



What is Python for Data Science?

The wide community of Python users have created quite a few relevant and useful tools specifically for Data Science and Analysis. This has made analysing, managing, or playing around with data, as easy as ever.

This course revolves around providing you with hands-on knowledge and practical experience with data structures, arrays, Python lists, functions, control flow etc. By the end of this course, you will have developed an aptitude for using Python with Data Science tasks and applications!

Why use Python for Data Science?

The easy-to-understand syntax of Python makes even the complex code simple and easy to write and read. Python works like a charm while implementing algorithms.

SciPy, an open-source Python library, is specifically used for scientific and technical computing. The library includes modules for linear algebra, integration, interpolation, signal and image processing among many others. Python facilitates easy integration of data analysis tasks with web applications.

Why you must learn Python for Web Application?

- The flexibility of Python makes it easy to learn and use.
- The demand for Python in Data Science has considerably grown in recent times and the trend is here to stay!
- The professionals and developers who want to make a mark for themselves in the field of Data Science are all using Python for data analysis and other applications.
- You will also learn to use NumPy and pandas that contribute to your technical expertise.

Python For Data Science Curriculum

1)Data Science Overview

- Introduction to Data Science
- Different Sectors Using Data Science
- Purpose and Components of Python

2) Data Analytics Overview

- Data Analytics Process
- Exploratory Data Analysis(EDA)
- EDA-Quantitative Technique
- EDA - Graphical Technique
- Data Analytics Conclusion or Predictions
- Data Analytics Communication
- Data Types for Plotting

3) Statistical Analysis and Business Applications

- Introduction to Statistics
- Statistical and Non-statistical Analysis
- Major Categories of Statistics
- Statistical Analysis Considerations
- Population and Sample
- Statistical Analysis Process
- Data Distribution
- Dispersion
- Histogram
- Correlation and Inferential Statistics

4) Python Environment Setup and Essentials

- Anaconda
- Installation of Anaconda Python Distribution (contd.)
- Data Types with Python
- Basic Operators and Functions

5) Mathematical Computing with Python (NumPy)

- Introduction to Numpy
- Activity-Sequence it Right

- Demo 01-Creating and Printing an ndarray
- Class and Attributes of ndarray
- Basic Operations
- Activity-Slice It
- Copy and Views
- Mathematical Functions of Numpy

6) Scientific computing with Python (Scipy)

- Introduction to SciPy
- SciPy Sub Package - Integration and Optimization
- SciPy sub package
- Demo - Calculate Eigenvalues and Eigenvector
- SciPy Sub Package - Statistics, Weave and IO

7) Data Manipulation with Pandas

- Introduction to Pandas
- Understanding DataFrame
- View and Select Data Demo
- Missing Values
- Data Operations
- File Read and Write Support
- Pandas Sql Operation

8) Machine Learning with Scikit-Learn

- Machine Learning Approach
- Supervised Learning Model Considerations
- Scikit-Learn
- Supervised Learning Models - Linear Regression
- Supervised Learning Models - Logistic Regression
- Unsupervised Learning Models
- Pipeline
- Model Persistence and Evaluation

9) Natural Language Processing with Scikit Learn

- NLP Overview

- NLP Applications
- NLP Libraries-Scikit
- Extraction Considerations
- Scikit Learn-Model Training and Grid Search

10) Data Visualization in Python using matplotlib

- Introduction to Data Visualization
- Line Properties
- (x,y) Plot and Subplots
- Types of Plots

11) Web Scraping with BeautifulSoup

- Web Scraping and Parsing
- Understanding and Searching the Tree
- Navigating options
- Modifying the Tree
- Parsing and Printing the Document

12) Python integration with Hadoop MapReduce and Spark



- Why Big Data Solutions are Provided for Python
- Hadoop Core Components
- Python Integration with HDFS using Hadoop Streaming
- Python Integration with Spark using PySpark