

Python for Excel

Lecturer: Gianluigi Pertusi

Language

English

Course description and objectives

Nowadays, huge datasets are available to everybody and they are often too large to fit into a spreadsheet. Although Microsoft introduced a few tools to analyse big datasets and to automate repetitive procedures, these tools are very specific and can be used only in the Microsoft environment. Another way to solve these problems involves the use of Python, which is a modern and widespread programming language. In this course Python will be used to support Excel in data analysis, overcoming Excel's limitations and simplifying data cleaning and processing.

Upon successful completion of this course, students will be able to use Python to:

- Read data from Excel files
- Analyse data using Python's libraries
- Create data report and charts through Python's libraries
- Print the result of their analysis on Excel files

Audience

The course is aimed at:

- students willing to improve their skills in data analysis
- students who need additional tools to perform quantitative analyses on large datasets
- students who want to exploit Python's potential to simplify data analysis and overcome Excel's limitations

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

- their second or third year bachelor programs, or

- any of the Bocconi Master programs, or
- any PhD programs

Prerequisites

Having passed the Computer Science curricular exam or having a good experience in coding with Python is highly recommended.

Duration

12 hours

Teaching mode

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

Calendar

Lecture	Date	Time	Room
1	Mon 26/06/2023	16.30 – 18.00	N31 (Velodromo)
2	Mon 26/06/2023	18.15 - 19.45	N31 (Velodromo)
3	Tue 27/06/2023	16.30 – 18.00	N31 (Velodromo)
4	Tue 27/06/2023	18.15 - 19.45	N31 (Velodromo)
5	Wed 28/06/2023	16.30 – 18.00	N31 (Velodromo)
6	Wed 28/06/2023	18.15 - 19.45	N31 (Velodromo)

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 References
1	Introduction to custom modules The datetime class Data analysis with pandas <ul style="list-style-type: none"> - DataFrames and Series - Data manipulation 	Ch. 3 Ch. 5
2	Data analysis with pandas <ul style="list-style-type: none"> - Combining DataFrames - Data aggregation - Filtering DataFrames <i>Exercises</i>	Ch. 5
3	Importing and exporting DataFrames with pandas Time series analysis with pandas <ul style="list-style-type: none"> - DatetimeIndex - Time zones - Time series manipulations 	Ch. 5 Ch. 6
4	Time series analysis with pandas <ul style="list-style-type: none"> - More time series manipulations Plotting tools <i>Exercises</i>	Ch. 5 Ch. 6
5	Excel files manipulation <ul style="list-style-type: none"> - pandas - openpyxl Useful tools for Excel reporting <ul style="list-style-type: none"> - os module 	Ch. 7 Ch. 8
6	<i>Final exercise: Excel reporting</i>	

Software used

Python 3.9 (or greater)

Microsoft Excel 2019 (Office365)

Suggested bibliography

Felix Zumstein, *Python for Excel – A modern Environment for Automation and Data Analysis*, O'Reilly, 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.

Spreadsheet paths

This course can be intended as part of a wider path:

