Lecture 4: numpy, Git/GitHub

LING 1340/2340: Data Science for Linguists

Na-Rae Han
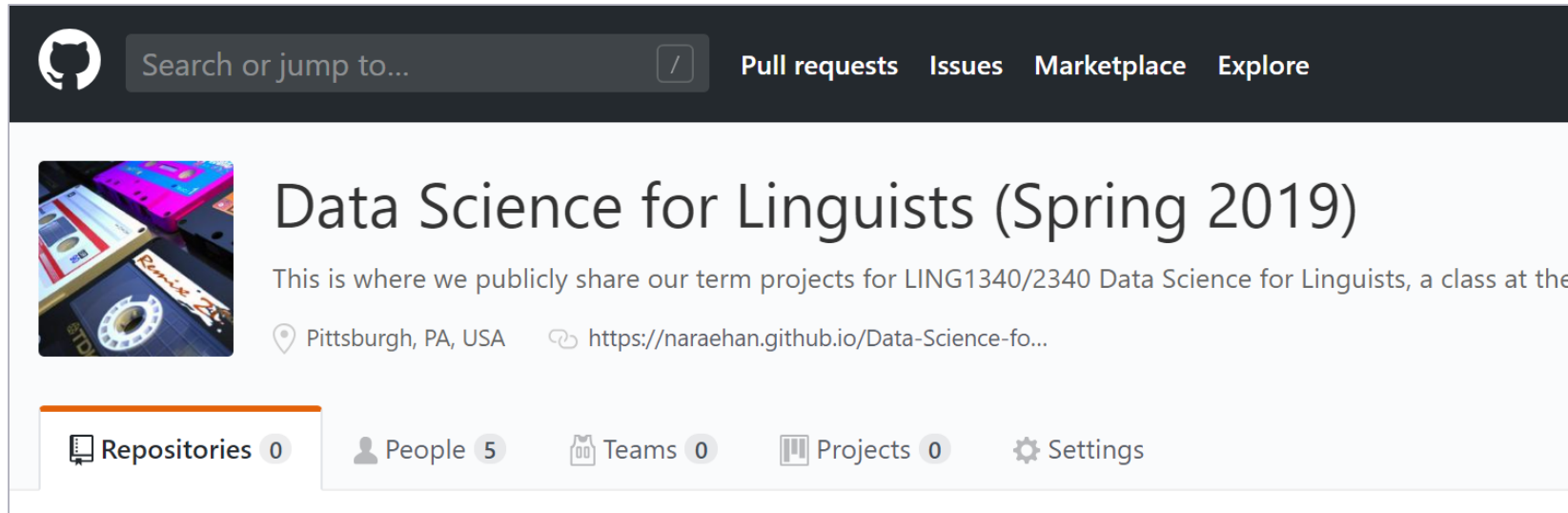
Objectives

▸ **GitHub class organization**

◆ Housekeeping our repos and forks

▸ **To-do2 review: study notes in JNB**

▸ **Python's numpy library**

▸ **Tools:**

◆ Git and GitHub

◆ Jupyter Notebook

Introducing… GitHub Class Organization
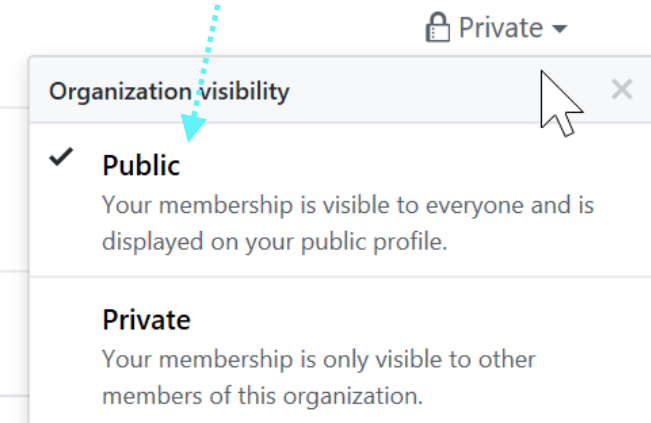
▸ https://github.com/Data-Science-for-Linguists-2019

Accept invitation from "notifications" page

Click "People", and make your membership public

▸ 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.

Just when you thought we've settled down...

▸ CHANGE 1: I'm **transferring** these two repos from my personal ("naraehan") account to class organization.

- ◆ HW1-Repo
- ◆ Class-Exercise-Repo

ⓘ 🔒 GitHub, Inc. (US) | https://github.com/naraehan/HW1-Repo

ⓘ 🔒 GitHub, Inc. (US) | https://github.com/Data-Science-for-Linguists-2019/HW1-Repo

What about your forks?
No problem --
GitHub will automatically
re-direct pulls.

▸ CHANGE 2: I am also switching them to **private**.

This *is* a problem.

Switching to private: aftermath

▶ When a GitHub repo is made private, it effectively goes into hiding.

▶ All its forks will lose track of it, getting stranded. (They will also remain public.)

Switching to private: aftermath

▶ When a GitHub repo is made private, it effectively goes into hiding.

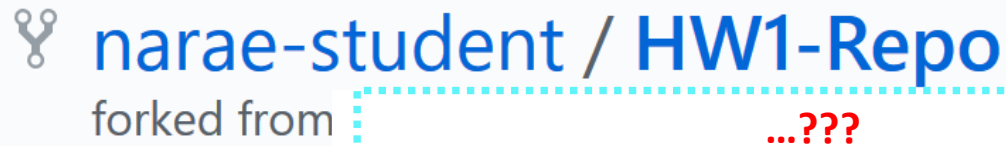▶ All its forks will lose track of it, getting stranded. (They will also remain public.)

narae-student / **HW1-Repo**
forked from   …???

This is BAD!

▶ So, unfortunately, you have to RE-FORK.

Re-forking, etc.   (It is worth the effort!)

▸ On GitHub

  ◆ **Delete** your forks: HW1-Repo and Class-Exercise-Repo.

  ◆ Head to our GitHub class organization. **Fork** the two repositories.

▸ On your laptop

  ◆ Your local repos' remote setting (`git remote -v`):

    ◆ The "**origin**" URL is still the same.

    ◆ The "**upstream**" URL will need updating:

  ◆ In HW1-Repo/ folder, execute:

```
git remote set-url upstream https://github.com/Data-Science-for-Linguists-2019/HW1-Repo.git
```

    ◆ In Class-Exercise-Repo/ folder, execute:

```
git remote set-url upstream https://github.com/Data-Science-for-Linguists-2019/Class-Exercise-Repo.git
```

Back to Class-Exercise-Repo

▶ https://github.com/Data-Science-for-Linguists-2019/Class-Exercise-Repo

▶ To-do2

◆ Who has the best study notes?

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 on Tuesday.**

   ◆ Study pandas, make your own JNB.

   ◆ Try two spreadsheet files, one your choice. (Make it small!)

▶ **Start thinking about project ideas**