JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -2 EXAMINATION- 2025

BBA-V Semester (BBA)

COURSE CODE(CREDITS): 25BBWHS532 (4) MAX. MARKS: 25

COURSE NAME: Data Analytics using Python

COURSE INSTRUCTORS: Dr. Nishant Jain MAX. TIME: 1.5 Hour

Note: *(a)All questions are compulsory.*

(b) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems.

(c) For all the questions, support your answers with Python code.

---

1. Describe the functions in NumPy like reshape(), flatten(), and transpose(), highlighting their roles in manipulating array structures; provide specific use cases in data analysis where these functions are crucial, along with example snippets. [3, CO1]
2. Detail how statistical operations such as mean, median, standard deviation, and correlation are performed with NumPy, emphasizing their importance in data exploration and preprocessing. [4, CO1]
3. Explain the importance of data normalization and standardization, demonstrating how NumPy can apply min-max scaling and z-score standardization. [3, CO4]
4. Cover common data cleaning tasks with pandas, including handling missing values, duplicates, and data type inconsistencies, using functions like dropna(), isnull(), fillna(), drop_duplicates(), and astype(). [3, CO1]
5. Discuss the advantages of using categorical data types in pandas for memory efficiency and performance, illustrating conversions and categorical operations. [3, CO2]
6. Explain pandas' data selection and filtering techniques, such as .loc[], .iloc[], and boolean indexing, with scenarios for each. [3, CO1]
7. Propose a detailed visualization plan using Matplotlib for datasets with categorical and numerical variables, justifying plot choices to explore relationships and distributions. [3, CO2]
8. Design a case study where you analyze a real-world dataset (e.g., sales data, healthcare data, social media metrics). Outline the step-by-step visualization plan using Matplotlib to extract meaningful insights, including initial exploratory plots, intermediate analysis, and final presentation visuals. [3, CO3]