

Basic Python for Data Analysis

In this course we will cover the basics of working with the Python programming language and the Interactive Python notebooks that are normally used for scientific data analysis. You will get an introduction to the basic workflow of data science and the most frequently used libraries for data analysis and visualization.

Who is this course for?

This course is meant for beginners who have never before done any programming for data analysis or for people who have some experience in R and want to acquire new tools.

How can this course benefit your career?

- You will get the basics of the python language and Jupyter Lab
- You will learn how to apply Python tools to your needs for data analysis
- You will learn how to get quick insights from your data
- You will learn how to clean and pre-process your data before analysis
- You will learn to do basic visualisations to better understand your datasets
- You will see how to create a valid statistical model
- You will learn how to apply python notebooks for your scientific publishing process

How can I prepare for this course?

If you prefer to bring your own computer, please install the Anaconda Distribution of python 3.8, you can download it for free at <https://www.anaconda.com/products/individual>

Instructors: Luis Rodil-Fernandez, Leticia G. Leon and David B.H. van Zessen

| | Day 1 (14 Sept) | Day 2 (15 Sept) | Day 3 (16 Sept) |
|--|--|---|---|
| theme | <i>essentials</i> | <i>libraries & workflows</i> | <i>bioinformatics</i> |
| morning block #1 9.30 - 11.00 | working with python notebooks in Jupyter Lab, computational thinking and basic notions of python | overview of basic libraries, working with libraries and managing them | practical examples of working with pandas, numpy and matplotlib |
| 11.00 - 11.15 | short break | short break | short break |
| morning block #2 11.15 - 12.30 | basic notions of python (cont.), flow control | basic workflows for data manipulation | working with biological data |
| 12.30 - 13.30 | lunch break | lunch break | lunch break |
| afternoon block 13.00 - 16.00 | data structures, functions | introduction to pandas and numpy | biostatistics, statistical modelling |
| 16.00 - 17.00 | exercises | exercises | exercises |