

Lecture 5: numpy, Git/GitHub

LING 1340/2340: Data Science for Linguists

Na-Rae Han

Objectives

- ▶ To-do3 review: study notes in JNB, git and GitHub
- ▶ Python's numpy library hands-on practice

- ▶ Tools:
 - ◆ Git and GitHub
 - ◆ Jupyter Notebook

First thing to do every class

```
MINGW64:/c/Users/narae/Documents/Data_Science
narae@X1Yoga MINGW64 ~
$ cd Documents/Data_Science/

narae@X1Yoga MINGW64 ~/Documents/Data_Science
$ pwd
/c/Users/narae/Documents/Data_Science

narae@X1Yoga MINGW64 ~/Documents/Data_Science
$ ls
Class-Exercise-Repo/  languages/

narae@X1Yoga MINGW64 ~/Documents/Data_Science
$ ls -la
total 12
drwxr-xr-x 1 narae 197121 0 Jan 10 14:30 ./
drwxr-xr-x 1 narae 197121 0 Jan  8 18:33 ../
drwxr-xr-x 1 narae 197121 0 Jan 10 14:30 Class-Exercise-Repo/
drwxr-xr-x 1 narae 197121 0 Jan  8 18:34 languages/

narae@X1Yoga MINGW64 ~/Documents/Data_Science
$ |
```

```
pwd
cd dir1/dir2
cd ..
cd
ls
ls -la
```

Hit **TAB** for auto-completion.

Up **↑** / Down **↓** arrow to use previous command

Ctrl + c to cancel

Last step: Sync your repos
(Class-Exercise-Repo especially)

Back to Class-Exercise-Repo

- ▶ <https://github.com/Data-Science-for-Linguists-2024/Class-Exercise-Repo>
- ▶ To-do3
 - ◆ Git and GitHub work flow: are you getting the hang of it?
 - ◆ Who has the best study notes?

numpy practice with the Simpsons

Activity 2
40 minutes



► In Class-Exercise-Repo, activity2/ 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 4: due Monday.
 - ◆ 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!
 - ◆ Details on "Term Project Guidelines" page