[This question paper contains 8 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 2002 **H**

Unique Paper Code : 62345627

Name of the Paper : Data Science using Python

Name of the Course : **B.A. Programme**

Semester : VI

Duration : 3 Hours Maximum Marks : 75

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Question No. **1** is compulsory.
3. Attempt any **five** questions from Question No. **2** to **8**.
4. Parts of a question must be answered together.

P.T.O.

Section A

1. (a) Describe the characteristics of the CSV format. (2)

(b) Define slicing operation in string of Python? Write Python code for slicing for fetching first name and last name from full name of a person and display it. (3)

(c) "Python is a first choice for data scientists". Discuss. (4)

(d) Write Python program to find the Greatest Common Divisor (GCD) of two positive numbers. (4)

(e) Dictionary dict contains the following elements:

```
dict={"Phy":94,"Che":70,"Bio":82,"Eng":95}
```

Write the output for the following code.

```
dict.update({"Che":72,"Bio":80})
print(dict)
```

(4)

(f) Explain sampling in terms of data science? (2)

(g) Explain range() function with suitable examples. (2)

(h) What is covariance? What is the importance of covariance in data analysis? Explain with the help of example. (4)

Section B

2. (a) Describe Arithmetic Operators, Logical Operators and Bitwise Operators in detail with examples. (6)

(b) Evaluate the expression according to the precedence of operators in Python:

$y = (10 + 2 * 3) ** 2/8 + 4\% \ 3 - 2 * 5 // 3$ (4)

3. (a) Explain the identifiers, keywords, statements and variables in Python programming language with examples. (4)

(b) Countries = ["USA", "Canada", "Mexico", "Brazil", "Argentina", "Peru", "Chile"]

Write the output for the following code.

```
del countries[2]
countries.remove("Brazil")
countries.pop(4)
print(countries)
```

(3)

(c) Explain the concept of scope and lifetime of variables in Python programming language with an example.

(3)

4. (a) Write the output for the following Python code

```
duList = [1,2,3, 'Delhi University', 'Data Science']
for i in range (7,9):
    duList.append(i)
print(duList)
```

(2)

(b) Write a program to print following patterns.

(i)
```
*
* *
* * *
* * * *
```

(ii)
```
$ $ $ $
$ $ $
$ $
$
```

(4)

(c) Discuss the following dictionary methods with an example.

(i) get()

(ii) keys()

(iii) pop()

(iv) update()

(4)

5. (a) List any three methods to handle missing values Pandas. Explain with example.

(b) Write a program in Python to create a datafram df with three columns x, y and z with the value given below :

x: 5, 7, 8, 9, 6, 7, 10

y: 2, 4, 6, 8, 10, 12, 14

z: 1, 1, 2, 3, 5, 8, 13

Calculate mean, median and standard deviation. (4)

6. (a) Find the linear regression coefficients using gradient decent method for the following data when learning rate = 0.1. Carry out the process for 2 iterations. (Note: x is independent and y is dependent variable)

x	y
0	2
1	3
2	5
3	4
4	6

(2)

(b) Write an iterative function in Python, fact(n) that returns the factorial of a given number n.

(4)

(c) Compare bar graph, box-plot and histogram with respect to their applicability in data visualization.

(4)

7. (a) Write a Python program using NumPy to compute the variance of three arrays x, y, and z. The arrays are defined as follows :

```
x = np.array([4,10,6, 8,2,7])

y = np.array([3,2,7, 9, 5,1])

z = np.array([5,1, 8,2,6,4])
```

(4)

(b) What are the differences between list, tuple and set in Python? Explain with the help of suitable example. (6)

8. (a) What isthe use of scatter-plot in data visualization? Can we draw trendline using scatter-plot? Explain with the help of example. (4)

(b) Illustrate the following Set methods with an example.

(i) To find the intersection of a and b.

(ii) To find union of a and b.

(iii) To find if a is subset of b. (6)