Lecture 4: numpy, Git/GitHub

LING 1340/2340: Data Science for Linguists

Na-Rae Han
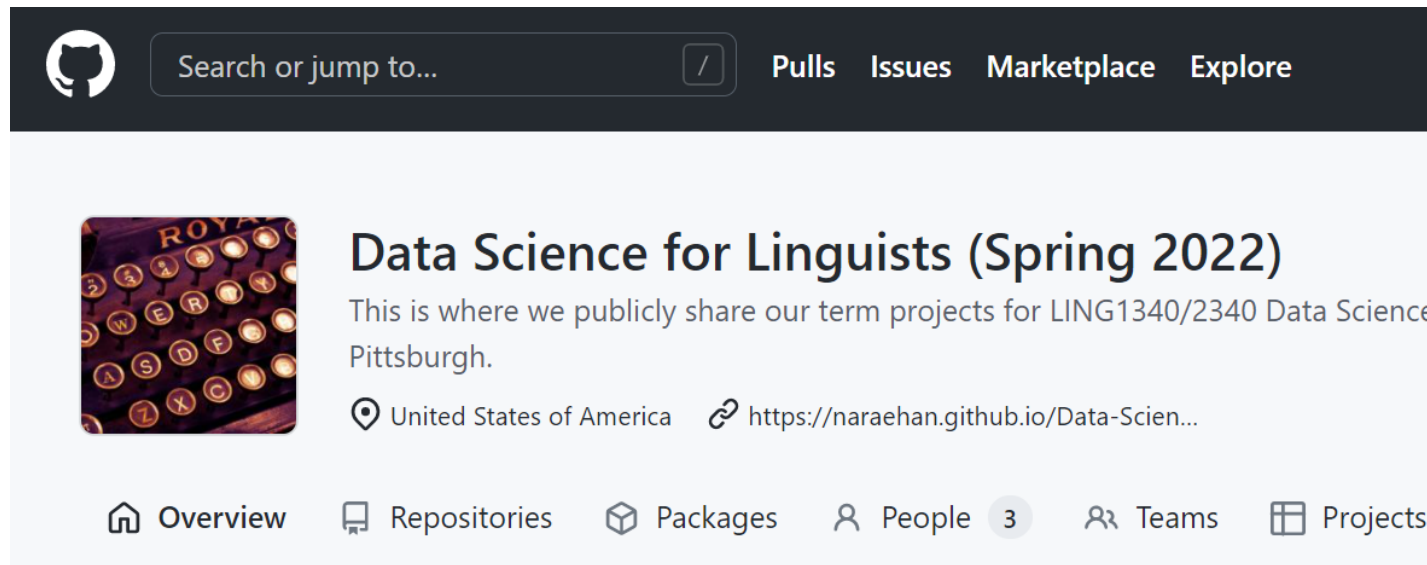
Objectives

▶ Term project overview: what's expected?

▶ To-do2 review: study notes in JNB, git and GitHub

▶ Python's numpy library hands-on practice

▶ Tools:

  ◆ Git and GitHub
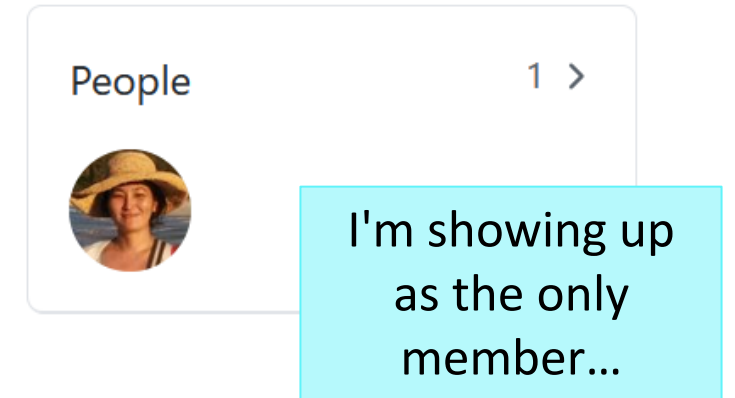
  ◆ Jupyter Notebook

GitHub Class Organization
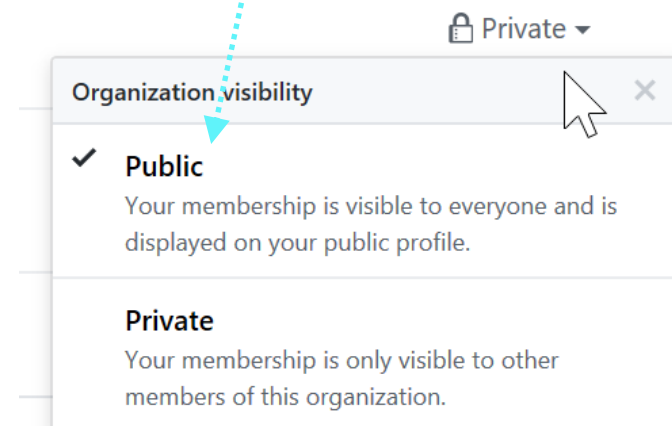
▸ https://github.com/Data-Science-for-Linguists-2022

▸ So we can:

- ◆ have everyone in one spot.
- ◆ have all class materials in one spot.
- ◆ have everyone's term project in one spot.
- ◆ share *private* repos as a group.

Back to Class-Exercise-Repo

▶ https://github.com/Data-Science-for-Linguists-2022/Class-Exercise-Repo

▶ To-do2

◆ Git and GitHub work flow: are you getting the hang of it?

◆ Who has the best study notes?

Term Project overview

▸ We went over the term project guidelines:

- https://naraehan.github.io/Data-Science-for-Linguists-2022/project

▸ Also talked about what makes a good/successful project, ways to come up with project ideas

numpy practice with the Simpsons

▶ In Class-Exercise-Repo, `activity1/` folder:

- ◆ You will find `numpy_simpsons_BLANK.ipynb`

- ◆ Make a copy for yourself as `numpy_simpsons_YOURNAME.ipynb`:

  - ◆ `cp xyz_BLANK.ipynb xyz_narae.ipynb`

- ◆ Let's get to work!

`cp file1 file2`
makes a copy of file1 as file2.

Wrapping up

- To-do 3: due next Tuesday.
  - Study pandas, make your own JNB.
  - Try two spreadsheet files, one your choice. (Make it small! Less than 100 rows!)

- Start thinking about project ideas!