MODULE 0 / UNIT 0

OVERVIEW OF STATISTICAL COMPUTING

Instructors

- **Instructor**
  - Name: Hyun Min Kang
  - Email: hmkang@umich.edu
  - Room: M4531 SPH II

- **Graduate Student Instructor**
  - Name: Tingting Zhou
  - Email: tkzhou@umich.edu

Class schedule

- **Class schedule**
  - Tuesday and Thursday 8:30-9:00am
  - No class on Thursday October 19th.
  - Midterm on Tuesday October 24th, in class
  - Final Exam on Friday December 15th, 10:30am – 12:30pm

- **Location**
  - Lectures : 1755 SPH I
  - Office hours: ???

Last chance!

- If you haven't voted yet, please do so now in 2 minutes!!

goo.gl/hVeU5z

Tentative Office Hours

(May change by poll results)

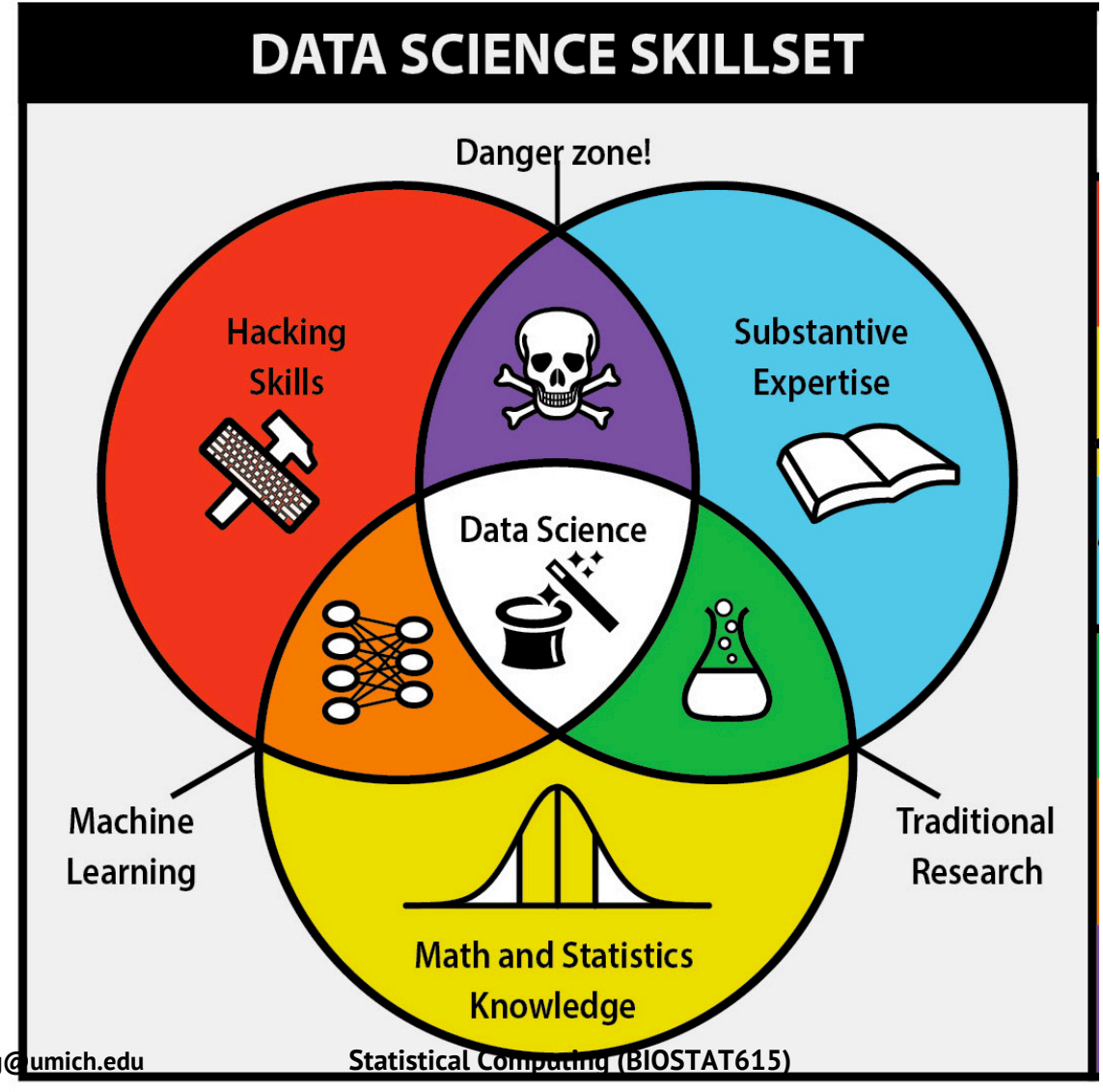
- **Instructor Office Hour**
  - Monday 4-5pm at M4531 SPH II

- **GSI Office Hours**
  - Tuesday 3-4pm at M4117 SPH II
  - Thursday 10-11am at M4117 SPH II

Goal of BIOSTAT615 in Fall 2017

Successful students should know upon completion..

- **Basic programming skills**
  - How to write code with python, R, and C++
  - How to interface R and C++
  - How to create an R package.
- **Basics of algorithms & data structure**
  - Divide and conquer, dynamic programming, and graph algorithms
  - List, tree, hash table, and graphs
- **Basics of numerical methods**
  - Randomization
  - E-M algorithm & MCMC methods


DATA SCIENCE SKILLSET
Danger zone!
Hacking Skills
Substantive Expertise
Data Science
Machine Learning
Traditional Research
Math and Statistics Knowledge

Eligibility to take BIOSTAT615

- **Students are qualified if they know…**
  - Basics of linear algebra
    - matrix multiplication, inverse, determinant, …
  - Multivariate calculus
    - partial derivative, integration, …
  - Basics of statistics
    - common distributions, linear model, …
  - How to do write basic code with R & python

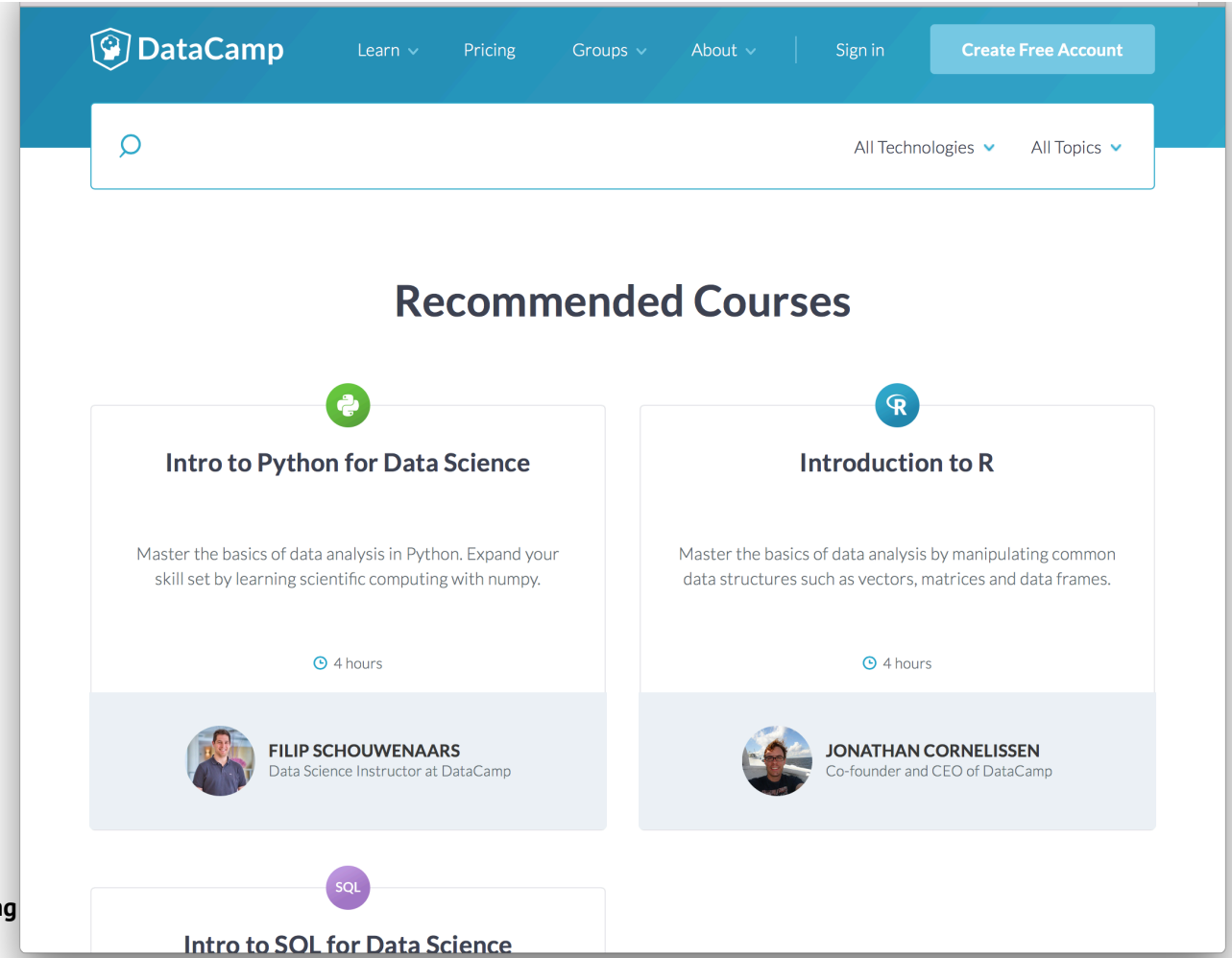
What if I miss one of the eligibilities?

- **If you do not know algebra and calculus…**
  - Not taking this class is strongly recommended.

- **If you do not yet know basics of statistics**
  - Take BIOSTAT601 (or equivalent) in parallel.

- **If you do not have experience in R or python at all…**
  - You need to spend **EXTRA** time to learn R and python yourself
  - … through **DataCamp.com**

What is DataCamp?



Access to DataCamp premium contents

- By the courtesy of DataCamp, students taking BIOSTAT615 will have access to all courses DataCamp provides.

- Students without R/python experiences should complete the following courses *within the first two weeks*
  - Intro to Python for Data Science
  - Intermediate Python for Data Science
  - Introduction to R
  - Intermediate R
  - UNIX : https://www.codecademy.com/learn/learn-the-command-line
- If your final grade is B+ or lower, having completed these courses will be slightly beneficial towards your grade.

Grading policy : distribution

- Homework (40%)
- Canvas Quiz (10%)
- Midterm exam (20%)
- Final exam (30%)
- Extra-credit project (??)

Grading policy : Homework

- Assignments will be announced approximately every other week
- Discussion on assignments between students are encouraged
- Students MUST NOT SHARE any piece of their code with others
- Plagiarism will NOT be tolerated and will be graded as ZERO.
- Late homework will NOT be graded.

Grading policy : Canvas Quiz

- At the end of each lecture, a set of Canvas quiz will be available.
- The due is BEFORE the next lecture.
- Scores will NOT contribute to the grade,
  only the record of completion will be counted towards grade.

Grading policy : Midterm and Final Exams

- **Closed book exam.**
- **Multiple choices and fill-in-blank questions only.**
- **Dates**
  - Midterm : October 24th (Tue) in class
  - Final : December 15th (10:30am-noon)
- **If midterm is missed due to approved medical reasons or other inevitable circumstances, final exam score will be extrapolated.**

Grading policy : Extra-credit project

- **Open to everyone, but only useful in boosting A to A+**
  - Without extra-credit project, the best achievable grade is A.
  - If your grade is A (without the project), and you finished the extra-credit project, then you may qualify for A+.
  - If your grade is A- or lower (without the project), finishing the extra-credit project will NOT affect to your final grade.
- **Choose a topic:**
  - Example problems that will be posted at the end of September
  - or a choice on your own (approved by the instructor)
- **Due is December 7th (Thu), 2017**
  - Submit your code, a brief document, and slides
  - Selected teams will present to the class in the last lecture (Dec 12).

Textbooks (not required, but recommended)

1. **[CLRS] Introduction to Algorithms**
   - Cormen, Lieserson, Rivest, and Stein
   - MIT Press, 3rd Edition (2009)

2. **[PTVF] Numerical Recipes**
   - Press, Teukolsky, Vetterling, and Flannery.
   - Cambridge University Press, 3rd Edition (2007)

3. **[RK] Statistical Computing in C++ and R**
   - Randall and Kupresanin.
   - CRC Press (2011)

Regarding class start/end time

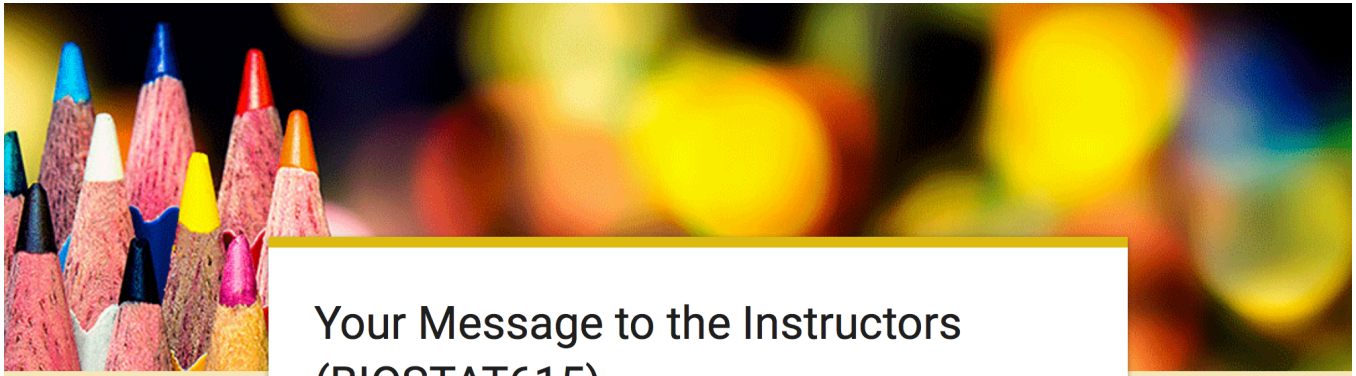
- **We will start at 8:35am** *(5 minutes earlier than Michigan time)*
  - First 5+ minutes will be used to review the material from the previous lecture, and the Canvas quiz results
  - Feel free to come in later than 8:35am if needed to. (especially if the Canvas quizzes were too easy for you ☺)

- **.. and will end no later than 10:00am.**

- **If this policy cause problems due to your special circumstances, feel free to send messages to the instructor.**

Express your thoughts anonymously.

- https://goo.gl/vR4M99



Your Message to the Instructors (BIOSTAT615)

This is an ANONYMOUS form for sending your messages -- including complaints, suggestions, compliments, or any other thoughts -- to the instructors of BIOSTAT615 Fall 2017 semester. Your identity will be protected, and your message will be notified to and read by the instructor.

* Required

Write your message to the instructor below. *

Your answer

Typically, the instructor will talk about your request during the lecture to suggest possible resolutions. Please check if you do NOT want such a feedback.

Any questions on the logistics of the course?