

Welcome to BIS180L

author: Julin Maloof date: March 30, 2015

Course Personnel

Instructor: Julin Maloof jnmaloof@ucdavis.edu

Teaching Assitant: Kristen Beck kbeck@ucdavis.edu

What this course is about

incremental: true

The goal of this course is to introduce you to the *tools* and *thinking* required for bioinformatics analysis.

- Introduction to Linux command line
- Introduction to R
- Hands-on experience analyzing genomics data
- Principles of good analysis

This is a computer-based class. No bench work.

Why R and linux?

incremental:true * Both R and Linux have command-line interfaces * Antiquated? * NO! Written code provides **flexibility, creativity, and power** not available in any other way * Linux (or Unix / Mac) * Outstanding built-in tools for data crunching * Provides access to hundreds of bioinformatics programs * R * Powerful statistical, data processing, and graphical capabilities * Many bioinformatics packages are developed in R

And

[coding is cool](#)

“Every student deserves the opportunity to learn computer programming. Coding can unlock creativity and open doors for an entire generation of American students. We need more coders — not just in the tech industry, but in every industry.” – Mark Pincus, CEO and Founder, Zynga

“Coding is engaging and empowering. It’s a necessary 21st Century skill.” – Jan Cuny, Program Officer, National Science Foundation

“Code has become the 4th literacy. Everyone needs to know how our digital world works, not just engineers.” – Mark Surman, Executive Director, The Mozilla Foundation

“If you can program a computer, you can achieve your dreams. A computer doesn’t care about your family background, your gender, just that you know how to code. But we’re only teaching it in a small handful of schools, why?” – Dick Costolo, CEO, Twitter

Course Schedule

- **Tuesdays, Thursdays**
- Lecture 1:10 - 2:00
- Lab 2:10 - 5:00
- Often lecture will be shorter than 1 hour. We will start lab work as soon as lecture is over.

- **Fridays**
- TA Office hours 12:10 - 1:00 (in this room)
- Discussion 1:10 - 2:00
 - Varied use
 - Q & A
 - Student presentations
 - Keep on working

Course History

- Third year of the course
- Previous two years team taught
- This year it is all me...
- Some flexibility required...

(Tentative) Course Outline

incremental: true * Week 1–2: * Linux fundamentals * Markdown, git repositories * Sequence analysis and BLAST * Week 3: * R fundamentals * Weeks 4–5: * Illumina short reads * Calling SNPs * GWAS

- Weeks 6–7:
- RNAseq
- ChIPseq
- Week 8:
- Genetic Networks
- Week 9:
- Metagenomics
- Week 10
- TBD

Course Grading

- 45% Lab assignments
- 25% Take home midterm (Available May 1, **Due May 5, 1:10 PM**)
- 25% Take home final (Available June 2, **Due June 8, 9:00 AM**)

Do your own work

incremental: true

Developing code is an interactive process. Both your friends and the web can be excellent resources.

However Any direct copying of text or code **from any source** is considered plagiarism in the context of this course.

If you receive inspiration or ideas from an external source **give attribution**

Course Website

- Main website: http://jnmaloof.github.io/BIS180L_web/
- Lab instructions
- Course Schedule
- Helpful Links (coming soon!)
- Reading assignments (coming soon!)
- Smartsite
- Gradebook
- Chatroom

Why are we here?

incremental: true And what is bioinformatics? * OK so it is a requirement, but what else?

Three principles of bioinformatics

transition: rotate

1. Clear documentation
2. Reproducible results
3. Documents/Data in open (non-proprietary) formats
4. This is essential for achieving 1 and 2

Today’s Lab

type: section transition: rotate

1. Get a virtual linux machine running
2. Learn a little markdown
3. Learn about linux

Virtual linux machine

incremental: true

- The computer lab machines run Windows
- Bioinformatics on a Windows machine is painful or worse (R is OK)
- Solution: virtual machine!
- Use [VirtualBox](#) to run a virtual linux machine
- Your virtual machine is pre-loaded on a flash drive
- You can download virtualbox to your personal computer and run the machine at home

Other virtual machine notes

incremental:true

As detailed in the lab notes for today:

- Use the USB 3 (Right hand) USB ports
- It is imperative that you properly shutdown the virtual machine before removing the flash drive
- Data will be backed up in the cloud using MEGA
- If you are already running linux, or want to try installing the relevant software on your Mac, I can give you installation notes. **But we will not be able to help you troubleshoot installation problems**

Markdown

Markdown is a text-based formatting system for quickly and easily generating nicely formatted output.

It helps achieve all three guiding principles:

1. Clear documentations
2. Reproducibility
3. Open formats

Markdown vs docx

What is we want to produce this:

A markdown example

bulleted list

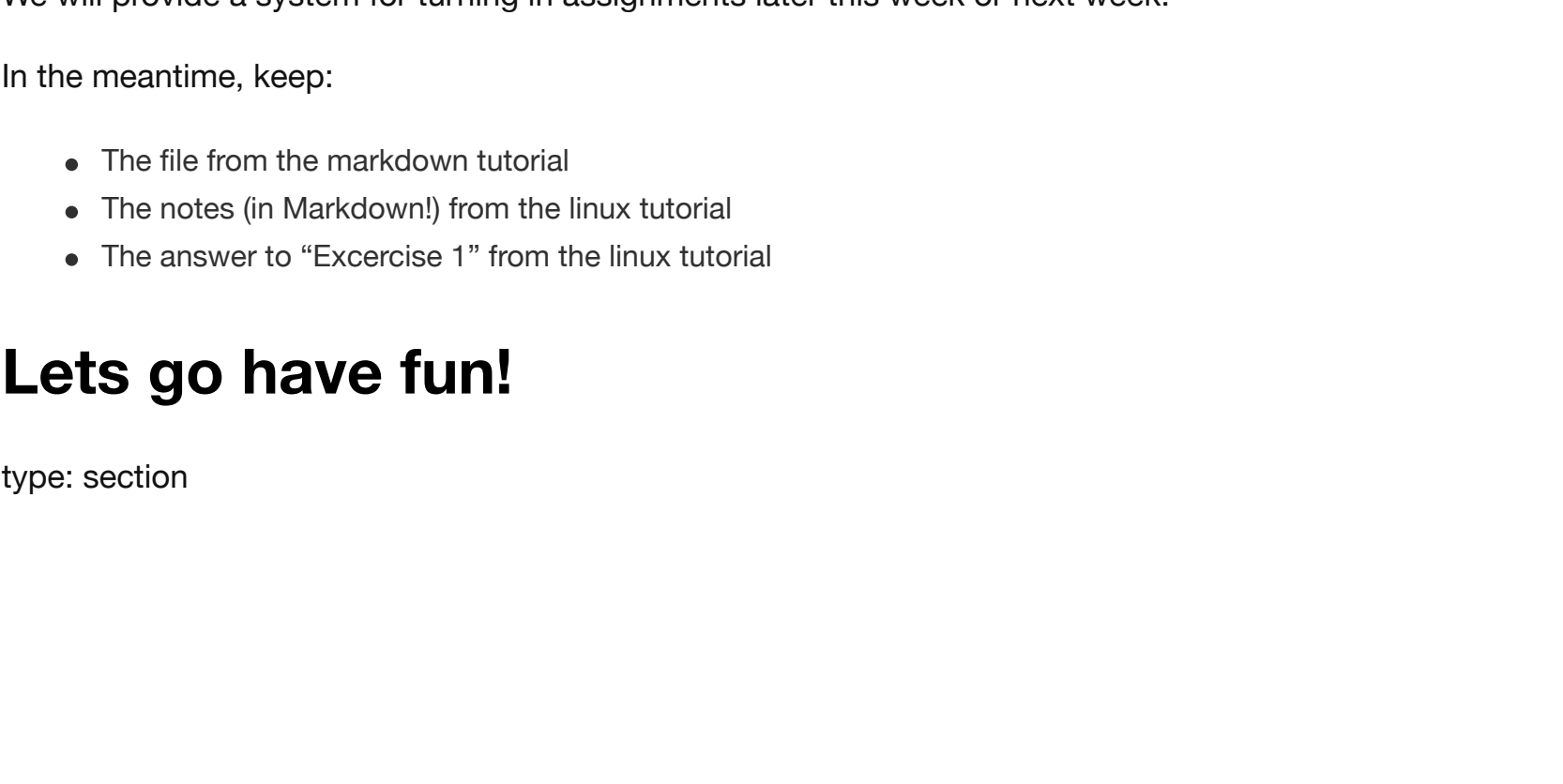
- point 1
- point 2
- point 3

some **bolded** or *italicized* text

code chunk

```
#simple code
x <- 1:10
y <- 1:10
x + y
```

format



Assignments to turn in for this lab

We will provide a system for turning in assignments later this week or next week.

In the meantime, keep:

- The file from the markdown tutorial
- The notes (in Markdown!) from the linux tutorial
- The answer to “Exercercise 1” from the linux tutorial

Lets go have fun!

type: section