

#PRG 6

```
# install.packages("ggplot2")
# library(ggplot2)

rm(list = ls())

normal_data <- rnorm(1000, mean = 50, sd = 10)
uniform_data <- runif(1000, min = 30, max = 70)

hist(normal_data,
      main = "Histogram: Normal Distribution",
      xlab = "Value",
      col = "lightblue",
      border = "black")

hist(uniform_data,
      main = "Histogram: Uniform Distribution",
      xlab = "Value",
      col = "lightgreen",
      border = "black")

plot(density(normal_data),
      main = "Density Plot: Normal Distribution",
      xlab = "Value",
      col = "blue",
      lwd = 2)

plot(density(uniform_data),
      main = "Density Plot: Uniform Distribution",
      xlab = "Value",
      col = "green",
      lwd = 2)
```

PRG 7

```
library(ggplot2)
rm(list = ls())

x <- c(1,2,3,4,5,6,7,8,9,10)
y <- c(2,4,5,7,9,10,11,13,14,16)
data <- data.frame(x, y)

model <- lm(y ~ x, data = data)
summary(model)

ggplot(data, aes(x = x, y = y)) +
  geom_point(color = "blue", size = 3) +
  geom_smooth(method = "lm", col = "red", se = FALSE) +
  ggtitle("Simple Linear Regression") +
  xlab("Independent Variable (x)") +
  ylab("Dependent Variable (y)")
```

Prg 8

```
library(pROC)

set.seed(123)
data <- data.frame(age=sample(20:60,100,TRUE),
                    income=sample(20000:80000,100,TRUE),
                    buy=sample(c(0,1),100,TRUE))

model <- glm(buy ~ age + income, data=data, family=binomial)

prob <- predict(model, type="response")

roc_obj <- roc(data$buy, prob)
plot(roc_obj, col="blue", main="ROC Curve")
auc(roc_obj)
```

#Prg 9

```
library(ggplot2)

set.seed(123)
data <- data.frame(x1=rnorm(100,5,2),
                    x2=rnorm(100,10,3),
                    x3=rnorm(100,15,4))

pca <- prcomp(data, scale.=TRUE)
summary(pca)

ggplot(data.frame(pca$x), aes(PC1, PC2)) +
  geom_point(color="blue") +
  ggtitle("PCA - Dimensionality Reduction")
```

#Prg 10

```
library(class)
library(caret)

data(iris)
set.seed(123)
train_index <- sample(1:nrow(iris), 0.7*nrow(iris))
train <- iris[train_index,]
test <- iris[-train_index,]

pred <- knn(train[,1:4], test[,1:4], train$Species, k=3)
confusionMatrix(pred, test$Species)
```

#Prg 11

```
library(arules)

data <- list(
  c("milk", "bread", "butter"),
  c("bread", "butter"),
  c("milk", "bread"),
  c("milk", "bread", "butter", "jam"),
  c("bread", "jam")
)

trans <- as(data, "transactions")
rules <- apriori(trans, parameter = list(supp = 0.4, conf = 0.6))
inspect(rules)
```

#Prg 12

```
library(ggplot2)

set.seed(123)
data <- data.frame(x=rnorm(100,5,1), y=rnorm(100,10,2))
kmodel <- kmeans(data, centers=3)
data$cluster <- as.factor(kmodel$cluster)

ggplot(data, aes(x, y, color=cluster)) +
  geom_point(size=3) +
  ggtitle("K-Means Clustering")
```

#Prg 13

```
data <- c(10, 12, 15, 18, 20, 22, 25, 100)
Q1 <- quantile(data, 0.25)
Q3 <- quantile(data, 0.75)
IQR <- Q3 - Q1
lower_bound <- Q1 - 1.5 * IQR
upper_bound <- Q3 + 1.5 * IQR
clean_data <- data[data >= lower_bound & data <= upper_bound]
print("Original Data:")
print(data)
print("Cleaned Data (Outliers Removed):")
print(clean_data)
```