

	PRODUCT SHEET	Identification: Trainevo-Ficheprod-01
		Revision Date: 11/2021
		Version number: 01

Course Title: Master Data Analysis with Pandas, Matplotlib Python

This training allows you to learn how to collect, clean and analyze data with Python using Datascience libraries.

Target audience

This course is for anyone who wants to learn how to analyze data with Python.

Prerequisite:

Basics in statistics.

Reference:	Category:	Duration:
Trainevo-0 92022-IT-08	IT	5 days
Language(s):	Training organization:	Number of participants (min/max):
French		8

Goal(s)

Collect different types of data.

Clean data using the Pandas library, Numpy.

Visualize data using Matplotlib, Seaborn libraries etc ...

Make descriptive / Multivariate statistics, Regressions.

Create dashboards, etc...

Program

Python:

- Reminders about Python
- Reminders on descriptive statistics

Libraries in data science 1:

- Data manipulation with pandas library, numpy
- Types of data
- Tools for loading data
- Describe and transform data
- Describe and transform unstructured data
- Temporal data processing
- Extract descriptive statistics

Data science libraries 2:

- Data visualization with Matplotlib and seaborn
- Scatter plot
- Constructions of histograms, bar charts etc ...
- Motion graphics
- Installation and use of Cartopy (Cartography)
- Interactive charts
- Scrape data on a web page
- Dashboard creation

Handling missing data:

- Causes of missing data
- Identify and manage missing data

Introduction to Machine Learning:

- Linear regression
- Logistic regression

Application exercises.

Data analysis project.

Method

This course takes place in three stages:

- Presentation of concepts,
- Examples & Application Exercises (Data Analysis),"
- Exercises & quizzes – correction of exercises.

20% theory & 80% practice.

Evaluation:

Make a project containing all the tools discussed in class.

Submission of a certificate of attendance.