

ANNEXURE I

Course Name: Certificate Course in Data Analysis using Python

Course Objective: The objective of this course is to create expertise in Python Programming to develop data science applications.

Prerequisite: Candidates should be proficient in Computer Fundamentals, Basic Database and Programming Concepts with Logical Approach.

Course Duration: 80 Hrs (8 hours/ day for 2 Weeks)

Teaching Schema

S. No.	Modules	Hours
1	Introduction to Python programming language	15
2	Understanding the concept of data structures in Python and their implementation	30
3	Pandas basics for data manipulation in Python	15
4	Data visualization on using matplotlib and seaborn libraries	10
5	Project	10
	Total	80

Detailed Course Content

- 1. Introduction to Python programming language:**
- 2. [2]** Installation and setup,
 - Creating and saving a script file
 - basics of Python, • data types,
 - Variables.
 - Syntax and comments in Python: Understanding the syntax and commenting the code
 - Python string manipulation: String data type, string indexing, slicing, concatenation, and formatting [2]
- [2] Introduction to Python operators:**
 - Arithmetic operators,
 - comparison operators,
 - logical operators, • Assignment operators.

3. Understanding the concept of data structures in Python and their implementation

Python lists,

- tuples,
- sets,
- dictionaries

Conditional statements and loops in Python:

- if-else,
- while loop, `for` loop,
- nested loops.

Introduction to Python functions:

- Defining functions, arguments, return statement, and scope of variables.

Introduction to Data Manipulation

- Introduction to NumPy

NumPy Package in Python, Importing NumPy, creating different arrays using NumPy, Array Functions and Methods, Different Mathematical Functions, Different Matrix Operations, Random Numbers, Generate Numbers between a range. **Pandas**

basics for data manipulation in Python:

- Understanding the Pandas library, series, and data frame operations.
- Concept of Series in Pandas, Creating Series using Pandas, Different Series Attributes, Series vs List, Series Operations, Series from CSV File. Exploratory data analysis,
- Reading files
- Data cleaning in Python: handling missing values and filling NA
- Data preparation and pre-processing
- Data feature engineering: handling categorical data
- Data validation techniques in Python
- Data feature engineering: removing columns and rows from raw data

4. Data visualization on using matplotlib and seaborn libraries

- Scatter plot
- Line plot
- Bar plot
- Histogram
- Box plot
- Pie plot

5. Final Project: showcase all your skills in an end-to-end data analysis project. You'll pick the dataset, do the data munging, ask the research questions, visualize the data, draw conclusions, and present your results

