

# DATA SCIENCE WITH PYTHON

Learn and perform data mining and machine learning efficiently, using Python in this self-paced course.

Data science is one of the most sought after jobs of the century, and there is a huge talent/skill gap of over half a million developers. The main elements of data science are acquiring, storing, cleaning, processing and visualizing data and this course is for those who want to get an understanding of how to perform each of these tasks. If you've got some programming or scripting experience, this course will teach you the techniques used by real data scientists in the tech industry - and prepare you for a move into this lucrative field. This comprehensive course includes hands-on Python code examples you can use for reference.

Learn more about this topic now. View the webinar [Python - The World's Most Useful Language](#)

## What You'll Learn

- Set up your data science toolbox using a Python scientific environment on Windows, Mac, and Linux
- Get data ready for your data science project
- Manipulate, fix, and explore data in order to solve data science problems
- Set up an experimental pipeline to test your data science hypotheses
- Choose the most effective and scalable learning algorithm for your data science tasks

- Optimize your machine learning models to get the best performance
- Visualize the results of your analysis using Python's Matplotlib library
- Train efficient machine learning models using Decision Trees and Random Forests

### Who Needs to Attend

If you know the ins and outs of Python and are exposed to data analysis, then this course will help you with practical approach in dealing with data mining, machine learning and tools like Apache Spark, NumPy, Pandas and Matplotlib efficiently.

### On-Demand Outline

- A glance at Python's essential libraries
- The data science process
- Data loading and pre-processing with pandas
- Data pre-processing with Numpy; Introducing EDA
- The detection and treatment of outliers
- Visualization and presentation using seaborn, ggplot and matplotlib recommenders system
- data mining and machine learning techniques  
machine learning using spark
- A/B testing and T-tests