Opened: Thursday, 2 October 2025, 12:00 AM
Due: Wednesday, 8 October 2025, 11:59 PM

Assignment: Mini Project on GitHub (15 points)

Your task is to create a small data science project and publish it on GitHub. The focus is **reproducibility** and demonstrating **git skills**, not complexity. Work individually. Submission: the link to your public GitHub repo.

Requirements:

- **Reproducibility (3 pts):** Follow the principles of reproducibility discussed in class: clear folder structure, raw data included, all analysis steps included in code, output created automatically. Goal is that an independent peer should be able to reproduce it with no (or minimal) modifications.
- **Environment (3 pts):** Project should be easy to reproduce (ideally with one click). You should define the environment or requirements for this.
- **Documentation (3 pts):** Provide a well-structured README.md using proper Markdown formatting. The README should briefly describe your project, data, folder structure, requirements, steps to reproduce and other necessary notes.
- **Analysis (3 pts):** Perform a very basic analysis. Minimal cleaning if needed, then one descriptive statistic or a simple graph is enough. Code must be clean and easy to follow.
- **Version Control (3 pts):** Use Git properly with at least 3 meaningful commit messages (additional to the init).

Notes:

- The goal is not to write complex code but to demonstrate the main concepts we covered in class.
- If you need support, use the provided class materials. Filling in the gaps on your own is part of the learning process. Copying or the use of AI tools is not allowed and will result in point deductions.
- Completing this exercise carefully will also help you prepare for the exam.

[Add submission]

Submission status

Submission status	No submissions have been made yet
Grading status	Not graded
Time remaining	4 days 2 hours remaining
Last modified	-
Students' comments	❯ [Comments (0)](#)