190

23BDS402

UG PROGRAM (4 YEARS HONOURS) WITH SINGLE MAJOR AT THE END OF FOURTH SEMESTER

DATA SCIENCE - DATA VISUALIZATION USING PYTHON

(COMMON FOR B.Sc DATA SCIENCE & B.C.A DATA SCIENCE)

(B.Sc HONOURS-MAJOR)

(w.e.f. Admitted Batch 2023-24)

Time: 3 Hours **Maximum: 70 marks**

Section – A

Answer any Five questions. **5x4=20**

1. What is Data Science and what are its key components?
2. What is vectorized computation in pandas and why it is advantages for data processing?
3. What is hierarchical indexing in Pandas and how it simplifies complex data structures?
4. What is 'spilt-apply-combine' strategy in data aggregation and how does it work?
5. What are techniques for cleaning categorical data using Pandas?
6. Describe the purpose and features of Jupyter notebook in python-based data science workflows.
7. What are the methods used in pandas for handling missing data?
8. How can Pandas be used for plotting data?

Section – B

Answer ALL questions. **5x10=50**

9. (a) Discuss in detail the motivation behind using Python for Data Science.

(Or)

(b) Evaluate the roles of machine learning libraries within the Data science process.

10. (a) Discuss in detail the functionalities of Pandas for summarizing data.

(Or)

(b) Evaluate the methods available in Pandas for reading and writing data in various file formats.

11. (a) Discuss the various methods for combining and merging datasets in Pandas.

(Or)

(b) Evaluate the techniques for creating advanced categorical and numeric plots using Python visualization tools.

12. (a) Discuss the use of pivot tables and cross tabulation for summarization.

(Or)

(b) Explain the methods of handling time zones, working with periods and performing period arithmetic in Python.

13. (a) Describe the advanced Group By methods available in Pandas.

(Or)

(b) Explain the concept of method chaining in Pandas, discuss its syntax and benefits.