Python Data Analysis Evaluation Test

Duration: 3 Hours  
Total Marks: 100  
Topics: NumPy, Pandas, Matplotlib

# Section A: NumPy (30 Marks)

Q1. Create a NumPy array of shape (5, 5) with random integers from 1 to 100. Write code to find: (6 marks)

a. The maximum value and its position

b. The mean and standard deviation of the array

Q2. Write a NumPy program to replace all odd numbers in a given array with -1 without changing the original array. (6 marks)

Q3. Given two arrays, A and B, write a code to check if they are equal (element-wise). (6 marks)

Q4. Create a 2D NumPy array and perform slicing to extract the middle 3x3 block. (6 marks)

Q5. Generate a 1D array of 20 random numbers. Normalize the array using Min-Max scaling. (6 marks)

# Section B: Pandas (40 Marks)

Q6. Load a CSV file using Pandas. Display the first 5 and last 5 rows. Also, print column data types. (5 marks)

Q7. Write code to: (10 marks)

a. Filter rows where a column 'Age' is greater than 30

b. Group data by 'Department' and compute average salary

Q8. Write Pandas code to fill missing values in a DataFrame with the mean of the respective column. (5 marks)

Q9. Merge two DataFrames on a common column and perform an inner join. Show a sample output. (5 marks)

Q10. Given a DataFrame, remove duplicate rows and reset the index. (5 marks)

Q11. Create a new column 'Bonus' as 10% of 'Salary'. Also, sort the DataFrame by Salary descending. (5 marks)

Q12. Write a function that takes a DataFrame and returns the top 3 departments with highest total salary. (5 marks)

# Section C: Matplotlib (30 Marks)

Q13. Create a line plot showing sales data over 12 months. Add title, xlabel, ylabel, and grid. (6 marks)

Q14. Draw a bar chart to compare number of employees in different departments. Add colors and labels. (6 marks)

Q15. Create a pie chart showing percentage distribution of expenses (Rent, Food, Travel, Utilities, Other). (6 marks)

Q16. Plot a histogram of 1000 normally distributed random numbers. Customize bins and color. (6 marks)

Q17. Create a subplot with 1 row and 2 columns: first for a bar plot, second for a line plot. (6 marks)