

MaDaS
Master in Data Science for Complex Economic Systems

Introduction to Python

Matteo Calabrese
Chiara Salvemini

Learning Objectives

The module is an introduction to the Python programming language. At the end of this module, students should be familiar enough with Python language to read and write non-trivial Python code, as well as to exploit specific Python packages in particular those related to scientific computing and treatment of economic datasets.

Course Content

The module is an introduction to the Python programming language and mainly deals with the following topics:

Monday, 10th September 2018 (6h)

1. Introduction to Python language
2. *Hello World*, variables, structures and functions
3. Specific packages: Numpy, Matplotlib, SciPy

Tuesday, 11th September 2018 (6h)

4. Dataframes & Pandas
 - csv format, I/O dataframe
 - Pandas packages
 - *How to*: descriptive statistics
 - *How to*: regressions
 - *How to*: plots and figures from dataframe

Thursday, 13th September 2018 (4h)

5. Discussion day
 - exercises correction
 - R & Python

Course Methodology

The course will be held in the computer lab. Students will be taught how to write their own code through concrete examples. Students are encouraged to actively interact in class and will be asked to work on problem sets assigned during the lessons.

Course Materials

Slides for theoretical parts of the lessons will be made available to the students, exercises will be developed using online, notebook systems such as colab.research.google.com. All materials will be available online, and we encourage students to download it and use it on-the-fly during the course hours.

IMPORTANT FOR STUDENTS: all materials work online, especially exercises will be implemented on an (online) notebook format, so no additional software needs to be installed on local machines. However, to avoid any problem related to a slow internet connection in the computer lab, we advice students to install on their own machine a running version of Python like Anaconda (see refs.) for Windows. For Mac and Unix OS just download the software and follow the standard installation. We stress again that during the course we shall use the online version of our codes, having a running python in local is just a precaution.

Reference

- Michael Dawson, Python Programming for the Absolute Beginner
- Allen Downey, Think Python. How to Think Like a Computer Scientist (available online for free at greenteapress.com/thinkpython/thinkpython.pdf)
- Wes McKinney, Python for Data Analysis
- Software:
 1. <https://www.python.org/downloads/>
 2. <https://www.anaconda.com/download/>

Many code examples will be presented during the course.

Course Evaluation

Students will be evaluated (pass/fail) on the basis of group projects that will be individually discussed in detail with each of them. Projects will be assigned during the course.

About the Instructors

Matteo Calabrese is a Ph.D. doctorate in Astrophysics, and he is currently working at the Astronomical Observatory of the Autonomous Region of Aosta Valley. His interests however do not cover only stars and galaxies, but complex systems in general. He is applying machine learning techniques to study how materials' surfaces degrade in time, in the context of artistic and cultural heritage sites.

Chiara Salvemini got a double degree in Physics of Complex Systems from the University of Turin and from the university Paris Diderot. She is currently a research fellow in a joint-venture between the University of Aosta Valley and the Astronomical Observatory. Chiara is implementing a web-crawling system to collect data and prices from popular hotels-booking sites, to quantify the connection between revenue management and the perceived price fairness.