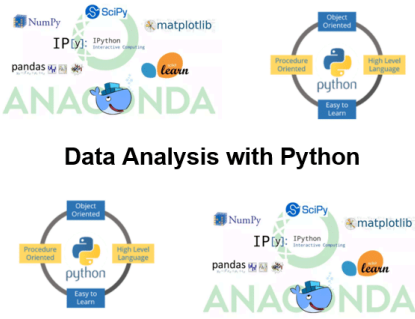
Data Analysis with Python

Duration: 4 days **Price: € 2650**

Open Schedule



Audience Course Data Analyse with Python

The course Data Analysis with Python is intended for data analysts who want to use Python and the Python libraries in Data Analysis projects.

Prerequisites training Data Analyse with Python

To participate in this course knowledge of and experience with any programming language or package such as SPSS, Matlab or VBA is desirable. The course starts with a discussion of the principles of the Python programming language.

Realization course Data Analyse with Python

The theory is discussed on the basis of presentation slides. Illustrative demos clarify the concepts. The theory is interchanged with exercises. The Anaconda distribution with Jupyter notebooks is used as a development environment. Course times are from 9:30 to 16:30.

Official Certificate Data Analysis with Python

After successful completion of the course participants receive an official certificate Data Analysis with Python.

Content Course Data Analysis with Python

In the course Data Analysis with Python you will learn how to use the Python language and Python libraries in Data Analysis projects.

Python Overview

The course Data Analysis with Python starts with a bird's eye view of the Python syntax aspects that are important in Data Analysis projects. Variables, data types, functions, flow control, comprehensions, classes, modules and packages are discussed. The operation of the Jupyter notebooks, the IPython shell and installing Python packages in Anaconda are also treated.

Numpy

Next the course Data Analysis with Python pays attention to the NumPy package with which large data sets can be processed very efficiently. NumPy's ndarray object and its methods are treated and attention is paid to the different array manipulation techniques with broadcasting and vectorized operations.

Pandas

Then use of the Pandas library for data analysis is on the schedule of the course Data Analysis with Python. The pandas library introduces two new data structures in Python that use Numpy and are therefore fast. The data structures are DataFrame and Series and extensive details are given on how to use them for data analysis when inspecting, selecting, filtering, combining and grouping data.

MatPlotLib

Also discussed in the course Data Analysis with Python is the MatPlotlib library, which is closely integrated with NumPy and is a very powerful tool for creating and plotting complex data relationships.

Scikit-Learn

Finally attention is paid to the essentials of the Scikit-Learn library for modeling. The course Data Analysis with Python uses many practical examples and shows how one- and two- and three-dimensional data sets can be visualized.

Modules Course Data Analysis with Python

Module 1 : Python Language Syntax	Module 2 : Functions and Modules	Module 3 : Classes and Objects
Python Features	Pass by Value and Reference	Creating Classes
Running Python	Scope of Variables	Creating and Using Objects
Anaconda Distribution	EFAP principle	Accessing Attributes
IPython Shell	What are comprehensions?	Property Syntax
Interactive and Script Mode	Lambda Operator	Constructors and Destructors
Python Data Types	Filter, Reduce and Map	Encapsulation
Numbers and Strings	List comprehensions	Inheritance
Sequences and Lists	Set and Dictionary comprehensions	super Keyword
Sets and Dictionaries	Creating and Using Modules	Checking Relationships
Python Flow Control	import Statement	issubclass and isinstance
Exception Handling	from…import Statement	Overriding Methods
Module 4 : Numpy	**Module 5 : Pandas**	**Module 6 : Data Manipulation**
NumPy Numerical Types	Pandas DataFrame	Indexing Data Frames
Data Type objects	Import Data	.loc and .iloc Accessor
dtype attributes	Inspect Data	Slicing and Indexing a Series
Slicing and Indexing	Data Visualization	Filtering with Boolean Series
Array comparisons	DataFrame Data Types	Zeros and NaNs
Manipulating array shapes	Indexing and selection	all and any Nonzeros
Stacking and Splitting arrays	Data operations in pandas	Using map Function
any(),all(), slicing, reshape()	Missing Data	Hierarchical Indexing
Manipulating array shapes	Hierarchical Indexing	Rearranging Data
Methods of ndarray	Plotting with Pandas	Reshaping by Pivoting
Views versus copies	Combining Datasets	Transformation and Aggregation
ravel(),flatten(),transpose()	Exploratory Data Analysis	Grouping Data
Module 7 : MatplotLib	**Module 8 : Time Series**	**Module 9 : SciKitLearn Essentials**
Simple Plots	Indexing Pandas Time Series	SkiKit Learn library
Plot format String	Reading and Slicing Times	Machine learning essentials
Subplots	Using a DatetimeIndex	Supervised and Unsupervised
Histograms	Reindexing the Index	Feature matrix
Logarithmic Plots	Separating and Resampling	Target array
Scatter plots	Rolling mean and Frequency	Estimator API
Fill between	Resample and Roll with it	Hyperparameters
Legend and Annotations	Manipulating Time Series	Fit method
Three Dimensional Plots	Method chaining and Filtering	Predict method
Contour Plots	Missing values and Interpolation	Model Selection
Transformations	Time Zones and Conversion	Linear Regression
Projections	Plotting Time Series	Logistic Regression

SpiralTrain BV
Standerdmolen 10, 2e verdieping
3995 AA Houten

info@spiraltrain.nl
www.spiraltrain.nl
Tel.: +31 (0) 30 – 737 0661

Locations
Houten, Amsterdam, Rotterdam, Eindhoven, Zwolle, Online