Friedrich-Alexander-Universität
Technische Fakultät

Department of Computer Science
Computer Science 6
(Data Management)

Knowledge Discovery in Databases with Exercises
Winter Semester 2025/2026

Exercise Sheet 1:
Introduction to Python and pandas

About this Exercise Sheet

This exercise sheet is a gentle introduction to the technical tools that we will use repeatedly in the exercise throughout the semester: Python and pandas.

In contrast to all other exercises, we only recommend participating in this exercise if you have no previous experience with Python and/or pandas or if you do not feel confident using them.

Preparation

Before participating in the exercise, you must prepare the following:

1. **Install Python and pip on your computer**
   - Install Python 3.8 or higher on your computer. A good guide on the installation process can be found at `https://realpython.com/installing-python/`.
   - If your Python installation doesn't come with pip (the package installer for Python), install pip on your computer. You can find more information on the installation process at `https://pip.pypa.io/en/stable/installation/`.
2. **Download provided additional files**
   - Download `Additional-Files-Student.zip` from StudOn
   - Extract it to a folder of your choice.

Exercise 1: Getting started

Before we can start with the actual exercise, we have to perform some basic steps. These will be similar for all Python based exercises:

1. **Install required Python packages**
   - Open a terminal and navigate to the folder where you extracted the files.
   - Run the command `pip install -r requirements.txt` within the extracted additional files folder to install the required Python packages.
2. **Start the Jupyter Notebook server**
   - Run the command `jupyter notebook` within the extracted additional files folder to start the Jupyter Notebook server[1].
   - There should be a new tab in your browser with the Jupyter Notebook interface.

Exercise 2: Get to know Python

This exercise utilizes a Jupyter Notebook:

1. Open `Python.ipynb` in the Jupyter Notebook interface.
2. Take a look at the tasks (blue boxes) in the notebook and try to solve them.

Exercise 3: Get to know pandas

This exercise utilizes a Jupyter Notebook:

1. Open `pandas.ipynb` in the Jupyter Notebook interface.
2. Take a look at the tasks (blue boxes) in the notebook and try to solve them.

[1] If you have problems starting the Jupyter Notebook server, try `python -m notebook` as an alternative command.