

Computational Structures in Data Science

Ask Me Anything



Announcements

- Final Exam Next Week
- Seating assignments out this week (we'll be in 2-3 rooms)
- Remember to wrap up assignments

Final Exam Information

- Time: Tues 11:30 AM -2:30 PM
 - We'll be split in 3 rooms; emails to be sent later week
 - Confirmations on accommodations, alt times sent soon.
- Similar to Midterm!
 - Fall 2022 was done on Gradescope, but is very different. Topically good practice.
- Clobber Policy:
 - If you improve upon your midterm score, your midterm score goes up
- Everything in class is in scope, but focus on post-midterm topics.

Join CS/DS Ed Groups

<https://eecs.link/ed> -- EECS 101 Ed

<https://eecs.link/data-ed> DATA 001 Ed

Computational Structures in Data Science

My Background



- 2015: L&S CS Graduate
 - I TA'd CS10 for 3 years
 - Summer Instructor Summer 2015
- 2015-2016: MS CS (CS Education area)
 - Built some autograding stuff for CS10.
- 2016-2021 (ish)
 - Engineer at Gradescope
- 2019-Now
 - Lecturer at Berkeley

Don't have Expensive Hobbies...

- Photography (<https://mball.co>)
- Travel
- I do enjoy programming for fun!
 - <https://github.com/cycomachead/uc-salary-data>
 - [Interactive Notebook](#)
 - <https://github.com/agenel/heatmap-apple-activities>
 - I don't have an image of my walks, but this is a fun thing to do with data.
- Coffee, Cocktails:
 - Reading: James Hoffman, Garret Richard (Tropical Standard), Dave Arnold (Liquid Intelligence), Alex Day+David Kaplan (Cocktail Codex)

Computational Structures in Data Science

(Serious-ish) Ed Questions



Question

What made you want to study and eventually work in + teach computer science and data science? Was it what you always wanted to do, or did you ever consider any other fields of interest? Also, how did you realize you wanted to teach in general?

I got hooked into teaching!

Question

How did you end up working at Gradescope?

Started as a Berkeley project! Got to know the founders and used the early versions.

Turned down Microsoft to stay in CA and do a startup! It was fun, but don't do it for the money!!

Question

Question

How is Data 100 different from CS 61B? Should I take one before the other

Any advice on succeeding in future data science courses (like Data 100) and what resources/programs here would you recommend to gain experience in the field?

Computational Structures in Data Science

Fun Ed Questions



Question

Would you rather be a human with mango thoughts or a mango with human thoughts?

Question

Kiss, Marry, Kill: `map()`, `filter()`, `reduce()`

Cameras?

Do you still shoot with DSLR cameras and any more recent developing albums?

- Technically, not DSLRs anymore, but Canon R5, R3 and Leica Q2,
 - 1970's Minolta SR-T101 film camera, but don't use it as much as I want
- I use my iPhone a lot.

Reddit

- How much Reddit karma do you have? Why?

About ~6500, mostly comments.

Question

what are your top berkeley/bay area eats?

Dalian (N. Shattuck), Dough Zone will probably be up there

Cheaper: Artichoke Basil's Pizza (It's better at 2AM...)

Fancy: Ippuku (Center), Comal

Bay Area: Gott's Roadside, Mendo Farms, Sightglass Coffee,
Andytown Coffee (SF only)

Question

Are there going to be data c88c stickers? :'D

Computational Structures in Data Science

(Un)solicited Advice



Sam Bankman-Fried Is Found Guilty of 7 Counts of Fraud and Conspiracy

The case against the founder of the failed FTX exchange had come to symbolize the excesses of the volatile cryptocurrency industry.

[Share full article](#) [Share](#) [Bookmark](#) [466](#)



The FTX founder Sam Bankman-Fried, shown in February, was convicted on Thursday after a federal jury in Manhattan deliberated for more than four hours. Hiroko Masuike/The New York Times

Other Corporate Scandals... Bears doing Good [NYT]

Who's Who in the Elizabeth Holmes Trial



Erin Woo

📍 Reporting from San Jose, Calif.

Tyler Shultz and **Erika Cheung** are former Theranos employees and were whistleblowers. They worked at the start-up in 2013 and 2014. Shultz is a grandson of George Shultz, a former secretary of state who was on the Theranos board.

—
Jeff Kravitz/FilmMagic, via Getty Images



Computational Structures in Data Science

THANK YOU!



CONGRATULATIONS!!

Computational Structures in Data Science

Wrap Up

Berkeley
UNIVERSITY OF CALIFORNIA



Is Your Brain Full Yet?

- Data: values, literals, operations,
- Functions, Variables
- List, Tuples, Dictionaries
- Functions and lambdas
- Conditional Statements
- Iteration: list comp, for, while
- Higher Order Functions
 - Map, Filter, Reduce
- Recursion
- Abstract Data Types
- Mutation
- Class & Inheritance
 - Attributes, Methods, etc.
- Linked Lists
- Trees
- Iterators & Generators
- SQL / Declarative Programming
 - Queries, filtering, aggregations

Computational Structures in Data Science

Wrap Up

Berkeley
UNIVERSITY OF CALIFORNIA

Computational Structures in Data Science

Final Review



Approaching The Exam

- Skim the topics (~1 min)
- Handle the "easy"(est) questions first
- Read the whole question first!
- Read the text
- Read the doctests!
- What techniques might be applicable?
 - Pattern matching is OK
- Draft a solution on scratch paper!
- Write yourself notes

THANK YOU!

COME JOIN COURSE STAFF!



Isabelle Ng she/her/hers

@ DSP isabelle.ng@berkeley.edu

Hi there! I am a senior CS/DS/Music Major from San Jose, CA and this is my 5th semester being a TA for c88c/61a. Excited to meet you all and have an awesome semester.



Alicia Wang she/her/hers

@ awwang629@berkeley.edu

Hi! I am a junior studying Data Science and Cognitive Science. I love playing badminton and traveling! Excited to meet everyone!



Grace Xie she/her/hers

@ gracexie@berkeley.edu

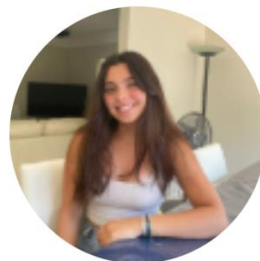
Hello! My name is Grace. I'm a senior majoring in MCB and Data Science :0 I love reading sci-fi and baking in my free time.



Reema Rafifar she/her/hers

reemarafifar@berkeley.edu

Hi everyone! I'm Reema, a junior majoring in Neuroscience 🧠 & Data Science 📖 Apart from studying the brain and staring at code, I love cafe hopping and watching movies (on film!!) I can't wait to get through C88C with you!



Mira Wagner she/her/hers

@ DSP mirawagner@berkeley.edu

Hi! I am a junior majoring in data science and linguistics. I love reading, especially mysteries, swimming and baking! Excited for this semester :)



Dhruv Syngol he/him/his

@ dhruvsyngol@berkeley.edu

Hey everyone, I'm a junior studying Data Science and Economics, originally from the Chicago Suburbs! I love to play pickleball, explore cafes and restaurants, and go on hikes! Super excited for this semester!



Grace Baek she/her/hers

gracebaek@berkeley.edu

Hi! I'm Grace, a senior majoring in Computer Science and Economics and this is my 3rd semester teaching C88C. In my free time, I like to bake and try new cafes :) Super excited to meet everyone!



Cynthia Shao she/her/hers

cynthia_shao@berkeley.edu

Hey! I'm Cynthia, a sophomore in data science. You can find me dancing in The 510, jogging around campus, and finding/making/eating food at 1AM. See you in session :)



Maryam Akelyan she/her/hers

maryam.akelyan@berkeley.edu

Hi! I'm Maryam, a junior majoring in Data Science and MCB! In my free time I like to watch sitcoms and sci-fi shows. Super excited to meet you all! :)



Thompson Zhou he/him/his

chuanjunzhou@berkeley.edu

Hi! I'm Thompson, a third year double majoring in Applied Math and Data Science. During my free time, I like to read, play video games, and follow sports.



Esha Bansiya she/her/hers

esha_bansiya@berkeley.edu

Hi! I'm Esha, a sophomore studying data science and city planning! I like to read, go on hikes, and explore Berkeley and SF!



Orazaly Kabdrakhmanov he/him/his

kabdrakhmanov@berkeley.edu

Hi there! I'm Orazaly and I am a senior majoring in Data Science. I like playing tennis and video games(CS2). Hope you enjoy the class!

Computational Structures in Data Science

The One Big Thing...

Berkeley
UNIVERSITY OF CALIFORNIA

Abstraction

- Detail removal

“The act of leaving out of consideration one or more properties of a complex object so as to attend to others.”
- Generalization

“The process of formulating general concepts by abstracting common properties of instances”
- Technical terms: Compression, Quantization, Clustering, Unsupervised Learning



Henri Matisse “Naked Blue IV”

Keep on Programming

LOTS of Courses to Follow Up! [[Courses](#)]

- Explore all the DS connectors
 - Data 88E (Econ)
 - Stat 88 / DATA C88S
- Data 100: Principles of DS
- Data 101: Data Engineering
- Data 102: Data, Inference, and Decisions
- Data 104: Human Context and Ethics
- Data 140: Probability
- INFO – Course Schedule Varies!
 - INFO C103 in Spring: "History of Information"
 - Some Database courses, web development, etc.

CS Courses

- CS61B: (conventional) data structures, statically typed production languages.
- CS61C: computing architecture and hardware as programmers see it.
- CS70: Discrete Math and Probability Theory.
- CS170, CS171, CS172, CS174: “Theory” —analysis and construction of algorithms, cryptography, computability, complexity, combinatorics, use of probabilistic algorithms and analysis.
- CS161: Security
- CS162: Operating systems.
- CS164: Implementation of programming languages
- CS168: Introduction to the Internet
- CS160, CS169: User interfaces, software engineering
- CS176: Computational Biology

CS Courses Part 2

- * CS182, CS188, CS189: Neural networks, Artificial intelligence, Machine Learning
- CS184: Graphics
- CS186: Databases
- CS191: Quantum Computing
- CS195: Social Implications of Computing
- EECS 16A, 16B: Designing Information Systems and Devices
- EECS 126: Probability and Random Processes
- EECS149: Embedded Systems
- EECS 151: Digital Design
- CS194: Special topics. (e.g.) computational photography and image manipulation, cryptography, cyberwar.
- Plus graduate courses on these subjects and more.

EE Courses Are There Too

- EE105: Microelectronic Devices and Circuits.
- EE106: Robotics
- EE118, EE134: Optical Engineering, Photovoltaic Devices.
- EE120: Signals and Systems.
- EE123: Digital Signal Processing.
- EE126: Probability and Random Processes.
- EE130: Integrated Circuit Devices.
- EE137A: Power Circuits.
- EE140: Linear Integrated Circuits (analog circuits, amplifiers).
- EE142: Integrated Circuits for Communication.
- EE143: Microfabrication Technology.
- EE147: Micromechanical Systems (MEMS).
- EE192: Mechatronic Design.

Personal Course Recommendations

- INFO C103
- Computational Photography – especially if you like photoshop
- HCE Courses
- History, Poli Sci, Psychology
- LS110 Brilliance of Berkeley (Spring)
 - <https://curricularconnections.berkeley.edu/ls110/>
- An 'experiment' of sorts, but seems cool

What's next? (Otherwise)

- Programming contests
- Hackathons
- More paradigms and languages: the web
- The open-source world: Go out and build something!
- Above all: Have fun!

And there's lots more to Python!

What can you do with Python?

- Almost anything!
- Webapp backends
- Web scraping
- Natural Language Processing
- Data analysis
- Machine Learning
- Scientific computing
- Games

What can you do with Python?

- Almost anything! Thanks to libraries!
- Don't try to memorize libraries, learn how to learn how to use them!
- Webapp backends (Flask, Django)
- Web scraping (BeautifulSoup)
- Natural Language Processing (NLTK)
- Data analysis (Numpy, Pandas, Matplotlib)
- Machine Learning (FastAi, PyTorch, Keras)
- Scientific computing (SciPy)
- Games (Pygame)

Computational Structures in Data Science

THANK YOU!

(Again!)

Berkeley
UNIVERSITY OF CALIFORNIA