

Assignment Description

Foundations of Data Analysis

Summer Term 2025

Lecture: Prof. Dr. Claudia Plant

Assistance: Kevin Sidak

Tutor: Mikołaj Adamus

Unsupervised Lab Assignment (Lab 2)

The goal of this assignment is to practice implementing data mining techniques in Python and to get to know useful packages.

These slides contain important information about the format of the assignment, deadlines, rules and remarks on using AI tools. Please read these carefully.

If you have questions, technical problems or personal issues, please let us know as soon as possible. Feel free to post a question on the dedicated Moodle forum or contact our tutor **Miki** (mikolaj.adamus@univie.ac.at).

- 1 Submitting your solutions
- 2 Jupyter notebooks
- 3 Peer-Review
- 4 Grading
- 5 Rules - Academic honesty
- 6 Citing correctly
- 7 AI tools

Submit your solutions as a zip file on Moodle. The files you submit should be at the root of the zip file (not in folders), except the datasets in the data folder. The zip file should contain your solutions (notebooks) and the necessary datasets. In case you created a Python environment (e.g. with venv), please do not submit the environment folder.



Submission and Peer-Review (Lab 2)

Submission Deadline: **25.06.2025 23:59**

The tasks are in Jupyter notebooks. Please solve the programming exercises and reports in those.

In each notebook you can find which non-native Python packages you are allowed to use. Please **do not import additional non-native packages**. You may use any native package from the standard library such as [itertools](#) or [collections](#).

Before the submission restart the kernel and run all the cells to check if your notebooks run without any problem. **If we cannot run your code, points will be deducted according to how difficult it is to fix it.**

Please use relative paths to open the datasets and submit the datasets with your notebooks.

If you are new to Jupyter, please watch the Python tutorial videos posted on Moodle in the Tutorials section.

After the submission deadline, the **peer review process** on Moodle will start and you will be assigned two other students' work to evaluate **until 02.07.2025 23:59**.

The questions will be available at the same place as the submission:



[Submission and Peer-Review \(Lab 2\)](#)

Although the feedback you receive will not impact your grade, they can offer useful ideas how to improve your work.

If you did not submit any solutions, you will not be assigned to do the peer-review automatically. If you still wish to do it and get some points, please write an email.

For the Lab assignment **you can get a total of 100 points.**

In the notebooks you will find tasks for 94points in total, but you can reach maximum **90 points for your submission.**

You can get **10 points for your peer-reviews** based on the quality of your answers.

After the deadline, we will grade your solutions as soon as we can but it will take some time. If you need the results early for an important reason (e.g. visa application), please write an email.

You are required to solve the tasks individually. **Sharing your solutions or copying those of other students is forbidden** and will lead to an 'X' in your transcript of records.

You are allowed to look up similar problems on the Internet to get you some ideas but **you need to solve the tasks yourself based on the ideas discussed in the lectures, copying solutions is not allowed.**

You need to mark and cite (details on the next slide), **which external sources you used for which part of your assignment.** Not doing so will lead to an 'X' in your transcript of records.

Please note that we will make use of plagiarism and code checking tools (e.g. Turnitin).

It should be clear, for which exact parts of your code/report you used external materials. Put the citation right next to the marked code/text part:

- ▶ To cite a printed material, state the full title of it, the authors and the relevant sections or page numbers.
- ▶ To cite an online material, state the full URL of its website and the relevant parts of it. ("I used the Internet." is not a citation.)

Including references show that you spent time with solving the tasks correctly and you can also refer to them if plagiarism is suspected in your answers. (= It is for your benefit!)

AI tools are not perfect. On the contrary, they often **return texts with very low quality, false information and parts from third sources without citation**. We therefore strongly discourage using them.

“ChatGPT sometimes writes plausible-sounding but incorrect or nonsensical answers.” - notes OpenAI [on their blog](#). Those who cannot decide whether the outputs are correct (e.g. students) should therefore not rely on such tools.

To encourage you to solve the tasks yourself, **we will** assess the quality of your reports and **deduct extra points if your answers consequently do not answer the questions**. If you solve the tasks yourself paying attention to the task description, this will not affect you. However, **if you copy solutions from AI tools, you will most likely submit irrelevant and nonsensical solutions**.