

Data Analysis with Python

At a Glance

In this course you will learn about: Data Acquisition How to Obtain Basic Insight from a Dataset Data.

LEARN TO ANALYZE DATA WITH PYTHON

Learn how to analyze data using Python. This course will take you from the basics of Python to exploring many different types of data. You will learn how to prepare data for analysis, perform simple statistical analyses, create meaningful data visualizations, predict future trends from data, and more!

You will learn how to:

- Import data sets
- Clean and prepare data for analysis
- Manipulate pandas DataFrame
- Summarize data
- Build machine learning models using scikit-learn
- Build data pipelines

Data Analysis with Python is delivered through lecture, hands-on labs, and assignments. It includes following parts:

- Data Analysis libraries: will learn to use Pandas DataFrames, Numpy multi-dimensional arrays, and SciPy libraries to work with a various datasets. We will introduce you to pandas, an open-source library, and we will use it to load, manipulate, analyze, and visualize cool datasets. Then we will introduce you to another open-source library, scikit-learn, and we will use some of its machine learning algorithms to build smart models and make cool predictions.

COURSE SYLLABUS

Module 1 - Importing Datasets

- Learning Objectives
- Understanding the Domain
- Understanding the Dataset
- Python package for data science
- Importing and Exporting Data in Python

- Basic Insights from Datasets

Module 2 - Cleaning and Preparing the Data

- Identify and Handle Missing Values
- Data Formatting
- Data Normalization Sets
- Binning
- Indicator variables

Module 3 - Summarizing the Data Frame

- Descriptive Statistics
- Basic of Grouping
- ANOVA
- Correlation
- More on Correlation

Module 4 - Model Development

- Simple and Multiple Linear Regression
- Model Evaluation Using Visualization
- Polynomial Regression and Pipelines
- R-squared and MSE for In-Sample Evaluation
- Prediction and Decision Making

Module 5 - Model Evaluation

- Model Evaluation
- Over-fitting, Under-fitting and Model Selection
- Ridge Regression
- Grid Search
- Model Refinement

REQUIREMENTS

- Some Python experience is expected
- Python for Data Science