

Roll No:

--	--	--	--	--	--	--	--	--	--

The Assam Royal Global University, Guwahati

Royal School of Applied and Pure Sciences

B.Sc. Mathematics 2nd Semester

Semester End Examination, June 2023

Course Title: Introduction to Data Science

Course Code: MAT012C203

Time: 3 Hours

Maximum Marks: 70

Note: Attempt all questions as per instructions given.

The figures in the right-hand margin indicate marks.

Section – A

1. Attempt **all** questions. (Maximum word limit 50) 2 x 8

- a. Distinguish between population and sample.
- b. Mention the uses of bar chart and pie chart.
- c. The average of 3, 5, $x+7$, 10, 12 is 8. Find x .
- d. A distribution has AM = 23, Mode = 24 and SD = 4. Show the nature of the distribution
- e. If X and Y be two correlated variables, then write the formula of product moment correlation coefficient.
- f. A bivariate distribution has $r = 0.73$ and $n = 144$. Show the reliability of r .
- g. Write the formulae to find mean and standard deviation in excel.
- h. Write the process to find the repeating values in excel.

Section – B

2. Attempt **any two** of the following: 6 x 2

- a. Distinguish between questionnaire and schedule.
- b. "Data can be arranged in the form of individual series, ungrouped frequency distribution and grouped frequency distribution"- Explain with examples.
- c. Draw a pie-chart for the following information:

Participants	Number of participants
Faculty	100
Admin staff	30
Workers	30
Students	200

3. Attempt **any two** of the following: 7 x 2

- a. What is meant by measures of central tendency? Write the essential characteristics of measures of central tendency. Why is arithmetic mean called best measure of central tendency?
- b. Marks obtained by the students in an examination are given below. Evaluate mean and mode.

Marks	20	25	30	35	40	45	50
No. of students	4	8	14	24	16	6	3

P.T.O

- c. Calculate standard deviation for the following distribution of marks:

Marks	20-30	30-40	40-50	50-60	60-70	70-80	80-90
No. of Students	5	15	18	25	14	9	4

4. Attempt **any two** of the following:

7 x 2

- a. What is Karl Pearson's coefficient of correlation? Mention the properties of correlation coefficient. Prove that correlation coefficient is symmetric.
- b. Calculate the rank correlation coefficient between the evaluators from the following data:

Score by 1 st Evaluator	15	17	14	16	18	19	20
Score by 2 nd Evaluator	10	12	16	18	14	18	8

- c. Explain why there are two lines of regression in a bivariate distribution. Also write the uses of the two regression equations.

5. Attempt **any two** of the following:

7 x 2

- a. Mention the steps to draw box plot in excel and write its' uses.
- b. "Regression can be done in excel" – Explain the steps for the same.
- c. Give a comparative summary between excel and tableau.