Publication date : 04/12/2024

Course : Python, Pandas advanced: data analysis with powerful techniques

The Python library for data analysis and synthesis

Practical course - 2d - 14h00 - Ref. PND

Price : 1500 CHF E.T.

4,9 / 5

You've already discovered Python, Pandas: the library for data analysis. If you'd like to be guided by an expert in the field to explore this library, which is essential to all data science projects, this training course is for you.

Teaching objectives

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

- Master the Pandas library for data analysis
- Understanding the subtleties of groupbys
- Handling pivot tables and cross-tabulations
- Learn how to speed up calculations with Pandas
- Learn about best practices in Data Science

Intended audience

Developers, engineers and anyone analyzing data with development skills.

Prerequisites

Mastery of Python

Course schedule

PARTICIPANTS

Developers, engineers and anyone analyzing data with development skills.

PREREQUISITES

Mastery of Python

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.

1 An overview of the Pandas bookshop

- A reminder of the basics of Pandas.
- Read data files (csv, excel, SQL, parquet).
- Dataset description and simple statistical analysis.
- Implement different analyses and visualizations depending on the type of data.
- Handling missing data.
- Date manipulation for Time Series.
- String management.
- Implementation of Data Science best practices.

Hands-on work

Set up a virtual environment for Data Science, read csv and xls files, brief statistical analysis and description of datasets.

2 Mastering the subtleties of groupbys

- Groupbys for understanding modalities in datasets.
- Single-index groupby with classic aggregation functions.
- Customize aggregation functions.
- Groupby with multiple indices.
- Difference between apply and transform functions.
- Reminders on anonymous functions.

Hands-on work

On 2 economic datasets, practice groupby and data visualization. Creation of a toy dataset and use of groupby.

3 Pivot tables and cross-tabulations

- Aggregation functions and pivot tables.
- Contingency matrix.
- Cross-tabulation.

Hands-on work

Use 2 sets of economic data to apply pivot tables and cross-tabulations.

4 Table joins

- Notions of axes.
- Concatenation.
- Merge according to one or more keys.
- Joint in relation to indices.

Hands-on work

On 2 economic data sets, put into practice the different types of joins.

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 to review your request and its feasibility.

5 Faster computing with Pandas

- Loop over rows and columns.
- Back to basics with NumPy.
- Examples of multiprocessing with the Modin library.
- Examples of multiprocessing with the Numba library.

Hands-on work
On a large dataset, put into practice the various concepts covered in the course.

Dates and locations

REMOTE CLASS
2026 : 4 June, 17 Sep., 19 Nov.