

Python Data Analysis: Pandas, EDA & Matplotlib

Learn to use Python, Pandas and Matplotlib to clean, explore and visualise data, from basics to a final project.

5 Days

40 HOURS

13

MODULES

35

TOPICS

COURSE OVERVIEW

This course begins with Python fundamentals, covering data types, control flows, operators, and core structures such as lists, tuples, sets, and dictionaries, along with file handling, strings, iterators, generators, regular expressions, and object-oriented programming. It then introduces NumPy, Pandas, and Matplotlib for array operations, DataFrame slicing, and initial data visualisation. The focus moves to data preprocessing and exploratory data analysis, including cleaning, handling missing and categorical data, and univariate, bivariate, and multivariate analysis using scatter plots, pair plots, and box plots. It closes with advanced Pandas techniques, further data visualisation methods such as line plots, bar plots, pie charts, and heatmaps, and a project applying these skills to a complete data analysis workflow.

WHO THIS COURSE IS FOR

Data Analyst

Business Analysts

Data Scientist

LEARNING OBJECTIVES

- Master Python fundamentals for data manipulation and analysis
- Explore data types, control flows, and operators in Python
- Gain proficiency in data pre-processing and cleaning techniques
- Perform exploratory data analysis using Pandas and NumPy
- Develop skills in data visualization with Matplotlib

13 Modules, 35 Topics

01

Introduction to Python and Basics

TOPICS

- Definition & Applications
- Features, Versions & Working
- Anaconda & Different IDEs for Python
- Introduction to IDE's - Jupyter Notebook, Spyder & Google Colab

02

Data Types and Control Flows

TOPICS

- Literals, reserved words and input functions
- **Data Types:** int, float, bool, str
- **Decision Control Flows:** If / Nested If / If-else / If-elif-else
- **Control Flow Loops:** While loop, For loop, While-else, For-else
- **Operators:** Arithmetic, Relational or Comparison, Logical

03

Lists, Tuples, Sets, and Dictionaries

TOPICS

- Bitwise Operators, Assignment Operators, Ternary Operator
- **List:** Ways of Accessing Values, Traversing Elements, List Operations, List Methods, Membership Operator, List Comprehension
- **Tuples:** Creating Tuples, Ways of Accessing Values, Tuple Vs Immutability, Tuple Comprehension
- **Sets:** Creating Sets, Ways of Accessing Values, Manipulating and Accessing Sets, Set Operations
- **Dictionary:** Why Dictionary, creating a Dictionary, Accessing Values, Updating Dictionaries, Functions of Dictionary

04

File Handling and Strings

TOPICS

- **File Handling:** Types of Files, Opening and Closing Files, Writing, Appending, and Reading Files
- **Strings:** String Literals, Single (') & Double Quotes (""), Triple Quotes ("""), Raw Strings ("r'...' ") and Operations on strings
- **Dictionary:** Accessing Values, Updating Dictionaries, Functions of Dictionary

05

Iterators & Generators

TOPICS

- Iterator vs Iterable, Containers, Generators in Python

06

Regular Expressions

TOPICS

- Uses of Regular Expressions - Text Analytics, import re, Character Classes, Backslash, Alteration, Quantifiers

07

OOPS Concept

TOPICS

- Class, Classes and Object, Creating Object, Accessing Objects, Need and Use of Self, Class Method, __init__() constructor

08

Introduction to NumPy, Pandas & Matplotlib

TOPICS

- Introduction to NumPy, Install NumPy
- Array Creation, Array Reshaping, Indexing, Operations
- Introduction to Pandas, Slicing Data, Slicing DataFrame
- Data Visualization with Matplotlib

09

Introduction to Data Preprocessing

TOPICS

- Filtering DataFrame, Transforming DataFrame, Advanced Indexing
- Data Cleaning & Data Preprocessing

10

Exploratory Data Analysis (EDA)

TOPICS

- Data Cleaning Techniques, Handling Missing Data, Handling Categorical Data
- Introduction to EDA, 2D Scatter-plot, 3D Scatter-plot, Pair plots
- Univariate, Bivariate, and Multivariate Analysis, Box-plot
- Variance and Standard Deviation, Median, IQR (Interquartile Range)

11

Advanced Pandas and Data Visualization

TOPICS

- Advanced Pandas, Advanced Indexing, Data Preparation
- Handling Missing Data, handling Categorical Data, Data Cleaning

12

Data Visualization

TOPICS

- Introduction to Data Visualization, Plotting with Matplotlib
- Scatter Plots, Line Plots, Bar Plots, Pie Charts, Heatmaps

13

Project Work

TOPICS

- Problem Statement, Data Collection, Data pre-processing (Exploratory Data Analysis), Feature Engineering (optional), Data visualizations (Pandas, NumPy & Matplotlib), Project Final Outcome & findings

WHAT IS INCLUDED

Everything You Need to Succeed

- Koenig course material for every module
- Delivery by certified instructors and subject matter experts
- Hands-on exercises throughout the course
- Course completion certificate from Koenig

ABOUT KOENIG

Who We Are

Established in 1993, Koenig Solutions is a globally recognised training provider with 35+ technology partnerships, 5,000+ training programmes and a network of certified instructors and subject matter experts, supporting individuals and enterprise teams worldwide.

FLEXIBLE DELIVERY

Learn the Way That Suits You

Classroom Training

Live Online Classes (ILO)

Fly-Me-A-Trainer (FMAT)

Flexi (Self-Paced)

1-on-1 Training

Ready to Get Started?

Talk to a learning advisor about schedules, group training and customised delivery.

Visit koenig-solutions.com

✉ info@koenig-solutions.com