



W200 Introduction to Data Science Programming

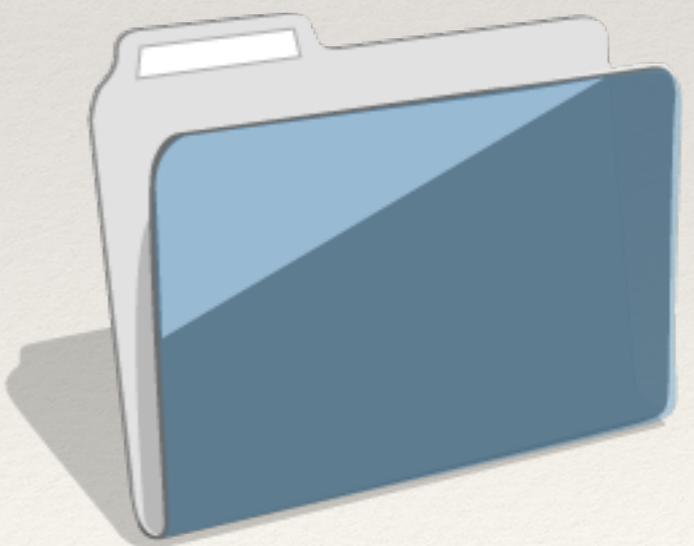
Introductions, Course Info & Tools

Week 01

version: Dec 31, 2019

Agenda

- ❖ Welcome introductions
- ❖ Course Overview
- ❖ Mechanics
- ❖ GitHub & git
- ❖ Shell Scripts



Brief Introductions

- ❖ Quick intro: your name, where you live and work
- ❖ What do you do with computing?
- ❖ For what do you want to use data science skills?
- ❖ Introducing the Breakout Rooms. We have team-based projects in this course, so let's spend a few minutes chatting in smaller groups.



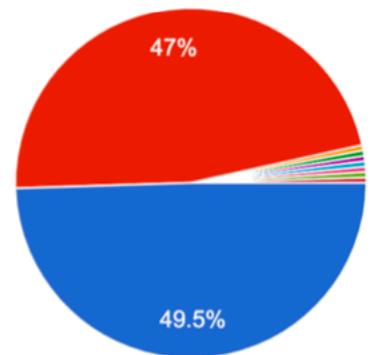
1. Course Overview

- ❖ Pretty fast-moving course and students join the course with a range of computing skills, programming experiences from neophyte to pro.
- ❖ Our goal is to establish a strong base for all students. Sharing your knowledge with other students helps all.
- ❖ Feel free to ask questions of each others, the instructors, and especially of the TAs.
- ❖ There's a couple of ways of communicating with W200-group, Slack, Office Hours, and TA appointments.

1. Overview: student computing

What kind of computer will you be using for this course?

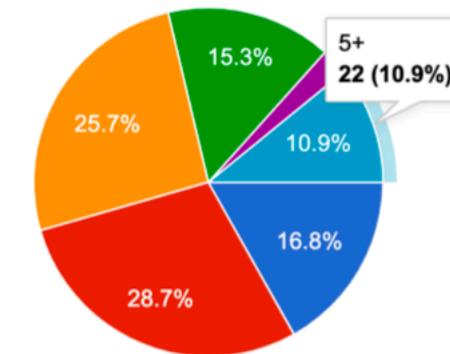
202 responses



- Windows
- Mac
- Linux
- I'll be using a virtual Linux machine on my Windows 10 laptop
- Both a windows desktop and a mac...
- Windows for now but potentially swi...
- Windows and Ubuntu
- I have access to all three
- Mac & Windows depending on locat...

How many programming classes have you taken in a school environment?

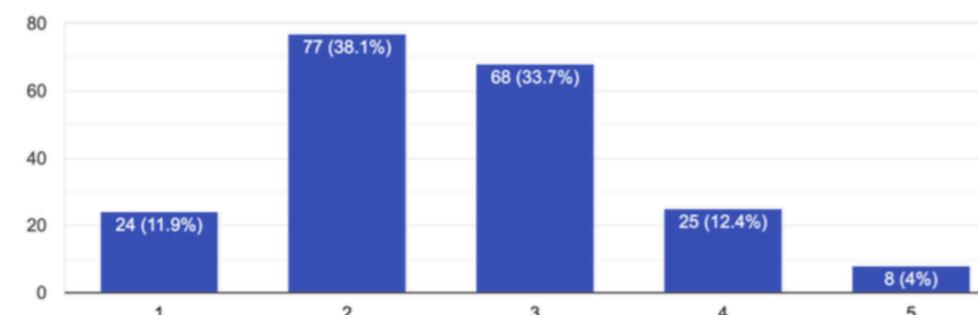
202 responses



- 0
- 1
- 2
- 3
- 4
- 5+

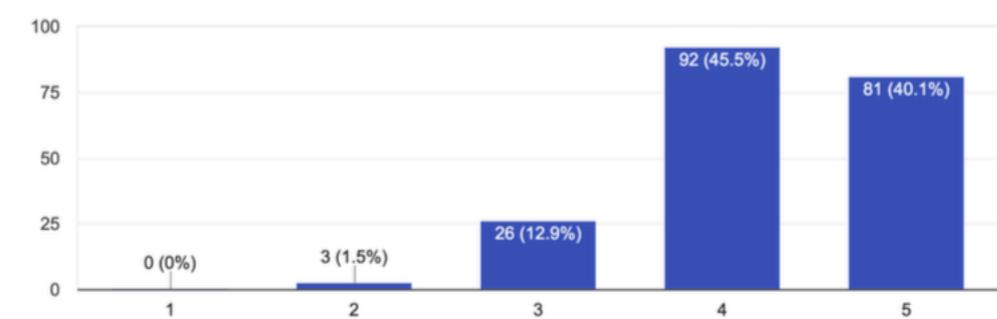
How would you rate your experience with programming?

202 responses



How would you rate your experience with excel?

202 responses



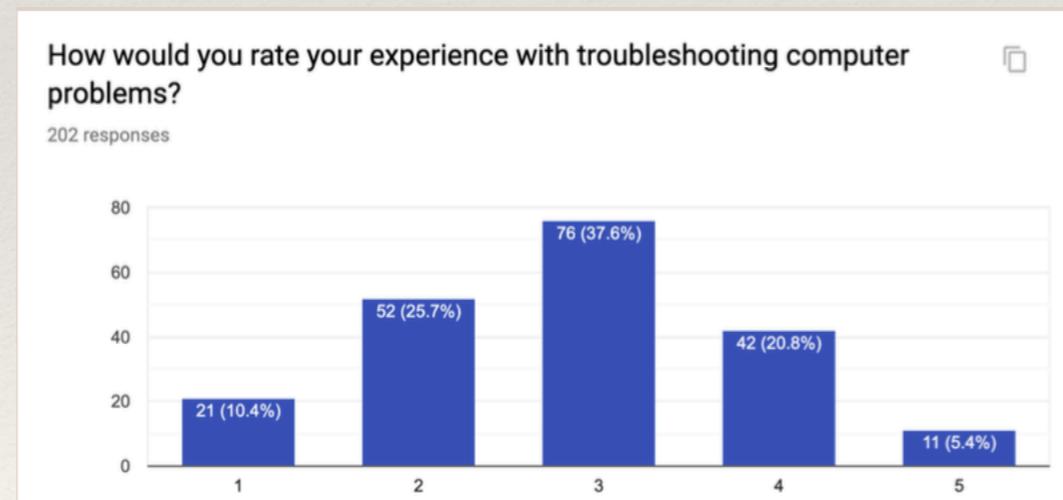
Students with less coding experience tend to enjoy significant experience in other areas.

1. Overview: Breakout Rooms

- ❖ There's a lot of individual effort in learning to code for DataSci; and there's breakout rooms to give everyone a chance to participate.
- ❖ During these breakouts we work on a weekly Activity notebook - encourage and mentor each other.
- ❖ This is true to life as a data scientist ... at times you'll gossip code and jointly share techniques.

1. Overview: Getting help

- ❖ Much of the time you'll have to discover answers on your own (that's how it is in computing).
- ❖ Learning to interpret and respond to errors are important skills. You'll be expected to deal with files in a domain that we haven't covered. Focus on the data!



2. Course Curriculum (1 of 2)

- ❖ Week 1: Getting to know the development environment
- ❖ Weeks 2-6: Python Fundamentals
 - ❖ Starting out, sequence types & dictionaries; control and algorithms; functions, complexity
- ❖ Weeks 7-8: Object-Oriented Design & Programming
 - ❖ Classes, OOD, OOP
 - ❖ Individual OOP program

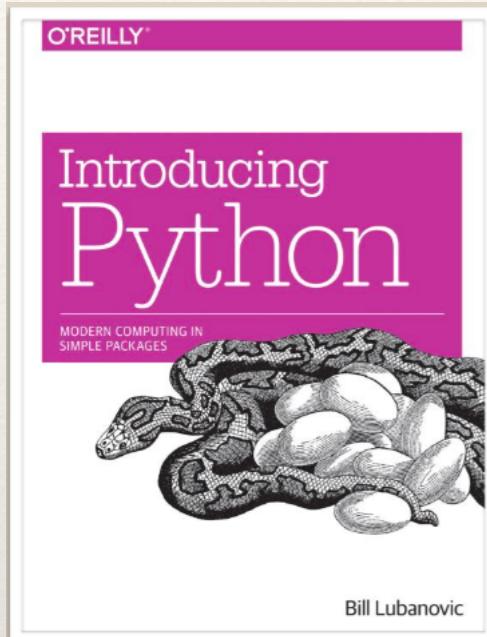
2. Course Curriculum (2 of 2)

- ❖ Weeks 9-14: Data Analysis
 - ❖ Text & Binary Data
 - ❖ Numpy (vectors)
 - ❖ Pandas (Data Frames)
 - ❖ Matplotlib, Seaborn (Visualization)
 - ❖ Advanced Pandas (Aggregation & Group Operations_
 - ❖ Testing
 - ❖ Team Data Analysis Project & Presentation

3. Grading

- ❖ 30% - Homework (that you submit to your GitHub)
- ❖ 40% - Projects (20% for the individual OOP; 20% for the team data analysis project)
- ❖ 20% - Online Exams (Midterm 10%, 24-hour window; Final is 10%, 48-hour window) [Don't wait 'til the last moment to start!]
- ❖ 10% - Participation

4. Materials: Asynch and Synch

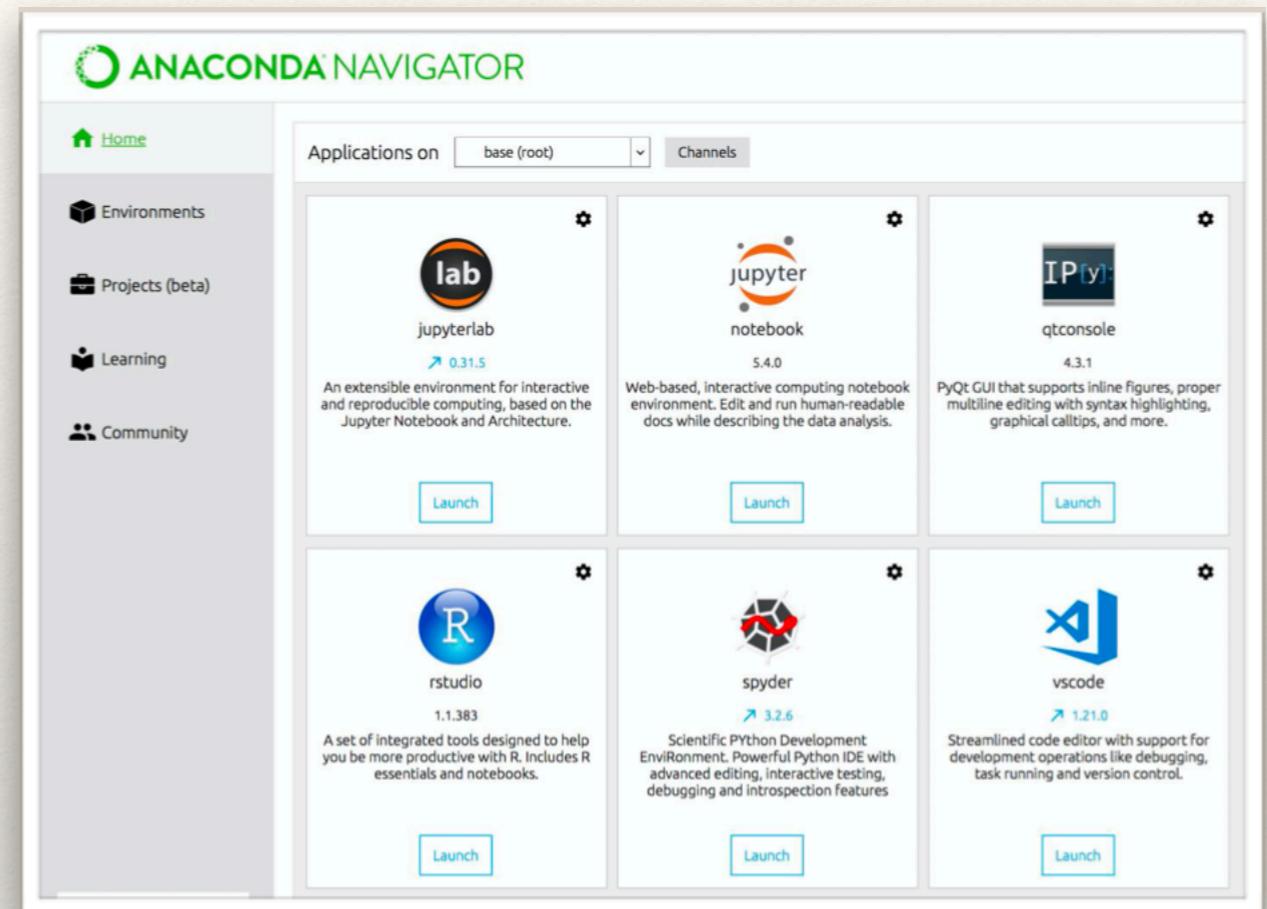


Other python and data sci texts are in the Resources folder.

- ❖ ISVC: view/study the asynchronous materials *before* the live session. Then attend one of the weekly live meetings.
- ❖ study.net: read/study before our live session
- ❖ GitHub: homework assignments are posted *after* the live session. Students *pull down* from this site. *Push* your completed homework to your individual SUBMISSIONS folder.

4. Programs & Tools

- ❖ Anaconda
- ❖ Jupyter
- ❖ Python (version 3)
- ❖ Bash shell*
- ❖ Git
- ❖ Text Editor (Atom, NotePad+, BBEdit)



Windows 10 has PowerWindow; earlier editions may need GitBash

4. Resources

Syllabus: https://docs.google.com/document/d/1MKp-RddgD9ZvaOK1w4LCIdGYivhwldtIV5NU_bGLvts

Live Calendar: <https://docs.google.com/spreadsheets/d/1jNmL4-K6F7OfTTAtscraFrQHdo-8-KepdhUbEaFGXSY>

Homework Assignments:

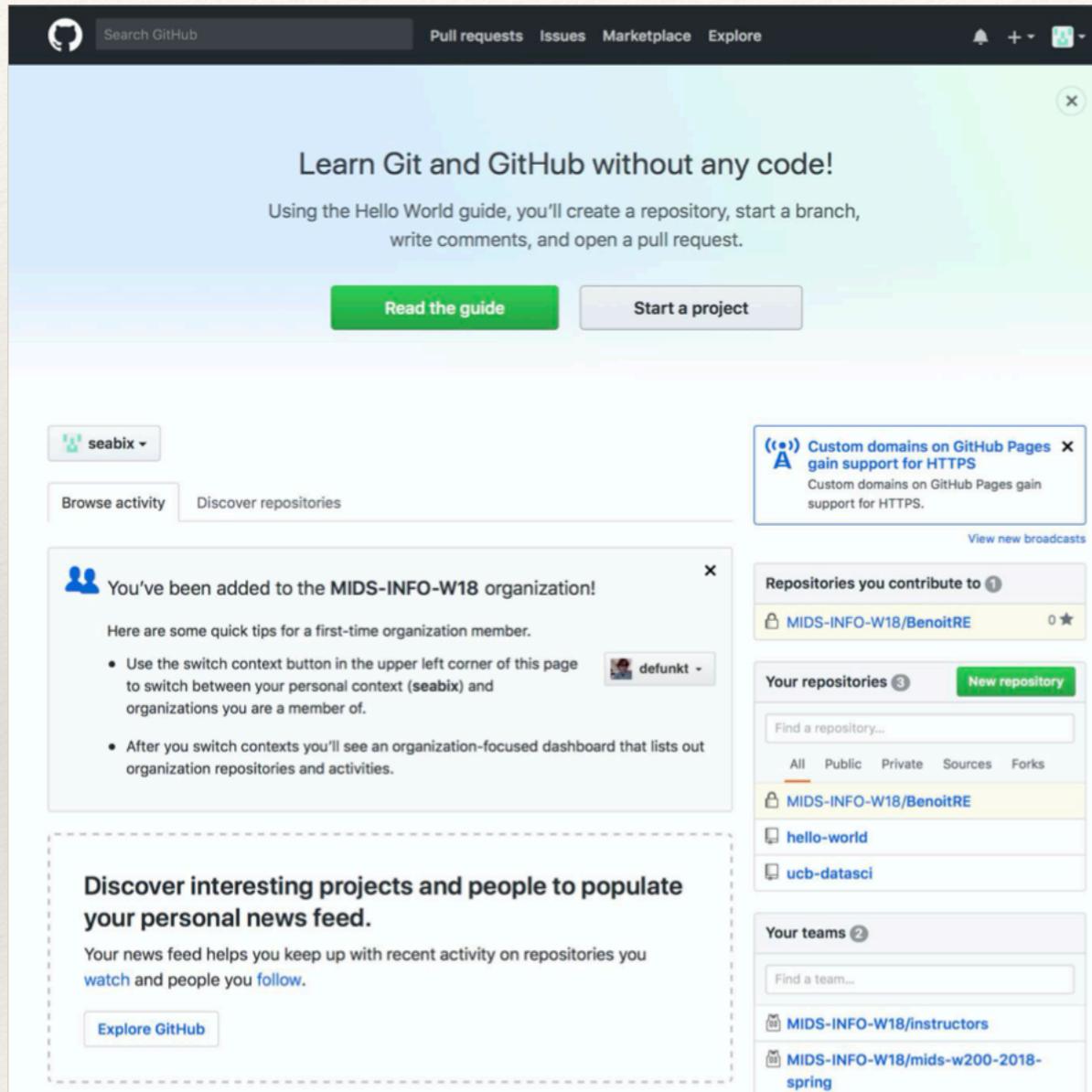
- https://github.prod.oc.2u.com/UCB-INFO-PYTHON/assignments_upstream_spring20

Communication to instructors & other students: w200-python-2020-spring@googlegroups.com

Communication with instructors only: mids-python-instructors@googlegroups.com

Slack Channels for information communication: ucbischool.slack.com channel #w200-python

4. Github

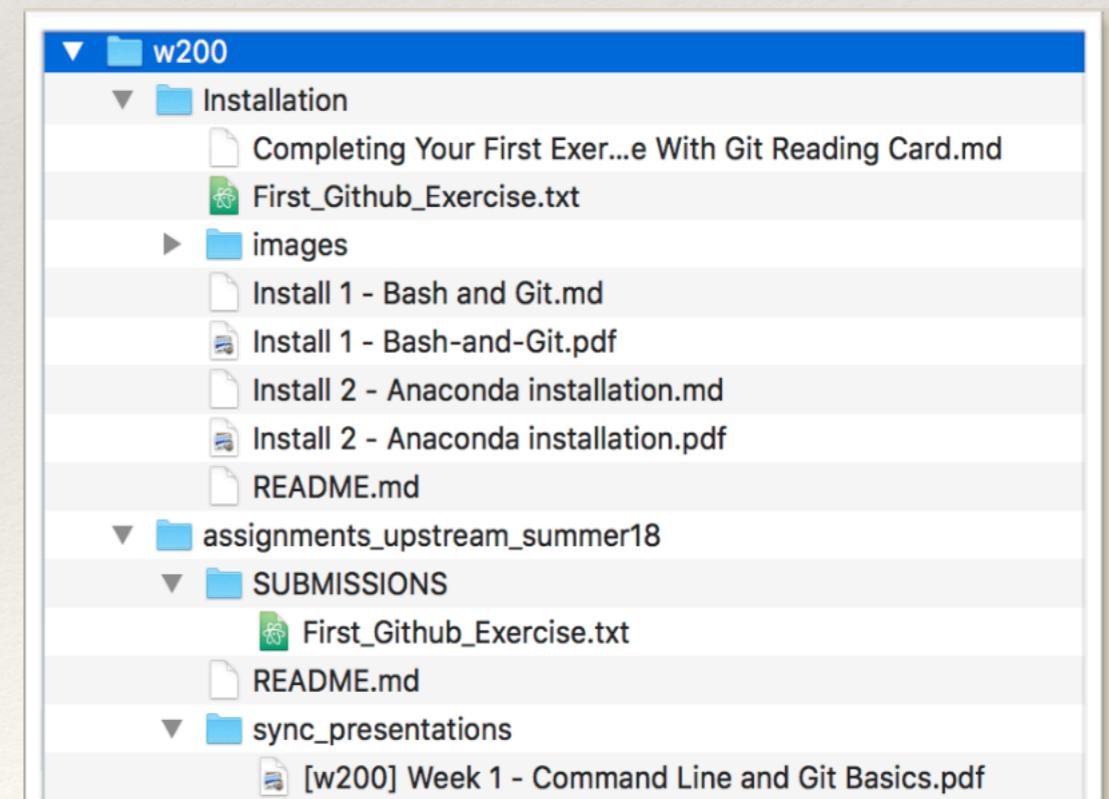


- ❖ File sharing & Version Control Site:
 - ❖ <https://github.prod.oc.2u.com>
- ❖ Please “Read the Guide”
- ❖ Press “Discover...” search for UCB-INFO-PYTHON/Installation
- ❖ Read “Completing your First Exercise with Git Reading Card”; follow the instructions.

4. More about Git

Folder structure:

- ❖ /Users/tom/w200
- ❖ /Users/tom/w200/FirstnameLastnameREPO/
- ❖ /Users/tom/w200/FirstnameLastnameREPO/
SUBMISSIONS/
- ❖ Let's visit GitHub for a demonstration



<https://guides.github.com/activities/hello-world/>

Create a “Hello, World!” page

Get started in an empty working directory (for example, `work`, if you downloaded the file from the previous step) and create an empty directory named “hello”, then create a `hello.html` file in it with the following contents.

RUN:

```
mkdir hello  
cd hello  
touch hello.html
```

FILE: HELLO.HTML

```
Hello, World!
```

Add the page to the repository

Now let's add the “Hello, World” page to the repository.

RUN:

```
git add hello.html  
git commit -m "First Commit"
```

You will see ...

RESULT:

```
$ git add hello.html  
$ git commit -m "First Commit"  
[master (root-commit) 911e8c9] First Commit  
 1 files changed, 1 insertions(+), 0 deletions(-)  
  create mode 100644 hello.html
```

your repository

get assignments here

fun stuff

TomJonesREPO

assignments_upstream_spring20

github-playground₁₆

5. Shell Scripts

- ❖ Create a shell script using a text editor
- ❖ Set the file permissions (rwx,rx,rx; 0755)
- ❖ Run the script.

```
#!/bin/sh
# a comment
echo Bonjour # a comment in French
echo "欢迎"
```



```
$ chmod 755
hello.sh
$ ./hello.sh
```

- ❖ Optional: see these resources (10 Concepts; Ox-OS.pdf) about Unix and Operating Systems

6. Wrap-up

- ❖ Enjoy the course - it's a challenge but keep at it!
- ❖ If you're working along for > 1 hour, take a break and have a "sanity check." [Really.] Shoot an email to colleagues, TAs, instructors (the w200 group is a good place).

Welcome and enjoy Introduction to Data Science Programming

Gunnar, Chris, Rob, Chris W., Mark, & Gerry

