

Python for Data Scientists - 64 academic hours

Workshop Overview:

This course provides a strong, hands-on foundation in Python programming tailored specifically for Data Scientists. Participants will learn how to work efficiently with Python as a core analytical tool, focusing on data manipulation, exploratory data analysis (EDA), and practical workflows used in real-world data science projects.

The course is taught in a Jupyter Notebook / Google Colab environment and emphasizes hands-on practice, enabling participants to move from basic Python syntax to applied data analysis using industry-standard libraries.

What Participants Gain?

- Writing clean, readable, and well-structured Python code for data analysis
- Confidently working in Jupyter Notebooks and Google Colab environments
- Cleaning, manipulating, and exploring data using pandas and NumPy
- Applying EDA techniques to uncover patterns, trends, and insights
- Working with multiple data sources: files, databases, and web APIs
- Building a strong foundation for advanced Data Science, Machine Learning, and AI courses

Target Audience

- Aspiring Data Scientists
- Junior Data Scientists
- Data Analysts looking to strengthen their Python skills
- Professionals with analytical or quantitative backgrounds transitioning into data science roles

Prerequisites

- Basic understanding of data concepts (tables, variables, basic statistics)
- Familiarity with analytical thinking
- No prior Python programming experience is required

Workshop Syllabus

Section 1 – Python

- The working environment
- Data types
- Data structures (list, dictionary, etc.)
- Flow control (if-else, for-in, etc.)
- Textual interface
- Functions (inc. lambda)
- Working with files
- Object-Oriented Programming (OOP) basics
- Python API's
 - Python Standard Library
 - Modules and packages
 - datetime
 - Regular expressions

Section 2 – EDA

- Pre-processing with *pandas*
 - Basic concepts
 - Indexation and filtering
 - Aggregations and advanced manipulations
- Mathematical packages (*scipy*, *numpy*)
- Visualization packages (matplotlib, seaborn)
- Working with data resources (JSON files, databases & web)