**SAVEETHA SCHOOL OF ENGINEERING**

**SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES**

**ITA 0443 - STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEM**

**DAY 3 – LAB ASSESSMENT**

**Reg No:**192021033

**Name:**Puli Bhavani

1. (i) Write a function in R programming to print generate Fibonacci sequence using    
Recursion in R  
 . fib <- function(n) {

if (n <= 0) return(0)

if (n == 1) return(1)

return(fib(n - 1) + fib(n - 2))

}

fib\_sequence <- sapply(0:10, fib)

print(fib\_sequence)  
  
(ii) Find sum of natural numbers up-to 10, without formula using loop statement.

sum <- 0

for (i in 1:10) {

sum <- sum + i

}

print(sum)  
  
(iii) create a vector 1:10 and Find a square of each number and store that in a  
separate list.

numbers <- 1:10

squared\_numbers <- sapply(numbers, function(x) x^2)

print(squared\_numbers)  
  
  
2.    (motor trend car road test) comprises fuel consumption, performance and  10 aspects of automobile  
design for 32 automobiles. It comes pre-installed  with  package in R.  
 (i)Find the dimension of the dataset

data<-mtcars

dim(data)

(ii)Give the statistical summary of the features.

data<-mtcars

summary(data)

(iii)Print the categorical features in Dataset  
(iv)Find the average weight(wt) grouped by Engine shape(vs)  
(v)Find the largest and smallest value of the variable weight with respect to Engine shape    
  
3.Use ggplot package to plot below EDA questions label the plot accordingly  
(i)Create weight(wt) vs displacement(disp) scatter plot factor by  Engine Shape(vs)

data<-mtcars

plot(data$wt, data$disp,

xlab = "Weight (wt)",

ylab = "Displacement (disp)",

main = "Weight vs Displacement Scatter Plot")  
(ii) Create horsepower(hp) vs mileage (mgp) scatter plot factor by  Engine Shape(vs)

data<-mtcars

plot(data$hp, data$mpg,

xlab = "Weight (wt)",

ylab = "Displacement (disp)",

main = "Weight vs Displacement Scatter Plot")  
  
(iv)In above plot , Separate columns according to cylinders(cyl) size  
(v) Create histogram plot for horsepower (hp) with bin-width size of 5  
data<-mtcars

hist(data$hp

main = "Histogram of Horsepower",

xlab = "Horsepower (hp)",

breaks = seq(min(hp), max(hp), by = 5),

col = "lightblue" )  
  
  4. Performing Logistic regression on dataset to predict the cars Engine shape(vs) .  
(i)Do the EDA analysis and find the features which is impact the Engine shape and use this for model.  
(ii) Split the data set randomly with 80:20 ration to create train and test dataset and create  
logistic model  
(iii)Create the Confusion matrix among prediction and test data.  
  
5. (I) Write R Program to create 15 x15 matrix filled with random numbers between -10 to 10, numbers can repeat. set random seed value to 328

set.seed(328)

mat <- matrix(runif(225, min = -10, max = 10), nrow = 15, ncol = 15)

print(mat)

Output:

[,1] [,2] [,3] [,4]

[1,] 4.650193 8.5982745 -3.4791402 -0.9142146

[2,] 5.601852 6.4087972 3.4444817 9.0329200

[3,] -6.086451 -5.1788499 6.4367553 -8.4596305

[4,] 6.433422 7.8384104 -8.7492383 1.4152955

[5,] -2.468834 9.0066129 -0.3196906 3.7282971

[6,] 9.778266 1.2960666 -9.1130314 -5.1491268

[7,] -5.717901 -9.0550294 -7.7964704 -6.8597079

[8,] 4.567306 0.6858991 -8.5369708 -8.8627821

[9,] 8.818495 9.7029622 8.2729891 8.5557580

[10,] -2.018106 7.3818216 -5.1239983 -3.1757205

[11,] 5.414889 -7.0444924 7.7643401 7.4922010

[12,] -4.013556 2.6575028 9.9340655 8.3782684

[13,] -9.099600 8.9458338 9.7816288 -7.1914213

[14,] 3.426013 5.1860794 2.8406790 -9.3579626

[15,] 1.237188 -3.3426812 5.9635083 -3.1029011

[,5] [,6] [,7] [,8]

[1,] 1.4040338 -9.073830 7.622729 6.32264633

[2,] 3.1132033 -2.539307 -5.609092 9.98325313

[3,] -0.7715112 -3.913295 -6.907849 -1.30280815

[4,] -6.0152877 6.185320 1.160646 -9.19050957

[5,] 4.4253976 -1.610087 6.796707 -8.37721239

[6,] 7.2769148 -9.877884 -3.176692 0.74530938

[7,] -9.5029370 6.250492 4.651562 1.27931052

[8,] -3.4467673 -4.356717 4.653412 4.62269311

[9,] -3.7369353 -3.566020 7.290522 1.82961073

[10,] -3.7975630 -9.436655 -5.156104 9.45244279

[11,] 6.6073701 -9.558082 -3.138002 5.84233446

[12,] -4.8960865 8.595799 -5.243082 -8.78539019  
    (ii)Write R Program to display Lower Diagonal and upper Diagonal matrix

set.seed(328)

mat <- matrix(runif(225, min = -10, max = 10), nrow = 15, ncol = 15)

print(mat)

print("Lower Diagonal:")

print(mat[lower.tri(mat)])

print("Upper Diagonal:")

print(mat[upper.tri(mat)])  
   (iii)Write R Program to count 0's in the matrix and check the matrix is sparse matrix or not

set.seed(328)

mat <- matrix(runif(225, min = -10, max = 10), nrow = 15, ncol = 15)

print(mat)

zeros <- sum(mat == 0)

is\_sparse <- ifelse(zeros >= 0.9 \* prod(dim(mat)), "The matrix is sparse", "The matrix is not sparse")

print(paste("Number of zeros in the matrix:", zeros))

print(is\_sparse)

   (iv) Write R code to remove outliers. Here the outliers are negative numbers. replace the negative values with positive values

set.seed(328)

mat <- matrix(runif(225, min = -10, max = 10), nrow = 15, ncol = 15)

print(mat)

mat[mat < 0] <- abs(mat[mat < 0])

print(mat)  
   (v) Find the mean median and mode of the values corresponding to column

set.seed(328)

mat <- matrix(runif(225, min = -10, max = 10), nrow = 15, ncol = 15)

mean\_col <- colMeans(mat)

median\_col <- apply(mat, 2, median)

mode\_col <- sapply(1:ncol(mat), function(x) {

u <- unique(mat[, x])

u[which.max(tabulate(match(mat[, x], u)))]

})

# Printing the result

print(paste("Mean of the values corresponding to column:", mean\_col))

print(paste("Median of the values corresponding to column:", median\_col))

print(paste("Mode of the values corresponding to column:", mode\_col))  
   (vi)Find the mean median and mode of the values corresponding to row

set.seed(328)

mat <- matrix(runif(225, min = -10, max = 10), nrow = 15, ncol = 15)

mean\_row <- rowMeans(mat)

median\_row <- apply(mat, 1, median)

mode\_row <- sapply(1:nrow(mat), function(x) {

u <- unique(mat[x, ])

u[which.max(tabulate(match(mat[x, ], u)))]

})