M.Sc. DEGREE EXAMINATION, NOVEMBER 2024
BRANCH I - MATHEMATICS
THIRD SEMESTER

COURSE   :   ELECTIVE
PAPER   :   DATA ANALYSIS USING R
SUBJECT CODE   :   23MT/PE/DR23
TIME   :   40 MINUTES   MAX. MARKS: 20

THEORY

Q. No.	SECTION A ($10 \times 1 = 10$) **Answer ALL questions**	CO	KL
1	The statistical measure that gives the degree of spread of numeric observations around their arithmetic mean is _____ a) Mean b) Mode c) Variance d) None of these	1	1
2	The monthly salary (Rs.) of 10 employees of a firm are 2500, 2700, 2400, 2300, 2550, 2650, 2750, 2450, 2600, 2400. The arithmetic mean is _____ a) 2530 b) 2550 c) 2750 d) None of these	1	1
3	If the number of observation is _____, then the median is the arithmetic mean of the two middle observations. a) finite b) odd c) even d) None of these	1	1
4	The output of the statements : $x = 4$; $y = 7$; $p = x > y$; p is _____ a) TRUE b) FALSE c) "logical" d) None of these	1	1
5	The function that is used to return the count of objects present in the dataset is a) length() b) vector() c) print() d) sqrt()	1	1

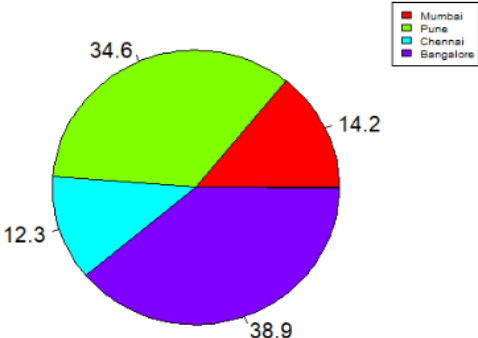| 6	If the mean, median, and mode are not equal in a distribution , then it is called _____
	    a) Symmetrical, b) Asymmetrical
	    c) Trimodal distribution d) None of these
		1	1

7	The value of _____ gives the degree of relationship between two variables. a) Skew b) covariance c) Variance d) None of these	1	1
8	The function _____is used to check the data type a) Factor() b) class() c) Class() d) factor()	1	1
9	The best plot to draw a data set with intervals is _____ a) Box plot b) Histogram c) Barplot d) None of these	1	1
10	To execute the dimension of a matrix, the function _____ is used. a) dimname() b) dim() rowname() d) None of these	1	1

Q. No.	SECTION B ($5 \times 2 = 10$) **Answer ANY FIVE questions**	CO	KL
11	Differentiate between sample and population.	2	2
12	What is one-tailed test?	2	2
13	State any two logical operators.	2	2
14	Define a matrix.	2	2
15	Define Histogram.	2	2
16	Consider a moderately skewed distribution in which mean and median are 35.4 and 34.3 respectively. Calculate the value of mode.	2	2

****

M.Sc. DEGREE EXAMINATION, NOVEMBER 2024
BRANCH I - MATHEMATICS
THIRD SEMESTER

COURSE   :   ELECTIVE
PAPER   :   DATA ANALYSIS USING R
SUBJECT CODE :   23MT/PE/DR23
TIME   :   140 MINUTES   MAX. MARKS: 80

PRACTICAL

Q. No.	SECTION C ($2 \times 20 = 40$)	CO	KL
	Answer ANY TWO questions		
1	a) Create the data frame "data" by defining the vectors $x1 = 1$ to 6, $x2 = 1, 2, 2, 3, 1, 2$ and $x3 =$ "F","B","C","E","A","D". b) i) Create a numeric variable and force it to be integer, ii) Force string "5.27" to an integer execute its data type.	3	3
2	Create a color pie chart in percentage for the following data: <table><tr><td>Mumbai</td><td>Pune</td><td>Chennai</td><td>Bangalore</td></tr><tr><td>23</td><td>56</td><td>20</td><td>63</td></tr></table> 	3	3

3	Analyse the following pairs of dataset:	3	3
	<table><tr><td>X_1</td><td>10, 8, 13, 9, 11, 14, 6, 4, 12, 7, 5</td></tr><tr><td>X_2</td><td>8.04, 6.95, 7.58, 8.81, 8.33, 9.96, 7.24, 4.26, 10.84</td></tr><tr><td>Y_1</td><td>20, 18, 23, 19, 21, 24, 6, 14, 22, 17, 5</td></tr><tr><td>Y_2</td><td>9.14, 8.14, 8.74, 8.77, 9.26, 8.10, 6.13, 3.10, 9.13</td></tr></table> Execute mean, variance, correlation between the pairs (X_1, Y_1) and (X_2, Y_2) and interpret the results.		
Q. No.	**SECTION D ($2 \times 10 = 20$)** **Answer ANY TWO questions**	**CO**	**KL**
4	Plot the scatter diagram and excute the covariance coefficient between the amount of fertilizer and yield of potatoes for the data: <table><tr><td>Amount (X)</td><td>0</td><td>4</td><td>8</td><td>12</td></tr><tr><td>Yield(Y)</td><td>8.34</td><td>8.89</td><td>9.16</td><td>9.50</td></tr></table>	4	4
5	a) Store four of your friends names as strings and perform the functions paste(), noquote(). b) Create a vector of length 10 and convert the numbers into string using the function tostring(). (5+5)	4	4
6	Consider two vectors u $< - c(8, 9, 10)$ v $< - c(1, 2, 3)$. and carry out the operations w1 = $\frac{(2 * u + v)}{10}$ and w2 = $(u + 2) * (u - 5) + v$ and $\frac{w1}{w2}$ and execute the results.	4	4
7	a) Create a sequence of values between -15 to 15 with a difference of 0.1. Execute Probability density function of normal distribution and plot the graph. b) Create a vector of 1000 random numbers with mean=90 and standard deviation =5 and the histogram with 50 bars. (5+5)	4	4

Q. No.	SECTION E (2 × 10 = 20) **Answer ANY TWO questions**	CO	KL
8	a) Create a sequence starting from 30 to 90 at an interval of 10. b) Create and store a vector that contains a sequence of length 7 from 3 to 6. c) Add and multiply the vectors of a) and b) d) Perform inner and outer multiply of vectors a) and b) (5+5+5+5)	5	5
9	Create Boxplot for the built-in dataset "mtcars" between the columns mpg and cyl in the dataset. Add title, label, new color to boxplot.	5	5
10	Use array() function to define a matrix with objects 12, suitable number of rows and columns of your choice and execute its dimension, row and column names.	5	5
11	Following are the time taken in seconds by 20 participants in a race: NA,NA,83,74,55,68,38,35,55,66,65,42,68,72,84,67,36,42,58 Execute the range, quartile deviation, variance and standard deviation.	5	5

****