

# Lecture 4: numpy

LING 1340/2340: Data Science for Linguists

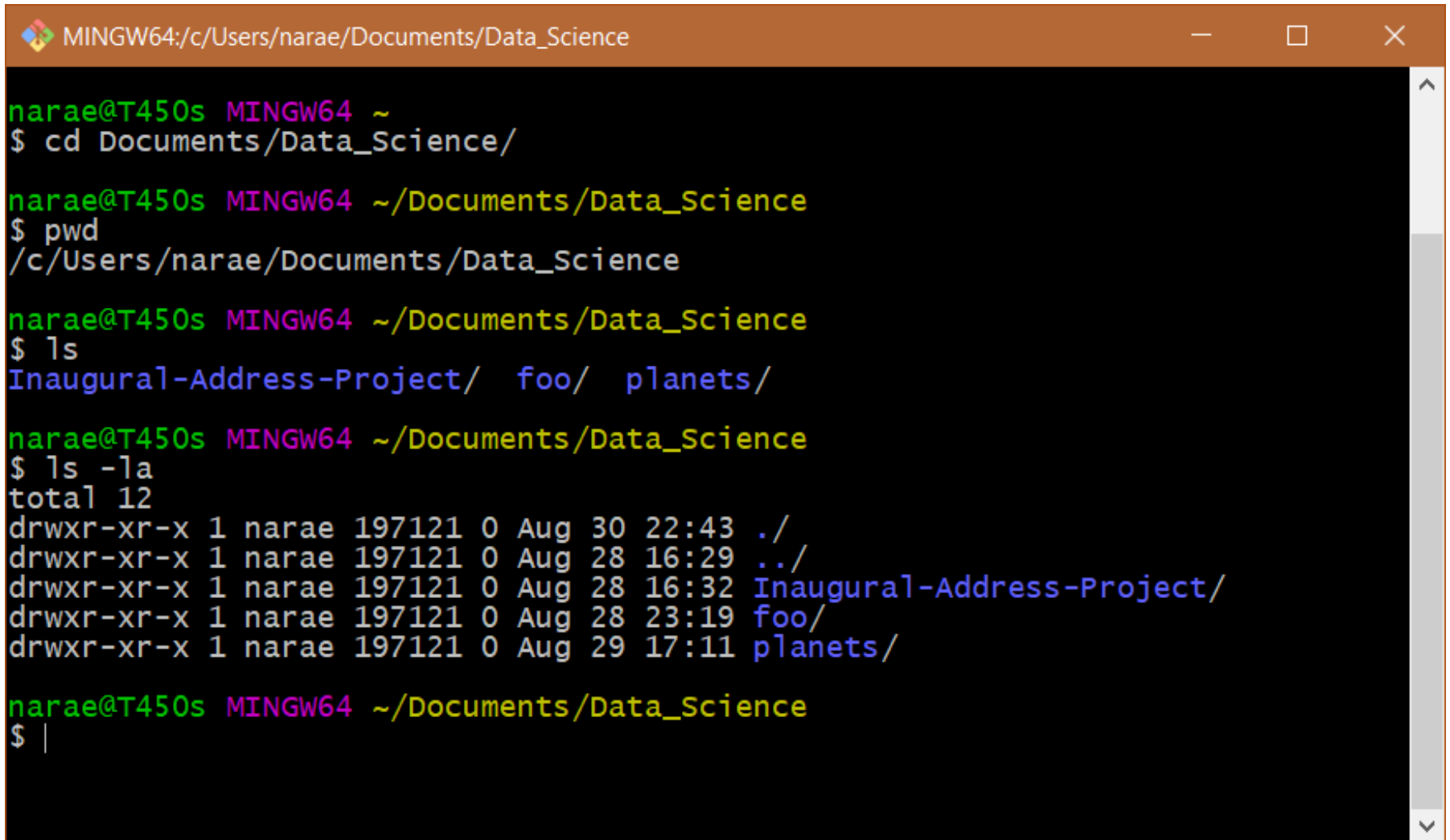
Na-Rae Han

# Objectives

---

- ▶ To-do 2: study notes in jupyter
- ▶ Python's numpy library
- ▶ GitHub class organization
  
- ▶ Tools:
  - ◆ Git and GitHub
  - ◆ Jupyter Notebook

# First thing to do every class



```
MINGW64:/c/Users/narae/Documents/Data_Science
narae@T450s MINGW64 ~
$ cd Documents/Data_Science/

narae@T450s MINGW64 ~/Documents/Data_Science
$ pwd
/c/Users/narae/Documents/Data_Science

narae@T450s MINGW64 ~/Documents/Data_Science
$ ls
Inaugural-Address-Project/  foo/  planets/

narae@T450s MINGW64 ~/Documents/Data_Science
$ ls -la
total 12
drwxr-xr-x 1 narae 197121 0 Aug 30 22:43 ./
drwxr-xr-x 1 narae 197121 0 Aug 28 16:29 ../
drwxr-xr-x 1 narae 197121 0 Aug 28 16:32 Inaugural-Address-Project/
drwxr-xr-x 1 narae 197121 0 Aug 28 23:19 foo/
drwxr-xr-x 1 narae 197121 0 Aug 29 17:11 planets/

narae@T450s MINGW64 ~/Documents/Data_Science
$ |
```

# Back to Class-Practice-Repo

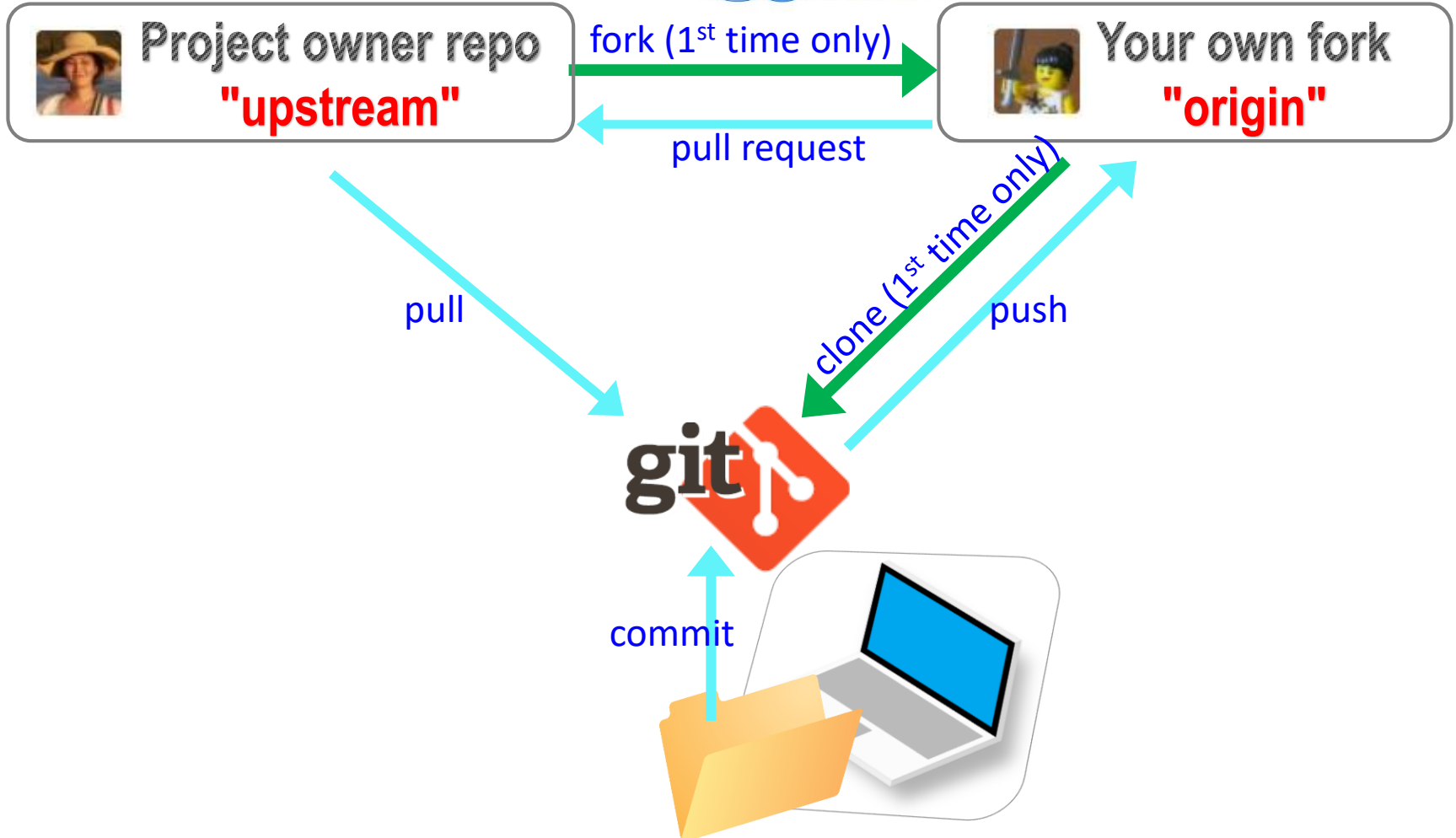
---

<https://github.com/naraehan/Class-Practice-Repo>

## ▶ To-do2

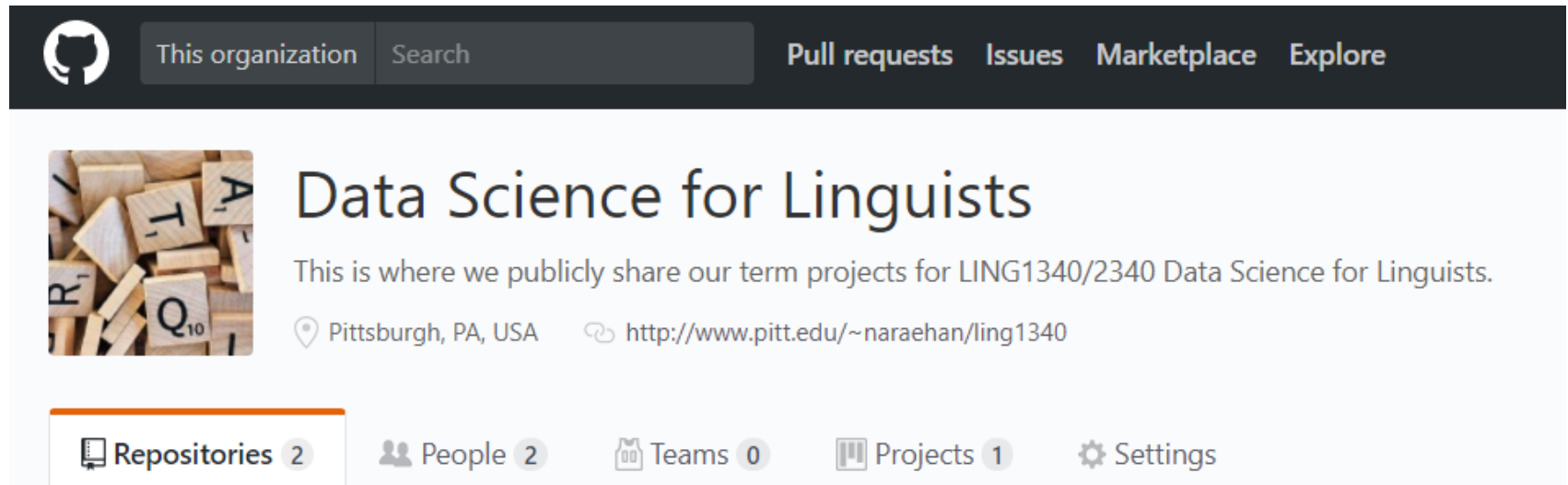
- ◆ Who has the best study notes?

# The fork triangle



# Introducing... GitHub Class Organization

▶ <https://github.com/Data-Science-for-Linguists>



The screenshot shows the GitHub organization page for 'Data Science for Linguists'. At the top, there is a dark navigation bar with the GitHub logo, a search bar, and links for 'Pull requests', 'Issues', 'Marketplace', and 'Explore'. Below this, the organization's profile is displayed. It features a profile picture of wooden blocks with letters, the organization name 'Data Science for Linguists', and a description: 'This is where we publicly share our term projects for LING1340/2340 Data Science for Linguists.' The location is listed as 'Pittsburgh, PA, USA' and the website as 'http://www.pitt.edu/~naraehan/ling1340'. At the bottom, there is a navigation bar with icons and counts for 'Repositories 2', 'People 2', 'Teams 0', 'Projects 1', and 'Settings'.

▶ 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...

---

- ▶ I'm transferring these two repos from my personal ("naraehan") account to class organization.
  - ◆ HW1-Repo
  - ◆ Class-Practice-Repo

## ▶ Before:

 [GitHub, Inc. \[US\] | https://github.com/naraehan/HW1-Repo](https://github.com/naraehan/HW1-Repo)

## ▶ After:

 [GitHub, Inc. \[US\] | https://github.com/Data-Science-for-Linguists/HW1-Repo](https://github.com/Data-Science-for-Linguists/HW1-Repo)

# Repo transfer: aftermath

---

## ▶ What now??

### ◆ How about your GitHub forks?

← Good news! GitHub automatically updates your fork's setting.

 narae-student / HW1-Repo  
forked from naraehan/HW1-Repo



 narae-student / HW1-Repo  
forked from Data-Science-for-Linguists/HW1-Repo

← Nothing to do; you can create pull requests like before.

### ◆ How about your local repo's remotes setting?

← You should update your "upstream" remote's URL before pulling.

```
git remote set-url upstream <new-git-URL>
```

(Apparently the old remote setting still works because GitHub will redirect pulls, but we want things tidier.)



# Updating remote's URL

---

```
git remote set-url upstream <new-git-URL>
```

```
narae@T450s MINGW64 ~/Documents/Data_Science/HW1-Repo (master)
$ git remote -v
origin https://github.com/narae-student/HW1-Repo.git (fetch)
origin https://github.com/narae-student/HW1-Repo.git (push)
upstream https://github.com/naraehan/HW1-Repo.git (fetch)
upstream https://github.com/naraehan/HW1-Repo.git (push)

narae@T450s MINGW64 ~/Documents/Data_Science/HW1-Repo (master)
$ git remote set-url upstream https://github.com/Data-Science-for-Linguists/HW1-Repo.git

narae@T450s MINGW64 ~/Documents/Data_Science/HW1-Repo (master)
$ git remote -v
origin https://github.com/narae-student/HW1-Repo.git (fetch)
origin https://github.com/narae-student/HW1-Repo.git (push)
upstream https://github.com/Data-Science-for-Linguists/HW1-Repo.git (fetch)
upstream https://github.com/Data-Science-for-Linguists/HW1-Repo.git (push)
```

# numpy practice

---

Activity 2  
30 mins



- ▶ In Class-Practice-Repo, activity2 folder:
  - ◆ You will find `numpy_practice.ipynb`
  - ◆ Rename it `numpy_practice_YOURNAME.ipynb`, or copy it out of the repo directory (where name does not matter)

# Becoming good at learning

---

- ▶ You should know by heart most of Python's basic built-in commands and syntax.
- ▶ Learning certain Python libraries doesn't necessarily mean you should know *everything* about their capabilities.
- ▶ You will however need to build a good conceptual understanding about the objects and functions they offer,
- ▶ ... and then become good at *looking stuff up* and *finding your way around*.

# Wrapping up

---

- ▶ To-do 3: due on Tuesday.
  - ◆ Study pandas, make your own notebook. Submit via Class-Practice-Repo.
- ▶ HW2: will be pushed back one class (due next next Tuesday)
- ▶ Office hours
  - ◆ Mon/Wed 2:30 -- 4pm. G17 CL
  - ◆ Also by appointment
- ▶ Learn:
  - ◆ pandas
  - ◆ Git, GitHub
  - ◆ Jupyter Notebook