**JECRC University, Jaipur**

**Assignment Unit IV**

**MCA - III Semester**

**Subject: Statistical Computing using R**

**Subject Code: MCA206A**

**Date of Release:**

**Last Date of Submission:**

**Part-A 10X1=10**

Q.1 Which of the following statement gives cumulative sum?  
a) cumsum(x,na=rm=TRUE)  
b) cumprod(x)  
c) cummax(x)  
d) cummin(x)

Q.2 Which of the following statement tells the row with the minimum value for every column?  
a) which.min(x)  
b) which.max(x)  
c) z=apply(x,1,which.min)  
d) z=apply(1,1,which.max)

Q.3 Which of the following function is used for plotting histogram?  
a) hist()  
b) histog()  
c) histg()  
d) histo()

Q.4 Which of the following will add the title “R language” to the graph?  
a) titleAdd( “R language”)  
b) title( “R language”)  
c) titleBar( “R language”)  
d) var(x, na.rm=TRUE)

Q.5  \_\_\_\_\_\_ finds K best paths in a given graph.  
a) kBestShortestPaths  
b) kcirt  
c) ktrees  
d) kmap

Q.6 \_\_\_\_\_\_\_\_ is used to view packages currently loaded.  
a) library()  
b) search()  
c) .libPaths()  
d) stringr()

Q.7\_\_\_\_\_\_\_\_ is used to view packages currently loaded.

a) library()

b) search()

c) .libPaths()

d) stringr()

Q.8 \_\_\_\_\_\_\_\_\_\_ read only parameter that returns the width and height of the current device surface in inches.

a) Din

b) Fin

c) Gin

d) Kin

Q.9Which level plotting commands generate figures?

a) Low

b) High

c) Both high and low

d) No levels

Q.10 Which of the following method make vector of repeated values?

a) rep()

b) data()

c) view()

d) read()

Q.11 For the population

Y<-c(1,2,3,4,5),Write the R command to find the mean?

a)mean{y}

b)mean[y]

c)mean(y)

d)means(y)

Q.12 Which of the following is multivariate version of lapply?

a) apply()

b) lapply()

c) sapply()

d) mapply()

Q.13 \_\_\_\_\_\_\_\_ loop over a list and evaluate a function on each element.

a) apply()

b) lapply()

c) sapply()

d) mapply()

Q.14\_\_\_\_\_\_\_\_\_\_\_\_ produces one-dimensional scatterplots.

a) xyplot

b) stripplot

c) barchart

d) bwplot

Q.15Which is the R command for obtaining 1000 random numbers through normal distribution with mean 0 and variance 1?

qnorm(0, 1, 1000)

rnorm(0, 1, 1000)

rnorm(1000, 0, 1)

norm(1000, 0, 1)

Q.16 \_\_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_\_ are types of matrices functions?

a) None of These

b) apply() and sapply()

c) apply() and lapply()

d) Both

Q.17 R Language functionality is divided into a number of \_\_\_\_\_\_\_\_

a)Documentation

b)Functions

c)Domains

d)Packages

Q.18What will be the output of the following R code?

> x <- list(a = 1:5, b = rnorm(10))

> lapply(x, mean)

a)$a

[1] 3

$b

[1] 0.1322028

b) $a

[1] 4

$b

[1] 0.1322028

c) $a

[1] 5

$b

[1] 0.1322028

d) $a

[2] 5

$b

[1] 3

Q.19 If a Programmer wants the output to be a list then \_\_\_\_\_\_\_\_\_\_ function is used.

a)Zapply

b)Sapply

C)Vapply

d) Lapply

Q.20 Which function helps you perform sorting in R language?

a) Library

b) Order

c)Simple

d)Inorder

**Part-B 5X2=10**

Q.1What does apply() do in R? Give Example

Returns a vector or array or list of values obtained by applying a function to margins of an array or matrix.

Syntax: apply(X, MARGIN, FUN, …)

Q.2 How do you make a graph in R programming?

The plot() function is used to draw points (markers) in a diagram.

The function takes parameters for specifying points in the diagram.

Parameter 1 specifies points on the **x-axis**.

Parameter 2 specifies points on the **y-axis**.

Q.3 Why do we use line charts in R ?

line graphs or line charts are used **to track variations over time**, which may be long-term or short-term. We can also use line graphs to compare changes over the same period for more than one group.

Q.4 Differentiate between Pie Chart, Bar Graph and Histogram?

A pie chart is **a circle that is divided into areas, or slices**. Each slice represents the count or percentage of the observations of a level for the variable. Pie charts are often used in business. Histogram refers to a graphical representation, that displays data by way of bars to show the frequency of numerical data. Bar graph is a pictorial representation of data that uses bars to compare different categories of data.

Q.5 What are the data frames? Write its significance in R-Language?

A data frame is a table or a two-dimensional array-like structure in which each column contains values of one variable and each row contains one set of values from each column.

Data Frames are **data displayed in a format as a table**. Data Frames can have different types of data inside it. While the first column can be character , the second and third can be numeric or logical . However, each column should have the same type of data.

**Part-C 5X6=30**

Q.1 Name different Types of Graphs in R Programming? Explain all with Example.

Bar plot or Bar Chart in R is used to represent the values in data vector as height of the bars. The data vector passed to the function is represented over y-axis of the graph. Bar chart can behave like histogram by using **table()** function instead of data vector.

***Syntax:****barplot(data, xlab, ylab)*

***where:***

* ***data****is the data vector to be represented on y-axis*
* ***xlab****is the label given to x-axis*
* ***ylab****is the label given to y-axis*

Pie chart is a circular chart divided into different segments according to the ratio of data provided. The total value of the pie is 100 and the segments tell the fraction of the whole pie. It is another method to represent statistical data in graphical form and **pie()** function is used to perform the same.

***Syntax:****pie(x, labels, col, main, radius)*

***where,***

* ***x****is data vector*
* ***labels****shows names given to slices*
* ***col****fills the color in the slices as given parameter*
* ***main****shows title name of the pie chart*
* ***radius****indicates radius of the pie chart. It can be between -1 to +1*

Histogram is a graphical representation used to create a graph with bars representing the frequency of grouped data in vector. Histogram is same as bar chart but only difference between them is histogram represents frequency of grouped data rather than data itself.

***Syntax:****hist(x, col, border, main, xlab, ylab)*

***where:***

* ***x****is data vector*
* ***col****specifies the color of the bars to be filled*
* ***border****specifies the color of border of bars*
* ***main****specifies the title name of histogram*
* ***xlab****specifies the x-axis label*
* ***ylab****specifies the y-axis label*

A Scatter plot is another type of graphical representation used to plot the points to show relationship between two data vectors. One of the data vectors is represented on x-axis and another on y-axis.

***Syntax:****plot(x, y, type, xlab, ylab, main)*

***Where,***

* ***x****is the data vector represented on x-axis*
* ***y****is the data vector represented on y-axis*
* ***type****specifies the type of plot to be drawn. For example, “l” for lines, “p” for points, “s” for stair steps, etc.*
* ***xlab****specifies the label for x-axis*
* ***ylab****specifies the label for y-axis*
* ***main****specifies the title name of the graph*

Box plot shows how the data is distributed in the data vector. It represents five values in the graph i.e., minimum, first quartile, second quartile(median), third quartile, the maximum value of the data vector.

***Syntax:****boxplot(x, xlab, ylab, notch)*

***where,***

* ***x****specifies the data vector*
* ***xlab****specifies the label for x-axis*
* ***ylab****specifies the label for y-axis*
* ***notch,****if TRUE then creates notch on both the sides of the box*

Q.2 What is Histogram? How do you make a Histogram in R?

Histogram is a graphical representation used to create a graph with bars representing the frequency of grouped data in vector. Histogram is same as bar chart but only difference between them is histogram represents frequency of grouped data rather than data itself.

***Syntax:****hist(x, col, border, main, xlab, ylab)*

***where:***

* ***x****is data vector*
* ***col****specifies the color of the bars to be filled*
* ***border****specifies the color of border of bars*
* ***main****specifies the title name of histogram*
* ***xlab****specifies the x-axis label*

***ylab****specifies the y-axis label*

Q.3 Write the steps how to Make BoxPlots in RStudio.

Box plot shows how the data is distributed in the data vector. It represents five values in the graph i.e., minimum, first quartile, second quartile(median), third quartile, the maximum value of the data vector.

***Syntax:****boxplot(x, xlab, ylab, notch)*

***where,***

* ***x****specifies the data vector*
* ***xlab****specifies the label for x-axis*
* ***ylab****specifies the label for y-axis*
* ***notch,****if TRUE then creates notch on both the sides of the box*

Q.4.Explain tapply() function with its parameter.

tapply() is used to apply a function over subsets of a vector. It is primarily used when we have the following circumstances:

1. A dataset that can be broken up into groups (via categorical variables - aka factors)
2. We desire to break the dataset up into groups
3. Within each group, we want to apply a function

# syntax of tapply function

tapply(x, INDEX, FUN, ..., simplify = TRUE)

The arguments to tapply() are as follows:

* x is a vector
* INDEX is a factor or a list of factors (or else they are coerced to factors)
* FUN is a function to be applied
* ... contains other arguments to be passed FUN
* simplify, should we simplify the result?

Q.5How do you apply a Dataframe in R Programming?

# Apply function to every value in dataframe

# Creating dataset

m <- c("Vikas","Varun","Deepak")

n <- c("Komal","Suneha","Priya")

# creating dataframe

df <- data.frame(A=m,B=n)

# Applying custom function to every element in dataframe

df[]<-data.frame(lapply(df,function(x) paste("Hello,",x,sep="")))

# display dataset

df

Text

Description automatically generated

**Part-D 2X10=20**

Q.1 Write a script that will print 'Is a Matrix' if the variable x is a matrix, otherwise "Not a

Matrix

**YE PURA NAHI MILA** Q.1 (a) What do you mean by Apply Function? Write the command in R console to specify the columns that needs to be excluded in the apply function.

Returns a vector or array or list of values obtained by applying a function to margins of an array or matrix.

syntax: apply(X, MARGIN, FUN, …)

(b) Write a Program Using R pie chart, demonstrate the percentage conveyance of various ways for

travelling to office such as walking, car, bus, cycle and train.

# Create data for the graph.

means<- c(9,3,6,12,2)

labels <- c("walking", "car", "bus", "cycle", "train")

# Plot the chart.

pie(means, labels)

Q.2. (a) Calculate the mean of ozone, solar radiation, and wind within each month using sapply

for air quality dataframe

data = airquality

print("Original data: Daily air quality measurements in New York, May to September 1973.")

print(class(data))

print(head(data,10))

sapply(data, mean)

**YE NAHI MILA** (b). Write the command in R console to obtain mean using tapply function by considering a vector having 10 normal and 10 uniform variables. Assume that these vectors have three groups.