Introduction to Python

Python is a general-purpose programming language popular among data scientists and statisticians. In this one-day introductory course, participants will learn to import, summarise and visualise their data. At each step, we avoid using "magic code", and stress the importance of understanding what Python is doing.

Course Outline

- **Introduction to Python**: A brief introduction to the Python language and development tools for writing Python
- **Data Types**: An introduction to common Python data types, such as lists and dictionaries as well as Pandas Series and DataFrame objects.
- **Visualisation**: Create standard statistical graphics such as scatter plots and histograms from your data.
- **Summary Statistics**: Apply common statistical summaries to your data sets using the power and speed of the numpy package.
- **Data Manipulation**: Become comfortable with manipulating your data, filtering, merging and joining, grouped aggregation.
- **Data I/O**: Load and save data from a variety of different data sources such as CSV files, Excel files and binary data storage formats.

Learning Outcomes

Session 1:

By the end of session 1 participants will…

- be comfortable working in the Jupyter notebook integrated environment.
- have developed a familiarity with Python's most common data types:
  - integers
  - floats
  - strings
  - booleans
  - lists
- understand the differences between a function, method, and attribute.
- have learned about Python's package system and how to import packages.
- be able to perform basic data manipulation tasks with numpy.

Session 2:

By the end of session 2 participants will…

- be familiar with Python's pandas DataFrame and Series objects.
- be able to manipulate, extract and summarise data with the pandas package.
- know how to use the matplotlib package to make data visualisations.

- understand how to read data into Python from multiple sources, and write analysed data back out.

This course does not include:

- Advanced data analysis, wrangling and manipulation techniques.
- Enough knowledge to compose **complex** visualisations and interactive plots, see our Python for Data Visualisation course.
- For loops, function writing or if statements, for more programming skills in Python see our Programming with Python course for this.
- Machine and deep learning modelling techniques, see our website for courses on this topic.

Attendee Feedback

- "The presenter really did an excellent job on explaining the content in the course and I believe that the actual content is very useful to know. I will definitely be using it"
- "It was well structured, the materials were clear and professionally produced, good enough to be followed on their own"