Introduction to Data Science

Course project planning

Onur Varol

@onurvarol
www.onurvarol.com
www.varollab.com

Sabancı Üniversitesi

VRLLab

Why individualized project?

- You can develop a portfolio of data science projects
    - Much more interesting in job / internship / graduate school applications
        - Time spent on the project will be useful to you
- You can learn about yourself and gain insights from your own data
- It’ much more fun for everyone to see these projects

Projects' "must-have" components

- Project will be done using YOUR OWN DATA
- Diverse set of techniques: EDA, visualization, machine learning models
- Github page for the project material
    - You don't need to share raw data, but analysis scripts needs to be there (Hint: search for .gitignore)
    - Website / short video / README file / report
- Participation to peer-evaluation

What to present

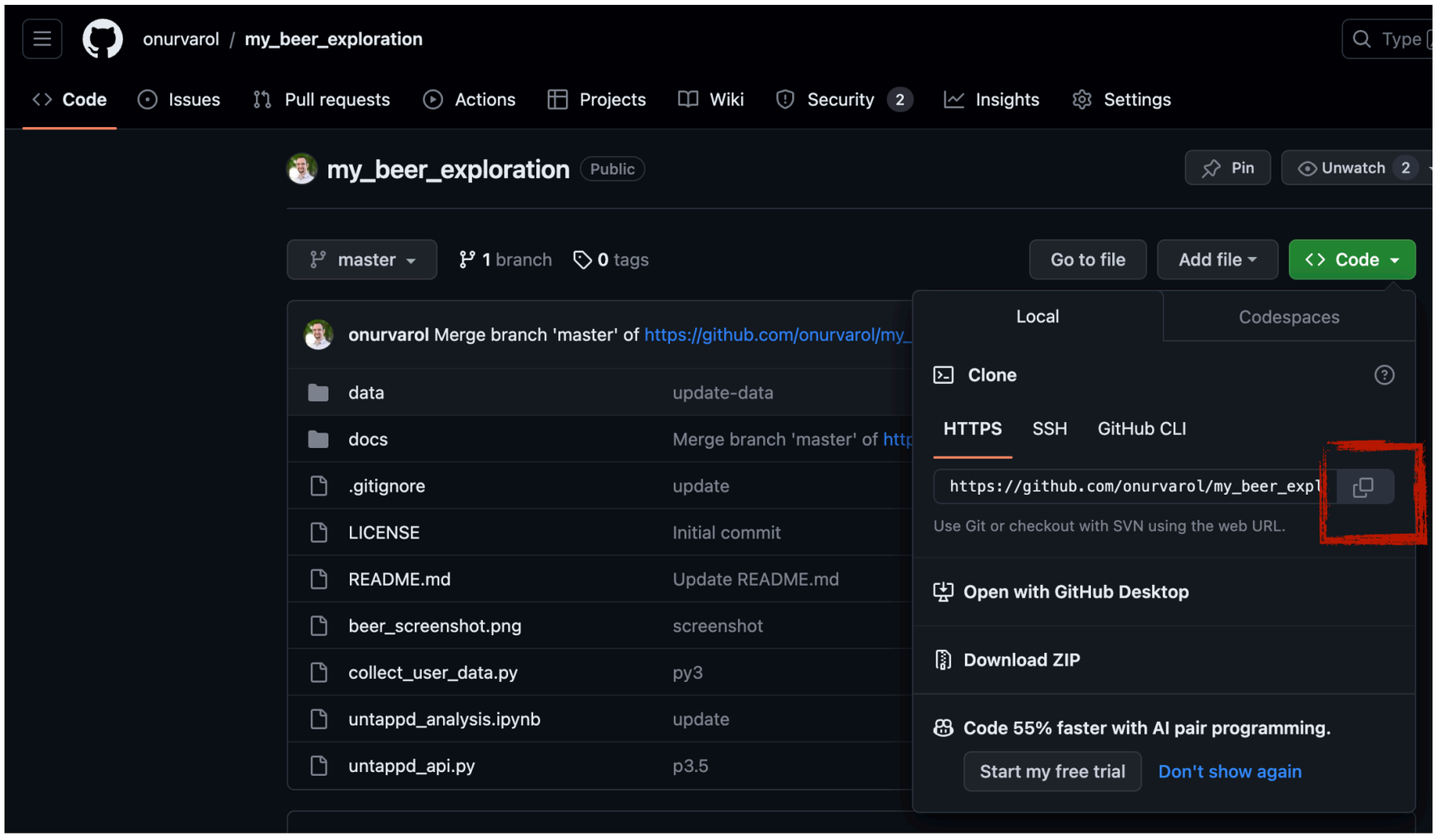
- Motivation: Why you are working on this project?
- Data source: Where did you get this data? How did you collect it?
- Data analysis: Techniques used, different stages of the analysis
- Findings: What you learned about yourself?
- Limitations and future work: What could be done better? Do you have any future plans about your project?

Evaluation

- Participation to project proposal (10%)
- Code and material review (20%)
- Quality of the project presentation (30%)
- Peer-evaluation
  - Grading other projects (10%)
  - Grades received from peers (30%)

What is the submission format

- We will **only** ask for a **public** Github repository URL.



- We will automate cloning Github repositories with a script. Mistakes in the URL format will cause you points!

Deadlines

- Since there are several hard deadline for grade submission, there will not be any late submission. **PLEASE PLAN ACCORDINGLY!**

- **Deadline for submitting the project material: 19 Jan.**

- **Deadline for peer-evaluations: 22 Jan.**

- Deadline for Onur to submit grades: 25 Jan. at 3PM