

# Data Visualization with Python

Lecturer: Ivan Renesto

## Course language

English

## Course description and objectives

Visualization of analytical results is probably one of the most important aspects that people want to highlight, either in a presentation or in a report. It is the result of complex data pipelines, which might require merging, transforming and wrestling data. The course covers some of the most famous libraries for data visualization in Python. We will cover the basics, using Matplotlib, and then move to more advanced libraries to create more sophisticated plots.

Upon successful completion of this course, students should be able to:

- perform basic data transformations suitable for data visualization;
- translate analytical results into charts and plots;
- be familiar with Matplotlib and related Python libraries.

## Audience

The course is targeted at:

- students who aim at improving their skills on data visualization using Python
- those who are curious on Data Visualization

However, note that the course is open to a restricted audience of Bocconi students. More specifically, only students enrolled in:

- second or third year bachelor programs,
- any of the Bocconi Master programs, or
- any PhD or SDA programs

## Prerequisites

Knowledge of Python programming language, having passed the curricular course 30424 Computer Science or having equivalent knowledge and skills.

## Guidelines

**Registration:**

You can sign up for the course only through the yoU@B student Diary, in the " **sign-up for various activities**" box (please note that the box appears only when registrations open. Before then it will not be visible).

You can only cancel your registration by Diary **no later** than the registration deadline for the course itself. No other ways of cancellation are allowed.

Registration will be confirmed a few days before the start of the course through a message posted in the yoU@B student Diary.

### Attendance:

- Attendance of **75% or more** of class hours: obtainment of the **Open Badge**
- Attendance of **less than 25%** of class hours: **blacklisting**

### Duration

12 hours

### Teaching mode

This course will be taught **in person** only. Distance mode will not be provided.

### Calendar

Lecture	Date	Time	Room
1	Wed 18/09/2024	18.15 - 19.45	2.1 (Sarfatti)
2	Thu 19/09/2024	18.15 - 19.45	2.1 (Sarfatti)
3	Tue 24/09/2024	18.15 - 19.45	2.1 (Sarfatti)
4	Thu 26/09/2024	18.15 - 19.45	2.1 (Sarfatti)
5	Tue 01/10/2024	18.15 - 19.45	3.3 (Sarfatti)
6	Thu 03/10/2024	18.15 - 19.45	1.2 (Sarfatti)

**Note:** lessons will be held in the traditional room and **all the students have to bring their own device.**

## Syllabus of the course

Lecture	Topics	Book reference
1	<p><b>Introduction to Visual Studio Code</b></p> <ul style="list-style-type: none"> <li>- Preliminaries</li> <li>- Walk through the development environment</li> </ul> <p><b>Introduction to Data Visualization</b></p> <ul style="list-style-type: none"> <li>- Importance of Data Visualization</li> <li>- Key Python libraries</li> <li>- Matplotlib history and architecture</li> <li>- Introduction to Jupyter Notebooks</li> <li>- Generating line plots</li> <li>- Exploring different types of plots</li> <li>- Read and import CSV files with Pandas</li> <li>- Generate simple plots using Pandas</li> </ul> <p><i>Exercises</i></p>	Ch. 1, 2, and 3
2	<p><b>Matplotlib Anatomy and Plot Customization</b></p> <ul style="list-style-type: none"> <li>- How to explore a dataset in Pandas</li> <li>- Building dataframes</li> <li>- Creating dataframes via dictionaries</li> <li>- Manipulate dataframes               <ul style="list-style-type: none"> <li>o select data via <code>iloc[ ]</code> and <code>loc[ ]</code></li> <li>o add columns</li> <li>o rename columns</li> <li>o delete columns</li> <li>o sort records</li> </ul> </li> <li>- Anatomy of a Plot               <ul style="list-style-type: none"> <li>o multiple-line plots</li> <li>o Figure and Axes</li> <li>o title and labels</li> <li>o markers</li> <li>o colors</li> <li>o legend</li> <li>o text</li> <li>o line styles</li> <li>o subplots</li> <li>o layout</li> </ul> </li> <li>- Customizing a Plot</li> </ul> <p><i>Exercises</i></p>	Ch. 4

Lecture	Topics	Book reference
3	<b>Pie and Bar Charts</b> <ul style="list-style-type: none"> <li>- Bar colors</li> <li>- Bar labels</li> <li>- Stacked bar chart</li> <li>- Grouped bar chart</li> <li>- Horizontal bar chart</li> <li>- Broken horizontal bar chart</li> <li>- Errorbars</li> <li>- Plotting categorical variables</li> <li>- Pie chart</li> <li>- Histograms</li> <li>- Set axes range (xlim, ylim)</li> <li>- Plots with different scales</li> </ul> <p><i>Exercises</i></p>	Ch. 4
4	<b>3D Plotting</b> <ul style="list-style-type: none"> <li>- Moving from 2D to 3D plots</li> <li>- plot( ), scatter( ), plot_surface( )</li> <li>- changing camera angle with init_view</li> <li>- surfaces modeling with Wireframe 3D plots</li> <li>- Data preparation with NumPy.meshgrid( )</li> <li>- 3D Charts: overview and use cases               <ul style="list-style-type: none"> <li>o Plot 2D data on 3D plot</li> <li>o 3D bar charts</li> <li>o 2D bar charts in different planes</li> </ul> </li> <li>- Exploring other 3D plots</li> </ul> <p><i>Exercises</i></p>	Ch. 4
5	<b>Practicing with other Data Visualization libraries</b> <ul style="list-style-type: none"> <li>- Seaborn               <ul style="list-style-type: none"> <li>o Understanding Axes and Figures</li> <li>o Colormaps and Hues</li> <li>o Basic plots</li> </ul> </li> <li>- Bokeh               <ul style="list-style-type: none"> <li>o Creating plots in Bokeh</li> <li>o Visualization examples</li> </ul> </li> <li>- GeoPandas</li> </ul> <p><i>Exercises</i></p>	Ch. 5, 6, 10
6	<b>Final Exercise</b>	

## Software used

Python, version 3.9+.

Python interpreter can be downloaded for free from here:

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

Microsoft Visual Studio Code (VS Code), version 1.81+.

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 (32-bit and 64-bit), macOS versions with Apple security update support, Linux Ubuntu Desktop 10.64, Debian 9, Red Hat Enterprise Linux 7, CentOS 7, Fedora 34.

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

### Suggested bibliography

Nelson D., **Data Visualization in Python**. *Explore and Manipulate Data and Create Engaging Interactive Plots with 9 Python Libraries*, StackAbuse, 2020-2021

### Available seats

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