Python for Data Analysis

Duration: 10 days
Prerequisites: Knowledge of any programming language

Day 1

Module 1: Introduction to Python

- Running Python • Hello, World!
- Literals
- Python Comments
- Variables
- Conceptual hierarchy and built-in object types.
- Core data types: int, str, float, bool
- Mutability and immutability.

Module 2: Functions and Modules

- Defining Functions
- Variable Scope
- Global Variables
- Function Parameters
- Returning Values
- Importing Modules Module

Day 2:

Module 3 Math:

- Arithmetic Operators
- Assignment Operators
- Built-in Math Functions
- The math Module
- The random Module

Module 4: Python Strings

- Quotation Marks and Special Characters
- String Indexing
- Slicing Strings
- Concatenation and Repetition
- Common String Methods
- String Formatting
- Formatted String Literals (f-strings)
- Built-in String Functions
- Introduction to Unicode
- ASCII vs Unicode
- Character Encodings
- Unicode Normalization

Day 3:

Module 5: Flow Control and Data Collections

- Conditional Statements
- Loops in Python
- Break and Continue
- Data collections (List, Tuples, Dictionaries and Sets)

Day 4 :

Module 6: Regular Expressions

- Introduction to Regular Expressions
- Overview of regex Syntax and Patterns
- Basic Syntax of Regular Expressions
- Meta Characters in Regular Expression

Module 7: Introduction to Python for Data Analysis

- Python's role in data analysis and analytics
- Jupyter Notebook/IDE setup/ Colab Notebook
- Error handling and debugging

Day 5:

Module 8: Numpy

- Numpy Arrays
- Array Indexing
- Array Slicing
- Array Broadcasting
- Mathematical operations and aggregations

Module 9 Data Acquisition and File Handling

- Reading from files (CSV, JSON, Excel)
- Writing to CSV, JSON, and Excel files
- Importing and exporting data

Day 6:

Module 10: Data Cleaning and Standardization

- Handling Missing and Erroneous Data
- Identifying null, duplicate, or inconsistent data
- Imputation techniques and outlier detection

Day 7:

Module 11: Data Transformation

- Reshaping, merging, and aggregating data with Pandas
- Applying string operations and datetime manipulations
- Standardization and scaling techniques

Module 12: Exploratory Data Analysis (EDA)

- Descriptive statistics
- Data summaries
- Grouping
- Data aggregation

Day 8:

Module 13: Data Visualization-Matplotlib

- Foundational Visualization Tools
- Creating charts with Matplotlib
- Line Plot
- Bar Plot
- Scatter Plot
- Enhancing plots
- Annotations
- Labels
- Legends

Day 9:

Module 14: Feature Engineering

- Introduction to Feature Engineering
- Importance of Feature Engineering in Machine Learning
- Differences between Raw Data and Features
- Types of Features
- Feature Creation
- Feature Selection
- Feature Transformation

Day 10:

Module 15: Statistics for Data Analysis

- Descriptive and Inferential Statistics
- Measures of central tendency and dispersion
- Probability distributions and hypothesis testing

Module 16: Time Series Analysis

- Introduction to Time Series
- Time Series Vs Cross sectional Data
- Stationarity Concept
- Autocorrelation & Partial Autocorrelation