

(i) Printed Pages : 3

Roll No. ....

(ii) Questions : 9

Sub. Code :

2	6	0	3	4
---	---	---	---	---

Exam. Code :

0	4	5	9
---	---	---	---

**M.Sc. Information Technology 1st Semester  
(2125)**

**Option (ii) DATA ANALYTICS-I**

**Paper-MS-71**

**Time Allowed : Three Hours]**

**[Maximum Marks : 80**

**Note :— Attempt ONE question from each Section and the entire compulsory question.**

**SECTION—A**

1. Differentiate between Structured and Unstructured data with examples. How does unstructured data pose challenges for data analysis and how do modern tools address these challenges? 16
2. What are the main components of Data Science? Explain the role of each component. How is classification in Data Science different from clustering? 16

**SECTION—B**

3. (a) Define Data Analytics. How is it used to extract insights from data? Give suitable example. 8

- (b) Why is EDA (exploratory data analysis) considered a crucial step before model building? How does data discovery contribute to business intelligence? 8
4. (a) Explain how each phase in the analytics life cycle contributes to the final output. 8
- (b) Discuss the importance of feature selection. How do Decision Trees perform feature selection automatically? 8

### SECTION—C

5. (a) Discuss how algorithmic thinking and Python programming together support data analysis. 8
- (b) Write a Python function to calculate the sum of digits of a number using recursion. 8
6. (a) What is a tuple? How does it differ from a list? Explain the structure of the for-loop in Python with an example. 8
- (b) How do mutable and immutable data types behave when passed into functions? Give example. 8

### SECTION—D

7. What are the main features of NumPy that make them suitable for data manipulation? How are masks used in NumPy for data filtering? Write a program to create a mask that selects all elements greater than a given threshold value. 16

8. Describe the main objects used in Pandas. How can indexing and selection be used to clean and preprocess data? Discuss the different techniques to handle missing data. 16

**(Compulsory Question)**

9. Answer the following :

- (a) List the commonly used programming languages and environments for data science applications.
- (b) Exemplify string slicing in Python.
- (c) How does wrapper technique work in evaluating features?
- (d) What is a dictionary in Python? How does it store data?
- (e) How are pivot tables useful for data summarization?

$$3+3+3+3+4=16$$