

# Python for Data Extraction and Data Mining Syllabus

## Learning Objectives

- to provide an overview of available data analysis tools in Python ecosystem
- to give knowledge about data analysis pipeline
- to practice how to use analytical tools in various tasks

## Expected Learning Outcomes

- understand the steps of the analytical process
- use basic Python modules for data analysis (numpy, pandas, matplotlib)
- perform exploratory data analysis
- select appropriate visualizations for data
- build predictive models for clusterization, regression and classification tasks
- prepare dataset before training the model
- select appropriate metric for model evaluation
- tune hyper-parameters of the model

## Course Contents

- **Introduction to Data Analytics in Python**  
*This is an introductory section that describes such key areas as the analytical process, how data is created, stored, accessed, and how the organization works with data. It also covers data analysis tools available in the Python ecosystem.*
- **Descriptive Analytics**  
*Descriptive analytics is a preliminary stage of data processing that includes exploratory data analysis and data visualization.*
- **Predictive Analytics**  
*This section covers basic machine learning tasks like clustering, regression and classification. It also includes the machine learning pipeline steps from feature engineering to metric selection and hyper-parameter optimization.*

## Assessment Elements

- Final Test
- Intermediate Tests
- Programming Assignments

## Interim Assessment

- Interim assessment (4 module)  
 $0.3 * \text{Final Test} + 0.3 * \text{Intermediate Tests} + 0.4 * \text{Programming Assignments}$