**ASSESSMENT-3**

**DT:10-02-23 ITA0443-STATISTICS WITH R PROGRAMMING**

**NAME:S.SHASHIKNATH**

**REG.NO:192125026**

**1. (i) Write a function in R programming to print generate Fibonacci sequence using**

**Recursion in R**

**CODE:**

fib <- function(n) {

if (n <= 1) {

return (n)

} else {

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

}

}

for (i in 1:10) {

print(fib(i))

}

**OUTPUT:**

[1] 0

[1] 1

[1] 1

[1] 2

[1] 3

[1] 5

[1] 8

[1] 13

[1] 21

[1] 34

**(ii) Find sum of natural numbers up-to 10, without formula using loop statement.**

**CODE:**

sum = 0

for (i in 1:10) {

sum = sum + i

}

print(sum)

**OUTPUT:**

[1] 55

**(iii) create a vector 1:10 and Find a square of each number and store that in a**

**separate list.**

**CODE:**

vec <- 1:10

squared\_vec <- sapply(vec, FUN = function(x) x^2)

print(squared\_vec)

**OUTPUT:**

**[1] 1 4 9 16 25 36 49 64 81 100**

**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**

**CODE:**

library(datasets)

data("mtcars")

dim(mtcars)

**OUTPUT:**

[1] 32 11

**(ii)Give the statistical summary of the features.**

**CODE:**

library(datasets)

data("mtcars")

summary(mtcars)

**OUTPUT:**

mpg cyl disp hp

Min. :10.40 Min. :4.000 Min. : 71.1 Min. : 52.0

1st Qu.:15.43 1st Qu.:4.000 1st Qu.:120.8 1st Qu.: 96.5

Median :19.20 Median :6.000 Median :196.3 Median :123.0

Mean :20.09 Mean :6.188 Mean :230.7 Mean :146.7

3rd Qu.:22.80 3rd Qu.:8.000 3rd Qu.:326.0 3rd Qu.:180.0

Max. :33.90 Max. :8.000 Max. :472.0 Max. :335.0

drat wt qsec vs

Min. :2.760 Min. :1.513 Min. :14.50 Min. :0.0000

1st Qu.:3.080 1st Qu.:2.581 1st Qu.:16.89 1st Qu.:0.0000

Median :3.695 Median :3.325 Median :17.71 Median :0.0000

Mean :3.597 Mean :3.217 Mean :17.85 Mean :0.4375

3rd Qu.:3.920 3rd Qu.:3.610 3rd Qu.:18.90 3rd Qu.:1.0000

Max. :4.930 Max. :5.424 Max. :22.90 Max. :1.0000

am gear carb

Min. :0.0000 Min. :3.000 Min. :1.000

1st Qu.:0.0000 1st Qu.:3.000 1st Qu.:2.000

Median :0.0000 Median :4.000 Median :2.000

Mean :0.4375 Mean :3.688 Mean :2.812

3rd Qu.:1.0000 3rd Qu.:4.000 3rd Qu.:4.000

Max. :1.0000 Max. :5.000 Max. :8.000

**(iii)**Print the categorical features in Dataset

**CODE:**

data("mtcars")

col\_classes <- sapply(mtcars, class)

categorical\_features <- mtcars[, col\_classes == "factor"]

print(categorical\_features)

**OUTPUT:**

mtcars$am <- factor(mtcars$am)

**(iv)Find the average weight(wt) grouped by Engine shape(vs)**

**CODE:**

library(dplyr)

data("mtcars")

result <- mtcars %>%

group\_by(vs) %>%

summarise(avg\_wt = mean(wt))

print(result)

**OUTPUT:**

# A tibble: 2 x 2

vs avg\_wt

<dbl> <dbl>

1 0 2.84

2 1 3.17

**(v)Find the largest and smallest value of the variable weight with respect to Engine shape**

**CODE:**

library(dplyr)

data("mtcars")

result <- mtcars %>%

group\_by(vs) %>%

summarise(min\_wt = min(wt), max\_wt = max(wt))

print(result)

**OUTPUT:**

# A tibble: 2 x 3

vs min\_wt max\_wt

<dbl> <dbl> <dbl>

1 0 1.51 3.22

2 1 2.31 3.73

**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)**

**CODE:**

library(ggplot2)

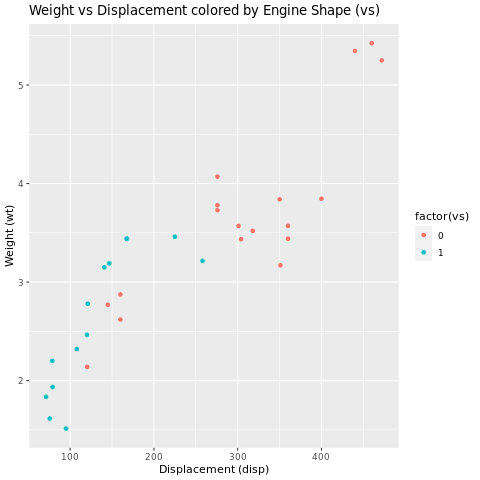
data("mtcars")

ggplot(mtcars, aes(x = disp, y = wt, color = factor(vs))) +

geom\_point() +

labs(x = "Displacement (disp)", y = "Weight (wt)", title = "Weight vs Displacement colored by Engine Shape (vs)")

**OUTPUT:**

****

**(ii) Create horsepower(hp) vs mileage (mgp) scatter plot factor by  Engine Shape(vs)**

**CODE:**

library(ggplot2)

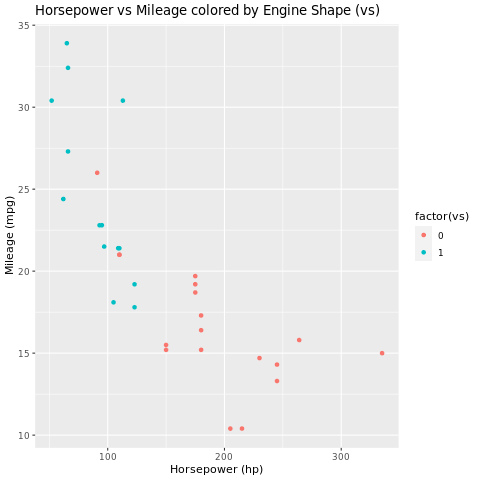
data("mtcars")

ggplot(mtcars, aes(x = hp, y = mpg, color = factor(vs))) +

geom\_point() +

labs(x = "Horsepower (hp)", y = "Mileage (mpg)", title = "Horsepower vs Mileage colored by Engine Shape (vs)")

**OUTPUT:**

****

**(iv)In above plot , Separate columns according to cylinders(cyl) size**

**CODE:**

library(ggplot2)

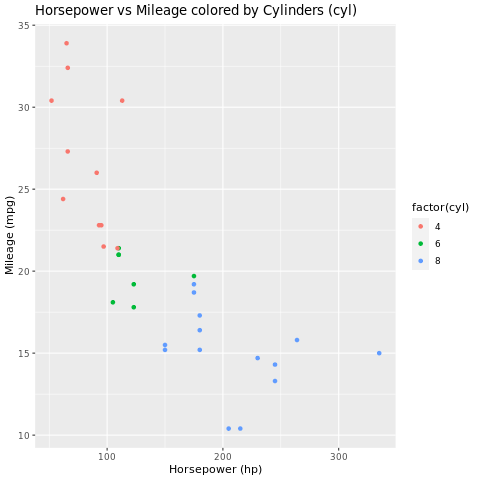
data("mtcars")

ggplot(mtcars, aes(x = hp, y = mpg, color = factor(cyl))) +

geom\_point() +

labs(x = "Horsepower (hp)", y = "Mileage (mpg)", title = "Horsepower vs Mileage colored by Cylinders (cyl)")

**OUTPUT:**

****

**(v) Create histogram plot for horsepower (hp) with bin-width size of 5**

**CODE:**

library(ggplot2)

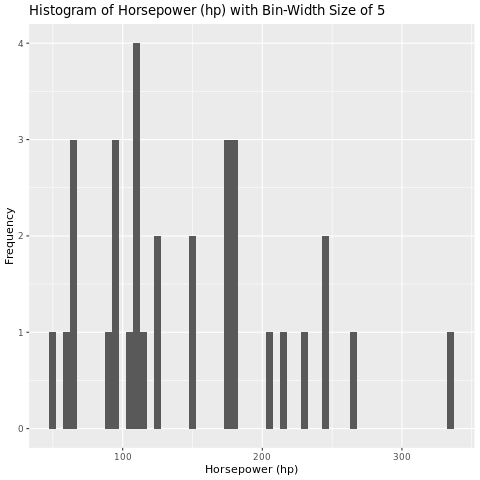
data("mtcars")

ggplot(mtcars, aes(x = hp)) +

geom\_histogram(binwidth = 5) +

labs(x = "Horsepower (hp)", y = "Frequency", title = "Histogram of Horsepower (hp) with Bin-Width Size of 5")

**OUTPUT:**

****

**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.**

**CODE:**

data("mtcars")

summary(mtcars)

library(ggplot2)

ggplot(mtcars, aes(x=wt, y=disp, color=factor(vs))) + geom\_point()

ggplot(mtcars, aes(x=hp, y=mpg, color=factor(vs))) + geom\_point()

ggplot(mtcars, aes(x=hp, y=mpg, color=factor(cyl))) + geom\_point()

ggplot(mtcars, aes(hp)) + geom\_histogram(binwidth=5)

cor\_matrix <- cor(mtcars)

**OUTPUT:**

mpg cyl disp hp

Min. :10.40 Min. :4.000 Min. : 71.1 Min. : 52.0

1st Qu.:15.43 1st Qu.:4.000 1st Qu.:120.8 1st Qu.: 96.5

Median :19.20 Median :6.000 Median :196.3 Median :123.0

Mean :20.09 Mean :6.188 Mean :230.7 Mean :146.7

3rd Qu.:22.80 3rd Qu.:8.000 3rd Qu.:326.0 3rd Qu.:180.0

Max. :33.90 Max. :8.000 Max. :472.0 Max. :335.0

drat wt qsec vs

Min. :2.760 Min. :1.513 Min. :14.50 Min. :0.0000

1st Qu.:3.080 1st Qu.:2.581 1st Qu.:16.89 1st Qu.:0.0000

Median :3.695 Median :3.325 Median :17.71 Median :0.0000

Mean :3.597 Mean :3.217 Mean :17.85 Mean :0.4375

3rd Qu.:3.920 3rd Qu.:3.610 3rd Qu.:18.90 3rd Qu.:1.0000

Max. :4.930 Max. :5.424 Max. :22.90 Max. :1.0000

am gear carb

Min. :0.0000 Min. :3.000 Min. :1.000

1st Qu.:0.0000 1st Qu.:3.000 1st Qu.:2.000

Median :0.0000 Median :4.000 Median :2.000

Mean :0.4062 Mean :3.688 Mean :2.812

3rd Qu.:1.0000 3rd Qu.:4.000 3rd Qu.:4.000

Max. :1.0000 Max. :5.000 Max. :8.000

[Execution complete with exit code 0]

**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**

**CODE:**

set.seed(328)

mat <- matrix(round(runif(225,-10,10)),15,15)

mat

**OUTPUT:**

[,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13]

[1,] 5 9 -3 -1 1 -9 8 6 9 1 -7 4 -2

[2,] 6 6 3 9 3 -3 -6 10 -3 7 -2 6 10

[3,] -6 -5 6 -8 -1 -4 -7 -1 9 6 10 -5 -8

[4,] 6 8 -9 1 -6 6 1 -9 -8 8 3 4 -3

[5,] -2 9 0 4 4 -2 7 -8 -5 2 -2 -1 -7

[6,] 10 1 -9 -5 7 -10 -3 1 8 4 -1 4 3

[7,] -6 -9 -8 -7 -10 6 5 1 0 -4 2 -1 -10

[8,] 5 1 -9 -9 -3 -4 5 5 -8 -6 9 -5 -7

[9,] 9 10 8 9 -4 -4 7 2 -7 8 10 -6 -4

[10,] -2 7 -5 -3 -4 -9 -5 9 8 7 -7 -8 -1

[11,] 5 -7 8 7 7 -10 -3 6 -6 2 9 -6 -9

[12,] -4 3 10 8 -5 9 -5 -9 -10 2 0 -1 -6

[13,] -9 9 10 -7 6 5 8 0 -4 6 -10 9 0

[14,] 3 5 3 -9 4 9 -6 9 -3 0 6 7 7

[15,] 1 -3 6 -3 -1 -10 7 5 1 4 4 1 -6

[,14] [,15]

[1,] -5 -6

[2,] 10 8

[3,] 5 4

[4,] -4 -9

[5,] 3 0

[6,] 8 2

[7,] -6 -5

[8,] 1 7

[9,] 4 -6

[10,] 6 -9

[11,] 3 -6

[12,] 2 -6

[13,] 10 1

[14,] 0 6

[15,] 4 8