

# National Testing Agency

**Question Paper Name:** Introduction to R  
**Subject Name:** Introduction to R  
**Creation Date:** 2018-12-02 17:35:46  
**Duration:** 180  
**Total Marks:** 150  
**Display Marks:** Yes  
**Share Answer Key With Delivery Engine:** Yes  
**Actual Answer Key:** Yes

## Introduction to R

**Group Number :** 1  
**Group Id :** 41652960  
**Group Maximum Duration :** 0  
**Group Minimum Duration :** 120  
**Revisit allowed for view? :** No  
**Revisit allowed for edit? :** No  
**Break time:** 0  
**Group Marks:** 150

## Introduction to R

**Section Id :** 41652960  
**Section Number :** 1  
**Section type :** Online  
**Mandatory or Optional:** Mandatory  
**Number of Questions:** 150  
**Number of Questions to be attempted:** 150  
**Section Marks:** 150  
**Display Number Panel:** Yes  
**Group All Questions:** No

**Sub-Section Number:** 1  
**Sub-Section Id:** 41652961  
**Question Shuffling Allowed :** Yes

**Question Number : 1 Question Id : 4165294886 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Which of the following is true of using R for statistical data analysis?

- a. R is entirely online
- b. R allows for creation of new modules and extensions to be built
- c. The source code of R is hidden
- d. R allows only data analysis and nothing else

**Question Number : 2 Question Id : 4165294887 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Which of the following statements is true about obtaining help in R?

- a. You can obtain help by calling an operator
- b. The command "help.start()" launches a general overview
- c. R users group is a closed group for developers only
- d. R does not come with in-built help

**Question Number : 3 Question Id : 4165294888 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Which of the following is true of R packages?

- a. You need special permission from R foundation to build R packages
- b. R packages are proprietary
- c. R packages contain functions and data as two elements
- d. To use packages, use resource("packagename")

**Question Number : 4 Question Id : 4165294889 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Which of the following statements is correct?

- a. To use a file of R codes and submit for processing, use the command, sink("filename.R")
- b. To input a csv file into R, use source("myfile.csv")
- c. To read an SPSS file into R use library(foreign); read.spss("spssfile.sav") (sav is the \_le extension for SPSS)
- d. It is not possible to export files from R to other programmes

**Question Number : 5 Question Id : 4165294890 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Which of the following statements is true of integrating R in research workflows?

- a. R cannot be used for anything beyond reading and manipulation of data
- b. R can be used to both read, store, analyse data, and write reports
- c. R cannot be combined with LaTeX
- d. R tables cannot be combined to a format to be read by standard spreadsheet programmes

**Question Number : 6 Question Id : 4165294891 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

How do you install a package (A, say) which is build with the help of another package(B, say)

- a. install.packages("A")
- b. install.packages("A", dependencies=TRUE)
- c. install.packages("A", dependencies="B")
- d. install.packages("A", dependencies="all")

**Question Number : 7 Question Id : 4165294892 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

A comma separated value(CSV) is

- a. an excel data file
- b. a text file separated by tab
- c. a text file separated by comma
- d. a fixed format file separated by comma

**Question Number : 8 Question Id : 4165294893 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The function `read.dta()` can be used to read any stata formatted file which required pre loading of package

- a. foreign
- b. mgcv
- c. epiDisplay
- d. spss

**Question Number : 9 Question Id : 4165294894 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To read a comma separated value format file in R, one has to use the function

- a. `read.table()` or `read.csv()`
- b. only `read.table()`
- c. only `read.csv()`
- d. `read.dta()`

**Question Number : 10 Question Id : 4165294895 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

What is the correct function to get a list of all objects in the workspace in R?

- a. The function: `ls()`
- b. The function: `list.objects()`
- c. The function: `lsobj()`
- d. The function: `list()`

**Question Number : 11 Question Id : 4165294896 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Which function do you use to find the class of an object in R?

- a. The function `as.class(x)` where x is the object name
- b. The function `which(x)` where x is the object name
- c. The function `class(x)` where x is the object name
- d. The function `help(x)` where x is the object name

**Question Number : 12 Question Id : 4165294897 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Read 15 data points from R console and save in a variable x, one can use the code

- a. `x=scan()`
- b. `x=scan("",what=numeric(),nmax=15)`
- c. `x=scan(what=numeric(),nmax=15)`
- d. `x=scan(nmax=15)`

**Question Number : 13 Question Id : 4165294898 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To load the foreign package one use the code

- a. `library(foreign)` or `require(foreign)`
- b. `library(foreign);or required(foreign)`
- c. only `library(foreign)`
- d. only `required(foreign)`

**Question Number : 14 Question Id : 4165294899 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To check whether an object x is numeric or not, one can use the function

- a. `as.numeric(x)`
- b. `is.integer(x)`
- c. `is.numeric(x)`
- d. `class(x)`

**Question Number : 15 Question Id : 4165294900 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To get the position of missing value of a numeric vector y, the code would be

- a. `is.missing(y)`
- b. `is.na(y)`
- c. `which(is.na(y))`
- d. `which(y == missing())`

**Question Number : 16 Question Id : 4165294901 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

someone mistakenly enter 0 in place of missing in variable y. which is the correct code to replace 0 by NA

- a. `y[is.na(y)]=0`
- b. `y[y==0]=NA`
- c. `y[y=0]=NA`
- d. `y[y==0]="NA"`

**Question Number : 17 Question Id : 4165294902 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Suppose you create a matrix of order 3 by 2 where the rows are (4,5),(6,7) and (8,9), the code is

- a. `matrix(c(4,5,6,7,8,9),nrow=3,ncol=2)`
- b. `matrix(c(4,5,6,7,8,9),nrow=3,ncol=2,byrow=T)`
- c. `matrix(c(4,5,6,7,8,9),nrow=2,ncol=3,byrow=T)`
- d. `matrix(c(4,5,6,7,8,9),nrow=2,ncol=3,byrow=F)`

**Question Number : 18 Question Id : 4165294903 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Suppose you have two matrices A (order 2x5) and B (order 5x2) to make the operation  $A+B'$  in R

- a. `A+tran(B)`
- b. `A+t(B)`
- c. `A+transpose(B)`
- d. `A+B`

**Question Number : 19 Question Id : 4165294904 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

If a R numeric object contain an alphabet or special character, R will treat the object as

- a. integer
- b. double
- c. logical
- d. character

**Question Number : 20 Question Id : 4165294905 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

If an object contains all numeric values but R reads as non-numeric, how do you convert that to numeric

- a. `class(x)=numeric()`
- b. `is.numeric(x)`
- c. `as.numeric(x)`
- d. `x=numeric()`

**Question Number : 21 Question Id : 4165294906 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Let  $A(m \times n)$  and  $B(m \times n)$  are two matrices. The appropriate R code for matrix multiplication of two matrices is

- a.  $A * B$
- b.  $A * t(B)$
- c.  $A \% \% t(B)$
- d.  $t(A) \% \% B$

**Question Number : 22 Question Id : 4165294907 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Suppose there is matrix (A) of order  $(10 \times 5)$ , one wants to have a scalar multiplication of the matrix by a column vector (x) of order  $(10 \times 1)$ . The R code would be

- a.  $A \% \% x$
- b.  $A * x$
- c.  $A . x$
- d.  $x \% \% A$

**Question Number : 23 Question Id : 4165294908 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Which one is correct R code to determine Pearson's Correlation Coefficient between x and y without using inbuilt 'cor()' function

- a.  $cor(x, y)$
- b.  $sum((x - mean(x)) * (y - mean(y))) / length(x)$
- c.  $sum((x - mean(x)) * (y - mean(y))) / sqrt(sum((x - mean(x))^2) * sum((y - mean(y))^2))$
- d.  $sum((x - mean(x)) * (y - mean(y))) / (sd(x) * sd(y))$

**Question Number : 24 Question Id : 4165294909 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

There is an estimated variance covariance matrix (S) of order  $(5 \times 5)$ . Someone wants to report SE of the five parameters. Diagonal elements of a variance covariance matrix are the variance estimates of the desired parameters. The right R code to get the SEs are

- a.  $sqrt(S)$
- b.  $diag(S)$
- c.  $sqrt(diag(S))$
- d.  $se(S)$

**Question Number : 25 Question Id : 4165294910 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To solve the linear system of equation  $Ax = b$  [A is matrix of order  $p \times p$  & x and b are vector of order  $(p \times 1)$ ], one can use the R code to find x

- a.  $solve(A, b)$
- b.  $inv(A) \% \% b$
- c.  $solve(A, x)$
- d.  $solve(b, A)$

**Question Number : 26 Question Id : 4165294911 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

For a multiple bar diagram of a cross table between Age group (age) and SES (ses), the correct R code is

- a.  $barplot(age, ses)$
- b.  $barplot(table(age, ses))$
- c.  $barplot(table(age, ses), beside = TRUE)$
- d.  $barplot(age, ses, beside = T)$

**Question Number : 27 Question Id : 4165294912 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To draw a density plot on histogram two important steps are

- Set "freq=TRUE" and call additional function 'lines()'
- Set "freq=FALSE" and call additional function 'lines()'
- Set "freq=TRUE" and call additional function 'line()'
- Set "freq=FALSE" and call additional function 'line()'

**Question Number : 28 Question Id : 4165294913 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To draw a boxplot of 'income' stratified by 'education' which are two columns of a data frame 'mydata', correct R code is

- boxplot(income~education)
- boxplot(income~education,data=mydata)
- boxplot(split(income,education))
- boxplot(split(income,education),data=data)

**Question Number : 29 Question Id : 4165294914 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To create two panel of plots side by side, one should set par as prior to call the relevant function for the plot

- par(panel=T)
- par(panel=c(1,2))
- par(mfrow=c(1,2))
- par(mfcol=c(1,2))

**Question Number : 30 Question Id : 4165294915 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To draw a regression line on the scatter plot of y against x, the R code is

- plot(x,y); abline(lm(y~x))
- plot(x,y); abline(y~x)
- plot(x,y); lines(x,y)
- plot(x,y); abline(h=0)

**Question Number : 31 Question Id : 4165294916 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Which of these is an example of a high level plotting function?

- lines()
- points()
- abline()
- plot()

**Question Number : 32 Question Id : 4165294917 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Function that is used to draw histogram by Lattice package is

- hist()
- histogram()
- Histogram()
- Hist()

**Question Number : 33 Question Id : 4165294918 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The low order function `abline(v=5)` draw

- a. vertical line at  $x=5$  on the existing plot
- b. diagonal line at  $x=5$  on the existing plot
- c. horizontal line at  $x=5$  on the existing plot
- d. tick at  $x=5$  on the existing plot

**Question Number : 34 Question Id : 4165294919 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The code for scatter plot using `ggplot2` is

- a. `ggplot(data=data,aes(x=bmi,y=birthweight))`
- b. `ggplot(data=data,aes(x=bmi,y=birthweight))+geom_point()`
- c. `plot(data=data,aes(x=bmi,y=birthweight))+geom_point()`
- d. `gplot(data=data,aes(x=bmi,y=birthweight))+geom_point()`

**Question Number : 35 Question Id : 4165294920 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The mean of a variable  $x=c(98,76,45,NA,56,67,87)$  by `mean(x)` is

- a. 71.5
- b. NA
- c. 75.2
- d. 68

**Question Number : 36 Question Id : 4165294921 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To calculate column mean of a numeric matrix `A` which contain some missing values, the code is

- a. `apply(A,2,mean)`
- b. `apply(A,2,mean,na.rm=T)`
- c. `sapply(A,2,mean,na.rm=T)`
- d. `tapply(A,2,mean)`

**Question Number : 37 Question Id : 4165294922 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The function `sapply()` returns a R object of

- a. vector
- b. matrix
- c. list
- d. data.frame

**Question Number : 38 Question Id : 4165294923 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The function `summary()` doesn't derive

- a. median
- b. quantile
- c. min
- d. sd

**Question Number : 39 Question Id : 4165294924 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The function `xtabs()` returns a

- a. table
- b. contingency table
- c. chi-square test
- d. OR

**Question Number : 40 Question Id : 4165294925 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

A data frame(mydata) includes income and region. Which code is not correct

- a. with(mydata,tapply(income,region,mean,na.rm=T))
- b. tapply(mydata\$income,mydata\$region,mean,na.rm=T)
- c. attach(mydata); tapply(income,region,mean,na.rm=T)
- d. tapply(income,region,mean,na.rm=T)

**Question Number : 41 Question Id : 4165294926 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

One derives the median income by country, which code is correct

- a. sapply(income,country,median,na.rm=T)
- b. sapply(income~country,median,na.rm=T)
- c. sapply(split(income,country),median,na.rm=T)
- d. sapply(cbind(income,country),median,na.rm=T)

**Question Number : 42 Question Id : 4165294927 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Suppose a data frame(data) include daily temperature(temp), RH(rh) and month of measurements. If you want derive monthly mean of both temperature and RH, the code is

- a. aggregate(cbind(temp,rh),month,data=data,mean)
- b. aggregate(cbind(temp,rh)~month,data=data,mean)
- c. aggregate(temp~month, rh~month ,data=data,mean)
- d. aggregate(temp+ rh~month ,data=data,mean)

**Question Number : 43 Question Id : 4165294928 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

let consider a data 'attitude' from library(stats), To subtract column median from each corresponding column, the right code is

- a. sweep(attitude,2,apply(attitude,2,median),FUN="-")
- b. attitude-apply(attitude,2,median)
- c. attitude- apply(attitude,1,median)
- d. sweep(attitude,1,apply(attitude,2,median),FUN="-")

**Question Number : 44 Question Id : 4165294929 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

What is output of as.Date('April 26,2001','%B %d,%Y')

- a. 26/04/2001
- b. 04/26/2001
- c. 26-04-2001
- d. 26-04-2011

**Question Number : 45 Question Id : 4165294930 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

What is output of ISOdate(2005,10,21,18,47,22,tz="GMT")

- a. "2005-10-21 18:47:22 GMT"
- b. "2005/10/21 18:47:22 GMT"
- c. "21/10/2005 18:47:22 GMT"
- d. "21-10-2005 18:47:22 GMT"

**Question Number : 46 Question Id : 4165294931 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**



Two data set can be merged by their respective ids by the function

- a. cbind()
- b. data.frame()
- c. merge()
- d. match()

**Question Number : 47 Question Id : 4165294932 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

An user defined function that returns the number of non-missing elements in an array

- a. count<-function(x) {length(x)}
- b. count<-function(x) {length(na.omit(x))}
- c. count<-function(x) {which(!is.na(x))}
- d. count<-function(x) {na.exclude(x)}

**Question Number : 48 Question Id : 4165294933 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To create a global variable, it needs to be defined before any user defined function is defined and anything assigned to that variable by

- a. '='
- b. '=='
- c. '<-'
- d. '<<-'

**Question Number : 49 Question Id : 4165294934 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Function that return a mean with default 2 decimal places, but allow to set the decimal place

- a. mymean=function(x) round(mean(x),2)
- b. mymean=function(x,sig) round(mean(x),sig)
- c. mymean=function(x,sig=2) round(mean(x,sig))
- d. mymean=function(x,...) round(mean(x,sig))

**Question Number : 50 Question Id : 4165294935 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

If y be the yield and g be the group indicator, the ANOVA table can be generated from

- a. aov(y~g)
- b. summary(aov(y~g))
- c. summary(lm(y~g))
- d. aov(y,g)

**Question Number : 51 Question Id : 4165294936 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To apply ANOVA post-hoc test, the code to be used

- a. TukeyHSD(aov(y~g))
- b. TukeyHSD(y~g)
- c. TukeyHSD(anova(y~g))
- d. TukeyHSD(anova(lm(y~g)))

**Question Number : 52 Question Id : 4165294937 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

How to test equal variance assumption in ANOVA?

- a. T test
- b. Fisher exact test
- c. Levene's Test
- d. Wald test

**Question Number : 53 Question Id : 4165294938 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

If y is continuous variable, and x and z are categorical variables. What is output of the following codes:

```
aovs=aov(y ~ x + z, data = d); TukeyHSD(aovs)
```

- a. Paired comparison of x
- b. Paired comparison of y
- c. Paired comparison of both x and z
- d. TukeyHSD plot only

**Question Number : 54 Question Id : 4165294939 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The output of the following R code of y variable and two categorical variables x(2 levels) & z( 3levels) : boxplot (y ~x + z) with generate

- a. One box-whisker plot
- b. Two box-whisker plots
- c. Six box-whisker plots
- d. Three box-whisker plots

**Question Number : 55 Question Id : 4165294940 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

A scatterplot with fitted line always fit

- a. linear regression
- b. multiple linear regression
- c. glm
- d. lme

**Question Number : 56 Question Id : 4165294941 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

For simple linear regression

- a. Both response and predictor should be categorical
- b. Response continuous but predictor categorical
- c. Response categorical but predictor continuous
- d. Both response and predictors should be continuous

**Question Number : 57 Question Id : 4165294942 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To fit a regression without intercept, the code is

- a. `lm(y ~x)`
- b. `lm(y ~x-1)`
- c. `lm(y ~-1+x)`
- d. `lm(y ~1)`

**Question Number : 58 Question Id : 4165294943 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

suppose fit is the fitted lm object to extract only regression coefficient

- a. coef(fit)
- b. summary(fit)
- c. reg(fit)
- d. beta(fit)

**Question Number : 59 Question Id : 4165294944 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To generate 95% confidence interval of regression coefficient

- a. confint(lm(y~x))
- b. interval(lm(y~x))
- c. confinterval(lm(y~x))
- d. ci(lm(y~x))

**Question Number : 60 Question Id : 4165294945 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To get regression diagnostic plots

- a. diag(lm(y~x))
- b. plot(lm(y~x))
- c. diagnostic(y~x)
- d. plot(residuals(lm(y~x)))

**Question Number : 61 Question Id : 4165294946 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Cook's distance is a measure to identify

- a. Outlier points in regression
- b. Missing values in regression
- c. Influential points in regression
- d. Leverage points in regression

**Question Number : 62 Question Id : 4165294947 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Which of the diagnostic plot of regression check the normality

- a. Residuals vs fitted
- b. Scale-location
- c. Cook's distance
- d. Normal-Q-Q plot

**Question Number : 63 Question Id : 4165294948 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The function `pwr.t.test()` is function for sample size calculation or power calculation for t test belongs to the R library

- a. power
- b. pwr
- c. epiDisplay
- d. samplesize

**Question Number : 64 Question Id : 4165294949 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Measure of goodness of fit for multiple linear regression is

- a. R<sup>2</sup>
- b. Adj R<sup>2</sup>
- c. Residual standard error
- d. P-value

**Question Number : 65 Question Id : 4165294950 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To fit the model  $y = a + b_1x_1 + b_2x_2 + b_3x_1x_2 + e$ , the appropriate R code is

- a. `lm(y ~ x1 + x2 + x1:x2, data = mydata)`
- b. `lm(y ~ x1 + x2 + x1:x2, data = mydata)`
- c. `lm(y ~ x1 * x2, data = mydata)`
- d. `lm(y ~ x1 + x2 + x1 * x2, data = mydata)`

**Question Number : 66 Question Id : 4165294951 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Suppose you have two models(model1 & model2) one has all predictors and another one has all but one dropped out. To compare both model fit, the r code is

- a. `anova(model1, model2)`
- b. `anova(model1); anova(model2)`
- c. `aov(model1); aov(model2)`
- d. `aov(model1, model2)`

**Question Number : 67 Question Id : 4165294952 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

For step wise linear regression R function is

- a. `STEP()`
- b. `stepwise()`
- c. `step()`
- d. `AICstep()`

**Question Number : 68 Question Id : 4165294953 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The Akaike Information Criterion is a measure of

- a. Association
- b. Goodness of fit
- c. Prediction power
- d. Precision

**Question Number : 69 Question Id : 4165294954 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Analysis of covariance model is used if

- a. Response is continuous, and predictors are categorical
- b. Both response and predictors are categorical
- c. Response is continuous, but predictors are mixed of continuous and categorical
- d. All are continuous

**Question Number : 70 Question Id : 4165294955 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To incorporate categorical variable in linear regression R convert them into

- a. Numeric variable
- b. Integer variable
- c. Dummy variable
- d. Character variable

**Question Number : 71 Question Id : 4165294956 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To fit linear mixed effects model one can use

- a. lmer() function from nlme library
- b. lmer() function lme4 library
- c. lme() function from lme4 library
- d. lmm() function lme4 library

**Question Number : 72 Question Id : 4165294957 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Suppose out of 760 birth, 360 babies belong to low birth weight categories, to estimate proportion low birth weight and its confidence interval

- a. prop(360,760)
- b. prop.test(360,760)
- c. prop.ci(360,760)
- d. p.ci(360,760)

**Question Number : 73 Question Id : 4165294958 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

A compact three way classified table can be generated by

- a. xtabs(~x+y+z)
- b. table(x,y,z)
- c. ftable(xtabs(~x+y+z))
- d. ftable(x,y,z)

**Question Number : 74 Question Id : 4165294959 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

$\phi$  contingency coefficient value 0.75 indicates

- a. strong positive association
- b. weak positive association
- c. moderate positive association
- d. mild positive association

**Question Number : 75 Question Id : 4165294960 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Contingency coefficient lies between

- a. 0 to 1
- b. -1 to 1
- c. 0.5 to 1
- d. 0 to 0.5

**Question Number : 76 Question Id : 4165294961 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Cramer's V is the measure of association between

- a. two measured variables
- b. two ordinal variables
- c. two nominal variables
- d. one nominal and one ordinal variable

**Question Number : 77 Question Id : 4165294962 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

suppose 45 cancer patients survived in 5 years out of 200. To test  $H_0: p=0.2$ , the R code is

- a. `prop.test(45,200)`
- b. `binom.test(45,200,p=0.2)`
- c. `binom.test(c(45,200),p=0.2)`
- d. `chisq.test(45,155)`

**Question Number : 78 Question Id : 4165294963 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The `ks.test()` function perform

- a. Kendall smith test
- b. Kolmogorov-Smirnov test
- c. Kendall-Smirnov test
- d. Kolmogorov-Smith test

**Question Number : 79 Question Id : 4165294964 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The function that derive empirical cumulative distribution function

- a. `ecdf()`
- b. `cdf()`
- c. `cum.df()`
- d. `df()`

**Question Number : 80 Question Id : 4165294965 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Friedman rank sum test is used for

- a. Nonparametric test for one way ANOVA
- b. Nonparametric test for two way ANOVA
- c. Nonparametric test for ANCOVA
- d. Nonparametric test for test of homogeneity

**Question Number : 81 Question Id : 4165294966 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

In build Newton's method can be found in the function

- a. `nlm()`
- b. `nlme()`
- c. `lme()`
- d. `lmer()`

**Question Number : 82 Question Id : 4165294967 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The in build Newton-Raphson method of finding root `maxNR()` can be found from the library

- a. `nlme`
- b. `mgcv`
- c. `glm`
- d. `maxLik`

**Question Number : 83 Question Id : 4165294968 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The newtonraphson() function belongs to

- a. Newtonraphson library
- b. gam library
- c. spuRs library
- d. nlme

**Question Number : 84 Question Id : 4165294969 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To apply optim() method one needs to define the

- a. Likelihood function
- b. Hessian matrix
- c. Gradient
- d. Design matrix

**Question Number : 85 Question Id : 4165294970 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Which one is not the part of the optimization methods supported by optim() function

- a. BFGS
- b. CG
- c. BCPE
- d. Brent

**Question Number : 86 Question Id : 4165294971 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Maximum likelihood Estimation function mle2() belongs to

- a. stats
- b. stats4
- c. nlme
- d. bbmle

**Question Number : 87 Question Id : 4165294972 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To derive 95% confidence interval of the parameters estimated by mle2(), one needs to derive

- a. Likelihood
- b. Profile likelihood
- c. Variance covariance matrix
- d. Hessian matrix

**Question Number : 88 Question Id : 4165294973 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

For a better visual estimate of whether the profile is quadratic, which one should be the correct option to the plot method of profile likelihood in in optimization problem

- a. absVal=FALSE
- b. absVal=TRUE
- c. absVal=0
- d. absVal=1

**Question Number : 89 Question Id : 4165294974 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

For likelihood ratio test on nested model, which function is commonly used on fitted model objects

- a. LR.test()
- b. anova
- c. aov
- d. likelihood()

**Question Number : 90 Question Id : 4165294975 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Which one is not true for R software

- a. R compiles and runs on Unix, Windows and MacOS.
- b. R is GNU General Public License software
- c. R has an effective data handling and storage facility
- d. C code cannot be compiled with in R

**Question Number : 91 Question Id : 4165294976 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Which one of the following R object does allow to store data of different classes but same length only?

- a. list
- b. matrix
- c. data frame
- d. array

**Question Number : 92 Question Id : 4165294977 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

What does R return in case of zero dividing zero

- a. NaN
- b. NA
- c. Inf
- d. 0

**Question Number : 93 Question Id : 4165294978 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Which one of the following function can convert multiple equal length vectors into a matrix

- a. list()
- b. array()
- c. cbind()
- d. apply()

**Question Number : 94 Question Id : 4165294979 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To access the name of the 5th person from a data frame(D), one can use the r code

- a. name[5]
- b. D\$name[5]
- c. D\$nam{5}
- d. name(5)

**Question Number : 95 Question Id : 4165294980 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**



To replace a binary variable(x) into a data frame (dat) by a variable having levels "Yes" and "No", the appropriate code is

- a. `dat$x=factor(dat$x,levels=c(0,1),labels=c("No","Yes"))`
- b. `levels(x)=c("No","Yes")`
- c. `dat$x=factor(dat$x, labels=c("No","Yes"))`
- d. `labels(x)= c("No","Yes")`

**Question Number : 96 Question Id : 4165294981 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Suppose you have three variables (x,y,z), you want to replace values of x by values of z and values of z by values x and y will be as is

- a. `x=z; z=x`
- b. `y=x; z=x; y=z`
- c. `a=x; x=z; z=a; y=y`
- d. `a=y; z=x; x=a; y=y`

**Question Number : 97 Question Id : 4165294982 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

If your variable 'age' has an entry '-' for missing, you want replace that by 'NA' to convert that into numeric variable

- a. `x[x="-"]=NA`
- b. `x[x=="-"]=NA`
- c. `x[x=="-"]='NA'`
- d. `x[x="-"]='NA'`

**Question Number : 98 Question Id : 4165294983 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Suppose you want to categorize your age variable by  $age \leq 18, 18 < age \leq 40$ , the correct code is

- a. `interval(age,c(18,40))`
- b. `cut(age,breaks=c(0,18,40,100),right=T)`
- c. `interval(age,breaks=c(0,18,40,100),right=T)`
- d. `cut(age,breaks=c(18,40),right=T)`

**Question Number : 99 Question Id : 4165294984 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Suppose you have two matrices  $A = (a_{ij})$  and  $B = (b_{ij})$  of same order, you want  $a_{56} = c_{56}$ ;  $C = BT$  (you suppose you have two matrices  $A = (a_{ij})$  and  $B = (b_{ij})$  of same order, you want  $a_{56} = c_{56}$ ;  $C = BT$  (you don't have C matrix). Choose the correct option

- a. `A[5,6]=B[5,6]`
- b. `A[5,6]=B[6,5]`
- c. `A[5,6]=C[5,6]`
- d. `A[5,6]=C[6,5]`

**Question Number : 100 Question Id : 4165294985 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Let a matrix  $A = \text{matrix}(c(56,78,45,34,23,12,67,53,47),3,3)$  & another matrix  $B = \text{matrix}(c(1:9),3,3)$ . To get  $A^{-1}B$ , choose the appropriate code

- a. `Inv(A)*B`
- b. `solve(A)%*%B`
- c. `inv(A)%*%B`
- d. `solve(A)*B`

**Question Number : 101 Question Id : 4165294986 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Suppose you have three matrices (A,B,C) of same order, you want to derive another matrix with element  $d_{ij} = (a_{ij}/b_{ji}) * c_{ij}$ , choose the right code

- a.  $D[i,j] = (A[i,j]/B[j,i]) * C[i,j]$
- b.  $D[i,j] = (A[i,j]/t(B)[i,j]) * C[i,j]$
- c.  $D = (A/t(B)) * C$
- d.  $D = (A/B) * C$

**Question Number : 102 Question Id : 4165294987 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To generate an identity matrix of order 5

- a. `I(5)`
- b. `diag(1,5)`
- c. `matrix(1,5,5)`
- d. `identity(5)`

**Question Number : 103 Question Id : 4165294988 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Let XX be the data frame of 5 numeric variables, The following R code output will be: `sqrt(colSums(sweep(XX,2,colMeans(XX),FUN="-")^2)/ncol(XX))`

- a. Means for each column
- b. SD for each column
- c. SD for each row
- d. Variance of each column

**Question Number : 104 Question Id : 4165294989 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

For a  $(n \times 1)$  vector X, the operation `X%*%t(X)` will return

- a. a scaler
- b. vector of order  $(1 \times n)$
- c. matrix of order  $(n \times n)$
- d. vector of order  $(1 \times n^2)$

**Question Number : 105 Question Id : 4165294990 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The dimension of a variance covariance matrix is

- a. No of observations  $\times$  no of parameters
- b. No of parameters  $\times$  no of parameters
- c. No of parameters  $\times$  no of observations
- d. No of observations  $\times$  no of observations

**Question Number : 106 Question Id : 4165294991 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Choose the correct answer

- a. Lower order graphics function work independent of higher order
- b. Lower order graphics function should always follow by higher order
- c. Higher order graphics function should always follow by lower order
- d. Higher order graphics function cannot work independent of lower order

**Question Number : 107 Question Id : 4165294992 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The function segments() is a

- a. Higher order graphics function can draw a line between two points
- b. Higher order graphics function can draw a cubic curve between two points
- c. Lower order graphics function can draw a line between two points
- d. Lower order graphics function can draw a cubic curve between two points

**Question Number : 108 Question Id : 4165294993 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Higher the choice of band width of a nonparametric density plot tends the curve to be

- a. negatively skewed
- b. positively skewed
- c. unchanged
- d. symmetric

**Question Number : 109 Question Id : 4165294994 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To draw a 45° diagonal line at (0,0) in scatter plot, one can write the lower order function

- a. abline(lm(y~x))
- b. abline(a=0,b=1)
- c. abline(h=0,v=0)
- d. abline(h=0,v=1)

**Question Number : 110 Question Id : 4165294995 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To change the name of levels of the group in a boxplot, the arguments should be used is

- a. labels
- b. names
- c. arg.names
- d. levels

**Question Number : 111 Question Id : 4165294996 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

In ggplot2 library, the variables in the data are mapped to the visual properties by

- a. geom()
- b. theme()
- c. aes()
- d. facet()

**Question Number : 112 Question Id : 4165294997 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To assign three different colour to three different levels of the grouping variable, the correct argument should be

- a. color=c(blue,red,green)
- b. col=c("blue","red","green")
- c. col=(blue,red,green)
- d. color=[blue,red,green]

**Question Number : 113 Question Id : 4165294998 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To draw a pie chart of gender, the right code would be

- a. `pie(sex)`
- b. `pie(table(sex))`
- c. `pie(levels(sex))`
- d. `pie(tab(sex))`

**Question Number : 114 Question Id : 4165294999 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To get the white background of a plot by `ggplot2`, a new layer can be added with function

- a. `facet()`
- b. `background()`
- c. `theme_bw()`
- d. `bw()`

**Question Number : 115 Question Id : 4165295000 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To change a symbol of a point used in a scatter plot, the argument should be used

- a. `mai`
- b. `type`
- c. `pch`
- d. `mar`

**Question Number : 116 Question Id : 4165295001 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The function `myfun=function(x){sd(x)/mean(s)}` returns

- a. IQR
- b. Mean deviation about mean
- c. Variance
- d. CV

**Question Number : 117 Question Id : 4165295002 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To derive mean deviation about median, the code would be

- a. `sum(abs(x-median(x,na.rm=T)),na.rm=T)/length(na.omit(x))`
- b. `sum(x-median(x,na.rm=T),na.rm=T)/length(na.omit(x))`
- c. `sum(abs(x-median(x,na.rm=T)),na.rm=T)/length(x)`
- d. `sum(x-median(x))/length(x)`

**Question Number : 118 Question Id : 4165295003 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To derive column median of a numeric data frame, only should use the function

- a. `apply()`
- b. `rapply()`
- c. `lapply()`
- d. `sapply()`

**Question Number : 119 Question Id : 4165295004 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

the function `strptime()` is

- a. to convert time from text
- b. date and time conversation function from and to character
- c. date and time conversation function from character
- d. date and time conversation function to character

**Question Number : 120 Question Id : 4165295005 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Suppose  $y$  is a date object in R, how do you extract year from that?

- a. `years(y)`
- b. `getyears(y)`
- c. `format(y,"%Y")`
- d. `format(y,"year")`

**Question Number : 121 Question Id : 4165295006 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To convert 4<sup>th</sup> Jan 2017 to a R date object

- a. `ISOdate(2017,1,4)`
- b. `ISODate(2017,1,4)`
- c. `as.Date(2017,1,4)`
- d. `format(c(2017,1,4),"%d/%m/%Y")`

**Question Number : 122 Question Id : 4165295007 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To get descriptive statistics stratified by multiple factors stored in a long format data frame, one can use

- a. `tapply()`
- b. `apply()`
- c. `aggregate()`
- d. `group_by()`

**Question Number : 123 Question Id : 4165295008 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To remove outlier dots from the boxplot one should set with `boxplot()`

- a. `notch = FALSE`
- b. `outline = FALSE`
- c. `plot=FALSE`
- d. `add=FALSE`

**Question Number : 124 Question Id : 4165295009 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

For a variable  $x$ :  $S=0$ ; for( $i$  in 1:100){ $S=S+x[i]$ };  $S/100$  will derive

- a. SD
- b. Median
- c. Mean
- d. Mode

**Question Number : 125 Question Id : 4165295010 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

For a variable  $x$ :  $S=SS=0$ ; for( $i$  in 1:100){ $S=S+x[i]$ ;  $SS=SS+x[i]^2$ };  $SS/100-(S/100)^2$  will derive

- a. Correlation coefficient
- b. Variance
- c. SD
- d. Mean deviation about mean

**Question Number : 126 Question Id : 4165295011 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Suppose you want to generate 10 consecutive dates from 1Jan2018, the code is

- a. `seq(01/01/2018,by="day",length=10)`
- b. `seq(01/01/2018, 01/01/2018,by="day")`
- c. `seq(01/01/2018, 01/01/2018,by=1)`
- d. `seq(ISOdate(2018,1,1),by='days',length=10)`

**Question Number : 127 Question Id : 4165295012 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

For a matrix A, the operation `sweep(A, 1, 5)` means

- a. Add 5 to diagonal elements of A
- b. Subtract 5 from each elements of A
- c. Subtract 5 from 1 column of A
- d. Add 5 to 1 row of A

**Question Number : 128 Question Id : 4165295013 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Suppose you want to merge two data sets of different number of rows [ $nrow(A) < nrow(B)$ ] by common ID variable. If you want in your merged data set all matched rows between A and B plus all mismatched rows of A, then the code is

- a. `merge(A,B,by.x="ID.A",by.y="ID.B",all=TRUE)`
- b. `merge(A,B,by.x="ID.A",by.y="ID.B",all=FALSE)`
- c. `merge(A,B,by.x="ID.A",by.y="ID.B",all.y=TRUE)`
- d. `merge(A,B,by.x="ID.A",by.y="ID.B",all.x=TRUE)`

**Question Number : 129 Question Id : 4165295014 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

For a design matrix X and response  $y(y = X\beta + e)$ , the operation `solve(t(X)%*%X)%*%t(X)%*%y` returns

- a. Regression error variance
- b. Regression coefficients
- c. Regression residuals
- d. Standard errors of regression coefficients

**Question Number : 130 Question Id : 4165295015 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Let `"%ABC%" = function(x,y){(x-y)^2/2}`, then `25%ABC%23=?`

- a. 2
- b. 5
- c. 3
- d. 4

**Question Number : 131 Question Id : 4165295016 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

For chi-square test of independence, the function to be used

- a. `tchisquare()`
- b. `chisquare.test()`
- c. `chisq.test()`
- d. `chisqtest()`

**Question Number : 132 Question Id : 4165295017 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

xtabs function retrun

- a. cross table, chisquare test and OR
- b. cross table, chisquare test
- c. cross table only
- d. none of the above

**Question Number : 133 Question Id : 4165295018 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

In the function `assocstats(x)` in package `vcd`, `x` has to be

- a. Continuous
- b. Discrete
- c. Categorical
- d. Ordinal only

**Question Number : 134 Question Id : 4165295019 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

ANOVA is analysis of variance where the outcome variable is continuous and the predictor variables are

- a. all continuous
- b. all categorical
- c. mixed of continuous and categorical
- d. mixed of ordinal and nominal

**Question Number : 135 Question Id : 4165295020 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To perform a right tail t-test, one has to set the argument of `t.test()` as

- a. `right=TRUE`
- b. `righttail=TRUE`
- c. `left=FALSE`
- d. `alternative="greater"`

**Question Number : 136 Question Id : 4165295021 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The observed probability of type-I is return by any testing of hypothesis function of R in terms of

- a. SE
- b. Test statistics
- c. P-value
- d. 95% CI

**Question Number : 137 Question Id : 4165295022 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

A test is statistically significance at 1% level if p-value derived by R is

- a.  $\leq 0.05$
- b.  $< 0.05$
- c.  $< 0.01$
- d.  $\leq 0.01$

**Question Number : 138 Question Id : 4165295023 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The code `boxplot(cut(mydata$x, breaks = c(0,70,75,90)))` plots

- a. One box-whisker
- b. Two box-whiskers
- c. Three box-whiskers
- d. Four box-whiskers

**Question Number : 139 Question Id : 4165295024 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To extract variance covariance matrix for regression coefficient from lm object, the R function is

- a. VAR()
- b. COV()
- c. vcov()
- d. varcov()

**Question Number : 140 Question Id : 4165295025 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Q-Q plot helps to check the one of following assumption

- a. Response is measured independently
- b. Response is linear on predictors
- c. Variance of response is constant for any values of predictors
- d. Response is normally distributed

**Question Number : 141 Question Id : 4165295026 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

To get a Q-Q plot with Q-Q line one can use the code

- a. qqnorm(y); qqline(y)
- b. qqnorm(y)
- c. qqline(y)
- d. qqnormline(y)

**Question Number : 142 Question Id : 4165295027 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Let  $fit = lm(y \sim x)$ , to get predicted values of y with its 95% CI, one can use the code

- a. predict(fit, interval = "confidence")
- b. predict(fit, interval = TRUE)
- c. predict(fit, se.fit = TRUE)
- d. predict(fit, ci = TRUE)

**Question Number : 143 Question Id : 4165295028 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

If the residuals plot shows pattern, this means

- a. Data violated normality assumption
- b. Data violated independent assumption
- c. Data violated equal variance assumption
- d. None of the above

**Question Number : 144 Question Id : 4165295029 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

For multiple linear regression object 'fit', prediction of response by individual predictor keeping other remain at constant level, one can use the code

- a. predict(fit)
- b. predict(fit, type = "terms")
- c. predict(fit, term = "all")
- d. predict(fit, predictor = "all")

**Question Number : 145 Question Id : 4165295030 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**



In multiple linear regression coefficient of determination is assessed by

- a.  $R^2$
- b.  $\text{Adj } R^2$
- c.  $\text{Partial } R^2$
- d. AIC

**Question Number : 146 Question Id : 4165295031 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Minimum power to be set to determine sample size is

- a. 80%
- b. 90%
- c. 70%
- d. 95%

**Question Number : 147 Question Id : 4165295032 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Which one from the summary of lm object can be treated as assumed constant variance of response

- a. F-statistic
- b. Residual standard error
- c. Std. Error
- d. None of the above

**Question Number : 148 Question Id : 4165295033 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

In linear regression with categorical predictor of 4 levels, algorithm include

- a. 4 dummy variables
- b. 3 dummy variables
- c. 2 dummy variables
- d. No dummy variable

**Question Number : 149 Question Id : 4165295034 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

Suppose for an interaction model with two predictors ( $X_1$  &  $X_2$ ) where  $X_1$  is continuous and  $X_2 = 1$  if yes; 0 otherwise. The regression coefficient of  $X_1, X_2$  and  $(X_1 X_2)$  are  $b_1, b_2$  and  $b_3$ . The true regression coefficient of  $X_1$  is

- a.  $b_1$
- b.  $b_1 + b_2$
- c.  $b_1 + b_2 X_2$
- d. None of the above

**Question Number : 150 Question Id : 4165295035 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical**

**Correct Marks : 1 Wrong Marks : 0**

The interpretation of slope is

- a. Average response at  $x = 1$
- b. Change of response at  $x = 1$
- c. Change of response per unit change in  $x$
- d. None of the above