Python Computing for Data Science

Files for Today:

http://goo.gl/dyKHxx = this PDF

http://goo.gl/XFiZYI = IPython Notebook

```
git clone https://github.com/profjsb/python-seminar.git
```

(if you dont have *git*, please set it up later)

Welcome to the ***Python Computing for Data Science*** Seminar

AY 250: Thursday 1-4pm (Hearst Field Annex B-1)

Instructor: Josh Bloom

GSI: Adam Morgan

Instructor+GSI email:
ucbpythonclass+seminar@gmail.com

Course email: python@bspace.berkeley.edu

Computational Science and Engineering

NSF

Award #0941742

Motivation:

short version

get you using Python to do cutting-edge research

long version

1) get you using Python to do cutting-edge research,
2) helping you realize that Python is a viable framework to do just about any 21st century problem well (and costs zero). "Super Glue"
3) fold you into the Python community so it benefits from having you part of it

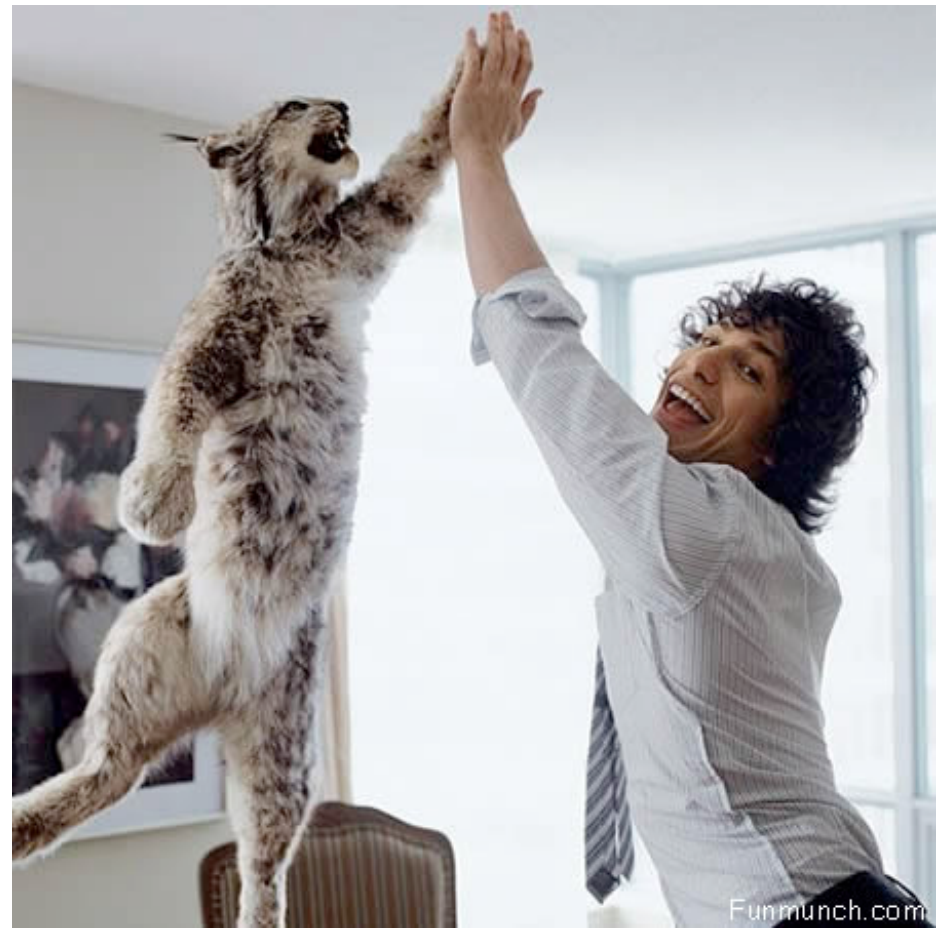
How we plan to do this:

- "formal" lectures on specialized topics each week by leading experts & local practitioners (Thursday)

- "breakout work sessions" interspersed within the lectures

- homework assignments based on week's lecture

- final project

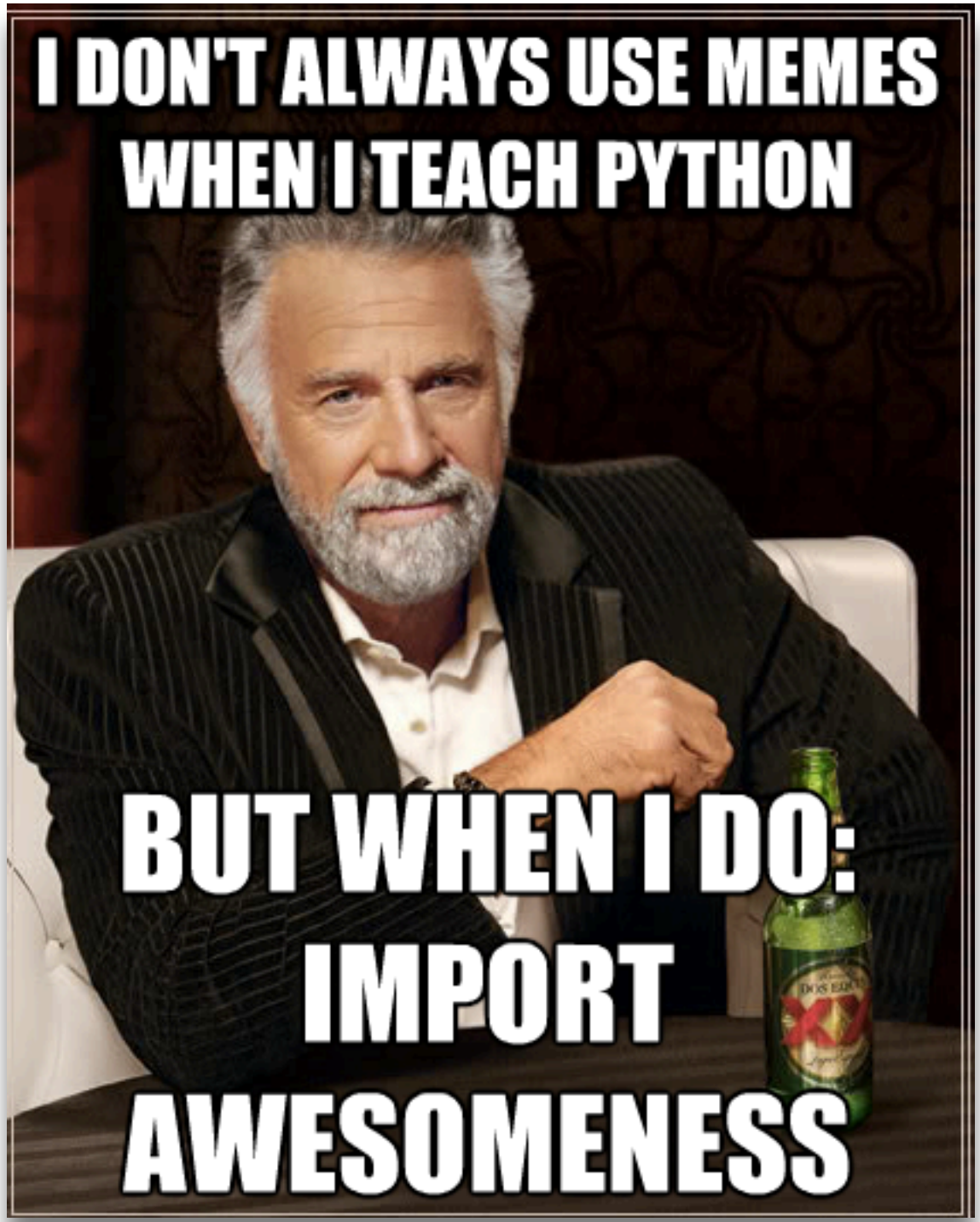


MONTY PYTHON

For awkward, forced laughter.

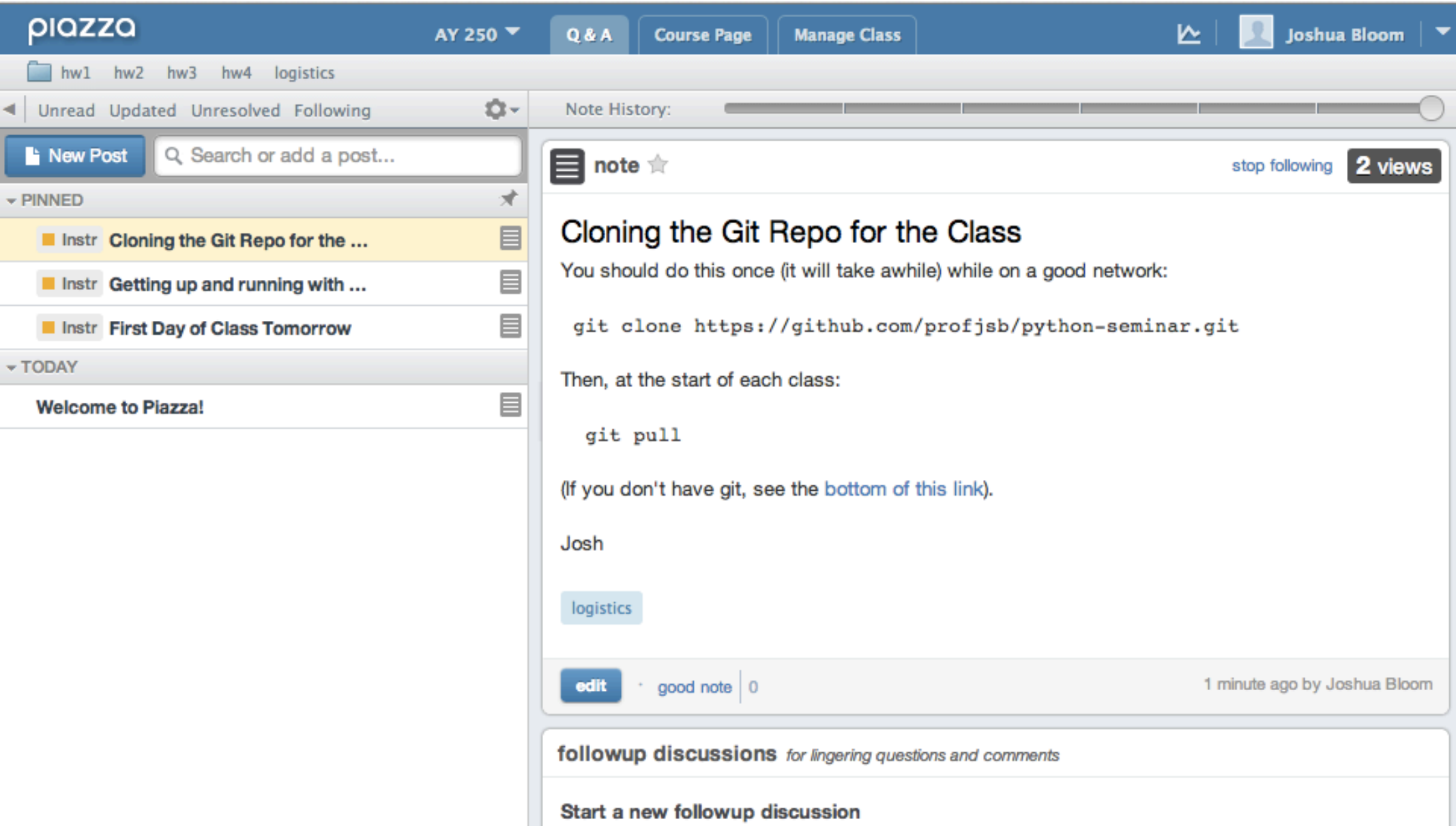
\o/ MotivatedPhotos.com

Bring me
a shrubbery!!
ICANHASCHEEZBURGER.COM

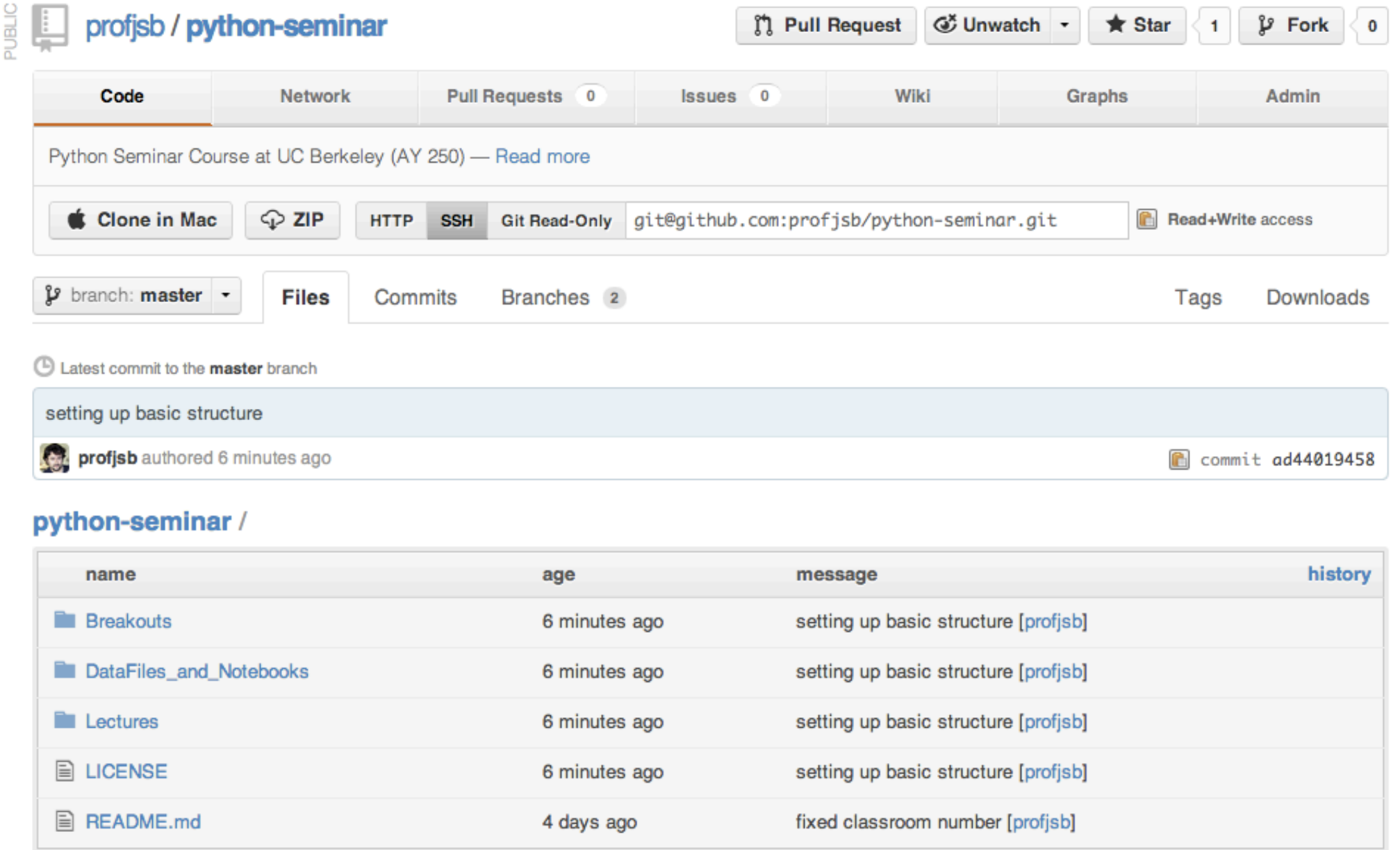

I DON'T ALWAYS USE MEMES
WHEN I TEACH PYTHON
BUT WHEN I DO:
IMPORT
AWESOMENESS

`piazza` for real-time/off-line interaction homework updates, solutions, ...

http://piazza.com/berkeley/fall2013/ay250



`github` is the main data portal for us...



Scientific Research Computing with Python

Files for Today:

http://goo.gl/dyKHxx = this PDF

http://goo.gl/XFiZYI = IPython Notebook

/DataFiles_and_Notebooks/00_AdvancedPythonConcepts = follow-along files

http://profjsb.github.io/python-seminar/

View on GitHub

Python Computing for Data Science

Undergraduate/Graduate Seminar Course at UC Berkeley (AY 250)

tar.gz .zip

+1 +10 Recommend this on Google

Hearst Field Annex B-1: Thursday 1 - 4 PM FALL 2013 (CCN #060800)

Synopsis

Python is becoming the *de facto* superglue language for modern scientific computing. In this course we will learn Pythonic interactions with databases, imaging processing, advanced statistical and numerical packages, web frameworks, machine-learning, and parallelism. Each

Course Schedule

Date	Content	Leader
Aug 29	Advanced Python Language Concepts (geared towards Boot Camp graduates)	Josh
Sep 5	scipy, numpy	Fernando
Sep 12	Advanced versioning, application building (optparse), debugging & testing	Josh
Sep 19	Advanced plotting, Notebooks and data vizualization, mayavi	Fernando
Sep 26	Machine Learning	Josh
Oct 3	Interacting with the world (xml-rpc, urllib, sending and receiving email, serial)	Josh
Oct 10	Database interaction, large datasets (HDF5)	Josh
Oct 17	Pandas and Timeseries Modelling	*Wes McKinney*
Oct 24	GUI (Tkinter, GTK, Traits)	Josh
Oct 31	Parallelism	Paul/Fernando
Nov 7	Web frameworks & RESTful APIs, Flask	Josh
Nov 14	Bayesian programming & Symbolic math	*Joey Richards*
Nov 21	Cython; wrapper around legacy code -- FORTRAN, C, etc	Paul
Nov 28	holiday	
Dec 5/Onward	final project work	

Workflow for a *typical* week

Wednesday:

email from week's instructor w/ special installation instructions, reading/tutorials

Thursday:

1:00 `cd python-seminar; git pull`

1:10 - 2pm: Intro topics Lecture

2 - 2:30pm: Breakout coding

2:30 - 3:30pm: Detailed topics lecture

3:30-4: Work on homework

Monday:

10am -12: Supervised help with homework
481 Evans (Time Series Center)

Wednesday:

Homework project due

Course Grade

10% participation in lectures/breakouts

60% Homeworks
there will be 11 assignments & we will drop your lowest score

30% Final Project, due Dec11
(no final exam)

Final Project

a) Build a substantial framework for doing something in your own research, based on at least two topics from different weeks. Something you will use for a long time...

e.g., image analysis package, hardware control software, a webservice that does some crunching under the hood, provide a parallelization of some algorithm or code you use, etc.

- or -

b) Contribute code/functionality to a major open-source Python project (IPython,scipy, Cython,numpy, matplotlib, etc.)

Prerequisites:

- working knowledge (or more) of the core Python language

- and/or -

Python BootCamp graduate

- installation of Python (2.7.3 - 2.7.5), scientific 3rd party packages (Anaconda distro), & git

- laptop for use in class and for homeworks

- tolerance for our terrible computer humor

http://www.pythonbootcamp.info/preparation/software