Specialization: Business Analytics

Course Code : 205 BA   Course Name: Business Analytics using R Programming

MCQ

Unit 1: Business Analytics Basics

Sr No	Question	Answer
1	**Which of these measures are used to analyse the central tendency of data?** a. Mean and Normal Distribution b. Mean, Median and Mode c. Mode, Alpha & Range d. Standard Deviation, Range and Mean	B
2	Five numbers are given: (5, 10, 15, 5, 15). Now, what would be the sum of deviations of individual data points from their mean? A) 10 B)25 C) 50 D) 0	D
3	A test is administered annually. The test has a mean score of 150 and a standard deviation of 20. If Ravi's z-score is 1.50, what was his score on the test? A) 180 B) 130 C) 30 D) 150	A
4	Business intelligence (BI) is a broad category of application programs which includes _____ a) Decision support b) Data mining c) OLAP d) All of the mentioned	A
5	Point out the correct statement. a) OLAP is an umbrella term that refers to an assortment of software applications for analyzing an organization's raw data for intelligent decision making b) Business intelligence equips enterprises to gain business advantage from data c) BI makes an organization agile thereby giving it a lower edge in today's evolving market condition d) None of the mentioned	A
6	BI can catalyze a business's success in terms of _____ a) Distinguish the products and services that drive revenues b) Rank customers and locations based on profitability c) Ranks customers and locations based on probability d) All of the mentioned	D
7	Which of the following areas are affected by BI? a) Revenue b) CRM c) Sales d) All of the mentioned	D
8	Business intelligence (BI) is a broad category of application programs which includes _____ a) Decision support b) Data mining c) OLAP d) All of the mentioned	D

9	Which of the following measures of central tendency will always change if a single value in the data changes? A) Mean B) Median C) Mode D) All of these	A
10	Strong assessment items are made up of five elements: a) Standard b) Stimulus c) Stem d) Key	A
11	A good question is----------------It focuses on recall of only the material covered in your lesson and aligns well with the overall learning objectives a) relevant. b) clear c) concise d) purpose	B
12	A good question is framed in a ----------, easily understandable language, without any vagueness. Students should understand what is wanted from the question even when they don't know the answer to it. a) clear b) relevant c) concise d) purpose	A
13	A good question is usually crisp and ------- --. It omits any unnecessary information that requires students to spend time understanding it correctly. The idea is not to trick learners but assess their knowledge. a) concise b) clear c) relevant d) purpose	A
14	1.____programming language is a dialect of S. a) B b) C c) R d) K	C
15	Point out the WRONG statement? a) Early versions of the S language contain functions for statistical modeling b) The book Programming with Data by John Chambers documents S version of the language c) In 1993 Bell Labs gave StatSci (later Insightful Corp.) an exclusive license to develop and sell the S language d) The book Programming with Data by IBM documents S version of the language	C
16	In 1991, R was created by Ross Ihaka and Robert Gentleman in the Department of Statistics at the University of a) John Hopkins b) California c) Harvard d) Auckland	D

17	Point out the wrong statement? a) R is a language for data analysis and graphics b) K is language for statistical modelling and graphics c) One key limitation of the S language was that it was only available in a commercial package, S-PLUS d) C is a language for data and graphics	A
18	Business analytics results in which of these? a. Evidence Based Decisions b. Data Driven Decisions c. Better Decisions d. All of these are correct	D
19	Which one of the following is not a type of Business Analytics? a. Descriptive Analytics b. Diagnostic Analytics c. Predictive Analytics d. Performance Analytics	D
20	What will be the output of the following R code snippet? > **paste**("a", "b", se = ":") a) "a+b" b) "a=b" c) "a b :" d) none of the mentioned	D
21	Point out the correct statement? a) In R, a function is an object which has the mode function b) R interpreter is able to pass control to the function, along with arguments that may be necessary for the function to accomplish the actions that are desired c) Functions are also often written when code must be shared with others or the public d) All of the mentioned	D
22	The_____function returns a list of all the formal arguments of a function. a) formals() b) funct() c) formal() d) fun()	A
23	What will be the output of the following R code snippet? > f <- **function**(num = 1) { + hello <- "Hello, world!\n"	A

| | ```
+        for(i in seq_len(num)) {
+            cat(hello)
+        }
+        chars <- nchar(hello) * num
+        chars
+ }
> f()

 a)
 Hello, world!
 [1] 14
 b) Hello,
 world![1] 15
 c) Hello,
 world![1] 16
 d) Error
``` | |
24	Point out the wrong statement? a) A formal argument can be a symbol, a statement of the form 'symbol = expression', or the special formal argument b) The first component of the function declaration is the keyword function c) The value returned by the call to function is not a function d) Functions are also often written when code must be shared with others or the public	**A**
25	You can check to see whether an R object is NULL with the_____function. a) is.null() b) is.nullobj() c) null() d) as.nullobj()	**A**
26	Which of the following code will print NULL? a) > args(paste) b) > arg(paste) c) > args(pastebin) d) > arg(bin)	**A**

27	What will be the output of the following R code snippet? > **paste**("a", "b", sep = ":") a) "a+b" b) "a=b" c) "a:b" d) a*b	**A**
28	What will be the output of the following R code snippet? > f <- **function**(a, b) { + **print**(a) + **print**(b) + } > f(45) a) 32 b) 42 c) 52 d) 45	**A**
29	What will be the output of the following R code snippet? > f <- **function**(a, b) { + a^2 + } > f(2) a) 4 b) 3 c) 2 d) 5	**A**

30	Which of the following is a base package for R language?
	   a) util
	   b) lang
	   c) tools
	   d) All of the above
		C

31	R comes with a _____ to help you optimize your code and improve its performance.
	   a) Debugger
	   b) Monitor
	   c) Profiler
	   d) None of the above
		A

32	debug() flags a function for_____mode in R mode.
	   a) debug
	   b) run
	   c) compile
	   d) None of the above
		B

33	_____ suspends the execution of a function wherever it is called and puts the function in debug mode a) recover() b) browser() c) Both of the above d) None of the above	**C**
34	A matrix is ___ dimensionsinal rectangular data set? a) 5 b) 4 c) 3 d) 2	**D**
35	The _____ function takes a vector or other objects and splits it into groups determined by a factor or list of factors. a) apply() b) split() c) isplit() d) mapply()	**B**
36	lapply function takes ___ arguments in R language a) 1 b) 3 c) 4 d) 5	**C**
37	___ is used to apply a function over subsets of a vector a) apply() b) lapply() c) mapply() d) tapply()	**D**
38	_____ applies a function over the margins of an array a) apply() b) lapply() c) tapply() d) mapply()	**A**

39	___function is same as lapply() in R b) apply() c) lapply() d) sapply() e) tapply()	C
40	___loop over a list and evaluate a function on each element a) apply() b) lapply() c) sapply() d) tapply()	A
41	___is proprietary tool for predictive analytics. a) R b) SAS c) SSAS d) SPSS	B
42	Data frames can be converted to a matrix by calling data.___ a) matr() b) mat() c) matrix() d) None of the above	C
43	Which of the following method make a vector of repeated values? a) rep() b) data() c) view() d) None of the above	B
44	R objects can have attributes, which are like___for the object a) metadata b) features c) expressions d) None of the above	A
45	Attributes of an object (if any) can be accessed using the___function. a) objects() b) attrib() c) attributes() d) None of the above	C

46	_____involves predicting a response with meaningful magnitude, such as quantity sold, stock price, or return on investment. a) Regression b) Clustering c) Summarization	A
47	_____provides needed string operators in R a) str b) forcast c) stringr	C
48	_____splits a data frame and results in an array (hence the da). Hopefully, you're getting the idea here. a) apply b) daply c) stats	B
49	System.time function returns an object of class_____which contains two useful bits of information. a) debug_time b) procedure_time c) proc_time	C
50	Which of the following will start the R program? a) $ R b) & R c) Rb d) None of the above	A

	Unit 2 : Analytical Decision Making	
1	The third step in decision making process is a linear predictions b dependent predictions c making predictions d independent predictions	**C**
2	The decision making step, which consists of organization goals, predicting alternatives and communicating goals is called a organization b alternation c planning d valuing	**C**
3	The fourth step in decision making process is a linear correlation b making decisions c implement decisions d evaluate performance	**B**
4	The costs that behaves as irrelevant costs in process of decision making are classified as a past costs b future costs c expected costs d sunk costs	**A**
5	Which of these is not a topic covered in a typical Business Analyst Aptitude Test? a. Analytical Thinking c. Data Interpretation b. Listening Skills d. Risk Management	**D**

6	If the test should be 30 minutes, Analytical Thinking is taken in how many minutes?
	      a. 5 c. 10
	      b. 7 d. 15
		C

7	Primary objective of a business analyst is to help businesses implement a. Business systems b. Business solutions c. Technology systems d. Technology solutions	**B**
8	Which business professional performs cost-benefit analyses of existing and potential customers a) Marketer b) Financial Analyst c) Business Analyst d) Sales Representative	**C**

9	1. A Use Case is a set of steps, typically defining interactions between a role, True of False a. True b. False	A
10	Any fact that the solution can assume to be true when the use case begins is what? a. A win b. A Failure c. A success d. A Precondition	C
11	A State Diagram is used for what? a. Which Events cause a transition between states b. Which events cause a success between states c. Allowable behaviour d, All	D
12	A Solution Requirement is comprised of two types of requirements what are they? a, Functional b. Hard c. Existing d. Non-Functional	A

13	Which of the following is used for Statistical analysis in R language?
	  a) Studio
	  b) RStudio
	  c) Heck
	  d) None of the above
		B

14	R functionality is divided into a number of _____
	  a) Packages
	  b) Functions
	  c) Domains
	  d) None of the above
		A

15	Which of the following is an example of vectorized operation as far as subtraction is concerned?
	> x <- 1:4
	> y <- 6:9
	  a) x+y
	  b) x-y
	  c) x/y
	  d) x*y
		B

Prof. Ujjval More

16	What would be the output of the following code?
	> x <- 1:4
	> y <- 6:9
	> z <- x + y
	> z
	    a) 7 9 11 13
	    b) 7 9 11 13 14
	    c) 9 11 13
	    d) Null
		A

17	What would be the output of the following code?
	> x <- 1:4
	> x > 2
	    a) FALSE FALSE TRUE TRUE
	    b) 1 2 3 4
	    c) 1 2 3 4 5
		A

18	What would be the value of the following expression?
	log(-1)
	    a) Warning in log(-1): NaNs produced
	    b) 1
	    c) Null
	    d) 0
		A

19	What will be the output of the following code?
	> g <- function(x) {
	+ a <- 3
	+ x+a+y
	c
	+ ## 'y' is a free variable
	+ }
	> g(2)
	
	    a) 8
	    b) 9
	    c) 42
	    d) Error
		D

20	What will be the output of the following code?	C
	function(p) {	
	params[!fixed] <- p	
	mu <- params[1]	
	sigma <- params[2]	
	## Calculate the Normal density	
	a <- -0.5*length(data)*log(2*pi*sigma^2)	
	b <- -0.5*sum((data-mu)^2) / (sigma^2)	
	-(a + b) } > ls(environment(nLL))	

	a) "data" "fixed" "param"
	    b) "data" "variable" "params"
	    c) "data" "fixed" "params"
	    d) None of the above	

21	Which of the following is a principle of analytic graphics?
	    a) Don't plot more than two variables at at time
	    b) Make judicious use of color in your scatterplots
	    c) Show box plots (univariate summaries)
	    d) Show causality, mechanism, explanation
		D

22	R is an_____programming language?
	    a) Closed source
	    b) GPL
	    c) Open source
	    d) Definite sourc
		C

23	Who developed R?
	    a) Dennis Ritchie
	    b) John Chambers
	    c) Bjarne Stroustrup
	    d) Bill Gates
		A

24	R was named partly after the first names of____R authors? a) One b) Two c) Three d) Four	**B**
25	Packages are useful in collecting sets into a____unit ? a) Single b) Multiple	**C**
26	Many quantitative analysts use R as their____tool? a) Leading tool b) Programming tool c) Both the above d) None of the above	**D**
27	Predictive analysis is the branch of_____analysis? a) Advanced b) Core c) Both the above d) None of the above	**B**
28	_____is used to make predictions about unknown future events? a) Descriptive analysis b) Predicitive analysis c) Both the above d) None of the above	**C**
29	How many steps does the predictive analysis process contained? a) 5 b) 6 c) 7 d) 8	D
30	Descriptive analysis tell about_____? a) Past b) Present c) Future	**A**
31	How many types of R objects are present in R data type? a) 4 b) 5 c) 6 d) 7	**C**

32	How many types of data types are present in R? a) 4 b) 5 c) 6 d) 7	A
33	Which of the following is a primary tool for debugging? a) debug() b) trace() c) browser() d) None of the above	B
34	Which function is used to create the vector with more than one element? a) Library() b) plot() c) c() d) par()	C
35	In R every operation has a_____call? a) System b) Function c) None of the above d) Both of the above	A
36	The_____in R is a vector. a) Basic data structure b) Basic datatypes c) Both d) None of the above	b
37	R is an interpreted language so it can access through_____? a) Disk operating system b) User interface operating system c) Operating system d) Command line interpreter	C
38	Vectors come in two parts____and____. a) Atomic vectors and matrix b) Atomic vectors and array c) Atomic vectors and list d) None of the above	A
39	How many types of atomic vectors are present? a) 3 b) 4 c) 5 d) 6	C

40	How many types of vertices functions are persent? a) 1 b) 2 c) 3 d) 4	**B**
41	_____and_____are types of matrices functions? a) Apply and sapply b) Apply and lapply c) Both	**C**
42	How many control statements are present in R? a) 6 b) 7 c) 8 d) 9	**A**
43	Which of the following finds the maximum value in the vector x, exclude missing values a) rm(x) b) all(x) c) max(x, na.rm=TRUE) d) x%in%y	**b**
44	Which of the following sort dataframe by the order of the elements in B a) a.x[rev(order(x$B)),] b) b.x[ordersort(x$B, c) c.x[order(x$B),] d) None of the above	**A**
45	_____initiates an infinite loop right from the start. a) Never b) Repeat c) Break d) Set	**B**
46	_____is used to skip an iteration of a loop. a) Next b) Skip c) Group d) None of the above	**A**

47	_____ programming language is a dialect of S. a) B b) C c) D d) S	A
48	In 1991, R was created by Ross Ihaka and Robert Gentleman in the Department of Statistics at the University of_____.Auckland a) Harvard b) California c) John Hopkins d) Dannies Rithie	A
49	Finally, in_____R version 1.0.0 was released to the public. a) 2000 b) 2005 c) 2010 **d)** 2012	D
50	R is technically much closer to the Scheme language than it is to the original _____ language. a) B b) S c) C d) C++	C

Unit-3 : **Fundamentals of R**		
1	They primary R system is available from the _____ a) CRAN b) CRWO c) GNU d) CRDO	C
2	Point out the wrong statement? a) Key feature of R was that its syntax is very similar to S b) R runs only on Windows computing platform and operating system c) R has been reported to be running on modern tablets, phones, PDAs, and game consoles d) R functionality is divided into a number of Packages	D
3	R functionality is divided into a number of _____ a) Packages b) Functions c) Domains d) Classes	A
4	Which Package contains most fundamental functions to run R? a) root b) child c) base d) parent	A
5	Which language is best for the statistical environment? a) C b) R c) Java d) Python	B
6	In order to use the R-related functionality in Dundas BI, you must have access to an existing _____ a) Console b) Terminal c) Packages d) R serve	D
7	The open source_____software is available for Unix, Linux, and Windows platforms. a) Rserve b) BServe c) CServe d) Dserve	A

8	Modification in Dundas BI is done _____ a) Directly b) Indirectly c) Need access to Server d) Not known	A
9	Is It possible to inspect the source code of R? a) Yes b) No c) Can't say d) Some times	A
10	_____function is used to watch for all available packages in library. a) lib() b) fun.lib() c) libr() d) library()	D
11	The longer programs are called _____ a) Files b) Structures c) Scripts d) Data	D
12	Scripts will run on _____ a) Script Editors b) Console c) Terminal d) GCC Compiler	A

13	What will be the output of the following R function?
	ab <- list(1, 2, 3, "X", "Y", "Z")
	dim(ab) <- c(3,2)
	print(ab)
	   a. 123
	      Xyz
	   b. Error
	   c. Xyz123
	   d. 123xyz
		A

14	What is the meaning of the following R function? x <- c(4, 5, 1, 2, 3, 3, 4, 4, 5, 6) x <- as.factor(x) a) x becomes a factor b) x is a factor c) x does not exist d) x is not a vector	A

15	What is the meaning of the following R function? print(sqrt(2)) a) 1.414314 b) 1.414214 c) Error d) 14.1414	**B**
16	What will be the output of the following R function? d <- date() a) Prints todays date b) Prints some date c) Prints exact present time and date d) Error	**C**
17	Which of the following commands will correctly read the above csv file with 5 rows in a dataframe? A) csv('Dataframe.csv') B) csv('Dataframe.csv',header=TRUE) C) dataframe('Dataframe.csv') D) csv2('Dataframe.csv',header=FALSE,sep=',')	**B**

18	R functionality is divided into a number of _____
	   a) Packages
	   b) Functions
	   c) Domains
	   d) None of the above
		A

19	The iris dataset has different species of flowers such as Setosa, Versicolor and Virginica with their sepal length. Now, we want to understand the distribution of sepal length across all the species of flowers. One way to do this is to visualise this relation through the graph shown below. Which function can be used to produce the graph shown above? A) xyplot() B) stripplot() C) barchart() D) bwplot()	**B**
20	**Which of the following command will help us to replace every instance of Delhi with Delhi_NCR in the following character vector?** C<-c("Delhi is","a great city.","Delhi is also","the capital of India.") A) gsub("Delhi","Delhi_NCR",C) B) sub("Delhi","Delhi_NCR",C) C) Both of the above D) None of the above	**C**

21	Which of the following commands will split the plotting window into 4 X 3 windows and where the plots enter the window column wise. A) par(split=c(4,3)) B) par(mfcol=c(4,3)) C) par(mfrow=c(4,3)) D) par(col=c(4,3))	B
22	Which of the following command will help us to rename the second column in a dataframe named "table" from alpha to beta? A) colnames(table)[2]='beta' B) colnames(table)[which(colnames=='alpha')]='beta' C) setnames(table,'alpha','beta') D) All of the above	D
23	A majority of work in R uses systems internal memory and with large datasets, situations may arise when the R workspace cannot hold all the R objects in memory. So removing the unused objects is one of the solution. Which of the following command will remove an R object / variable named "santa" from the workspace? A) remove(santa) B) rm(santa) C) Both D) None	C
24	"dplyr" is one of the most popular package used in R for manipulating data and it contains 5 core functions to handle data. Which of the following is not one of the core functions of dplyr package? A) select() B) filter() C) arrange() D) summary()	D
25	Sometimes as a Data Scientist working on textual data we come across instances where we find multiple occurrences of a word which is unwanted. Below is one such string. A<-c("I can use because thrice in a sentence because because is a special word.") A) gsub("because","since",A) B) sub("because","since",A C) regexec("because","since",A) D) None of the above	C
26	Imagine a dataframe created through the following code. Which of the following command will help us remove the duplicate rows based on both the columns? A) df[!duplicated(df),] B) unique(df) C) dplyr::distinct(df) D) All of the above	A

27	Which language is best for the statistical environment? a) C b) R c) Java d) Python	**B**
28	R has many functions regarding _____ a) Statistics, Biotechnology b) Probability, Microbiology c) Distributions, Physics d) Statistics, Probability, Distributions	**D**
29	A _____ is a variable that holds one value at a time. a) Scalar variable b) Duplex c) High d) Vector	**A**
30	Files containing R scripts ends with extension _____ a) .S b) .R c) .Rp d) .SP	**B**
31	_____ will divert all subsequent output from the console to an external file. a) sink b) div c) exp d) exc	**A**
32	The entities that R creates and manipulates are known as _____ a) objects b) task c) container d) packages	**A**
33	Collection of objects currently stored in R is called as _____ a) package b) workspace c) list d) task	**B**
34	What will be the output of the following R function? d <- date() d a) Prints todays date b) Prints some date c) Prints exact present time and date d) Error	**C**

35	What will be the output of the following R function? nchar() a) no. of characters b) first 5 characters c) last 5 characters d) Does not exist	**A**
36	What will be the output of the following R function? Sys.Date() a) Tomorrow date b) Present date c) Some date d) Yesterday date	**B**
37	R has how many atomic classes of objects? a) 1 b) 2 c) 3 d) 5	**D**
38	Numbers in R are generally treated as _____ precision real numbers. a) single b) double c) real d) imaginary	**B**
39	Which of the following can be considered as object attribute? a) dimensions b) class c) length d) all of the mentioned	**D**
40	_____ function returns a vector of the same size as x with the elements arranged in increasing order. a) sort() b) orderasc() c) orderby() d) sequence()	**A**
41	What will be the output of the following R code? > sqrt(-17) a) -4.02 b) 4.02 c) NaN d) 3.67	**C**
42	Which of the following is used for reading in saved workspaces? a) unserialize b) load c) get d) set	**B**

43	_____ is used for outputting a textual representation of an R object. a) dput b) dump c) dget d) dset	**A**
44	Which of the following argument denotes if the file has a header line? a) header b) sep c) file d) footer	**A**
45	Which of the following function is identical to read .table? a) read.csv b) read.data c) read.tab d) read.del	**A**
46	Point out the wrong statement? a) The grammar of the language determines whether an expression is complete or not b) The <- symbol is the assignment operator in R c) The ## character indicates a comment d) R does not support multi-line comments or comment blocks	**C**
47	Point out the wrong statement? a) : operator is used to create integer sequences b) The numbers in the square brackets are part of the vector itself c) There is a difference between the actual R object and the manner in which that R object is printed to the console d) Files containing R scripts ends with extension .R	**B**
48	What will be the output of the following R function? paste("Everybody", "is", "a" , "warrior") a) "Everybody", "is", "a" , "warrior" b) Everybody is a warrior c) Everybody", "is", "a" , "warrior d) "Everybody is a warrior"	**D**
49	_____ hosts many add-on packages that can be used to extend the functionality of R. a) CRAN b) GNU c) R studio d) 450	**A**
50	R runs on the _____ operating system. a) Linux b) Windows c) Ubuntu d) Any operating system	**D**

	Unit - 4 : Data types & Data Structures in R	
1	Accessing elements is achieved through a process called _____ a) Indexing b) Outdexing c) Highlighting d) Scrapping	A
2	Which are indexed by either row or column using a specific name or number? a) Datasets b) Data frames c) Data d) Functions	B
3	What should we use to access elements with a value greater than five? a) Subsetting commands b) Use functions c) Packages d) Interfaces	A
4	Lists can be created using the _____ function. a) Matrix.li b) Matrix.lists c) Lists.matric d) List	D
5	R is an_____programming language? a) Closed source b) GPL c) Open source d) Definite source	**C**

6	.Solve
	    varx<-23, 34->vary
	    print(varx+vary)
	    a. 57
	    b. 2334
	    c. 3423
	    d. 66
		A

7	find the output
	varx<-23, 34->vary
	print(varx == vary)
	    a. True
	    b. False
	    c. None of the above
	    d. Error
		B

8	Below, we have represented six data points on a scale where vertical lines on scale represent unit. Which of the following line represents the mean of the given data points, where the scale is divided into same units? A) A B) B C) C D) D	**C**
9	If a positively skewed distribution has a median of 50, which of the following statement is true? A) Mean is greater than 50 B) Mean is less than 50	**E**

10	Which of the following is a possible value for the median of the below distribution? A) 32 B) 26 C) 17 D) 40	**B**
11	Which of the following statements are true about Bessels Correction while calculating a sample standard deviation? Bessels correction is always done when we perform any operation on a sample data. Bessels correction is used when we are trying to estimate population standard deviation from the sample. Bessels corrected standard deviation is less biased. A) Only 2 B) Only 3 C) Both 2 and 3 D) Both 1 and 3	**C**
12	If the variance of a dataset is correctly computed with the formula using (n – 1) in the denominator, which of the following option is true? A) Dataset is a sample B) Dataset is a population C) Dataset could be either a sample or a population D) Dataset is from a census E) None of the above	**A**
13	What would be the critical values of Z for 98% confidence interval for a two-tailed test ? A) +/- 2.33 B) +/- 1.96 C) +/- 1.64 D) +/- 2.55	**A**
14	Studies show that listening to music while studying can improve your memory. To demonstrate this, a researcher obtains a sample of 36 college students and gives them a standard memory test while they listen to some background music. Under normal circumstances (without music), the mean score obtained was 25 and standard deviation is 6. The mean score for the sample after the experiment (i.e With music) is 28. What is the null hypothesis in this case? A) Listening to music while studying will not impact memory. B) Listening to music while studying may worsen memory. C) Listening to music while studying may improve memory. D) Listening to music while studying will not improve memory but can make it worse.	**D**

15	Studies show that listening to music while studying can improve your memory. To demonstrate this, a researcher obtains a sample of 36 college students and gives them a standard memory test while they listen to some background music. Under normal circumstances (without music), the mean score obtained was 25 and standard deviation is 6. The mean score for the sample after the experiment (i.e With music) is28. What would be the Type I error? A) Concluding that listening to music while studying improves memory, and it's right. B) Concluding that listening to music while studying improves memory when it actually doesn't. C) Concluding that listening to music while studying does not improve memory but it does. D) None of the above	**B**
16	Studies show that listening to music while studying can improve your memory. To demonstrate this, a researcher obtains a sample of 36 college students and gives them a standard memory test while they listen to some background music. Under normal circumstances (without music), the mean score obtained was 25 and standard deviation is 6. The mean score for the sample after the experiment (i.e With music) is After performing the Z-test, what can we conclude____? A) Listening to music does not improve memory. B)Listening to music significantly improves memory at p C) The information is insufficient for any conclusion. D) None of the above	**B**
17	A researcher concludes from his analysis that a placebo cures AIDS. What type of error is he making? A) Type 1 error B) Type 2 error C) None of these. The researcher is not making an error. D) Cannot be determined	**D**
18	What happens to the confidence interval when we introduce some outliers to the data? A) Confidence interval is robust to outliers B) Confidence interval will increase with the introduction of outliers. C) Confidence interval will decrease with the introduction of outliers. D) We cannot determine the confidence interval in this case	**B**
19	A medical doctor wants to reduce blood sugar level of all his patients by altering their diet. He finds that the mean sugar level of all patients is 180 with a standard deviation of 18. Nine of his patients start dieting and the mean of the sample is observed to 175. Now, he is considering to recommend all his patients to go on a diet. Note: He calculates 99% confidence interval. What is the standard error of the mean? A) 9 B) 6 C) 7.5 D) 18	**B**

20	--------------is function in R to get number of observation in a data frame a) n( ) b) ncol( ) c) nobs( ) d) nrow( )	**D**

21	A key property of vectors in R language is that
	   a. A vector cannot have attributes like dimensions
	   b. Elements of a vector can be of different classes
	   c. Elements of a vector can only be a character or numeric
	   d. Elements of a vector all must be of the same class
		D

22	The definition of free software consists of four freedoms (freedoms 0 through 3). Which of the following is NOT one of the freedoms that are part of the definition?
	
	   a. The freedom to study how the program works, and adapt it to your needs.
	   b. The freedom to improve the program, and release your improvements to the public, so that the whole community benefits.
	   c. The freedom to run the program, for any purpose.
	   d. The freedom to sell the software for any price.
		D

23	Point out the correct statement :
	
	   a) Blocks are evaluated until a new line is entered after the closing brace
	   b) Single statements are evaluated when a new line is typed at the start of the syntactically complete statement
	   c) The if/else statement conditionally evaluates two statements
	   d) All of the mentioned
		C

24	Which will be the output of following code ?
	x - 3
	 switch(6, 2+2, mean(1:10), rnorm(5))
	
	   a) 10
	   b) 1
	   c) NULL
	   d) All of the mentioned
		C

25	_____is used to continue an iteration of a loop.
	
	   a) next
	   b) skip
	   c) group
	   d) All of the mentioned
		A

26	Point out the correct statement : a) R has a number of ways to indicate to you that something's not right b) Executing any function in R may result in the condition "condition" is a generic concept for indicating that something unexpected has occurred d) All of the mentioned	D
27	. Which of the following is primary tool for debugging? a) debug() b) trace() c) browser() d) All of the mentioned	A
28	Point out the correct statement : a) Vectorizing the function can be accomplished easily with the Vectorize() function b) There are different levels of indication that can be used, ranging from mere notification to fatal error c) Vectorizing the function can be accomplished easily with the vector() function d) None of the mentioned	A
29	Functions are defined using the_____directive and are stored as R objects a) function() b) funct() c) functions() d) All of the mentioned	A
30	The_____function returns a list of all the formal arguments of a function a) formals() b) funct() c) formal() d) All of the mentioned	A
31	Which of the following is multivariate version of apply? a) apply() b) lapply() c) sapply() d) mapply()	D

32	Point out the correct statement : a) split() takes elements of the list and passes them as the first argument of the function you are applying b) You can use tsplit() to evaluate a function single time each with a same Argument c) Sequence of operations is sometimes referred to as "map-reduce" d) None of the mentioned	C
33	A function, together with an environment, makes up what is called a _____ closure. a) formal b) function c) reflective d) All of the mentioned	B
34	The_____function is used to plot negative likelihood. a) plot() b) graph() c) graph.plot() d) None of the mentioned	A
35	Which of the following is apply function in R? a) apply() b) tapply() c) fapply() d) rapply()	B
36	Point out the wrong statement? a) Functions in R are "second class objects" b) The writing of a function allows a developer to create an interface to the code, that is explicitly specified with a set of parameters c) Functions provides an abstraction of the code to potential users d) Writing functions is a core activity of an R programmer	A
37	The _____ function returns a list of all the formal arguments of a function. a) formals() b) funct() c) formal() d) fun()	A
38	You can check to see whether an R object is NULL with the _____ function. a) is.null() b) is.nullobj() c) null() d) as.nullobj()	A
40	Which package can be integrated with dplyr for large fast tables? a) Table b) Data, dplyr c) Data.table d) Dplyr.table	C

41	In the base graphics system, which function is used to add elements to a plot? a) Boxplot() b) Text() c) Boxplot() or Text() d) Treat()	C
42	What are the different types of sorting algorithms available in R language? a) Bubble b) Selection c) Merge d) All sorts	D
43	What will be the output of log (-5.8) when executed on R console? a) NAN b) NA c) Error d) 0.213	A
44	_____ is an indication that a fatal problem has occurred and execution of the function stops. a) message b) error c) warning d) message & warning	B
45	In 1991, R was created by Ross Ihaka and Robert Gentleman in the Department of Statistics at the University of _____ a) John Hopkins b) California c) Harvard d) Auckland	D
46	Warnings are generated by the _____ function. a) warning() b) error() c) run() d) message()	A
47	Finally, in _____ R version 1.0.0 was released to the public. a) 2000 b) 2005 c) 2010 d) 2012	A
48	R is technically much closer to the Scheme language than it is to the original _____ language. a) B b) C c) C++ d) S	D

Prof. Ujjval More   www.dimr.edu.in

49	Which of the following describes R language? a) Free b) Paid c) Available for free trial only d) Testing	**A**
50	The copyright for the primary source code for R is held by the _____ Foundation. a) A b) S c) C d) R	**D**

Unit 5 : Data Visualization

1	_____is a subset of _____ a) Information design, visual modality b) Information design, data visualization c) None of the answers are correct. d) Data visualization, information design	A
2	Which of the answers is an example of the kinesthetic modality? a) A speech b) A movie c) A picture d) The rain on our face	B
3	What area represents information in a graphical or pictorial form? a) Data design b) None of the answers are correct. c) Information design d) Data visualization	C
4	Which of the following is an example of a temporal data visualization? a) A Gnatt chart that is use in project management b) A histogram that represents proportions c) A matrix representing interconnecting data among various entities d) A 3D molecular rendering of a protein	D
5	By definition, Tableau displays measures over time as a a) Bar b) Line c) Histogram d) Scatter Plots	D
6	How do you identify a continuous field in Tableau? a) It is identified by a blue pill in the visualization b) It is identified by a green pill in a visualization c) It is preceded by a # symbol in the data window d) When added to the visualization, it produces distinct values	A
7	For creating variable size bins we use a) Sets b) Groups c) Calculated fields d) Table Calculations	B

8	Business intelligence (BI) is a broad category of application programs which includes _____ a) Decision support b) Data mining c) OLAP d) All of the mentioned	**D**
9	Which of the following measures of central tendency will always change if a single value in the data changes? A) Mean B) Median C) Mode D) All of these	**A**

10	Strong assessment items are made up of five elements:
	   a) Standard
	   b) Stimulus
	   c) Stem
	   d) Key
	   e) Distractors
		A

11	A good question is --------------- It focuses on recall of only the material covered in your lesson and aligns well with the overall learning objectives a) relevant b) clear c) concise d) purpose	**B**
12	A good question is framed in a-----------, easily understandable language, without any vagueness. Students should understand what is wanted from the question even when they don't know the answer to it. a) clear b) relevant c) concise d) purpose	**A**
13	____ programming language is a dialect of S. a) B b) C c) R d) K	**C**
14	Data visualization is also an element of the broader _____. A. deliver presentation architecture B. data presentation architecture C. dataset presentation architecture D. data process architecture	B

15	Point out the WRONG statement? a) Early versions of the S language contain functions for statistical modeling b) The book Programming with Data by John Chambers documents S version of the language c) In 1993 Bell Labs gave StatSci (later Insightful Corp.) an exclusive license to develop and sell the S language d) The book Programming with Data by IBM documents S version of the language	**C**
16	In 1991, R was created by Ross Ihaka and Robert Gentleman in the Department of Statistics at the University of _____ a) John Hopkins b) California c) Harvard d) Auckland	**D**
17	Point out the wrong statement? a) R is a language for data analysis and graphics b) K is language for statistical modelling and graphics c) One key limitation of the S language was that it was only available in a commercial package, S-PLUS d) C is a language for data and graphics	**A**
18	Business analytics results in which of these? a. Evidence Based Decisions b. Data Driven Decisions c. Better Decisions **d.** All of these are correct	**A**
19	Which one of the following is not a type of Business Analytics? a. Descriptive Analytics b. Diagnostic Analytics c. Predictive Analytics **d.** Performance Analytics	**D**
20	What will be the output of the following R code snippet? > paste("a", "b", se = ":") a) "a+b" b) "a=b" c) "a b :" d) none of the mentioned	**D**
21	Point out the correct statement? a) In R, a function is an object which has the mode function b) R interpreter is able to pass control to the function, along with arguments that may be necessary for the function to accomplish the actions that are desired c) Functions are also often written when code must be shared with others or the public d) All of the mentioned	**D**

22	Suppose there are 2 dataframes "A" and "B". A has 34 rows and B has 46 rows. What will be the number of rows in the resultant dataframe after running the following command? merge(A,B,all.x=TRUE) A) 46 B) 12 C) 34 D) 80	C
23	The very first thing that a Data Scientist generally does after loading dataset is find out the number of rows and columns the dataset has. In technical terms, it is called knowing the dimensions of the dataset. This is done to get an idea about the scale of data that he is dealing with and subsequently choosing the right techniques and tools. Which of the following command will not help us to view the dimensions of our dataset? A) dim() B) str() C) View() D) None of the above	C
24	Point out the wrong statement? a) A formal argument can be a symbol, a statement of the form 'symbol = expression', or the special formal argument b) The first component of the function declaration is the keyword function c) The value returned by the call to function is not a function d) Functions are also often written when code must be shared with others or the public	A
25	You can check to see whether an R object is NULL with the _____ function. a) is.null() b) is.nullobj() c) null() d) as.nullobj()	A
26	Which of the following code will print NULL? a) > args(paste) b) > arg(paste) c) > args(pastebin) d) > arg(bin)	A
27	What will be the output of the following R code snippet? > **paste**("a", "b", sep = ":") a) "a+b" b) "a=b" c) "a:b" d) a*b	A
28	What will be the output of the following R code snippet? > f <- function(a, b) { + print(a) + print(b) + } > f(45) a) 32 b) 42 c) 52 d) 45	A

29	What is the output of the command – paste(1:3,c("x","y","z"),sep="") ? A) [1 2 3x y z] B) [1:3x y z] C) [1x 2y 3z] D) None of the above	C

30	The number of accidents in a city during 2010 is
	   a) Discrete variable
	   b) Continuous variable
	   c) Qualitative variable
	   d) Constant
		A

31	The mean of a distribution is 23, the median is 24, and the mode is 25.5. It is most likelythat this distribution is:
	   a) Positively Skewed
	   b) Symmetrical
	   c) Asymptotic
	   d) Negatively Skewed
		A

32	Data collected by NADRA to issue computerized identity cards (CICs) are
	   a) Unofficial data
	   b) Qualitative data
	   c) Secondary data
	   d) Primary data
	   e) None of these
		C

33	Sum of dots when two dice are rolled is
	   a) A discrete variable
	   b) A continuous variable
	   c) A constant
	   d) A qualitative variable
		A

34	A chance variation in an observational process is
	   a) Dispersion/ Variability
	   b) Measurement error
	   c) Random error
	   d) Instrument error
		C

35	If a distribution is abnormally tall and peaked, then is can be said that the distribution is:
	   a) Leptokurtic
	   b) Pyrokurtic
	   c) Platykurtic
	   d) Mesokurtic
		A

36	The mean of a distribution is 14 and the standard deviation is 5. What is the value ofthe coefficient of variation?
	   a) 60.4%
	   b) 48.3%
	   c) 35.7%
	   d) 27.8%
		C

37	The first hand and unorganized form of data is called a) Secondary data b) Organized data c) Primary data d) None of these	C
38	Questionnaire survey method is used to collect a) Secondary data b) Qualitative variable c) Primary data d) None of these	
39	The data which have already been collected by someone are called a) Raw data b) Array data c) Secondary data d) Fictitious data	**C**
40	The grouped data is also called a) Raw data b) Primary data c) Secondary data d) Qualitative data	**C**
41	A constant variable can take values a) Zero b) Fixed c) Not fixed d) Nothing	**B**
42	A parameter is a measure which is computed from a) Population data b) Sample data c) Test statistics d) None of these	**A**
43	According to the empirical rule, approximately what percent of the data should liewithin $\mu \pm \sigma$? a) 75% b) 68% c) 99.7% d) 90% **e)** 95%	**E**
44	Primary data and_____data are same a) Grouped b) Secondary data c) Ungrouped d) None of these	**C**
45	Which one of the following measurement does not divide a set of observations intoequal parts? a) Quartiles b) Standard Deviations c) Percentiles d) Deciles e) Median	**B**

46	In descriptive statistics, we study
	   a) The description of the decision-making process
	   b) The methods for organizing, displaying and describing data
	   c) How to describe the probability distribution
	   d) None of the above
		A

47	Which of the following is not based on all the observations?
	   a) Arithmetic Mean
	   b) Geometric Mean
	   c) Harmonic Mean
	   d) Weighted Mean
	   e) Mode
		E

48	Which one is the not measure of dispersion.
	   a) The Range
	   b) 50th Percentile
	   c) Inter-Quartile Range
	   d) Variance
		B

49	When data are collected in a statistical study for only a portion or subset of allelements of interest we are using:
	   a) A sample
	   b) A Parameter
	   c) A Population
	   d) Both b and c
		A

50	In inferential statistics, we study
	   a) The methods to make decisions about the population based on sample results
	   b) How to make decisions about mean, median, or mode
	   c) How a sample is obtained from a population
	   d) None of the above
		A