Computational Structures in Data Science

Lecture 1:
Welcome to CS88!

Berkeley
UNIVERSITY OF CALIFORNIA

Michael Ball | UC Berkeley |

Useful Links

- https://c88c.org
- https://go.c88c.org/qa1

Goals today

- Introduce you to
  - the field
  - the course
  - the team
- Answer your questions

- Big Ideas:
  - Abstraction
  - Data Type

CS88 Team –Michael

- Michael Ball
  - ball@Berkeley.edu – You're best off by using Ed! ☺
  - 784 Soda Hall
  - http://michaelball.co – I don't update this much…
    - It was great procrastination when I was a CS student.
  - Office hours: tentatively Monday afternoon. "coffee chat"
    - My office hours should be conceptual, higher level advice. ☺
- Things I do:
  - Intro CS Research (Tools, curriculum)
  - Training TAs
  - Building Educational Software (Snap!, frmr Gradescope)
  - Tools for web accessibility

CS88 Team – Head TAs : cs88@berkeley.edu

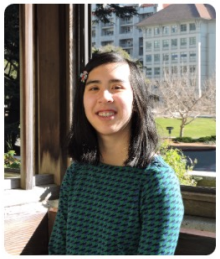
Head Teaching Assistants

Ethan Yoo Photo

Ethan Yoo [he/him]

ethanyoo7912@berkeley.edu

Third-year Applied Mathematics + Computer Science major. The only reason why I'm a math double is so I can avoid taking CS70 >_<.



Rebecca Dang [she/her]

rdang@berkeley.edu

Hey there! I'm a 3rd year EECS major from San Jose, CA. This is my 3rd semester on course staff and I'm super excited to meet you! Feel free to chat with me about 88, classes at Berkeley, professional development, Computer Science Mentors, guitar, books, movies, TV, music, and more :D

CS88 Team -- TAs

Angela Bi [she/her]

angelabi2003@berkeley.edu

Hello! I'm Angela and I'm a junior majoring in Data Science and minoring in Political Economy from the bay area. In my free time, I like illustrating for the Daily Cal, playing volleyball, listening to music, and watching anime. Feel free to reach out and looking forward to a great semester!!

John Teng [he/him]

johnteng9@berkeley.edu

Hi, I'm John, a third year CS major from Pennsylvania. I like playing video games, soccer, and working out. Looking forward to this semester!

Karim El-Refai [he/him]

karim.el-refai@berkeley.edu

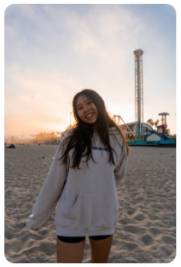
Fan of robots and philosophy and the philosophy behinds robots and soon the robots behind philosophy, but not the philosophy of the philosophy of robots but in fact the robots for the robots of philosophy. If you are confused, don't worry so am I :D

Liliana Gonzalez [she/her]

liliana@berkeley.edu

hi!! my name is lily and i am a third year from sacramento, ca. i love listening to music, shopping, drinking coffee, and going to concerts. please feel free to reach out with any questions - i'm looking forward to meeting you all! :)

CS88 Team -- TAs



Michelle Chen [she/her]

michelle.chenn@berkeley.edu

im michelle, cs+econ junior from singapore. i<3 traveling, eating, snowboarding, photography and my students :D

Ramya Chitturi [she/her]

ramya.chitturi@berkeley.edu

Hi! I'm Ramya, a junior majoring in CS and linguistics. I enjoy sci-fi/fantasy books, crosswords, rock music, museums, civic technology, and more! Excited to get to know you this semester :)

Sean Yang [he/him]

sean_yang@berkeley.edu

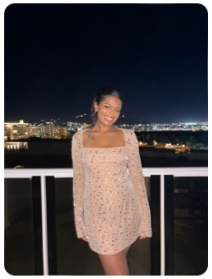
Hi my name is Sean, I'm a third year studying data science. Nice to meet you!

CS88 Team -- Tutors

Christy Quang [she/her]

christyquang@berkeley.edu

Hi hi! My name is Christy and I'm a third year CS & DS major from the East Bay. In my (limited) free time, I enjoy destroying people in Word Hunt, watching basketball (Warriors), listening to kpop and eating fruit! Excited for a fun semester in 88 and always feel free to reach out :D



Miha Bhaskaran [she/her]

athmiha@berkeley.edu

Hello hello! I'm Miha, a third year studying data science from the bay area. In my free time I love playing card games, critiquing all forms of media and going on spontaneous road trips. Looking forward to a fulfilling semester :)

Morgan Dehdashti [she/her]

mdehdashti@berkeley.edu

Hi! My name is Morgan and I'm a junior from socal majoring in Data Science and MCB. Looking forward to a great semester!

In The News

AI-Driven Misinformation 'Biggest Short-Term Threat to Global Economy'

The World Economic Forum's annual risks report, based on a survey of 1,300 experts, revealed that respondents believe the biggest short-term threat to the global economy will come from AI-driven misinformation and disinformation. This is a major concern, given that elections will be held this year in countries accounting for 60% of global gross domestic product. Other short-term risks cited by respondents include extreme weather events, societal polarization, cyber insecurity, and interstate armed conflict.

The Guardian; Larry Elliott (January 10, 2024)

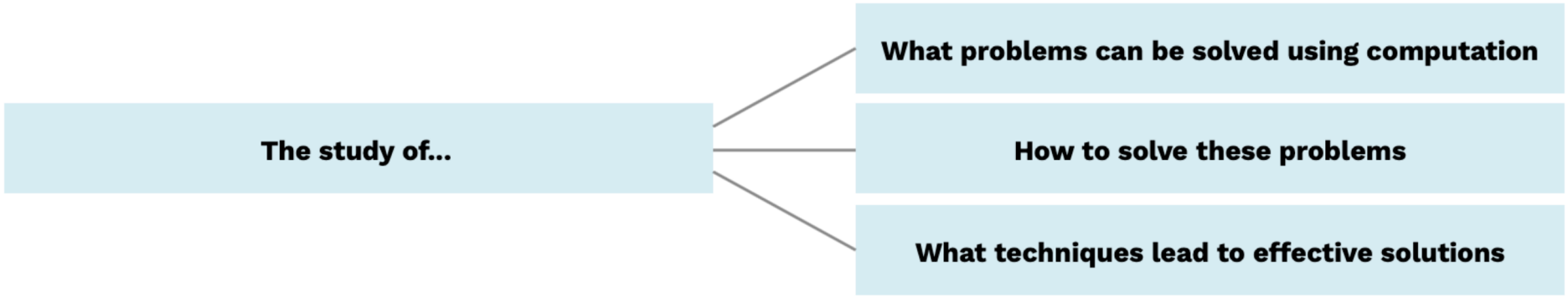
Computational Structures in Data Science

Computer Science and Data Science

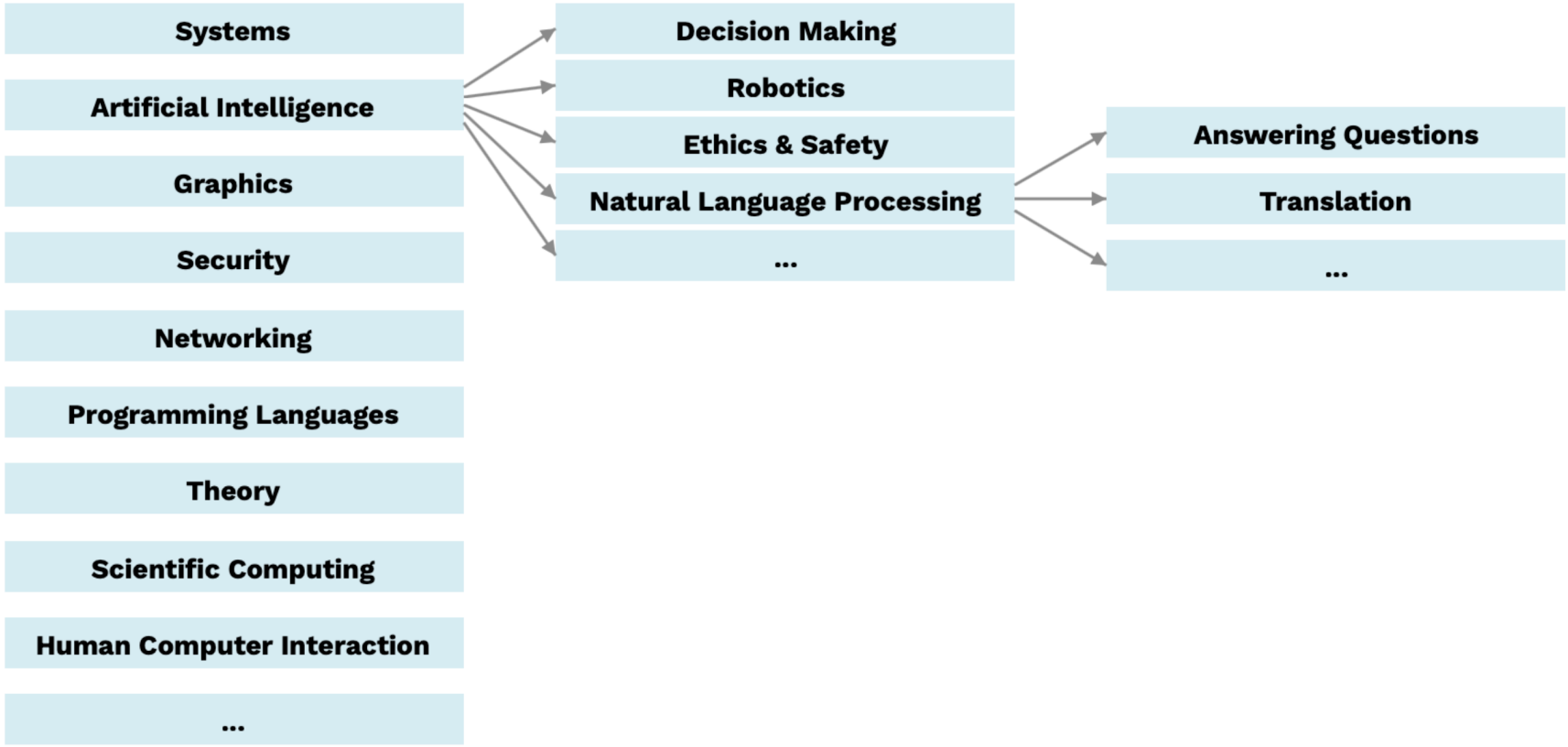
Berkeley
UNIVERSITY OF CALIFORNIA

Michael Ball | UC Berkeley |

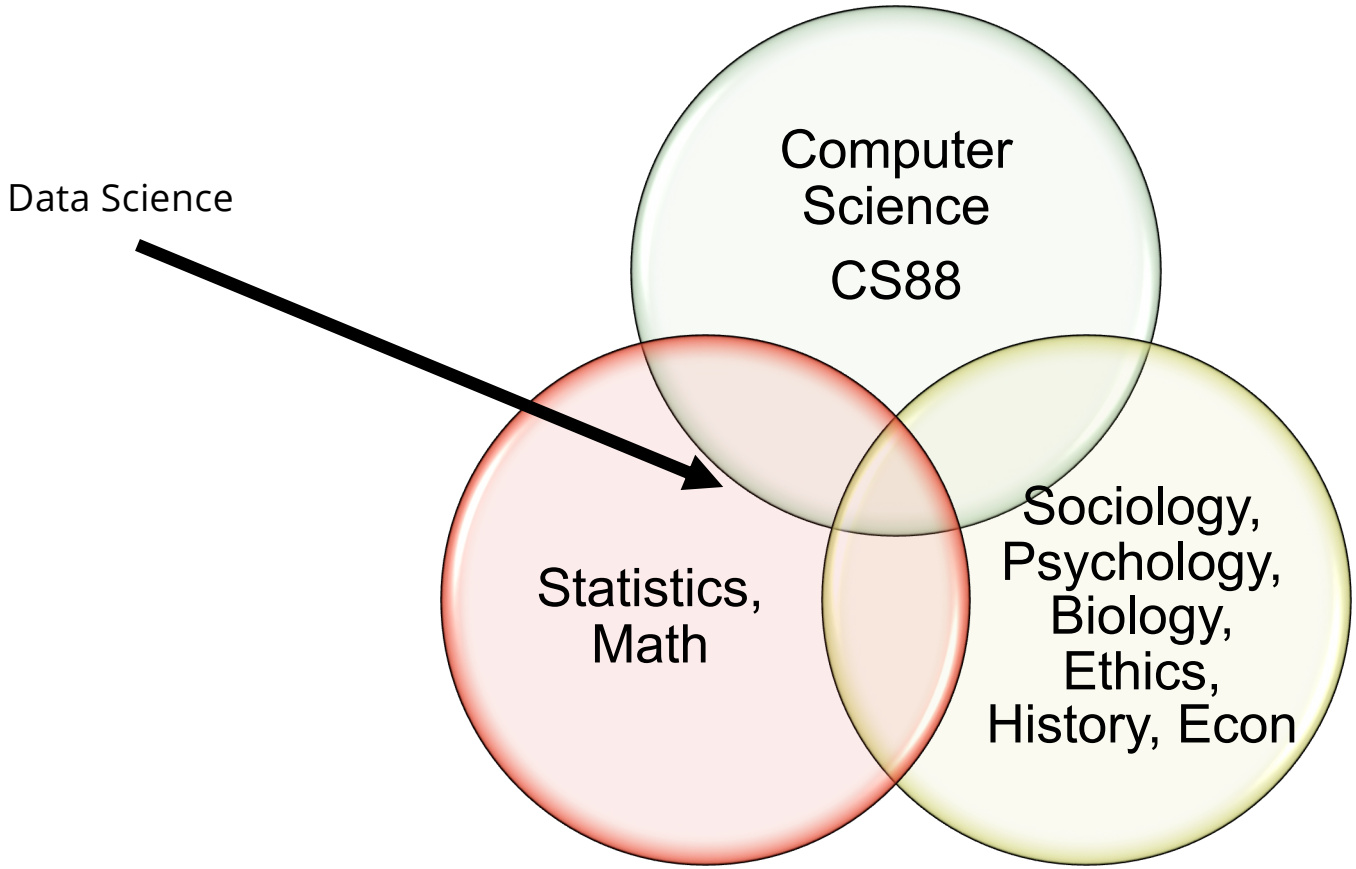
Computer Science



Computer Science, Some Ideas...Definitely Not Exhaustive!

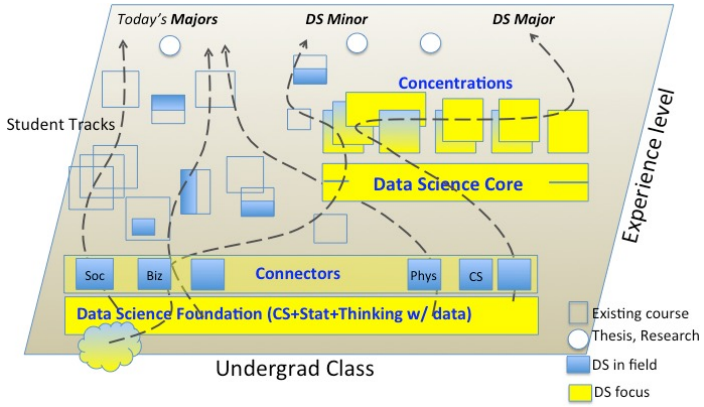


Computer Science & Data Science (One View)



Data 8 – Foundations of Data Science

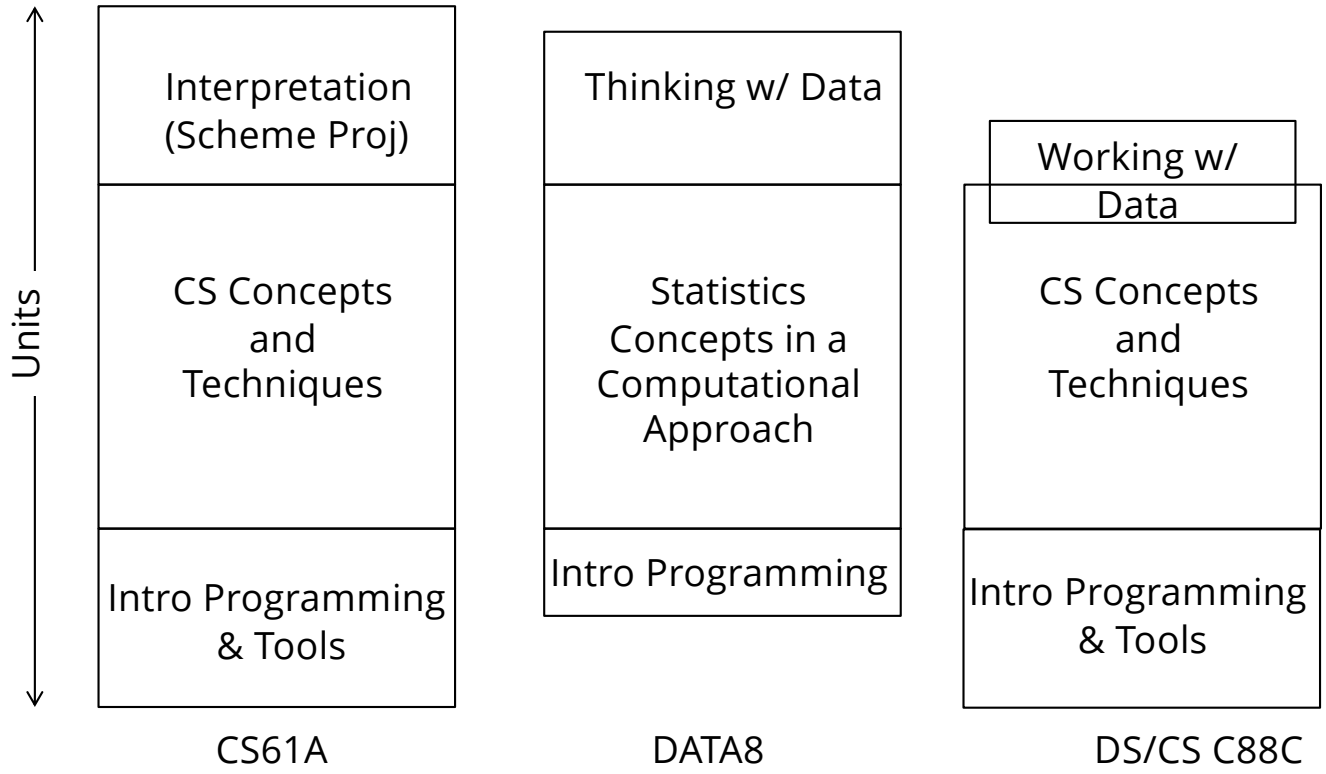
- Computational Thinking + Inferential Thinking in the context of working with real world data
- Introduce you to several computational concepts in a simple data-centered setting
  - Authoring computational documents
  - Tables
  - A LOT of statistics



CS88 – Computational Structures in Data Science

- Deeper understanding of the computing concepts introduced in DATA8
  - Hands-on experience => Foundational Concept
  - How would you create what you use in DATA 8?
- Extend your understanding of the structure of computation
  - What is involved in interpreting the code you write?
  - Deeper CS Concepts: Recursion, Objects, Classes, Higher-order Functions, Declarative programming, …
  - Managing complexity in creating larger software systems through composition
- Create complete (and fun) applications
- In a data-centric approach

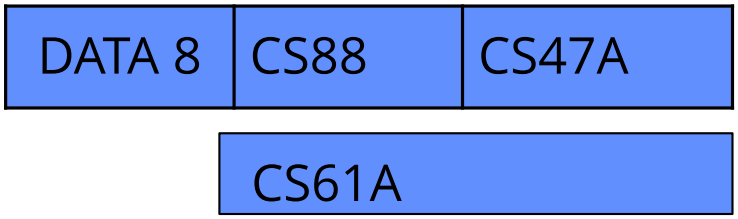
How does C88C relate to CS61A?



Opportunities for students

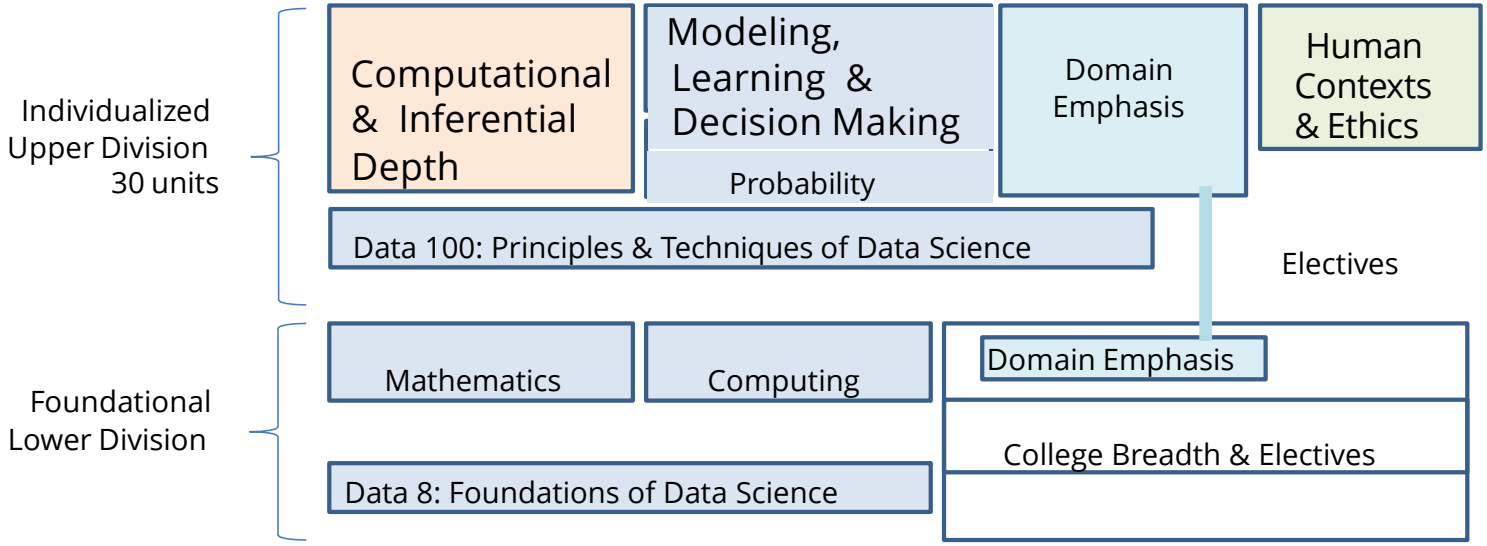
DS Minor

CS minor & DS Major



CS major:
CS47A path is not recommended, but there if you change your mind.

The Data Science Major



Computational Structures in Data Science

Success In C88C

Berkeley
UNIVERSITY OF CALIFORNIA

Course Culture

- Learning
- Community
  - **Collaboration**
  - Peer Instruction
- Respect
- A supportive course staff & environment
  - Lots of outside community, CS Mentors, HKN, others.

Collaboration

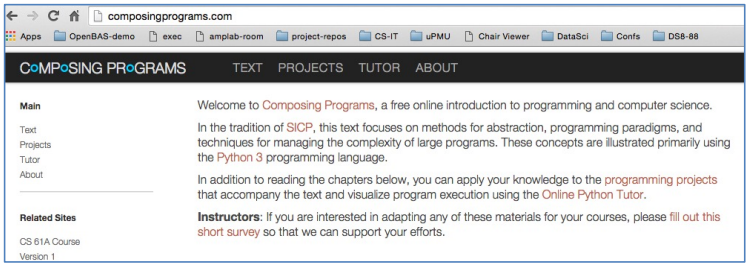
- Asking questions, discussing topics, helping each other is always encouraged!
  - When you're working with a partner, you are expected to share in the work.
- Collaboration has limits
  - Please don't read someone else's code
    - except if you have already turned in the assignment, or a TA/staff member is present.
  - You can help others, but not give the solutions.
- We have a very particular set of skills and we will use them.

So...We Know About ChatGPT

- You *may* use ChatGPT (and other tools) like a **tutor**
- **You may not ask about specific questions assigned to you!**
- Remember: It's a bot.
  - Bots are fallible!
  - Check its advice.
- Good Example:
  - "When would I prefer a for-loop over a while loop in Python?"
  - "What is a higher order function?"
- **Bad Example:**
  - "Provide a solution to the Fibonacci sequence using recursion"

Course Structure

- 2 lectures, 1 lab each week
- Lecture introduces concepts (quickly!), answers why questions.
- Lab provides concrete detail hands-on
- Homework (11) cements your understanding
- Projects (2) put your understanding to work in building complete applications
  - Maps
  - Ants vs Some Bees



Class Format

- Mon and Weds Lectures:
  - Each lecture has a series of short self-check questions
  - Lectures go quickly
- Labs are paced throughout the week. See the Ed post to pick a time.
- Labs are HANDS ON – Get help as you're trying the lab.
  - Labs are active
  - Find and make friends!
  - Do not ask for more lectures
    - Ask *why* something doesn’t make sense.

Class Format: Assignments

- Lecture Quizzes, 1 point, max 20.
  - 1 per lecture, “due” in ~4 days. (Partial credit after)
  - https://go.c88c.org/1
- Lab Work: 4 points, 11 labs, 1 drop
  - Start them during lab! You can probably finish some labs in 2 hours. Will be Python + some interactive questions.
- Homework: 8 points, 11 HW, 1 drop
  - Start early!
- Projects: 100 points between 2 projects
  - Start early! ”Checkpoint” assignments

Lab Attendance & Credit

- Labs are graded on correctness
- If you attend, you get 2/4 points
- You get the other 2/4 points by solving questions
  - Most labs will have 2-5 questions.
- If you can't attend, you solve all the lab questions.
  - **There are no "make ups" or excused absences.**
  - Solving half the questions shouldn't take 2 hours in most cases.
- If folks abuse the attendance system, we'll change it.

Extensions & Extenuating Circumstances

- https://go.c88c.org/extensions
- **Contact us early!!**
- Our goal is to have you do the work, but we can't manage things at 11:50PM
- If you need <= 3 days, **do not** submit the form.
- If you know you will be travelling, etc. Let us know ASAP.
-

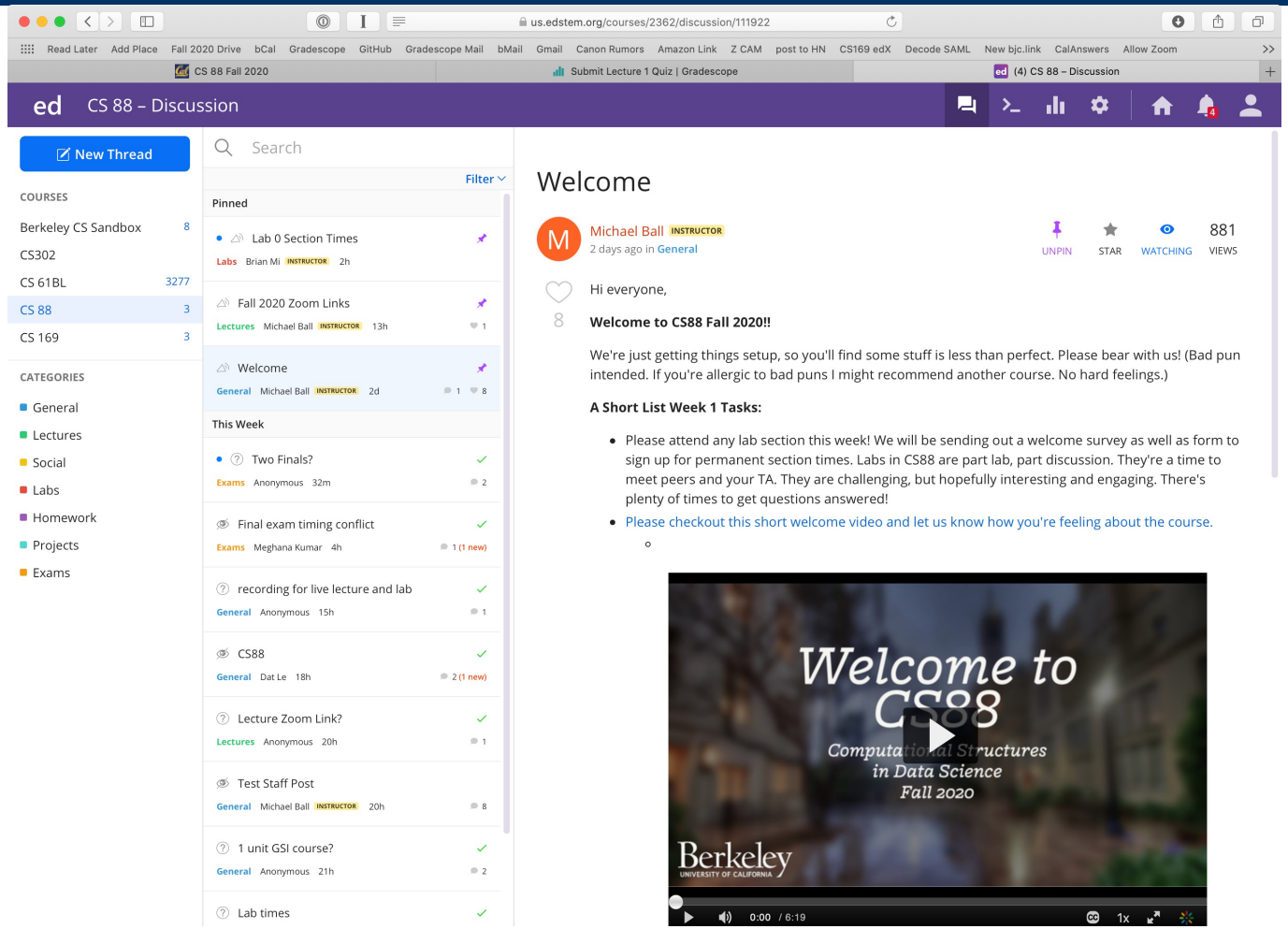
Class Format: Assignments

- Slip Days: 9 total
  - Use up to 3 on any assignment
  - We apply the in the order that's most beneficial!
    - i.e. use them on projects if you need!
  - Can be used for homework, labs, projects, but not project checkpoints.
- Slip Days take care of almost all special circumstances!
- We will release an exceptions / extensions request form.
- What if you go over slip days?
  - 25% deduction for each day over. Mathematically you can still earn 25% if you turn in something 3 days late.

Class Format: Exams

- 1 midterm and 1 final exam, in person
  - Remote exams proctored via Zoom
  - Required verification for alternate exams.
- Midterm 2 hours, early October
- Exam will be during the slot assigned by campus.
- 5 handwritten cheat sheets double-sided.
- **You don’t actually need 5 sheets!**

Ed For Class Discussion: Try it!



Where will we work?

- Your laptop
  - Using an editor and a terminal
- okpy – okpy.org
- c88c.org
- gradescope.com – HW, Lab, Lecture Self-Checks
- Ed Discssion: edstem.org
  - Can write and run (!!) python in you own posts!

Computational Structures in Data Science

Demo?

Berkeley
UNIVERSITY OF CALIFORNIA

Take Things 1 Step at a Time

- We interact with Python via the *Terminal*
- We type *programs* into files and into other programs.
- Everything you do in this class is safe!
- **Try and experiment!**

Your Tasks

- Lecture 1 "self-check" on Gradescope
- Attend Lab 0
- Attend OH

Welcome, and Good luck!

Questions?