




## Computational Structures in Data Science



UC Berkeley EECS  
Lecturer  
Michael Ball

### Lecture 1: Welcome to CS88!





UC Berkeley EECS  
Adj. Ass. Prof.  
Dr. Gerald Friedland

UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

1

## In The News

**Job Screening Service Halts Facial Analysis of Applicants But it's still using intonation and behavior to assist with hiring decisions.**

<https://www.wired.com/story/job-screening-service-halts-facial-analysis-applicants/>



HireVue, a leading provider of software for vetting job candidates based on an algorithmic assessment, said Tuesday it is killing off a controversial feature of its software: analyzing a person's facial expressions in a video to discern certain characteristics.


Job seekers screened by HireVue sit in front of a webcam and answer questions. Their behavior, intonation, and speech is fed to an algorithm that assigns certain traits and qualities.

UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

2

## Goals today

- Introduce you to
  - the field
  - the course
  - the team
- Answer your questions
- Big Ideas:
  - Abstraction
  - Data Type




UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

3

## CS88 Team – Gerald and Michael

- Michael Ball
  - [ball@Berkeley.edu](mailto:ball@Berkeley.edu) – You're best off by using Ed! ☺
  - 625 Soda Hall / Berkeley.zoom.us / my apartment
  - <http://michaelball.co> – I don't update this much...
    - » It was great procrastination when I was a CS student.
    - Office hours: tentatively Wednesday early afternoon.
- Things I do:
  - Intro CS Research (Tools, curriculum)
  - Training TAs
  - Building Educational Software (Gradescope)
  - Tools for web accessibility



UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

4

## CS88 Team – Gerald and Michael


- Gerald Friedland
  - [fractor@Berkeley.edu](mailto:fractor@Berkeley.edu) – You're best off by using Ed! ☺
  - 424 SD Hall / Berkeley.zoom.us
  - <https://www.ccs.berkeley.edu/~fractor> – I don't update this much either...
  - Office hours: tentatively Monday 2pm
- Things I do:
  - Multimedia Computing
  - Machine Learning Research/Data Science Measurements
  - Discussion Group at Berkeley Institute for Data Science (BIDS)

UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

5


## CS88 Team

### Head Teaching Assistants



Shreyas Karmali  
Office hours: Friday 10 am - 12 pm  
[skarmali@cs88.org](mailto:skarmali@cs88.org)


If everyone I'm super excited to meet you all this semester! In particular order, I prefer outside user concerns, cool new research, sharing your insights, and such over books. Feel free to take me down after recessions, cool study topics, and of course, CS 88!



Vandana Ganesh  
Office hours: Monday 10 am - 12 pm  
[vandana@cs88.org](mailto:vandana@cs88.org)


If everyone, I'm super excited to meet all of you this semester! This is my 3rd time teaching for CS, but my 25th time being involved with it all! I look forward to great questions & CS about me - I'm going on sabbatical, starting grad school (teaching), and going through my favorite to do, movies, and books. Feel free to come to me with recommendations and definitely feel free to reach out about CS 88!

### Teaching Assistants




Luna Chang  
Office hours: Thursday 1 - 2 pm  
[lchang@cs88.org](mailto:lchang@cs88.org)

If everyone I'm excited to meet with you all this semester! I'm a 2nd year CS major from the Bay Area. This is my first semester teaching. I'm also about to start my first year of grad school in Berkeley. I look forward to meeting you all!




Nicholas Ng  
Office hours: Tuesday 2 - 3 pm  
[nng@cs88.org](mailto:nng@cs88.org)

Hi! I'm Nicholas, and I'm a 3rd year Computer Science major from the San Diego Valley. I've played video games (I probably spent too much of my summer playing Warcraft), and every once in a while I'll go back for a bit to my roots (I hope). I look forward to meeting you all!



Sofia Shi  
Office hours: Tuesday 9 - 10 am  
[sshi@cs88.org](mailto:sshi@cs88.org)

Hi! I'm Sofia, a fourth-year EECS major from Santa Clara. I'm here this semester, working on my thesis, and I'm super excited to meet you all! I look forward to meeting you all!



Torrey Janson  
Office hours: Tuesday 1 - 2 pm  
[torreyj@cs88.org](mailto:torreyj@cs88.org)


Hi! I'm a second-year CS major from Golden Gate. I'm super excited to meet you all! I look forward to meeting you all!

6

1


## CS88 Team

### Tutors



Kevin Gu  
Office Hours: Monday 5 - 7 pm  
[kevingu@berkeley.edu](mailto:kevingu@berkeley.edu)

Hi everyone! I'm Kevin, a 4th year Data Science major and CS minor from Irvine, So Cal. I am so excited to be tutoring in this class again as I really like the content in this class. While I'm not analyzing data or programming/coding, I like to appreciate the humanities in general. Specifically, I love music, especially classical music or just any classics in general. I also love learning languages and enjoy talking/typing in French. Feel free to talk to me in any of these things, in either English, French (but speak SLOWLY please!), or even Python print statements and strings!



Minnie Chen  
Office Hours: Thursday 6 - 8 pm  
[chn\\_minnie\\_22@berkeley.edu](mailto:chn_minnie_22@berkeley.edu)


Hi, I'm Minnie and I'm a third year Civil Engineering major and EECS-Sustainability minor from San Gabriel Valley. When I'm lazy, I enjoy watching all kinds of movies/shows (from cheesy rom-coms to horrors and action) and when I'm not feeling like a couch potato, I love to do physical activities like hiking or basketball. Excited to meet everyone!

UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

7

## Course Structure

- 2 lectures, 1 lab each week
- Lecture introduces concepts (quickly!), answers why questions.
- Lab provides concrete detail hands-on
- Homework (12) cements your understanding
- Projects (2) put your understanding to work in building complete applications
  - Maps
  - Ants vs Some Bees



- Readings: <http://composingprograms.com>
  - Same as cs61a

UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

8

## Class Format

- Mon and Weds Lectures:
  - Some lectures will be pre-recorded, class time will be largely demo/Q&A
  - Each lecture has a series of short self-check questions
- 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.

UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

9

## Class Format: Assignments

- Lecture Quizzes, 1 point, max 20.
  - 1 per lecture, due in 1 week. (Half credit after)
- Lab Work: 4 points, 12 labs, 1 drop
  - Start them during lab! You can probably finish some labs in 2 hours. Will be Python + some interactive questions.
  - Out Tues, due next Tues Night.
- Homework: 8 points, 12 HW, 1 drop
  - Start early!
  - Out Weds, Due Next Friday Night

UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

10

## Class Format: Assignments

- Projects: 100 points between 2 projects
  - Start early! "Checkpoint" assignments
- Slip Days: 8 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 nearly all, but not all special circumstances!
- 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.

UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

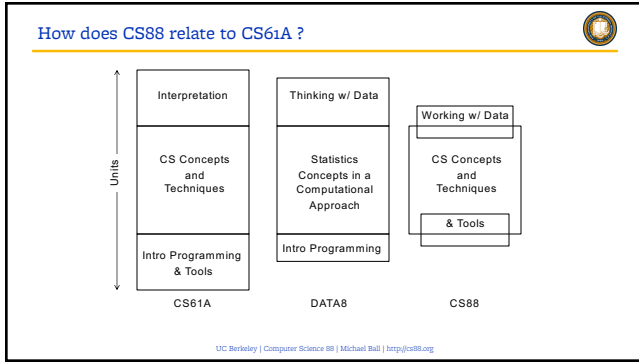
11

## Class Format: Exams

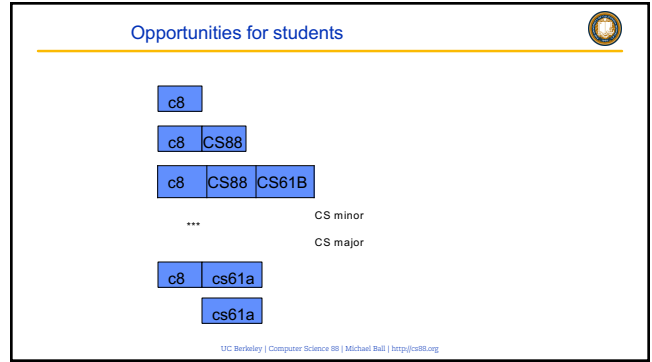
- 1 midterm and 1 final exam
- Midterm 2 hours, "Mid march", sometime before Spring Break; 2 hours.
- Exam will be during the slot assigned by campus.
- We will be proctoring via Zoom. Exact policies coming soon, but, essentially, you'll record yourself at home.
- Open book, but no collaboration.

UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

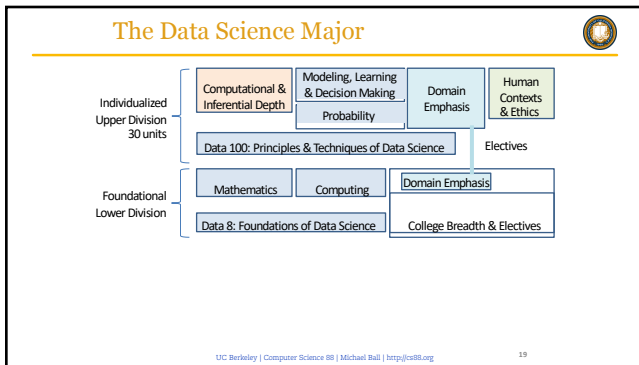
12



13



14



15

### Data Science

Nearly every field of discovery is transitioning from "data poor" to "data rich".

UC Berkeley | Computer Science 88 | Michael Hall | <http://cs88.org>

16

### A National Challenge

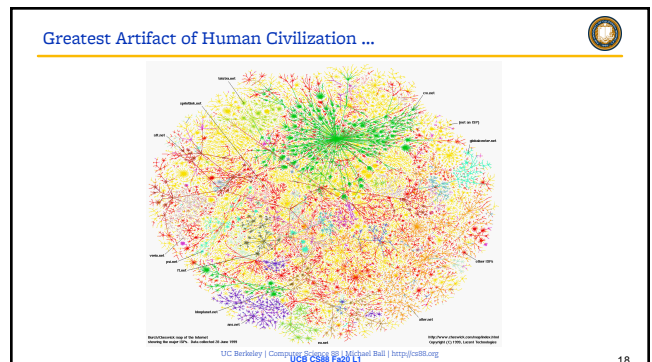
Big data: The next frontier for innovation, competition, and productivity.

Increasingly US jobs require data science and analytics skills. Can we meet the demand? The current shortage of skills in the national job pool demonstrates that business-as-usual strategies won't satisfy the growing need. If we are to unlock the promise and potential of data and all the technologies that depend on it, employers and educators will have to transform.

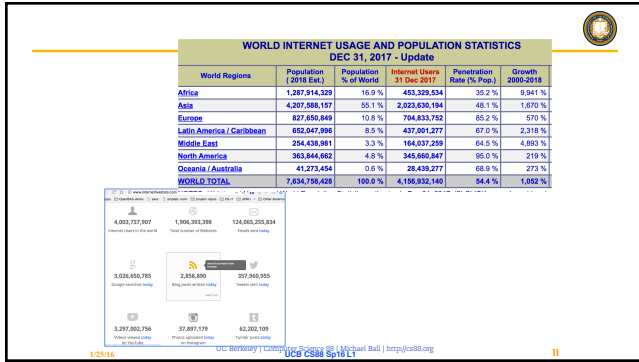
By 2021, 69% of employers expect candidates with DSA skills to get preference for jobs in their organizations. Only 23% of college and university leaders say their graduates will have those skills.

UC Berkeley | Computer Science 88 | Michael Hall | <http://cs88.org>

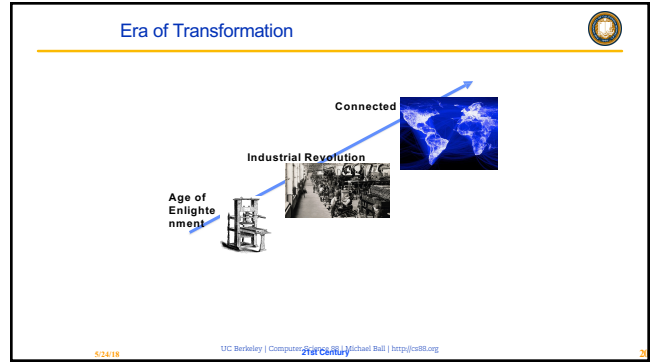
17



18



19



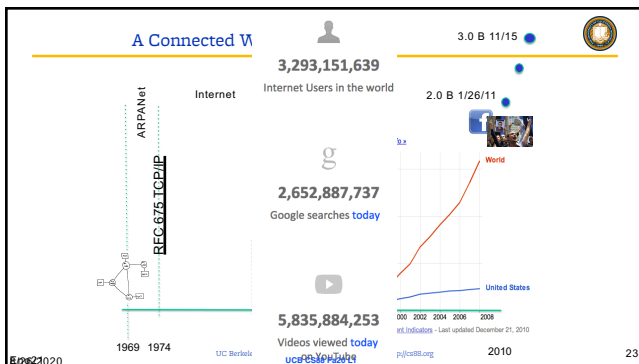
20

- ### A Connected World of Data
- The world's knowledge at our finger tips
  - *Digitalization* of life, industry and society
  - Intimately connected to billions of us, globally
  - Explosion of observational instruments
    - Genomics, Microscopy, Astronomical, ...
  - Vast Computational power to do analytics
  - Synthetic design exploration thru simulation
  - Machine reading of everything
  - Statistical machine learning algorithms to “discover” structure

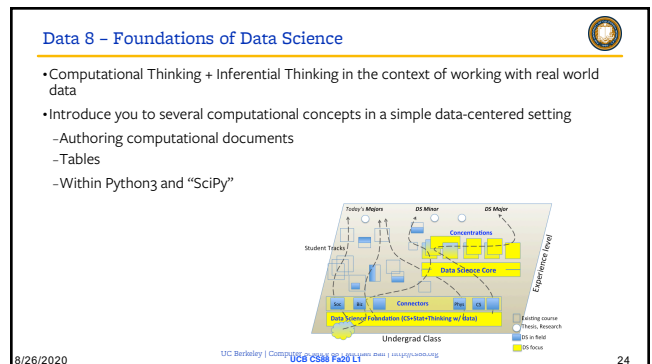
21

- ### What if I could ... ?
- 
- See the world's digital footprints?
  - Read everything that's ever been written?
  - Take it all in and dive down anywhere as far as the science can take me?
  - Learn the physical/chemical/biological /sociological/neurological... models from the data?
  - Explore billions of designs and pick the one I want?
  - ... ?

22



23



24

### CS88 – Computational Structures in Data Science



- Deeper understanding of the computing concepts introduced in c8
  - Hands-on experience => Foundational Concept
  - How would you create what you use in c8?
- 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

UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

25

### Course Culture

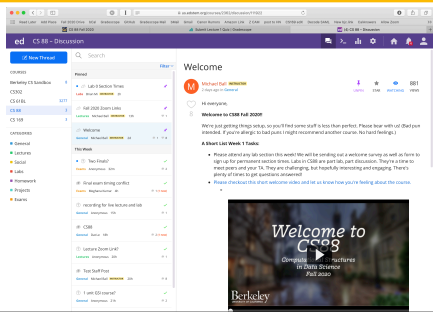
- Learning
- Community
- Respect
- Collaboration
- Peer Instruction

UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

26

### Ed For Class Discussion: Try it!



UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

27

### Where will we work?

- Your laptop
  - Using an editor and a terminal
- cs88.org
- datahub.berkeley.edu
  - Not as often, but an option
- us.edstem.org
  - Check out the “Workspaces”,
  - Can write and run (!!) python in you own posts!

UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

28

### Poll: Check In

- How has lab gone so far?
  - A. Labs have gone fantastic!
  - B. Labs have gone alright...
  - C. Labs have gone very well...
  - D. I haven't been to lab yet.

UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

29

### Poll: Check In

- Are you enrolled in Data 8?
  - A. I took it Spring 2020 or earlier
  - B. I took it Fall 2020 (“last semester”)
  - C. I’m taking it right now
  - D. I am trying to enroll in Data 8
  - E. I am not taking Data 8

UC Berkeley | Computer Science 88 | Michael Ball | <http://cs88.org>

30

### Poll: Check In



- Where are you right now?
- A. I made it to Berkeley!
- B. I'm somewhere in California
- C. I'm somewhere else in the US
- D. I'm somewhere internationally for the semester
- E. I've made it to Space where there is no COVID.

UC Berkeley | Computer Science 88 | Michael Hall | <http://cs88.org>

31

### Pro-student Grading Policies



- **EPA**
  - **Rewards good behavior**
  - **Effort**
    - » E.g., Office hours, doing every single lab, hw, reading Ed posts
  - **Participation**
    - » E.g., Raising hand in lec or discussion, asking questions
  - **Altruism**
    - » E.g., helping other students in lab, answering questions on Ed

UC Berkeley | Computer Science 88 | Michael Hall | <http://cs88.org>

32

### Your Tasks



- Lecture 1 Quiz On Gradescope
- Signup Genius form for lab times

**Welcome, and Good luck!**

UC Berkeley | Computer Science 88 | Michael Hall | <http://cs88.org>

41