

K-1116

Total Page No. : 4]

[Roll No.]

CDSA-102

**Certificate in Data Science and
Application Ist Semester
Examination Dec., 2023**

PROGRAMMING FOR DATA SCIENCE

Time : 2 Hours]

[Max. Marks : 100

Note :- This paper is of Hundred (100) marks divided into two (02) Sections 'A' and 'B'. Attempt the questions contained in these Sections according to the detailed instructions given there in. Candidates should limit their answers to the questions on the given answer sheet. No additional (B) answer sheet will be issued.

Section-A

Long Answer Type Questions 2×26=52

Note :- Section 'A' contains Five (05) Long-answer type questions of Twenty Six (26) marks each. Learners are required to answer any two (02) questions only.

K-1116

(1)

P.T.O.

1. (a) Discuss brief overview of Python and its key characteristics. [13]
- (b) Describe the steps involved in installing Python, and mention any considerations for different operating systems. [13]
2. Provide a definition of a function and explain its role in programming. What are the different functions available in python ? Discuss in detail.
3. Discuss about an overview of the applications and uses of R programming. Describe the steps involved in installing and loading a package in R.
4. Explain the steps involved in conducting a basic statistical analysis in R. Include loading a dataset, calculating summary statistics, performing hypothesis testing, and creating relevant visualizations using functions from base R and the ggplot2 package.
5. Explain the concept of Principal Component Analysis (PCA) in R. Discuss its application for dimensionality reduction and visualization of high-dimensional data.

Section–B

Short Answer Type Questions 4×12=48

Note :- Section 'B' contains Eight (08) Short-answer type questions of Twelve (12) marks each. Learners are required to answer any *four* (04) questions only.

1. Provide a definition of a dictionary and explain its purpose in Python. Give an example of creating an empty dictionary and explain its use.
2. What do you mean by a string and explain its significance in Python. Discuss how escape characters are used to represent special characters in strings. What is string interpolation in Python ?
3. Compare the performance of tuples and lists in terms of memory usage and access time.
4. What is the Global Interpreter Lock (GIL) in Python, and how does it affect multithreading ?
5. Discuss strategies for handling missing values when working with multiple variables in R. Provide examples of functions or techniques to address missing data.

6. Compare and contrast vectors and lists in terms of structure and functionality.
7. Describe how R supports object-oriented programming, including the use of classes and methods.
8. Define factors and discuss their role in representing categorical data in R.
