**********program 1 part 1 **********

```
#importing dataset
dataset = read.csv("trees.csv")

# summarizing dataset
summary(dataset)

#Plotting numeric values of the two columns
plot(dataset[,"Girth"],dataset[,"Height"],type="b",main="Plotting numeric
values",xlab="Girth",ylab="Height",col="red")
```

**********program 1 part 2 **********

```
dataset <- read.delim("annual.txt", header=TRUE, sep="\t")
dataset

maxTemp = c(58,43,70,66,70,90,89,87,82,77,52,77)
new_col <- data.frame(maxTemp)
dataset = cbind(dataset,new_col)
dataset
write.table(dataset, "annual_written.txt", quote=FALSE, sep="\t", row.names=FALSE)
```

**********program 2 **********

```
Buff_tail = c(10,1,37,5,12)
Garden_bee = c(8,3,9,6,4)
Red_tail = c(1,8,9,12,4)
Carden_bee = c(8,27,6,32,23)
Honey_bee= c(12,13,16,9,10)

bee_visits = data.frame(Buff_tail,Garden_bee,Red_tail,Carden_bee,Honey_bee)
bee_visits

plant_names = c("Thistle", "Vipers", "Golden Rain", "Yellowalfala","blackberry")
rownames(bee_visits) = plant_names
bee_visits
```

**********program 3 part 1 **********

```
Buff_tail = c(10,1,37,5,12)
Garden_bee = c(8,3,9,6,4)
Red_tail = c(1,8,9,12,4)
Carden_bee = c(8,27,6,32,23)
Honey_bee= c(12,13,16,9,10)

bee_visits = data.frame(Buff_tail,Garden_bee,Red_tail,Carden_bee,Honey_bee)
bee_visits

plant_names = c("Thistle", "Vipers", "Golden Rain", "Yellowalfala","blackberry")
rownames(bee_visits) = plant_names
bee_visits

bee_matrix = data.matrix(bee_visits)
bee_matrix
rownames(bee_matrix) = plant_names
bee_matrix
```

**********program 3 part 2 **********

```
num<-list(10,20,30,40)
num
Str<-list("A","B","C")
Str
Realn<-list(3.14,9.8,0.33)
Realn
combined_list<-list(num,Str,Realn)
combined_list

names(combined_list) = c("Numeric","Strings","Real numbers")
combined_list
```

**********program 4 part 1 **********

```
library("gcookbook")
library(ggplot2)
cabbage_exp
ggplot(cabbage_exp,aes(x=Date,y=Weight,fill=Cultivar)) +
geom_bar(position="dodge",stat="identity") + geom_text(aes(label=Weight),vjust=1.5,
position=position_dodge(.9), colour = "white",size=5)
```

**********program 4 part 1 **********

```
library("gcookbook")
BirthWt
birthwt = read.csv("birthwt.csv")
summary(birthwt)

ht_names = unique(birthwt$ht)

boxplot(birthwt$age,birthwt$ht,col="yellow",names =ht_names,border = "red",
      notch = TRUE,main="Birthwt boxplot",xlab="heights",ylab="Age")
```

**********program 5 part 1 **********

```
library('ggplot2')
mtcars
mt <- mtcars[order(mtcars$mpg),]
mt
grps <- as.factor(mtcars$cyl)
my_cols <- c("RED", "BLUE", "DARK GREEN")

dotchart(mtcars$mpg, labels=row.names(mtcars), cex = 0.8, xlab="mpg",pch = '*',main = "Miles
Per Gallon Of Car" ,groups = grps, gcolor =my_cols, color = my_cols[grps])
```

**********program 5 part 2 **********

```
library("gcookbook")
summary(heightweight)
```

```
plot(heightweight$ageYear,heightweight$heightIn,main="Scatterplot Matrix",col="darkgreen",
cex=0.4,xlab="Age Year",ylab="Height In")
abline(lm(heightweight$heightIn~heightweight$ageYear),col="Red")
pairs(~ageYear+heightIn,data=heightweight,main="Scatterplot Matrix")
```

**********program 6 **********

```
# Loading package
library(ClusterR)
library(cluster)

# Removing initial label of
# Species from original dataset
iris_1 <- iris[, -5]

# Fitting K-Means clustering Model
# to training dataset
set.seed(240) # Setting seed
kmeans.re <- kmeans(iris_1, centers = 3, nstart = 20)
kmeans.re

# Cluster identification for
# each observation
kmeans.re$cluster

# Confusion Matrix
cm <- table(iris$Species, kmeans.re$cluster)
cm

# Model Evaluation and visualization
plot(iris_1[c("Sepal.Length", "Sepal.Width")])
plot(iris_1[c("Sepal.Length", "Sepal.Width")],
    col = kmeans.re$cluster)
plot(iris_1[c("Sepal.Length", "Sepal.Width")],
    col = kmeans.re$cluster,
    main = "K-means with 3 clusters")

## Plotiing cluster centers
kmeans.re$centers
kmeans.re$centers[, c("Sepal.Length", "Sepal.Width")]

# cex is font size, pch is symbol
points(kmeans.re$centers[, c("Sepal.Length", "Sepal.Width")],
      col = 1:3, pch = 8, cex = 3)

## Visualizing clusters
y_kmeans <- kmeans.re$cluster
clusplot(iris_1[, c("Sepal.Length", "Sepal.Width")],
        y_kmeans,
        lines = 0,
        shade = TRUE,
        color = TRUE,
        labels = 2,
        plotchar = FALSE,
        span = TRUE,
        main = paste("Cluster iris"),
        xlab = 'Sepal.Length',
```

```
        ylab = 'Sepal.Width')
```

**********program 7 **********

```
count <- c(19,35,25,12,24,23,4,57)
speed <- c(12,13,15,19,24,34,9,44)
fw <- data.frame(count, speed)
names <- c("Taw", "Torridoe", "Ouse", "Exe", "Lyn", "Brook", "Ditch", "Fal")
rownames(fw) <- names
fw
fw.lm <- lm(count ~ speed, data = fw)
summary(fw.lm)
names(fw.lm) # components of linear regression
fw.lm$coefficients # gives slope and intercept
newypred <- fitted(fw.lm) # predict y values for each x value
newypred
residuals(fw.lm)
# plot x, y values
plot(fw$speed, fw$count, col = "red")
# plot the fitted line
abline(coef(fw.lm), lty = 1, col = "blue")
# plot residuals
plot(fw.lm, which = 1)
```

**********program 8 **********

```
install.packages("e1071")
library(e1071)
getwd()
setwd("/home/nikhil/Downloads")
df = read.csv("bayes.csv")
str(df)
class(df)
testset = data.frame(Age="<=30",Income="Medium",JobSatisfaction=
"No",Desire="Fair",Enrolls="No")
testset
df = rbind(df,testset)
df
train = as.data.frame(df[1:14,])
test = as.data.frame(df[15,])
train
test
Bm = naiveBayes(Enrolls ~ Age+Income+JobSatisfaction+Desire,train)
Bm
res = predict(Bm,test)
res
modell = naiveBayes(Enrolls ~ ., train, laplace=.01)
modell
results1<-predict(modell,test)
results1
```

**********program 9 **********

```
install.packages("rpart.plot")
library("rpart")
```

```
library("rpart.plot")
play<-c('yes','no','yes','no','yes','yes','yes','yes','yes','no')
outlook<-c('rainy','rainy','overcast','sunny','rainy','sunny','rainy','sunny','overcast','sunny')
Temperature<-c('cool','cool','hot','mild','cool','cool','cool','hot','mild','mild')
Humidity<-c('normal','normal','high','high','normal','normal','normal','normal','high','high')
Wind<-c('FALSE','TRUE','FALSE','FALSE','FALSE','FALSE','FALSE','FALSE','TRUE','TRUE')
play_decision<-cbind(play,outlook,Temperature,Humidity,Wind)
play_decision
play_decision=as.data.frame(play_decision)
play_decision
summary(play_decision)
fit=rpart(play ~ outlook + Temperature  +Humidity
+Wind,method="class",data=play_decision,control=rpart.control(minsplit=1),parms=list(split='info
rmation'))
fit
str(fit)
?rpart.plot
rpart.plot(fit,type=4,extra=1)
rpart.plot(fit,type=4,extra=2,clip.right.labs=FALSE,varlen=0,faclen=0)
```

**********program 10 **********

```
sfly<-c(26,23,33,6,3,4,20,2)
mfly<-c(4,5,12,9,15,10,8,22)
speed<-c(3,4,4,5,6,7,7,9)
ivert<-cbind(sfly,mfly)
spd<-cbind(speed)
matplot(spd,ivert,type='b',pch=1:2,col=1,lty=2:3,xlab='Speed',ylab='Invertebrate')
legend(x = 'topright',legend =c('Stonefly','MayFly'),col=1,pch=1:2,lty=2:3)
```