newDataSet)
table(newDataSet$Class)
table(newDataSet$diagnosis)
#removing NANs
dataset<-na.omit(newDataSet)</pre>
dataset<-dataset[-1]
dataset_dist<-dist(dataset[,-1])</pre>
    Implementing Hclust
hclust<-hclust(dataset_dist)</pre>
plot(hclust)
hClust2<-cutree(hclust,2)
matrix<-table(hClust2,dataset[,1])</pre>
accuracy<-sum(diag(matrix))/nrow(dataset)</pre>
accuracy
# remove
rm(list=ls())
#Load the "wisc_bc_ContinuousVar" from canvas into R and perform the
analysis
newDataSet<-read.csv("~/Downloads/wisc_bc_ContinuousVar.csv")</pre>
   Summarizing each column
summary(newDataSet)
table(newDataSet$diagnosis)
```

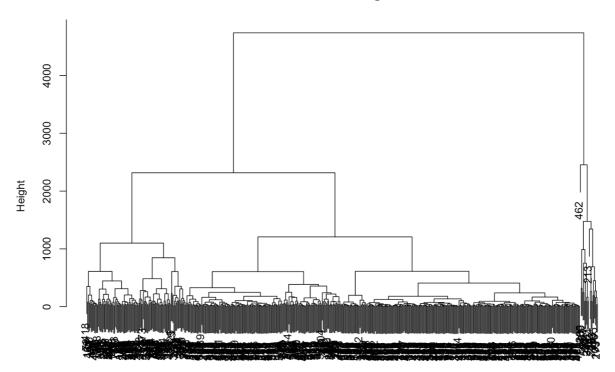
```
# removing NANs
dataset<-na.omit(newDataSet)
dataset<-dataset[-1]

# Kmean
Kmeans2<- kmeans(dataset[,-1],2,nstart = 10)
Kmeans2$cluster
matirx<-table(Kmeans2$cluster,dataset[,1])
accuracy<-sum(diag(matirx))/nrow(dataset) * 100
accuracy</pre>
```

OUTPUT -

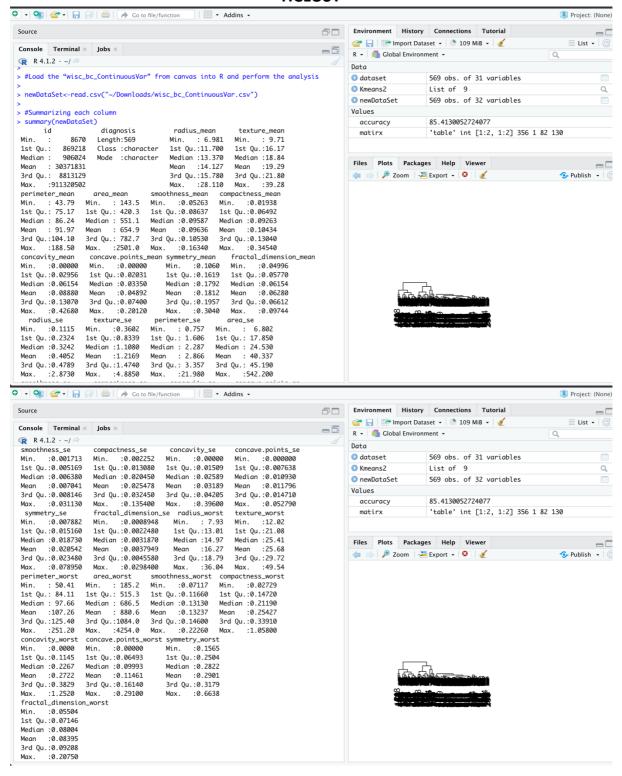
HCLUST PLOT-

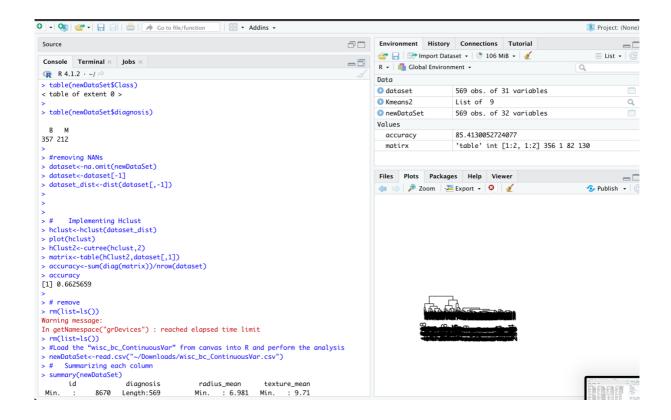
Cluster Dendrogram



dataset_dist hclust (*, "complete")

HCLUST





K-MEANS

