# IT coding and applications

## Prof. Andrea Pozzi; Prof. Emanuele Goldoni

***COURSE AIMS AND INTENDED LEARNING OUTCOMES***

Quantitative data analysis of complex systems in economics and social sciences require a multidisciplinary methodological training. Computer science, especially its algorithmic aspects, constitutes one of the fundamental building blocks of such preparation.

This course aims to provide students modern computational capabilities for quantitative data analyses and is organized into two modules: Basic Programming and IT concepts (module A - 30 hours of lecture and 20 hours of laboratory) and Introduction to algorithms and methods for Data Analysis (module B - 30 hours of lectures and 10 hours of laboratory).

In the first module students are introduced to computer programming, starting from the foundations of coding up to basic data structure and algorithm implementations. Programming laboratories based on the Python programming language.

The second module focuses on algorithms for data analysis and visualization by providing hands-on experience on a real dataset. In addition, the lab sessions will give an overview of basic implementation of algorithms for automated data processing and extraction of useful trends and indexes. At the end of the course students will be able to realize simple Python coding using popular toolkits to exploit basic data analysis capabilities.

*Intended learning outcomes*

* Understanding the components of computer programming
* Working knowledge of python programming language
* Familiarity with basic data structures and algorithm implementations
* Working knowledge of the steps involved in the development of an application for data analysis

***COURSE CONTENT***

***Module A***

1. Introduction to computational thinking;
2. Python basics: variables, data types, operators, expressions, flow control, loops;
3. Functions and modules;
4. Data structures: list, tuples, set, dictionaries;
5. File I/O and exception handling;
6. Introduction to Numpy, Matplotlib, and Pandas;
7. Introduction to Streamlit for data visualization;

*Module B*

1. Hands-on experience on plots and data visualization
2. Introduction to linear regression
3. Introduction to clustering algorithms for unsupervised machine learning
4. Introduction to classification algorithms for supervised machine learning

***READING LIST***

* Downey, A. (2012). *Think python*. " O'Reilly Media, Inc.".
* VanderPlas, J. (2016). *Python data science handbook: Essential tools for working with data*. " O'Reilly Media, Inc.".
* Lecture notes and online contents

***TEACHING METHOD***

The course will include lectures and class exercises based on traditional teaching and teach by example principles. It is strongly advised to attend to lectures for working on case studies and examples, and for revising materials.

The course also involves lectures and exercise sessions using personal notebooks/PC-labs.

Active participation, and ongoing personal study are required.

***ASSESSMENT METHOD AND CRITERIA***

The final examination comprises two written assessments, each covering the material presented in Modules A and B, respectively. Both assessments will include a combination of open-ended and multiple-choice questions, as well as coding exercises. Each assessment will contribute 50% to the overall grade. The highest possible score for each assessment is 16 points, with a total score of at least 31 points required to achieve honors (Laude).

***NOTES AND PREREQUISITES***

Prof. Andrea Pozzi covers the first module while Prof. Emanuele Goldoni the second one.

A very basic knowledge on computer systems is required. Attendance is strongly recommended.