

Course code: **PYTHON/ANA**

Course title: **Introduction to data analysis in Python**

Days: 3

Description:

Course intended for

The course is designated for data analysts who wish to go beyond the basic analytical tools and begin to use a wide range of analytical libraries available in Python. The course is also addressed to programmers who wish to begin their adventure with data analysis. The course allows for decreasing the distance between the profession of an analyst and a programmer. In addition, participants may better understand the use of analytical tools in the production language environment.

Course objective

The participants get familiarised with Python and the use of various analytical tools available in that language. The course features use of a combination of various tools, which in turn enables execution of more complex analyses by using an extensive range of techniques. Furthermore, algorithms in Python may very often be easily transferred to the production environment in a relatively short time.

Course strengths

The course is conducted by people dealing with the problem of data analysis using Python in their everyday work, holding practical experience in that area. Therefore, the course often goes beyond the available, though often scattered, materials. Besides, the programme is regularly updated because of fast development of solutions discussed during the course.

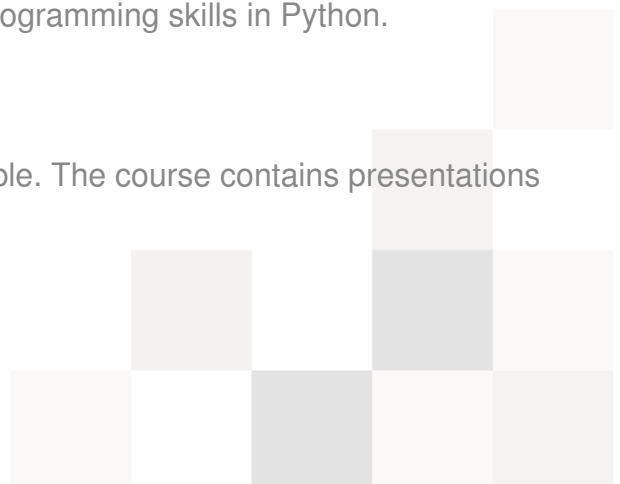
Requirements

The course requires basic programming skills in any language (a version including an introduction to Python - 1 additional day) or basic programming skills in Python.

Course parameters

3 working days, 3*7 working hours, group 8-10 people. The course contains presentations and coding workshops.

Course curriculum:



- Introduction
 - Purpose of the course
 - Python
 - History
 - Characteristics
 - Comparison with other languages
 - Why Python?
- Basic environment installation
 - Available versions
 - Python distributions
 - Python installation
 - Libraries installation
 - PyPI
 - Setuptools, Pip
 - VirtualEnv
 - interpreter
 - basic
 - IPython
- Introduction to Python (optional day)
 - Basic characteristics of the language
 - Syntax
 - Data structures



- Instructions
- Functions
- Objects
- Modules and packages
- Basic library
- IPython - interactive Python
 - Operating modes
 - Basic interactive command execution
 - History of commands
 - IPython Notebook
 - Advanced use
- Numerical data analysis with NumPy
 - Boards, vectors, matrixes
 - Functions
 - Data analysis
 - Recording and reading files
 - Linear algebra
- Introduction to Pandas
 - Data structures
 - Basic operations
 - Working with data
 - Basic statistics
- Reading, recording and storing data



- Text files
- Binary files
- Internet sources
- Databases
- Data processing
 - Cleaning
 - Joining
 - Rotation
 - Transformations
 - Working with text data
- Data visualisation
 - Matplotlib
 - Concept
 - Basic charts
 - Advanced functions
 - Image saving
 - Pandas
 - Review of available visualisations
 - Other packages
- Data aggregation and grouping
 - Grouping
 - Aggregation
 - Grouped data transformations



- Time series
 - Date and time
 - Basic time series
 - Ranges, frequencies and displacements
 - Time zones
 - Periods
 - Sampling and frequencies
 - Charts
 - Time windows

