



Registration Number:

Date & session:

ST. JOSEPH'S UNIVERSITY, BENGALURU -27
M.Sc. (STATISTICS – II SEMESTER
SEMESTER EXAMINATION: APRIL 2023
(Examination conducted in May 2023)
ST 8521 – INTRODUCTION TO DATA SCIENCE
(For current batch students only)

Time: 1 Hour

Max Marks: 25

This paper contains One printed page and one part
Each question carries FIVE marks

PART-A

Answer any FIVE of the following

(5 × 5 = 25)

1. a) Briefly describe the procedure to import files from excel to R. (1)
b) Create a function to determine BMI of an individual. (2)
c) How does one create graphs on Excel. (2)
2. a) Define
i) Data Frame (1)
ii) Vector (1)
iii) Lists (1)
b) What is Big Data? Give 2 examples of the same. (2)
3. a) Define the 3 V's of Big data. (3)
b) Explain the two graphical methods of finding outliers. (2)
4. What are the main techniques and tools used in exploratory data analysis, and how can they be applied to gain insights and understanding from a dataset? (5)
5. Explain clustering and its applications. (5)
6. a) Why is it important to model a data? (2)
b) Briefly describe Support vector machine. (3)
7. Write a note on Random Forest. Explain its relevance in real world.

ST 8521_A_23