SUBJECT – INFORMATICS PRACTICES

Section A: MCQs (1 mark each)

1. Which library is used for data manipulation in Python?
   a) NumPy
   b) Pandas
   c) Matplotlib
   d) Seaborn
2. What is the output of `len(df)` in Pandas?
   a) Number of columns
   b) Number of rows
   c) Total elements
   d) Data type
3. Which function is used to read a CSV file?
   a) read_excel()
   b) read_csv()
   c) open_csv()
   d) csv_read()
4. Which of the following creates a Series?
   a) pd.DataFrame()
   b) pd.Series()
   c) pd.Array()
   d) pd.List()
5. What does `df.head()` return?
   a) Last 5 rows
   b) First 5 rows
   c) Random rows
   d) Column names
6. Which function is used to check null values?
   a) isnull()
   b) notnull()
   c) dropna()
   d) fillna()
7. NumPy is mainly used for:
   a) Data visualization
   b) Numerical operations
   c) File handling
   d) Web development
8. Which operator is used for element-wise comparison in Pandas?
   a) ==
   b) =
   c) !=
   d) Both (a) and (c)
9. What is the default index start value in Pandas?
   a) 0
   b) 1
   c) -1
   d) None

10. Which function removes missing values?
    a) fillna()
    b) dropna()
    c) remove()
    d) delete()

□ Section B: Assertion–Reason Questions (1 mark each)

For each question, choose the correct option:
A. Both Assertion and Reason are true and Reason is the correct explanation
B. Both are true but Reason is not correct explanation
C. Assertion is true but Reason is false
D. Assertion is false but Reason is true

1. Assertion: Pandas DataFrame is 2-dimensional.
   Reason: It consists of rows and columns.
2. Assertion: `df.tail()` shows first rows.
   Reason: It displays last rows of DataFrame.
3. Assertion: NumPy arrays are faster than Python lists.
   Reason: NumPy uses contiguous memory allocation.
4. Assertion: `fillna()` deletes missing values.
   Reason: It replaces missing values.
5. Assertion: Index in Pandas must always be numeric.
   Reason: Index can also be string labels.

□ Section C: Short Answer Questions (2–3 marks)

1. Define Pandas Series with an example.
2. What is a DataFrame? Give its structure.
3. Differentiate between Series and DataFrame.
4. Explain `head()` and `tail()` functions.
5. What are missing values? How can they be handled?
6. Write steps to read a CSV file in Pandas.
7. Explain the use of `loc[]` and `iloc[]`.
8. What is NumPy array? How is it different from list?
9. Write syntax to create a DataFrame from dictionary.
10. Explain Boolean indexing with example.

□ Section D: Long Answer Questions (4–5 marks)

1. Explain the features of Pandas library in detail.
2. Write a Python program to:
    - Create a DataFrame of student details
    - Display first 3 rows
    - Add a new column
3. Explain different ways to handle missing data in Pandas.
4. Compare NumPy arrays and Python lists with examples.
5. Write steps and code to:
    - Read a CSV file
    - Display summary using `describe()`
    - Filter data based on condition
6. Explain indexing and slicing in Pandas with examples.
7. Write a program to demonstrate:
    - Creating Series
    - Accessing elements
    - Performing arithmetic operations
8. Explain data cleaning process in Pandas.

□ Section E: Case Study Based Questions

A DataFrame contains employee data: EmpID, Name, Salary, Department.

1. Write code to display employees with salary > 50,000.
2. Find average salary of employees.
3. Add a new column “Bonus”.
4. Handle missing salary values.