22BDAE215 Reg No : .....

CHOICE BASED CREDIT SYSTEM

M.Sc. BIG DATA ANALYTICS SECOND SEMESTER DEGREE EXAMINATION MAY 2024

Data Analytics using Python

Duration:3 Hours Max Marks:70

PART A

I. Answer any FOUR of the following **(4×5= 20)**

1) Explain with suitable example the concept of python functions.
2) Develop a python code for the function dropna() with the expected output.
3) Create a numpy array X with 10 even numbers between 1 and 20. Split this array into 3 separate lists X1, X2, X3 and print them.
4) Explain the split() and compile() functions associated with Regular Expressions.
5) Elaborate on the parameter "color" in the function plot().

PART B

II. Answer any FIVE questions selecting at least one question from each unit: **(5×10= 50)**

UNIT-I

6) Explain the concept of statements and expressions used in python programming language.
7) Elaborate on break and pass statements with suitable examples.

UNIT-II

8) 1. Explain the terms Series and Dataframe in Pandas with an example for each.
   2. Consider any five departments in an educational institution. Create a dictionary containing details of department names and the total student strength of each department in the current semester. Create a series using this dictionary in Pandas and display it.
9) Explain any five aggregate functions in numpy library with examples.

UNIT-III

10) With illustrations and python code, explain the concepts of correlation and covariance.
11) Create dataframe1 with details of Subject_ID, Firstname and Lastname. Create dataframe2 with details of Subject_ID, test_ID. Write the code to execute the following and display what is the expected output on executing each of them.
    1. Merge the two dataframes along the Subject_ID
    2. Join the two dataframes along columns.
    3. Join the two dataframes along rows.

UNIT-IV

12) Develop a Python program to create a pie chart of medals, achievements in sports events of five most successful countries in 2020 Summer Olympics. Add a suitable title to the chart. Display the medal details suitably.

13) Develop a Python program:

1. To create a line graph of number of seminars attended by 5 students in a class. Each of the data is recorded as numpy array of integers.
2. Add suitable title and labels to the graph.

****