**MODEL PAPER**

PYTHON PROGRAMMING FOR DATA ANALYSIS

**BSc - Data Science - Semester III**

**Time: 3 hrs Total Marks: 60**

**SECTION - A**

**Answer any five of the following questions** **5 X 4 = 20 M**

1. What is the difference between Data Analysis and Data Analytics?
2. What are Essential Python libraries?
3. What is a NumPy array?
4. What is vectorized computation in NumPy?
5. What is a DataFrame in Pandas?
6. How can you read a CSV file using Pandas?
7. How do you handle missing data in Pandas?
8. What is the purpose of merging datasets in Pandas?
9. What is the use of the Patsy library in Python modeling?
10. Write about any two Python libraries used for data visualization.

**SECTION - B**

**Answer the following questions** **5 X 8 = 40 M**

1. (a). What is data analysis? How is it different from data analytics? Why do we use Python for data analysis?  
    **(OR)**(b). What is a Python library? Name and explain any three libraries used in data analysis.
2. (a). What are Python’s built-in data types like list, tuple, and dictionary? Give examples.  
    **(OR)**(b). What is NumPy? How are arrays created and used in NumPy?
3. (a). What is a DataFrame in Pandas? How is it different from a Series? Give examples.  
    **(OR)**(b). How do we read data from a CSV file and write it back using Pandas?
4. (a). What is missing data? How do we handle missing values in a dataset using Pandas?  
    **(OR)**(b). What is data wrangling? How do we join and combine two datasets in Pandas?
5. (a). What is Patsy in Python? How is it used to describe data models?  
    **(OR)**(b). What is data visualization? How can we make simple plots using Pandas or Seaborn?