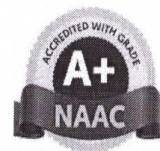


SHRI GNANAMBICA DEGREE COLLEGE: MADANAPALLE



(AUTONOMOUS)
Course 5: Python Programming for Data Analytics (MAJOR)
SEMESTER III
(W.E.F.2024-25)
Program: BSC (DS)



Hours per week: 4

Credits: 3

Course Objectives:

- To be able to program in Python.
- To know and understand the data analysis phases.
- To know the usage of all relevant libraries.

Course Outcomes:

Upon successful completion of the course, the student will be able to:

- Understand and learn all basic concepts of Python programming.
- Apply data analysis methods using Python.
- Become familiar with Python programming environments.

UNIT I

Introduction to Data Analysis and Python : What is Data Analysis, Differences between Data Analysis and Data Analytics, Introduction to Python language basics and Jupyter Notebook, Why Python for Data Analysis, Essential Python Libraries.

UNIT II

NumPy and Core Python Concepts: Built-in Data Structures, Functions, Files, and Operating System Interfaces, NumPy Basics: Arrays and Vectorized Computation, Universal Functions, Object-Oriented Programming with Arrays, File Input and Output with Arrays, Linear Algebra using NumPy.

UNIT III

Working with Pandas and Data Loading:

Getting Started with Pandas: Introduction to Pandas Data Structures, Essential Functionality, Summarizing, and Computing Descriptive Statistics, Data Loading, Storage, and File Formats: Reading and Writing Data in Text Format, Binary Data Formats, Interacting with Web APIs, Interacting with Databases.

UNIT IV

Data Cleaning and Wrangling: Data Cleaning and Preparation: Handling Missing Data, Data Transformation, String Manipulation, **Data Wrangling:** Join, Combine, and Reshape, Hierarchical Indexing, Combining and Merging Datasets, Reshaping and Pivoting.



c Mahesh Babu
CHAIRMAN
BOARD OF STUDIES
Shri Gnanambica Degree College (A)
MADANAPALLE - 517 325

UNIT V

Modeling, Plotting, and Visualization:

Modeling Libraries in Python: Interfacing between Pandas and Model Code, Creating Model Descriptions with Patsy, Introduction to Statsmodels,

Plotting and Visualization: Matplotlib API Primer, Plotting with Pandas and Seaborn, Other Python Visualization Tools,

REFERENCES:

1. Wes McKinney – *Python for Data Analysis*, O'Reilly Publications, Second Edition.
2. Charles R. Severance – *Python for Everybody: Exploring Data Using Python 3*.
3. John Zelle, Michael Smith – *Python Programming*, Second Edition, 2010.



c Mahesh Babu
CHAIRMAN
BOARD OF STUDIES
Shri Gnanambica Degree College (A)
MADANAPALLE-517 325

SHRI GNANAMBICA DEGREE COLLEGE: MADANAPALLE



(AUTONOMOUS)
Course 5: Python Programming for Data Analytics (MAJOR)- Practicals
SEMESTER III
(W.E.F.2024-25)
Program: BSC (DS)



Hours per week: 2

Credits: 1

List of Experiments

1. Use Matplotlib and plot an inline graph in Jupyter Notebook.
2. Implement basic commands of the Python Language.
3. Create Tuples and Lists, and illustrate slicing conventions.
4. Create and demonstrate built-in sequence functions.
5. Clean and transform elements using List, Set, and Dictionary Comprehensions.
6. Create a functional pattern to modify strings at a high level.
7. Write a Python program to cast a string to a floating-point number and handle improper inputs using Errors and Exception Handling (raise ValueError for invalid cases).
8. Create an N-dimensional array object and perform various operations on it.
9. Use arithmetic operations on NumPy arrays.
10. Perform Indexing and Slicing, Boolean Indexing, and Fancy Indexing operations using NumPy arrays.
11. Create an image plot from a two-dimensional array of function values using Matplotlib.
12. Implement basic array statistical methods such as sum, mean, std, var, min, max, argmin, argmax, cumsum, and cumprod. Also perform sorting using the sort() method.
13. Implement and demonstrate NumPy random functions.
14. Plot the first 100 values obtained from random walks.
15. Create a DataFrame using Pandas and retrieve specific rows and columns by performing indexing operations. Also perform transpose operations on it.
16. Implement methods of descriptive and summary statistics using Pandas.
17. Load and write data to and from different file formats, including Web APIs.
18. Implement data cleaning and filtering methods, including handling missing data using isna() and fillna() functions.
19. Transform data using functions or mapping techniques.
20. Rearrange data using the unstack() method of hierarchical indexing.



C Mahesh Babu
CHAIRMAN
BOARD OF STUDIES
Shri Gnanambica Degree College
MADANAPALLE - 517 325

SHRI GNANAMBICA DEGREE COLLEGE: MADANAPALLE

(AUTONOMOUS)

Course 5: Python Programming for Data Analytics (MAJOR)

SEMESTER III

(W.E.F.2024-25)

Program: BSC (DS)

Question Paper – Blue Print

Time : 3 Hrs

Marks : 70

PART-A

Answer any 4 of the 8. Each Question Carries 5 marks.

(4 x 5 =20)

1. Question
2. Question
3. Question
4. Question
5. Question
6. Question
7. Question
8. Question

PART-B

Answer one from each unit. Each Question Carries 10 marks.

(5X10=50)

UNIT 1

9. Question

OR

10. Question

UNIT 2

11. Question

OR

12. Question

UNIT 3

13. Question

OR

14. Question

UNIT 4

15. Question

OR

16. Question

UNIT 5

17. Question

OR

18. Question



c Mahesh Babu
CHAIRMAN
BOARD OF STUDIES
Shri Gnanambica Degree College (A)
MADANAPALLE - 517 325