

# Course : Python Data Science, manipulating and visualizing data

optional TOSA® certification

*Practical course - 4d - 28h00 - Ref. IYT*

*Price : 2630 CHF E.T.*

★★★★★ 4,6 / 5

BEST

Data analytics is an ever-expanding multidisciplinary field. It relies on scientific methods, algorithms and processes that Python has mastered thanks to a particularly rich ecosystem. Today, it has become the reference language for data analysis, whatever the format. Our training course enables you to get to grips with Python's tools, libraries and modules, and rapidly acquire data science skills using this language.

## Teaching objectives

At the end of the training, the participant will be able to:

- ✓ Get an overview of Python's scientific ecosystem
- ✓ Learn about the essential scientific libraries for data analytics
- ✓ Be able to manipulate large data sets with Python
- ✓ Understanding the benefits of datavisualization
- ✓ Visualizing data with Python

## Intended audience

Engineers, developers, researchers, data scientists, data analysts and anyone who wants to learn about the scientific world of Python.

## Prerequisites

Python language practice.

## Practical details

### Hands-on work

Numerous exercises are used to illustrate the topics.

### Teaching methods

Active teaching, feedback and demonstrations are provided by the trainer to enable participants to put them into practice more quickly.

### PARTICIPANTS

Engineers, developers, researchers, data scientists, data analysts and anyone who wants to learn about the scientific world of Python.

### PREREQUISITES

Python language practice.

### TRAINER QUALIFICATIONS

The experts leading the training are specialists in the covered subjects. They have been approved by our instructional teams for both their professional knowledge and their teaching ability, for each course they teach. They have at least five to ten years of experience in their field and hold (or have held) decision-making positions in companies.

### ASSESSMENT TERMS

The trainer evaluates each participant's academic progress throughout the training using multiple choice, scenarios, hands-on work and more. Participants also complete a placement test before and after the course to measure the skills they've developed.

## Course schedule

### 1 Introduction to the scientific Python ecosystem

- Overview of Python's scientific ecosystem: must-have libraries.
- Know where to find new bookstores and assess their sustainability.
- The main open source tools and software for data analysis.
- Why use scientific distribution, Anaconda.
- Understand the benefits of a virtual environment and know how to use it.
- IPython interpreter and Jupyter server.
- Best practices for getting your data analytics project off to a good start with Python.
- Scientific file formats and libraries for manipulating them.

#### Hands-on work

Setting up the development environment: installing Anaconda, creating a virtual environment, exporting and duplicating an environment, using Jupyter notebooks.

### 2 The SciPy Stack

- The SciPy Stack is the foundation of the essential scientific libraries on which all the others are based.
- NumPy: numerical calculation and linear algebra (vectors, matrices, images).
- SciPy, based on NumPy for statistics, functional analysis, geospatial analysis, signal processing, etc.
- Pandas: tabular data analysis (CSV, Excel, etc.), statistics, pivots, filters, search...
- Matplotlib: the essential data visualization library.

#### Hands-on work

Measure the performance of the NumPy installed by your Linux and that of Anaconda. Image processing with NumPy. First plots. Statistical analysis of CSV files. First mapping elements. Fourier transforms.

### 3 Visualization libraries

- Overview of Python visualization libraries: 2D/3D, desktop/web, statistics, cartography, big data...
- Desktop-oriented libraries: Matplotlib, Pandas, Seaborn.
- Web-oriented libraries: Bokeh, Altair, Plotly...
- 3D libraries: Plotly, pythreejs, ipyvolume...
- Cartographic libraries: Cartopy, folium, ipyleaflet, Bokeh, cesiumpy...
- Big data libraries: datashader, Vaex...

#### Hands-on work

Multiple exercises with some of the libraries presented. Big data, cartographic, 2D and 3D visualization.

#### TEACHING AIDS AND TECHNICAL RESOURCES

- The main teaching aids and instructional methods used in the training are audiovisual aids, documentation and course material, hands-on application exercises and corrected exercises for practical training courses, case studies and coverage of real cases for training seminars.
- At the end of each course or seminar, ORSYS provides participants with a course evaluation questionnaire that is analysed by our instructional teams.
- A check-in sheet for each half-day of attendance is provided at the end of the training, along with a course completion certificate if the trainee attended the entire session.

#### TERMS AND DEADLINES

Registration must be completed 24 hours before the start of the training.

#### ACCESSIBILITY FOR PEOPLE WITH DISABILITIES

Do you need special accessibility accommodations? Contact Mrs. Fosse, Disability Manager, at [psh-accueil@orsys.fr](mailto:psh-accueil@orsys.fr) to review your request and its feasibility.

#### 4 Data visualization

- The benefits of data visualization
- Using PyViz and the HoloViz ecosystem.
- Overview of SuperSet, Mayavi, Paraview and VisIt tools.

##### Hands-on work

Continue to use visualization libraries and manipulate tools.

#### 5 Scientific file formats and the handling of voluminous data

- Overview of the main scientific file formats: NetCDF, HDF5, GRIB, JSON, PARQUET, MATLAB, CGNS...
- Handle big data with Dask, Vaex and Xarray.

##### Hands-on work

Handle data exceeding GB, read and write NetCDF/HDF5 files. Visualization of climate data, satellite images, creation of videos/graphic animations.

## Dates and locations

### REMOTE CLASS

2026 : 23 June, 29 Sep., 24 Nov.