

# Python for Data Analysis

**Lecturer: Ivan Renesto**

## Course language

English

## Course description and objectives

Python is a widely used high-level, general-purpose, interpreted, dynamic programming language.

Through this course you will learn how to manipulate, process, and clean data with Python, using its data-oriented library ecosystem and tools that will lay the foundations to let you become an effective data analyst.

At the end of the course, participants will be able to:

- work with arrays and vectorized computation
- work with tabular or heterogeneous data
- plot and visualize data

## Audience

The course is open to students of the Master of Science Programs at Bocconi University and is part of the Enhancing Experience activities. It is aimed at:

- those who want to approach the world of data analysis;
- students who want to acquire the basic knowledge to develop future expertise in the area of Data Science;
- those who are interested in facing advanced topics in Python or are planning to be part of projects where to extract information from a data set.

Upon successful completion of the course (attendance of at least 75% of the scheduled hours and passing the final exam), students will get 2 credits and an Open Badge, sharable across the web (LinkedIn) or personal CV.

## Prerequisites

Knowledge of Python basics, having attended the curricular course 30424 Computer Science, or the extracurricular course: Python start, or having equivalent knowledge and skills.

## Duration

16 hours

## Teaching mode

This course will be delivered exclusively in **synchronous, in-person sessions**. Remote participation will not be offered.

The **final test** of the course will take place on the **last day of the course**.

## Calendar

Lecture	Date	Time	Room
1	Tue 24/03/2026	18.15 - 19.45	N30 (L. Del Vecchio)
2	Wed 25/03/2026	18.15 - 19.45	N30 (L. Del Vecchio)
3	Tue 31/03/2026	18.15 - 19.45	N30 (L. Del Vecchio)
4	Thu 02/04/2026	18.15 - 19.45	N30 (L. Del Vecchio)
5	Tue 14/04/2026	18.15 - 19.45	N30 (L. Del Vecchio)
6	Thu 16/04/2026	18.15 - 19.45	N30 (L. Del Vecchio)
7	Tue 21/04/2026	18.15 - 19.45	N30 (L. Del Vecchio)
8	Wed 22/04/2026	18.15 - 19.45	InfoAS04/05

## Syllabus of the course

Lecture	Topics	Book reference
1	<b>Introduction to Visual Studio Code</b> <ul style="list-style-type: none"> <li>- Preliminaries</li> <li>- Install Visual Studio Code</li> <li>- Walk through the development environment</li> <li>- Built-in data structures and sequences.</li> </ul> <i>Exercises</i>	Ch. 1, 2, and 3
2	<b>Arrays and vectorized computation</b> <ul style="list-style-type: none"> <li>- NumPy basics</li> <li>- Working with multidimensional array objects</li> <li>- Indexing, slicing, and transposing arrays</li> <li>- Array-Oriented Programming</li> <li>- Mathematical and statistical methods.</li> </ul>	Ch. 4
3	<b>Plotting and visualization</b> <ul style="list-style-type: none"> <li>- Data visualization using matplotlib</li> <li>- Figures and Axes</li> <li>- Saving figures to file</li> <li>- Sub-plots</li> <li>- Multiple line plots</li> <li>- Colors, line styles, axes limits, labels plot title, legend and other chart elements</li> <li>- Histograms.</li> </ul>	Ch. 9
4	<b>Data manipulation with pandas</b> <ul style="list-style-type: none"> <li>- Pandas basics</li> <li>- Introduction to Series, DataFrame, Index objects</li> <li>- Essential functionalities of pandas library</li> <li>- Summary statistics methods</li> <li>- Data visualization using pandas.</li> </ul> <i>Exercises</i>	Ch. 5
5	<b>Problem requiring data analysis</b> <ul style="list-style-type: none"> <li>- Data loading, storage and file formats</li> <li>- Dataset analysis</li> <li>- Reading and writing data in text format</li> <li>- Interacting with Web APIs</li> <li>- Interacting with Databases via pyodbc.</li> </ul> <i>Exercises</i>	Ch. 6

Lecture	Topics	Book reference
6	<b>Data Cleaning and Preparation</b> <ul style="list-style-type: none"> <li>- Handling missing data</li> <li>- Data formatting and string manipulation</li> <li>- Data transformation (normalization and binning)</li> <li>- Categorical values</li> </ul> <i>Exercises</i>	Ch. 7
7	<b>Exploratory Data Analysis</b> <ul style="list-style-type: none"> <li>- Descriptive statistics</li> <li>- GroupBy mechanics</li> <li>- The analysis of variance</li> <li>- Correlation between different variables</li> <li>- Pearson correlation and correlation heatmaps.</li> </ul> <i>Exercises</i>	Ch. 8, 10, 12
8	<b>Final Exam</b>	

## Software used

Python, version 3.9+. Current version is 3.13.5.

Python interpreter can be downloaded for free from here:

<https://www.python.org/downloads/>.

Microsoft Visual Studio Code (VS Code). Current version is 1.101.2.

Visual Studio Code is a free coding editor that helps to start coding quickly. It supports multiple programming languages, and the use of a Python web-based interactive computing platform (Jupyter Notebook).

Supported in: Windows 10 and 11 (64-bit and Arm64), macOS 11.0+ versions with Apple security update support (Intel chip, or Apple silicon), Linux Ubuntu, Debian, Red Hat, Fedora, or SUSE.

VS Code can be downloaded from here: <https://code.visualstudio.com/download>.

## Suggested bibliography

McKinney W., *Python for Data Analysis, second edition. Data Wrangling with Pandas, NumPy and IPython*, O'Reilly Media, 2017

## Available seats

This activity is limited to **110** participants and reserved to **students of the Master of Science Programs**. Registrations cannot be carried out once this number has been reached or after closing of the registration period.