



'समानो मन्त्रः समितिः समानी'

**UNIVERSITY OF NORTH BENGAL**

MDC 1st Semester Examination, 2023

**UPOAMDC11007-FUNDAMENTALS OF DATA SCIENCE**

Time Allotted: 2 Hours 30 Minutes

Full Marks: 60

*The figures in the margin indicate full marks.  
All symbols are of usual significance.*

**GROUP-A**

**Answer any four of the following**

3×4 = 12

1. What is data?
2. Name any two data formats used in data analysis.
3. Define API.
4. Explain any one multivariate statistical method.
5. What are the advantages of using decision tree algorithm in data science?
6. Define unsupervised learning.

**GROUP-B**

**Answer any four of the following**

6×4 = 24

7. Discuss different issue related to privacy of data.
8. Briefly explain “for loop” and “while loop” in R/Python.
9. Explain different ethical guidelines for handling data.
10. What are outliers? How can we find outliers in the data?
11. Explain various data cleaning methods.
12. Discuss the K-means clustering algorithm briefly.

**GROUP-C**

**Answer any two of the following**

12×2 = 24

13. Discuss the working of decision tree algorithm with a suitable example.
14. Explain the K-fold cross validation scheme with a suitable diagram.
15. Discuss various ways of obtaining data from web and APIs.
16. What is the significance of line fitting? Discuss the algorithm of least square method with an example.

—x—