

PYTHON FOR DATA ANALYSIS

Edition 2025

Designed & Developed by Swapnjeet S (Lead Data Analyst – 10.5 Years Experience)

Python is a versatile and powerful programming language widely used in data analysis for its rich ecosystem of libraries, scalability, and ease of use. It is instrumental in data cleaning, preparation, and exploratory data analysis (EDA) using libraries like pandas, numpy, matplotlib, and seaborn. Python enables data analysts to perform statistical analysis, create interactive visualizations, and automate data workflows.

Beyond basic analytics, Python empowers users to develop and deploy custom solutions tailored to business needs. Its compatibility with tools like Tableau, Power BI, and Excel enhances its utility in the data pipeline. For interactive work, Jupyter Notebooks offer an ideal environment for exploratory tasks, while deployment frameworks like Flask or Dash enable sharing insights through web applications. Whether for natural language processing, forecasting, or data integration, Python is a must-learn tool that remains at the forefront of modern data analysis workflows.

Module 1: Introduction

- What is Python?
- Data Roles that use Python
- Python for Data Analysis

Module 2: Jupyter Notebooks

- Jupyter Notebooks Intro & Install
- Launching Jupyter & Creating a Notebook
- The Jupyter Interface
- Comments & Markdown
- The Print Function & Function Help

Module 3: NumPy

- Introduction to NumPy
- NumPy Arrays
- Coding Exercise Check-in: Creating NumPy Arrays

- NumPy Indexing and Selection
- Coding Exercise Check-in: Selecting Data from Numpy Array
- NumPy Operations
- Check-In: Operations on NumPy Array

Module 4: Pandas

- Introduction to Pandas
- Series - Part One
- Check-in: Labelled Index in Pandas Series
- Series - Part Two
- Data Frames
- Pandas - Conditional Filtering
- Pandas - Useful Methods - Apply on Single Column
- Pandas - Useful Methods - Apply on Multiple Columns
- Pandas - Useful Methods - Statistical Information and Sorting
- Missing Data - Overview
- Missing Data - Pandas Operations
- Group By Operations
- Combining Data Frames
- Pandas - Text Methods for String Data
- Pandas - Time Methods for Date and Time Data
- Pandas Input and Output - CSV Files
- Pandas Input and Output - HTML Tables
- Pandas Input and Output - Excel Files
- Pandas Input and Output - SQL Databases

Module 5: Matplotlib

- Introduction to Matplotlib
- Matplotlib Basics
- Understanding the Figure Object
- Implementing Figures and Axes
- Figure Parameters
- Subplots Functionality
- Legends
- Colours and Styles
- Advanced Matplotlib Commands (Optional)

Module 5: Seaborn

- Introduction to Seaborn

- Scatterplots with Seaborn
- Distribution Plots - Understanding Plot Types
- Distribution Plots - Coding with Seaborn
- Categorical Plots - Statistics within Categories - Coding with Seaborn
- Categorical Plots - Distributions within Categories - Coding with Seaborn
- Comparison Plots
- Seaborn Grid Plots
- Seaborn - Matrix Plots

Module 6: Data Visualization

- Matplotlib and basic plotting
- Creating line plots, bar plots, scatter plots, histograms, etc.
- Customizing plots with labels, titles, legends, etc.
- Introduction to Seaborn for advanced statistical visualization

Module 7: Python Projects

- Application of Python and data analytics concepts to a real-world project
- Data cleaning, pre-processing, and analysis
- Visualization of insights
- Building and evaluating predictive models

Course/ Study Links:

Complete Python Basics:

- ❖ <https://www.youtube.com/watch?v=wUSDVGivd-8>

Python Projects

- ❖ **Hotel Bookings:** <https://www.youtube.com/watch?v=obJZ1rB7TKc&t=2074s>
- ❖ **EDA Project from scratch:** <https://www.youtube.com/watch?v=obJZ1rB7TKc&t=2074s>

Practice Websites:

- ❖ **Hacker Rank:** <https://www.hackerrank.com/domains/python>
- ❖ **Leetcode:** <https://leetcode.com/problemset/>

Interview Questions/ Notes/ PPT's/:

- ❖ **Data Analyst Complete Material:**
 - If already purchased please refer to Python Folder in the link you received.

- If not purchased use below link: https://topmate.io/data_tutorials/498634
- This material is designed to help you master Python concepts, build real-world expertise, and get fully prepared for interviews with confidence.