Reg. No.												

Question Paper Code	12336

B.E. / B.Tech. - DEGREE EXAMINATIONS, NOV / DEC 2023

Fifth Semester

Information Technology

20ITPW501 - STATISTICAL ANALYSIS USING R PROGRAMMING WITH LABORATORY

(Regulations 2020)

Duration: 3 Hours Max. Marks: 100

PART - A (10 × 2 = 20 Marks)

Answer ALL Questions

		Marks, K-Level, CO
1.	List out some features of R programming.	*2,K1,CO1*
2.	What is the usage of lapply() and sapply() functions.	*2,K1,CO1*
3.	Look at the code below and find the output: n <- 3 n <- 5 print (choose (k, n))	*2,K2,CO2*
4.	Write R function used to find mean, median, mode and standard deviation.	*2,K1,CO2*
5.	What is dotcharts?	*2,K1,CO4*
6.	What is the use of Wilcoxon signed-rank test?	*2,K2,CO4*
7.	Differentiate between confidence bands and prediction bands.	*2,K2,CO5*
8.	Mention how can you produce correlations and covariance in R.	*2,K1,CO5*
9.	What is the difference between lm() and glm() in R?	*2,K2,CO6*
10.	Why we need logistic regression?	*2,K2,CO6*

PART - B (5 × 13 = 65 Marks)

Answer ALL Questions

11.	a)	(i) How to create and access the R matrix. Perform addition and division operations on R matrices.	*7,K2,CO1*
		(ii) What is data frame in R and why it is a crucial data structure in data analysis?	*6,K2,CO1*
		OR	
	b)	(i) Discuss how to read and write data from files in R Program.	*7,K2,CO1*
		(ii) Discuss how to use the data editor module in R.	*6,K2,CO1*
12.	a)	How to use sample () function to select a random sample in R from both a vector and a data frame (i) with replacement and (ii) without replacement. Explain with example.	*13,K2,CO2*

OR

b) (i) What is a Histogram? List the parameters used by the hist() function in R. Draw histogram for the following data 19, 23, 11, 5, 16, 21, 32, 14, 19, 27, 39. *7,K2,CO2*

(ii) What is box plot? Show how a box plot is drawn in R with example. *6,K2,CO2*

13. a) How are the following functions different and/or similar? What do they take as arguments, and what do they return? Explain with example data set. *13,K2,CO4*

(i) tapply() (ii) aggregate() (iii) by() (iv) table()

OR

b) Illustrate some general graphical techniques that allows to display similar plots for several groups on the same page with example. *13,K2,CO4*

14. a) Describe the computation of various parametric and non parametric correlation measures in R. *13,K2,CO5*

OR

b) Consider the below vectors *13,K2,CO5*

x = c(1, 2, 3, 4, 5, 6, 7)

y = c(1, 3, 6, 2, 7, 4, 5)

(a) Implement Pearson Correlation test for x and y using R program.

(b) Find Spearman Correlation test statistics for x and y.

15. a) Write R program script to carry out one way ANOVA for the following data by verifying the underlying assumptions. *13,K3,CO6*

Treatment	Block I	Block II	Block III	Block IV
A	12	18	19	14
B	14	20	22	17
C	16	21	24	18

OR

b) Explain the differences and similarities in simple linear regressions, multiple regressions and polynomial regression. Explain how to perform polynomial regression in R. *13,K3,CO6*

PART - C (1 × 15 = 15 Marks)

16. a) Discuss in detail about the ways of finding the different quantiles with appropriate example using R Code. *15,K2,CO3*

OR

b) (i) Write R program to create pie chart for the following data *8, K3,CO3*

Housing -------600; Food --------300; Clothes ---------150; Entertainment---100; Others ---------200

(ii) How to plot multiple curves in same graph for a table data? Explain with example? *7, K2,CO3*