

Welcome to BCB/EEOB546X!

# Computational Skills for Biological Data

Instructors:

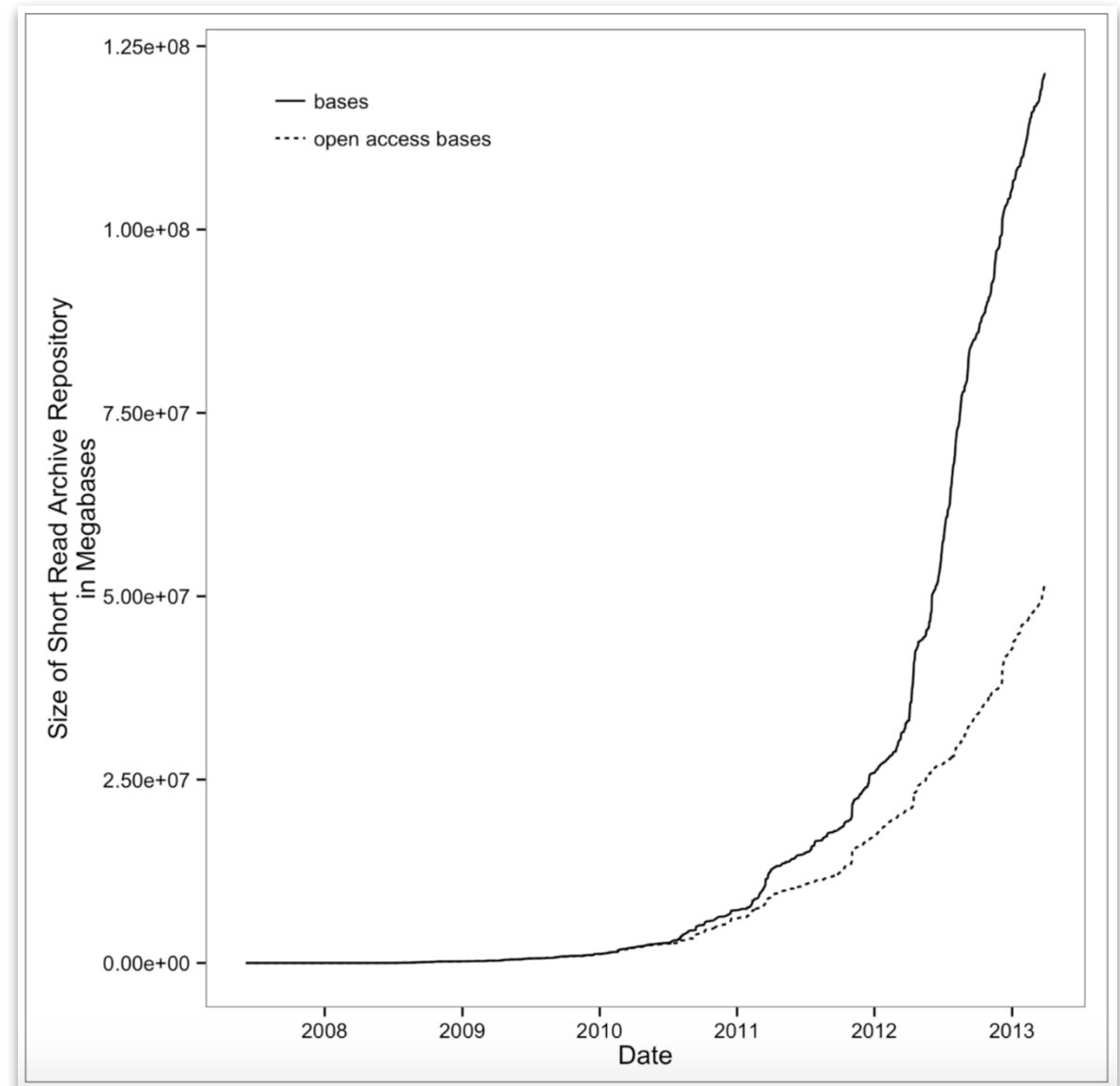
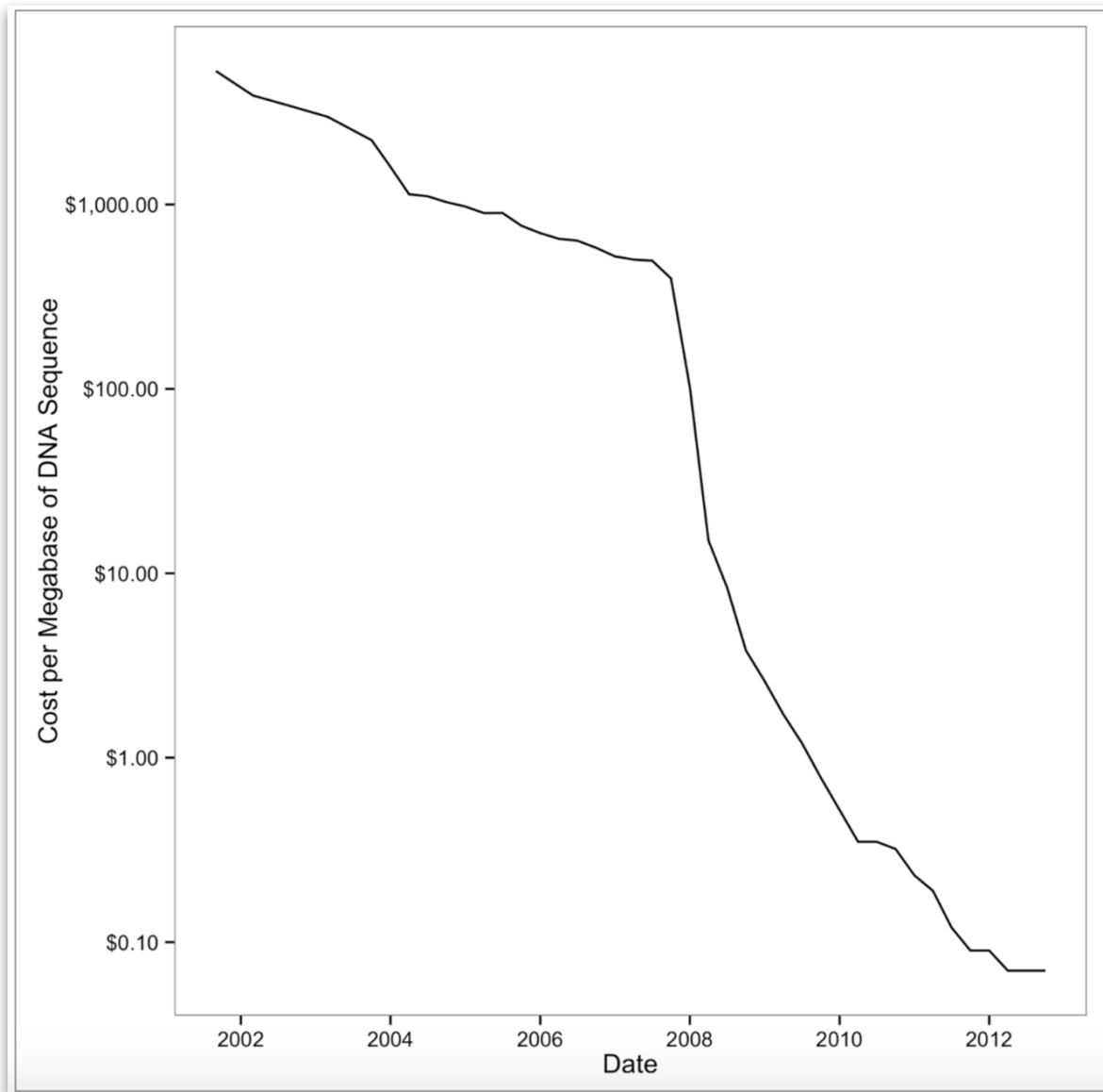
Matt Hufford

Tracy Heath

Dennis Lavrov

# Introduction and Basic Unix

## What motivated us to teach this class?

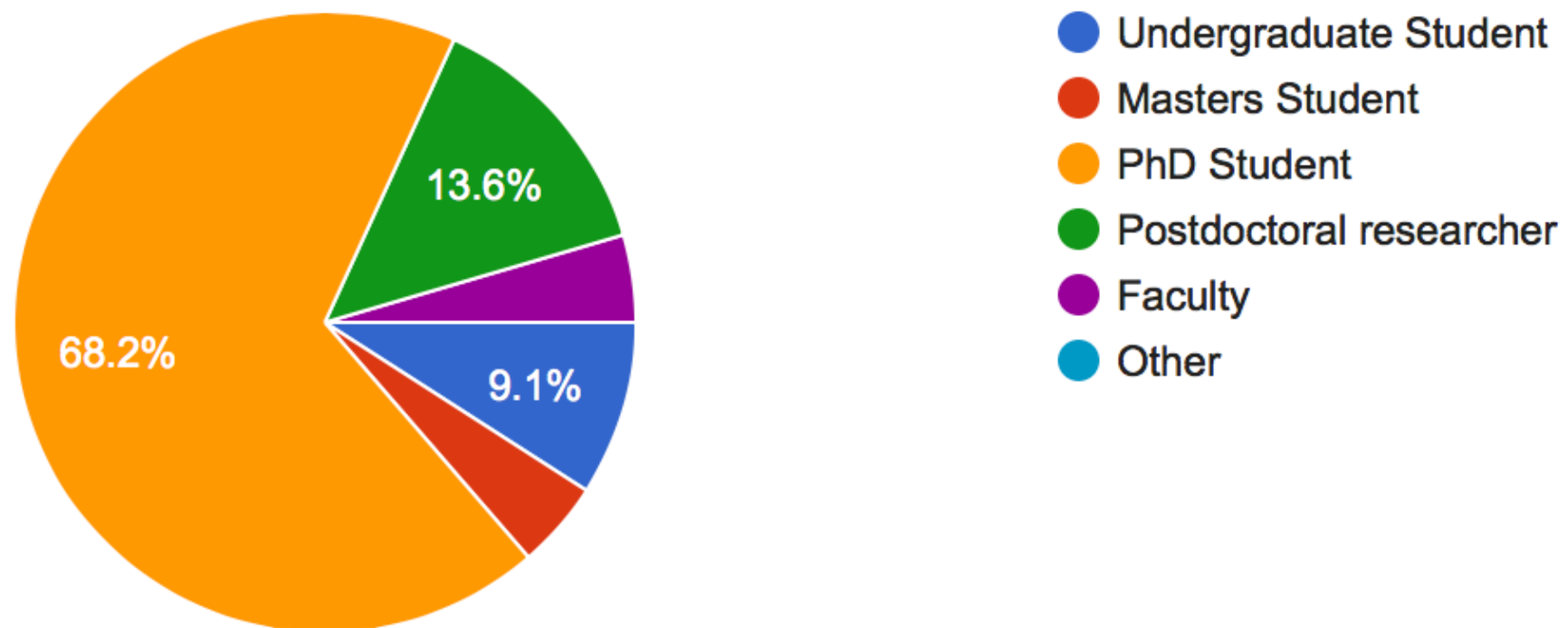


# Introduction and Basic Unix

What motivated you to take this class?

Course Make Up

What is your current academic level? (22 responses)



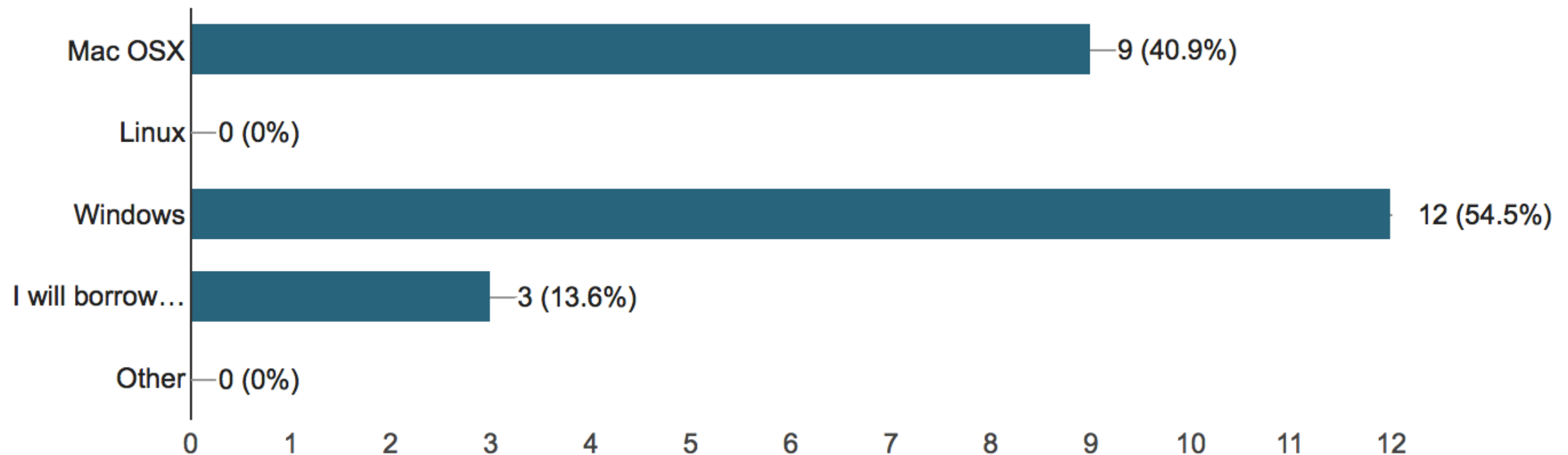
# Introduction and Basic Unix

What motivated you to take this class?

## Platform Use

If you own a laptop, what operating system do you have? (Please select all that apply.)

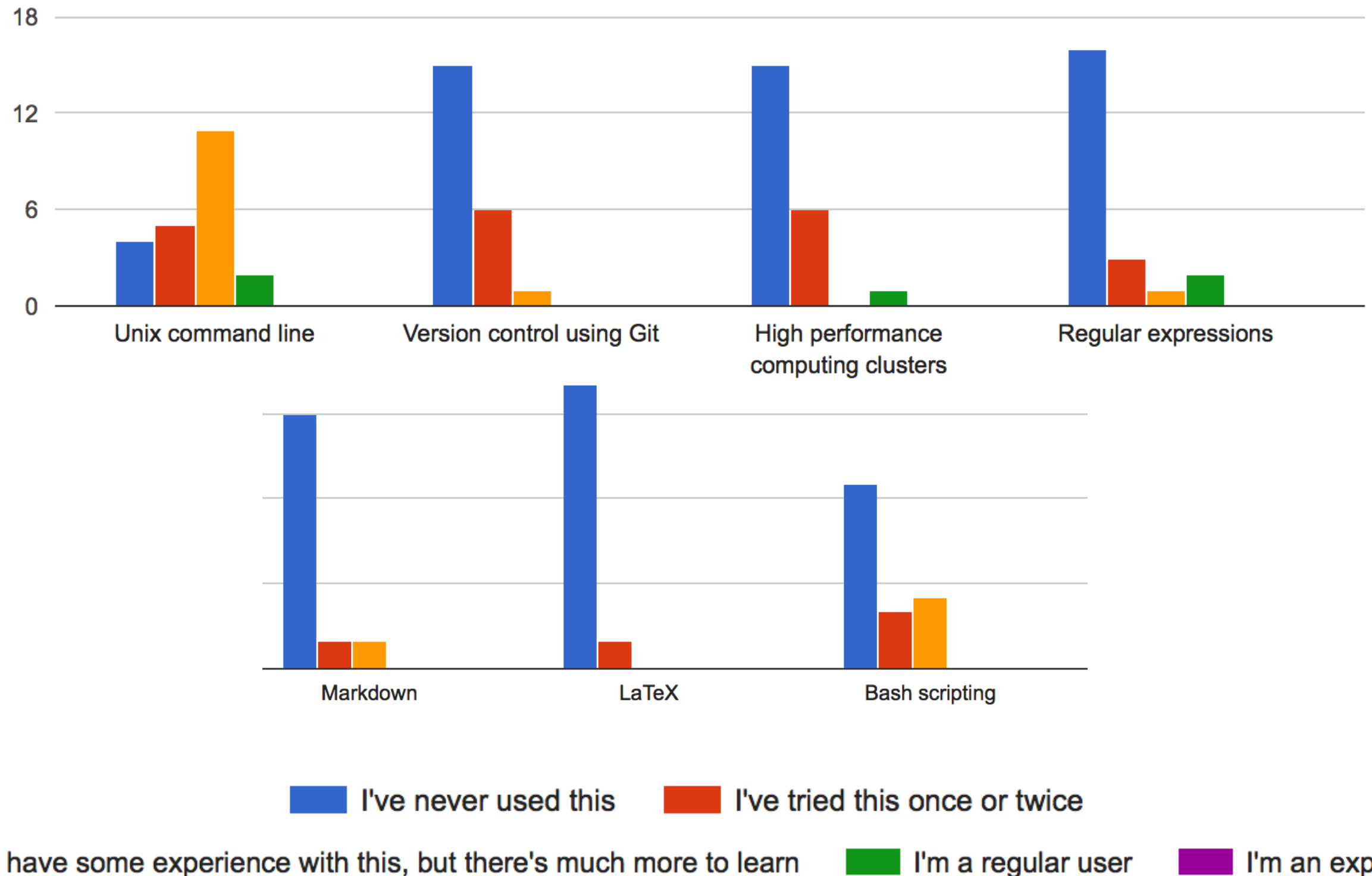
(22 responses)



# Introduction and Basic Unix

What motivated you to take this class?

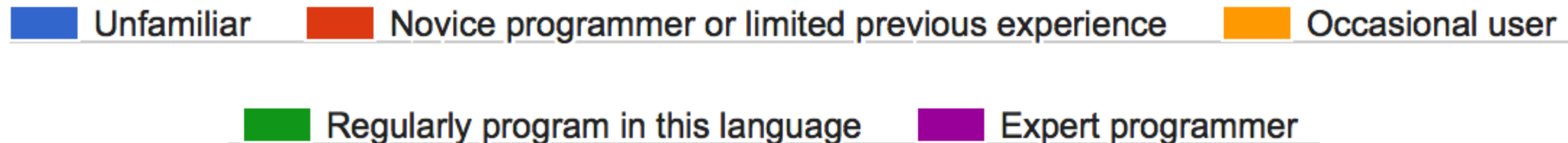
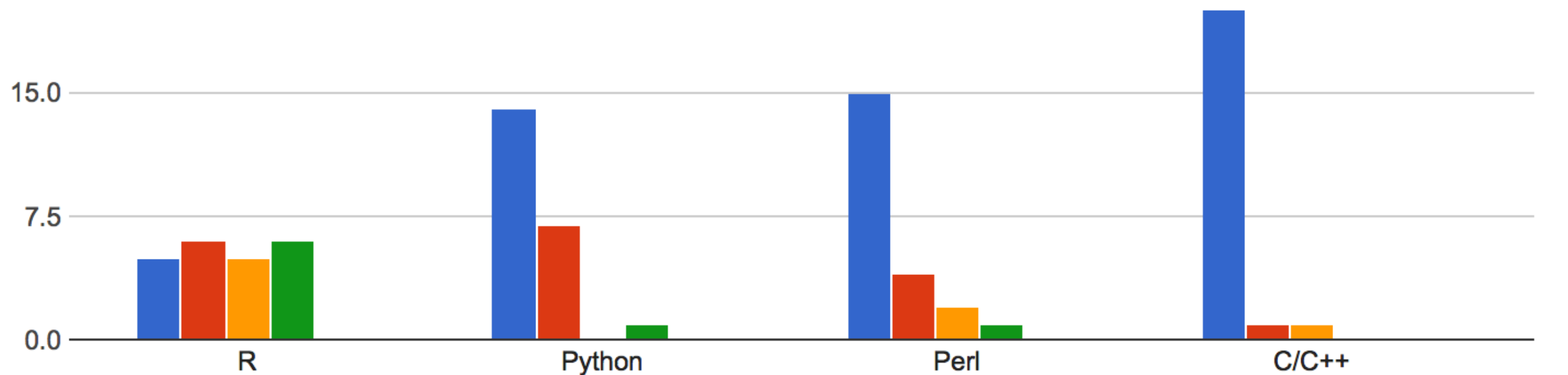
Familiarity with Topics



# Introduction and Basic Unix

What motivated you to take this class?

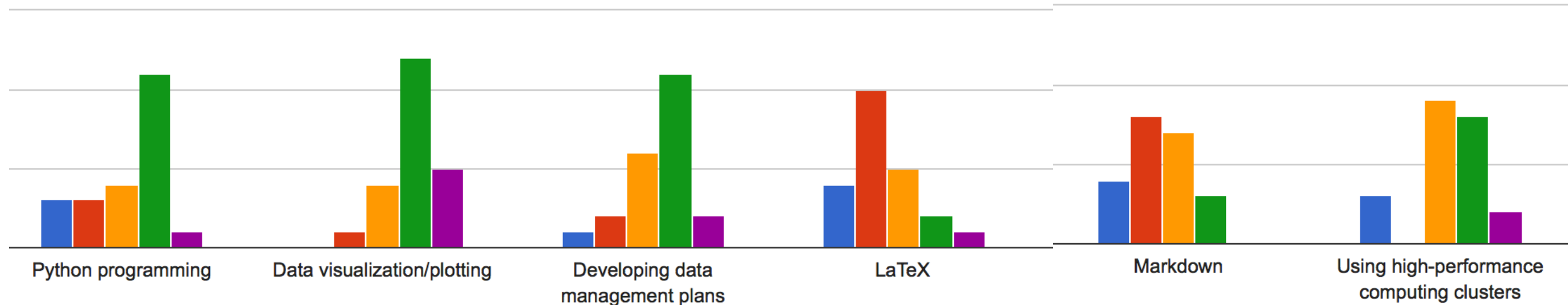
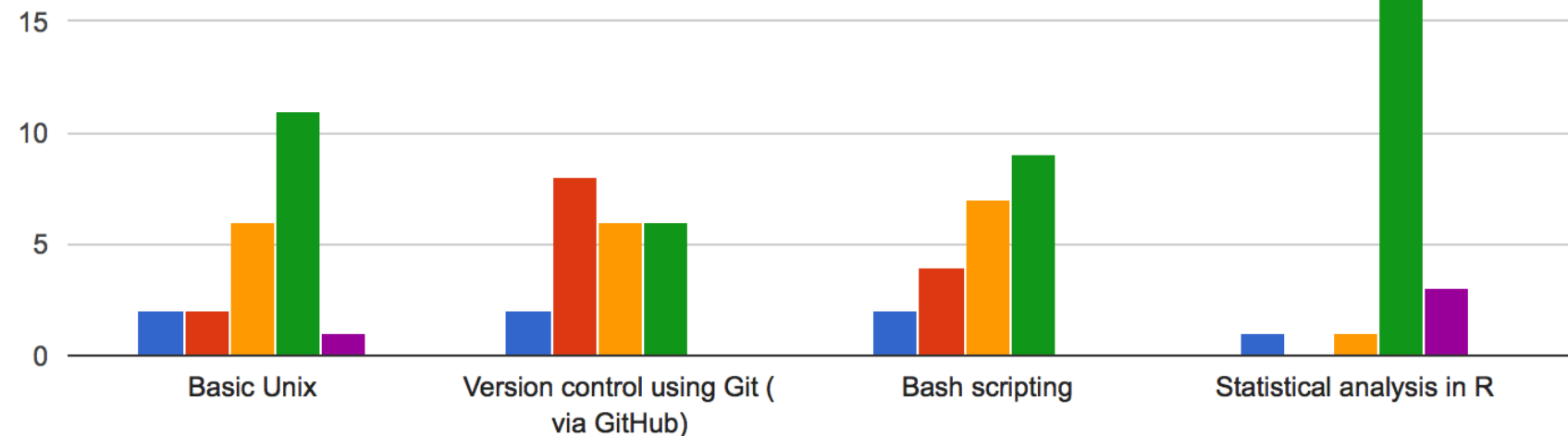
Familiarity with Coding Languages



# Introduction and Basic Unix

What motivated you to take this class?

## Topics of Interest



# Introduction and Basic Unix

What motivated you to take this class?

A few more take-homes...

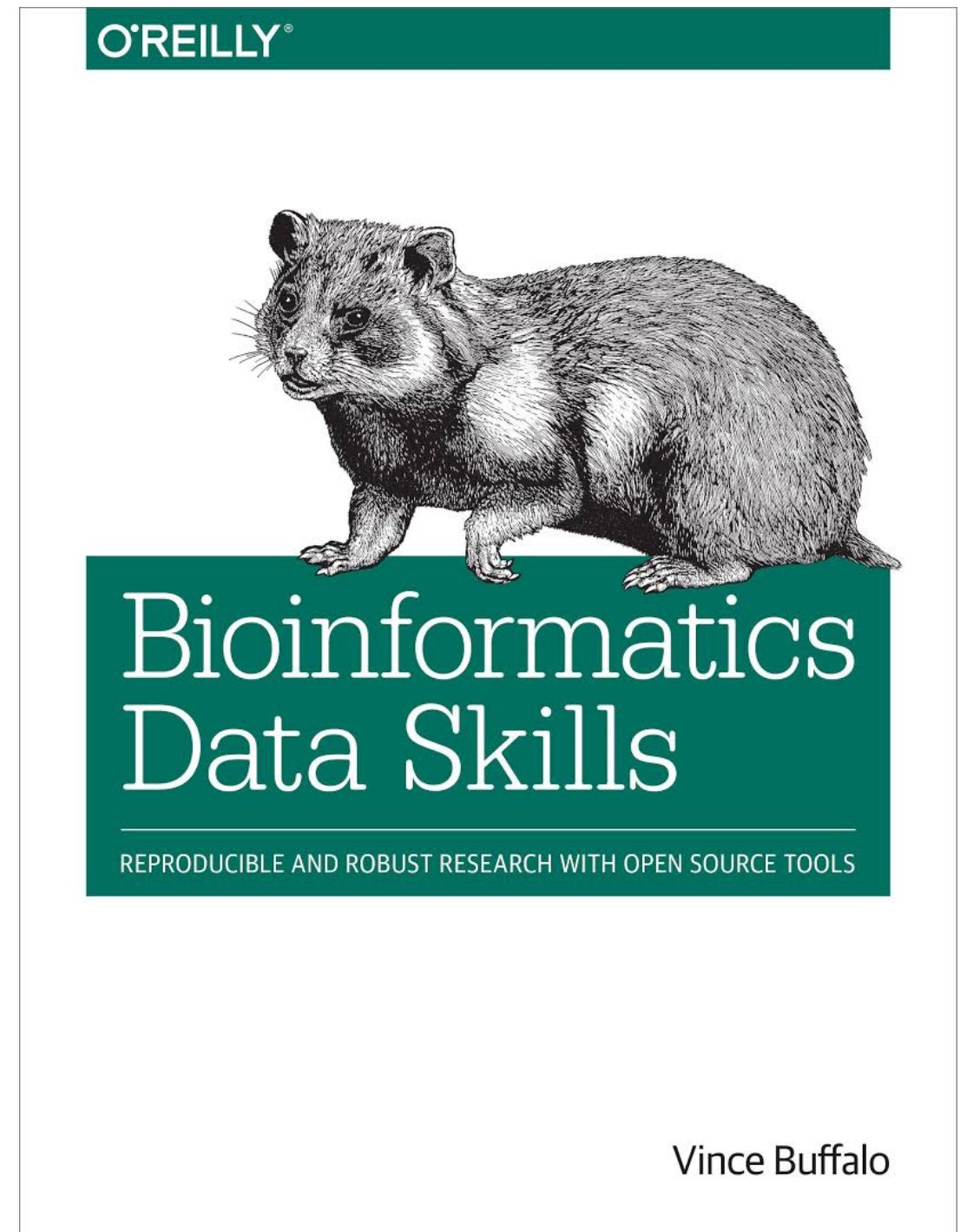
- ♦ Quite a lot of interest in creation, use, and maintenance of databases...we'll see what we can do about this
- ♦ You all are drinking from the data firehose!



# Introduction and Basic Unix

## Our Textbook

- ✦ Written to help address sudden need in biology to be able to handle Big Data
- ✦ Available through Amazon (hard copy), O'Reilly (hard copy and eBook), and ISU Library (eBook, FREE!!)



# Introduction and Basic Unix

## Our Objectives

By the end of this course, you should:

- Navigate through your computer, create and modify files and directories, and process data using basic Unix commands
- Become familiar with basic R syntax and data structures and implement these in data analysis and plotting.
- Utilize the Python scripting language for more sophisticated data processing.

# Introduction and Basic Unix

## Our Objectives

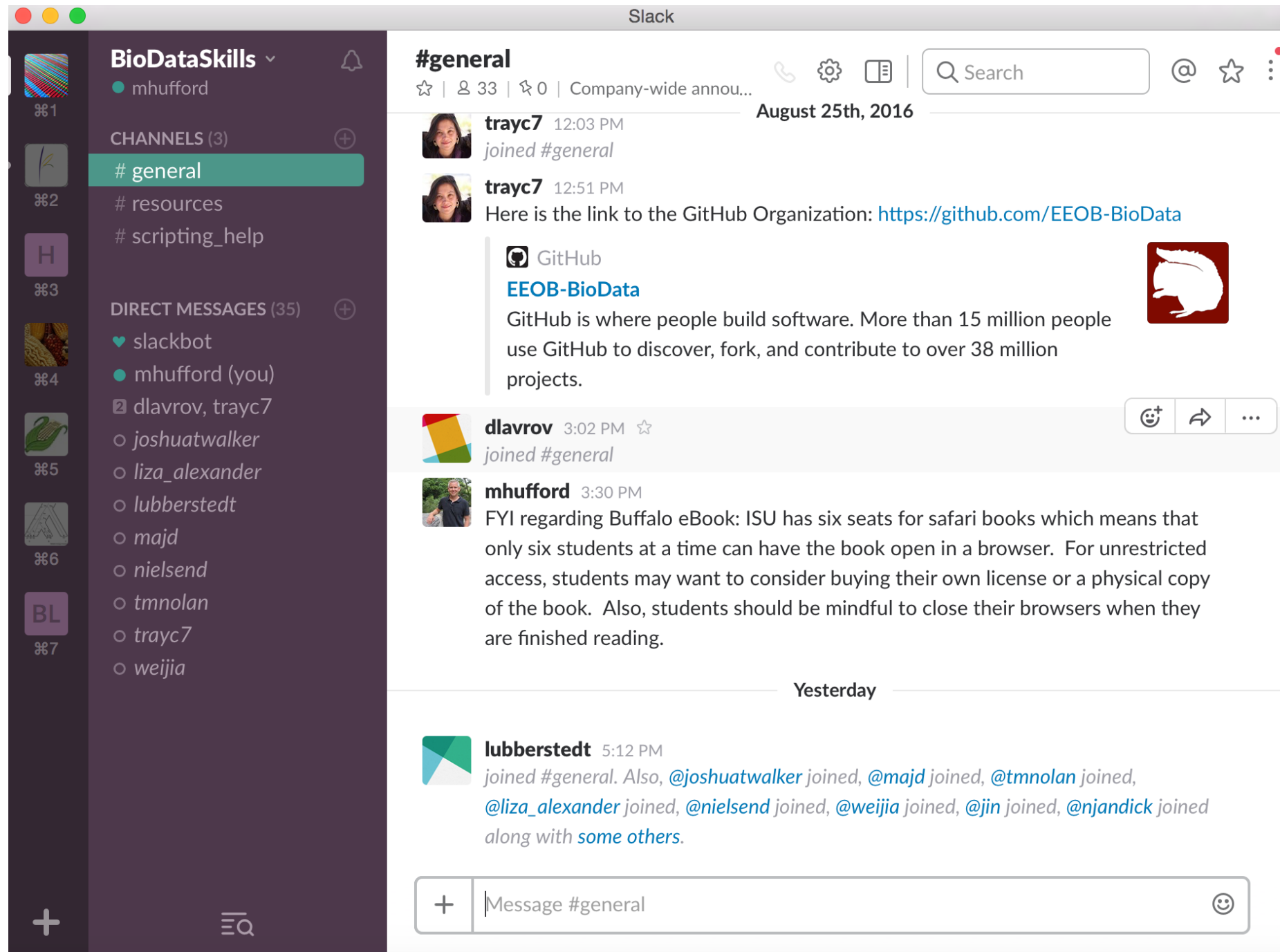
By the end of this course, you should:

- Become familiar with various genomic data types (range, sequence, and alignment data) and learn how to write scripts and analysis pipelines for working with these data.
- Become familiar with high performance computing resources at Iowa State as well as how and when to employ these resources.
- Explore additional resources/topics in computational biology including manuscript preparation in LaTeX and Overleaf and creation of NSF-style Data Management Plans.

# Introduction and Basic Unix

## How will we communicate?

### Slack



# Introduction and Basic Unix

What is our schedule?

Google Sheet

[https://docs.google.com/spreadsheets/d/  
1JBceaPuVd3BFrmCHOfKq64-MQeXMknZvTVWGGJpdDP44/  
edit?usp=sharing](https://docs.google.com/spreadsheets/d/1JBceaPuVd3BFrmCHOfKq64-MQeXMknZvTVWGGJpdDP44/edit?usp=sharing)

# Introduction and Basic Unix

## How will grades be assigned?

### Grading:

Assignment 1: Unix	15%
Assignment 2: R	15%
Assignment 3: Python	15%
Assignment 4: Data/Pipelines/HPC	15%
Group Project and Presentation	40%

# Introduction and Basic Unix

## Our Computational Goals for Today

1. Make sure everyone has a Unix solution
2. Installation of Git
3. Clone the Git repository for the textbook and the course
4. Work through a Basic Unix example

# Introduction and Basic Unix

## Where to from here?

1. If the basic Unix commands in our example were all new (and even if they weren't!), you should consider working through the Unix portions of these tutorials :

<https://sites.google.com/site/eeob563/computer-labs/lab-1>

[http://korflab.ucdavis.edu/Unix\\_and\\_Pperl/](http://korflab.ucdavis.edu/Unix_and_Pperl/)

2. If you haven't already, read Chapters 1-3 of Buffalo
  - For Chapter 1, create a text snippet in Slack with a few favorite points and any questions on points that were not clear, we'll talk through these on Thursday
  - We'll also discuss and work through examples from Chapters 2 & 3 on Thursday