Programming Methodologies for Data Analysis.
Luca Di Gaspero
Kevin Roitero

The course aims at providing a comprehensive working knowledge of different computer programming styles using the Python language. After introducing language basics, the course will cover specifically functional programming in Python, which is the fundamental tool of several data-processing libraries and frameworks. Moreover, a number of relevant modules from the Python standard library will be introduced, with a particular focus at the data-analysis and machine learning ecosystem. Also, we will cover best practices to use programming for executing frequent tasks and will investigate how to use Python in different scenarios, from small scripting tasks to medium-scale projects.

The course will be split in a set of conceptual lectures 18h and a set of guided practice 12h.

The detailed topics that will be covered by the course include:

- Basic data types, control flow, structured data types (tuples, sets, lists, dictionaries, strings), functions;
- Comprehensions and generators;
- Functional programming style and higher-order functions (map, filter, reduce);
- Input/Output (file manipulation, network access), Exceptions, modules and packages;
- Data representation and manipulation libraries (pandas, json, xml);
- Web scraping;
- Multi-processing and multithreading;
- Data visualization (seaborn / plotly) and dashboards (streamlit / dash).