Reg. No. : ....

Name : ....

Second Semester B.Sc. Degree Examination, September 2022

First Degree Programme Under CBCSS

Statistics

Foundation Course

ST 1221 : STATISTICAL METHODS II

(2018 & 2019 Admission)

Time : 3 Hours

Max. Marks : 80

SECTION – A

Answer **all** questions. **Each** question carries 1 mark.

1. What are the methods of studying correlation?

2. Define scatter diagram.

3. The product of the two regression coefficient is _____

4. What do you mean by curve fitting?

5. Write the normal equations of $y = mx + c$

6. Define data mining.

7. What are the different data objects in R?

8. Define OLAP.

9. What is the use of C function in R?

10. What is Microsoft EXCEL?

<div align="right">(10 × 1 = 10 Marks)</div>

SECTION – B

Answer **any eight** questions. **Each** question carries **2** marks.

11. Define positive and negative correlation.

12. Define regression.

13. Describe the terms (a) multiple correlations and (b) partial correlation.

14. Explain properties of correlation coefficient?

15. What are the merits of rank correlation coefficient?

16. What do you mean by predictive data mining?

17. Explain the usage of time series in data mining.

18. What is rank correlation?

19. Define principle of least squares.

20. Write main features of EXCEL.

21. How do you install a package in R?

22. How many data formats are available in EXCEL?

<div align="right">(8 × 2 = 16 Marks)</div>

SECTION – C

Answer **any six** questions. **Each** question carries **4** marks.

23. Distinguish between correlation and regression.

24. Explain properties of regression coefficients.

25. Fit a straight line $y = ax + b$ to the following data.

X	1	2	3	4	5
Y	14	13	4	6	2

Estimate the value of y when $x = 3.5$.

26. Explain different data mining tools..

27. What is artificial neutral networks?

28. What are the applications of data mining? Explain in detail?

29. Explain the advantages and disadvantages of R.

30. Can you write and explain some of the most common system in R.

31. How do you create charts and graphics in EXCEL?

(6 × 4 = 24 Marks)

SECTION – D

Answer **any two** questions. **Each** question carries **15** marks.

32. Calculate Pearson's coefficient of correlation for the following data

X 39 65 62 90 82 75 25 98 36 78

Y 47 53 58 86 62 68 60 91 51 84

33. Given that $3x + 5y = 19$ and $2x + y = 16$ are regression lines. Find (a) the correlation coefficient (ρ_{xy}) and (b) the values of means.

34. Derive Spearman's formula for rank correlation coefficient.

35. Describe the importance of R software in data analysis. Explain the graphical and statistical function of R.

(2 × 15 = 30 Marks)

_____

P – 1304