Python for Data Analytics – 15 Hours

Module 1: Python Fundamentals & Programming Logic (3 Hours)

Goal: Build a strong foundation in Python syntax, logic, and structure.

- Introduction to Python and Jupyter Notebook – Understand how Python is used in data analytics; working in Jupyter Notebook environment for interactive coding.
- Data Types, Variables, and Operators – Explore numeric, string, and Boolean types; arithmetic, assignment, and logical operators.
- Conditional Statements (if, else, elif) – Apply decision-making constructs to control flow based on conditions.
- Loops (for and while) – Automate repetitive tasks using for-loops and while-loops with examples.
- Functions and Lambda Expressions – Write reusable code blocks using def functions and concise lambda functions for one-liners.

Module 2: Core Python Structures & File Handling (3 Hours)

Goal: Master Python's built-in data structures and handle external files.

- Lists, Tuples, Sets, and Dictionaries – Create, access, and manipulate various collections with use-case comparisons.
- String Manipulation and Formatting – Work with slicing, formatting, and key string methods for cleaning or analyzing text data.
- File Handling (Reading and Writing) – Read from and write to .txt and .csv files using open(), read(), write(), and context managers (with keyword).

Module 3: pandas for Data Analysis (3 Hours)

Goal: Learn to manipulate and analyze structured data using pandas.

- Introduction to pandas: DataFrames and Series – Create and explore Series and DataFrames; understand their properties and usage.
- Importing and Exporting Data – Load data from CSV/Excel files and save cleaned datasets using read_csv(), to_csv(), etc.
- Data Cleaning – Handle missing values with fillna(), dropna(), and remove duplicates.
- Data Aggregation – Use groupby(), .agg(), and other summarization techniques for grouped insights.

Module 4: NumPy for Efficient Data Handling (2 Hours)

Goal: Use NumPy arrays for fast numerical operations.

- Understanding NumPy Arrays – Create and manipulate arrays for large-scale data operations.
- Indexing, Slicing, and Broadcasting – Efficiently access and modify array elements with intuitive slicing and broadcasting rules.
- Mathematical and Statistical Functions – Perform array-wise operations: mean, median, standard deviation, etc.

Module 5: Data Visualization with matplotlib and seaborn (2 Hours)

Goal: Create insightful charts and plots for data exploration.

- Introduction to matplotlib and seaborn – Understand the syntax and styling of both libraries.
- Common Plot Types – Create Line, Bar, Pie, Histogram, and Scatter plots for data analysis.
- Customization – Add titles, labels, legends, colors, grid lines, and annotations to enhance readability.

Module 6: Working with Datetime and Text Data (2 Hours)

Goal: Analyze and extract insights from dates and textual data.

- Handling Datetime Objects – Convert strings to datetime, extract parts (day, month, year), and perform calculations (differences, ranges).
- Extracting Insights from Strings – Use pandas string methods (str.contains(), str.split(), str.replace() etc.) for filtering, cleaning, and analyzing textual data.