141B: We're learning Python!

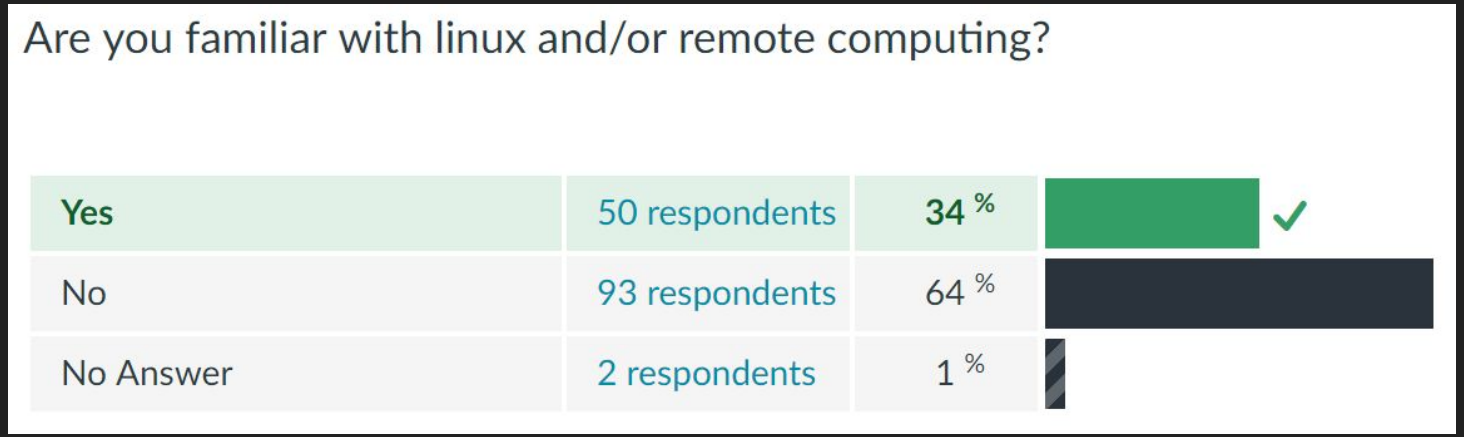
James Sharpnack

Programming Languages

Which of the following programming languages, if any, are you familiar with?

Python	67 respondents	**47 %** ✓
R	142 respondents	99 %
C/C++	71 respondents	49 %
Java	15 respondents	10 %
Ruby	1 respondents	1 %
Perl	6 respondents	4 %
Matlab	68 respondents	47 %
Other	16 respondents	11 %
No Answer	1 respondents	1 %

Linux


Are you familiar with linux and/or remote computing?
Yes
50 respondents
34 %
No
93 respondents
64 %
No Answer
2 respondents
1 %

Lesson 1 Time

Approximately, how long did Lesson 1 take you?

<30min	8 respondents	**11 %**	
30min-1hr	25 respondents	35 %	
1hr-2hr	27 respondents	38 %	
>2hr	10 respondents	14 %	
No Answer	1 respondents	1 %	

Why are you interested in this course?

1. Learn Python
2. What is Data Science?
3. Get a job (public policy, biostatistics, econ)
4. Visualization, D3, etc.
5. Web scraping, SQL, pandas, urllib
6. Machine learning, big data, data mining

"I previously took STA 141 and loved the unstructured, problem-solving aspect of it that is not offered in other statistics courses."

Why are you interested in this course?

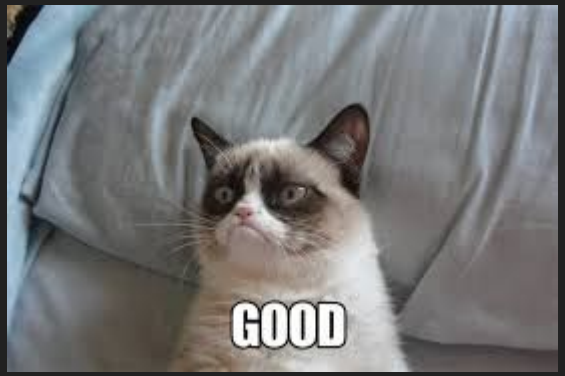
“In summer 2016, I had internship as a statistical data analyst. That internship taught me a lot. First, in real projects the **data are never clean**. It is up to us to do **data munging** part and get ready the data. Second, it is up to data scientist to retrieve data from **database systems** using SQL or any other appropriate language. Third, **the questions are never given a-priori** like in many statistical classes. Therefore, my main goal is to improve myself in these areas. To become the expert in data munging and posing questions. From my experience, this constitutes the majority of the work of data scientist. Moreover, I want to get more experience with **Python DS packages** such as pandas, numpy, scipy, matplotlib. I feel that python is more suitable for the data wrangling than R. Also, building my DS portfolio will be of primary importance for me to land a job as Data Scientist.”

Lacking clarity

1. What’s the difference between a text editor, terminal, python, ipython, jupyter?
2. Can’t see the video.
3. How to use the terminal
4. How to use Jupyter
5. Do we need to use python 2.7? (Yes, sorry)
6. Install was tough...

Lacking clarity

1. What's the difference between a text editor, terminal, python, ipython, jupyter?
2. Can't see the video.
3. How to use the terminal
4. How to use Jupyter
5. Do we need to use python 2.7? (Yes, sorry)
6. Install was tough...



Who am I? (How am I qualified to teach this?)

Ph.D. Carnegie Mellon University in Machine Learning & Statistics

Postdoc in Math department at UC San Diego

Research on Network Statistics, Machine Learning Theory

Picked up “Data Munging” in applied projects (Insurance, demography, biostatistics, marketing, learning software)

GIT what?