Python2
Data Analysis Using Python

Description:

The three-day course aims to familiarize students with the possibilities of working with data, to perform transformations, analysis and visualization in Python. Its content is to learn to use packages and modules for work with data. It focuses mainly on the Pandas, NumPy and Dask libraries and its possibilities in data processing and publication.

Students will be able to:

- Set up the environment and necessary components
- Work with data
- Perform data transformations
- Perform data visualizations

Course requirements:

- Basic work with computer and basics of programming
- Python fundamentals

This course is intended for:

Developers, users, analysts, data architects, managers and IT staff

Literature:

All participants will get internal OKsystem materials.

Hardware:

Classrooms are equipped with high-performance computers with Internet access and the possibility of wireless connection.

Syllabus:

Module 1: Introduction to Data Handling

- Lesson: Environment, tools, settings
- Lesson: Data Manipulation
- Lesson: Packages, modules, libraries

Module 2: Introduction to pandas

- Lesson: Installing pandas and other components

- Lesson: The Basics of Pandas
- Lesson: Objects
- Lesson: Data import and processing

Module 3: NumPy

- Lesson: NumPy Installation
- Lesson: Working with Data
- Lesson: Data transformation and data analysis

Module 4: Pandas Objects and Working with Data

- Lesson: Working with Data
- Lesson: Data Handling and Transformation
- Lesson: Reading and Writing Different Types of Data Formats
- Lesson: Cleaning and filtering data
- Lesson: Grouping, aggregating data and pivoting data
- Lesson: Working with time series

Module 5: Dask

- Lesson: Dask module introduction
- Lesson: Data manipulations and transformations

Module 6: Data Visualization

- Lesson: Preparation and configuration
- Lesson: Using pandas, mathplotlib and other modules
- Lesson: Visualizations
- Lesson: Configuring and editing visualizations