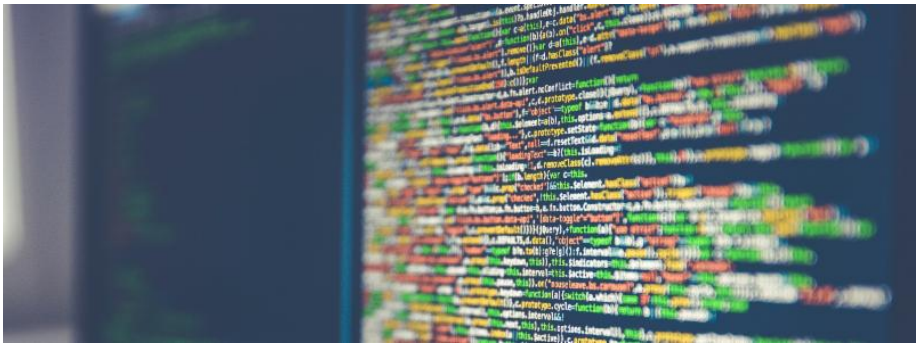
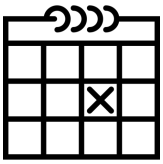
Volume 1 | Issue 6 - April 23, 2019



Data Science Newsletter

Greetings from Data Science Advising. Welcome to the sixth issue of our Newsletter! If you missed previous issues, make sure to check them out here.

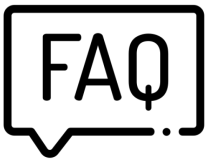
Major Updates



Drop-In Advising Updates

There will be **no Major Advisor drop-in advising April 29 - May 1**. Students should make appropriate plans to attend drop-ins the week before, or attend Peer Advising drop-ins.

Data Science Minor FAQ Updated



We have updated the Frequently Asked Questions page on our website with new information about the Data Science minor. At this time, we still cannot say anything about the specific technical elective courses for the minor.

All of the information we have about the minor is currently available on the FAQ page. If you have questions about choosing classes within the guidelines already provided, please contact the Data Science Peer Advisors at ds-peer-advising@berkeley.edu or during their drop-in hours at Moffitt Library, 3rd floor.

The Data Science Major Advisors are not able to see students with questions about the minor during the Spring 2019 semester.

Fall 2019: STAT 102 ("Data 102")!

We are delighted to announce the new course **Data, Inference, and Decisions (Stat 102)**. It will open for enrollment during Phase I. In Fall 2019, the instructors will be **Michael Jordan** (Stat and EECS) and **Fernando Perez** (Stat).

A large fraction of seats in the Fall 2019 pilot offering will be reserved for **senior Data Science majors**. The remaining seats will be reserved for **senior Statistics majors**.

For more information or to ask questions, please refer to the original post on the STAT 001 Piazza page: https://piazza.com/class/jua820aaxcq1o6?cid=9 or view this post on the Data 001 Piazza page: https://piazza.com/class/j7s01y165odq5?cid=1480

Fall 2019: CS 189, CS 61A, CS 61C, CS C8 with Changing Time/Location

A message from the EECS Department:

Unfortunately there have been some issues with getting rooms for some of our larger CompSci classes. We have been, and continue to work diligently with Campus Scheduling to resolve this issue as soon as possible. Be aware that the meeting day/time and room location for these rooms is **SUBJECT TO CHANGE!** (In other words, it VERY likely will change.)

What this means to you:

If you plan to enroll into any of the following classes, be aware that the class day/time/location information is **NOT ACCURATE.**

- CS C8 (Data 8)
- CS 61A
- CS 61C
- CS 189

We have **CLOSED** enrollment for these classes to prevent students from using phase 1 units in a class that may not fit their schedule.

How you will know when the information is correct:

- Watch for an **EECS 101 Piazza** post(s) saying the problem has been fixed.
- The class will show **OPEN** for enrollment.

Thank you for your patience as we work to resolve the issue.

*This message was posted to the EECS 101 Piazza on 4/12/19. **Subscribe here**, if needed. Questions should be directed to cs-enrollments@eecs.berkeley.edu or you can add questions to the original post. Do not reply to this email with questions.*

Student Resources

Data Science Peer Advising

If you aren't able to book an appointment with a Data Science Major Advisor or if you can't attend their drop-in hours, come drop by the Peer Advising drop-in hours! Or just drop by regardless :).

Peer Advisors can help you with:

- Information about the major or minor
- How to declare the major
- Choosing classes (i.e. course workload)
- Creating a 2-Year or 4-Year Program Plan
- Data Science affiliated student organizations
- Division of Data Sciences Student Opportunities
- Internships
- Research opportunities
- Choosing a domain emphasis (see Peer Advisor biographies for their select domain emphases)

Data Peer Consulting

DATA PEER CONSULTING

It's nearing the end of the semester! Do you need help with a final project or an honors thesis

that requires visualizing data, or programming in R, SQL, Python? Come drop by for Data Peer Consulting! Peer Consultants hold drop-in hours on the 3rd floor of Moffitt, see their schedule and short bios on their website.

Student Opportunities

Healthy Campus Food and Beverages Innovation Challenge

Healthy Campus Food and Beverages Innovation Challenge

The Coalition for Healthy Campus Food and Beverages will hold an hackathon/designathon to think out the next ten years of beverage service at UC Berkeley. This case competition challenges you to think at the intersection of health and business—considering human and planetary health as well as financial sustainability and corporate responsibility. UC Berkeley is near the end of an exclusive pouring rights contract with PepsiCo, and the results of this designathon will help shape the future of beverage options on Berkeley's campus.

Location: Haas Innovation Lab | **Register: https://bit.ly/2T2hoFv**

Friday, April 26, 4 – 6:30pm: Challenge Overview
Saturday, April 27, 8 – 4:30pm: Innovation Challenge Main Event

Follow the Facebook event for updates on prompt, prizes, and innovation experts! https://www.facebook.com/events/566255219871620/

Sponsors: Berkeley Food Institute, Berkeley Haas Center for Responsible Business, Food@Haas, Basic Needs Security Committee, FoodInno, FEED (Food, Equity, Entrepreneurship, and Development), Net Impact Berkeley, Berkeley Student Collective, and BFI Undergraduate Council.

Want to apply your food systems knowledge to addressing a challenge right here on the Berkeley campus? Participate in the **April 26-27 Healthy Campus Food and Beverages Innovation Challenge**. This event will think out the next ten years of beverage service at UC Berkeley by challenging students to think at the intersection of health and business—considering human and planetary health as well as financial sustainability and corporate responsibility. The results of this **hackathon/designathon** will **help shape the future of beverage options on Berkeley's campus.**

https://food.berkeley.edu/foodscape/healthy-campus-food-and-beverages-innovation-challenge/

RippleMatch - Get Matched with your Dream Job

RippleMatch is a tech software startup that matches ambitious and diverse students with potential first-round interviews based on the student's personality, work experience, and work preferences. Data Science students might find this tool particularly interesting given they work with employers such as Dell, IBM, BlackRock, Cardinal Health, and Yelp to name a few.

Students can start getting matched after taking **5 minutes** to upload their resume (**no cover letter needed**) and set their preferences here: **tinyurl.com/ripplejob**

Sign up is zero cost and takes 5 minutes, so all-in-all, RippleMatch is a risk-free way for students to increase their chances of landing a dream job or internship.

If you have any questions, please contact ranizhu@berkeley.edu.

Healthcare Data Science Summer Internship Opportunity at UCSF

University of California San Francisco

The laboratory of **Dr. John Fahy** (Pulmonary Division) at the University of California San Francisco is looking for a summer intern to take part in a project to develop image-processing algorithms for analyzing chest CT scans of patients with lung disease. The ultimate goal of the project is to **improve diagnosis and tailor treatment for the patients suffering from lung diseases**. The intern will work closely with a resident with machine learning expertise and will interface with collaborating clinicians including radiologists and pulmonologists.

We are looking for an enthusiastic individual with interest in applying his/her computer and data science skills to an important clinical problem. The position is geared towards

students with quantitative backgrounds, such as computer science, engineering or physics (and we can teach the clinical background). Master's level students are preferred; consideration to qualified upper-level undergraduates will be given as well.

Principal duties and responsibilities:

- Development and implementation of machine and deep learning algorithms in computer vision.
- Image processing and data curation of large radiology data sets
- Work with collaborating clinicians including radiologists and pulmonologists to implement algorithms in clinical settings

Qualifications:

- Coursework or research experience in data analysis, applied physical science, signal processing, modeling, machine learning, and deep learning.
- Experience in data science, specifically deep learning, is desirable but not required.
- The applicant should be familiar with relevant computer programming and should possess:

- Fluency in a scientific scripting language such as Python or Matlab
- Know general principles of computer programming
- Have the ability to learn common tools for data management and analysis including machine learning (i.e. Tensorflow, Torch, Caffé), GPU toolsets (CUDA), database software (SQL, Mongo, or variants), and general scientific computing
- Scientific visualization for communication with interdisciplinary teams

Location: UCSF Parnassus Heights campus

Interested? Would like more information?
Contact Irina Gitlin (irina.gitlin@ucsf.edu, 415-502-4328)

Digital Humanities Summer Minor

The Digital Humanities Minor can be completed in one or two summers. "Digital Humanities" means using computational methods to answer questions about humanity, and communicate your results in a rich visual format. Studies show that employers value humanistic skills as well as technical skills. The DH minor is your chance to bring the two together and use your own talents in new realms of discovery to add to your career portfolio.

The minor is five courses that can be completed this summer, or over this summer and next summer. You will receive $2000 upon completion of the minor.

Summer Courses

Ling 188: Linguistic Data (Session D)

How can we use **data science methods** to understand **human language**? Linguistics involves the study of language sounds, words, meanings, context, structure and change. This course provides students with the computational skills necessary to analyze linguistic data from these areas. We will draw on data from languages around the world and use computer programming and data visualization techniques from **Foundations of Data Science**.

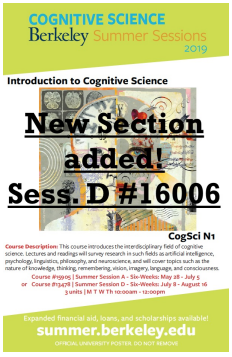
https://classes.berkeley.edu/content/2019-summer-linguis-188-001-lec-001

PH107: Violence, Social Justice, and Public Health (Session A)

This two unit course will examine **violence through the lens of the college campus**, paying particular attention to the types of violence more commonly seen on, or associated with, collegiate life. In particular, we will examine such topics as sexual violence/harassment, alcohol-related harm, suicide, and bias and identity-based violence. The class is **discussion-based, interactive**, and taught by people working in the field of violence prevention. It meets **Tuesday, Wednesday, and Thursday from 10am to noon**. Join us!

https://classes.berkeley.edu/content/2019-summer-pbhlth-107-001-lec-001

COGSCI N1: Introduction to Cognitive Science (Session D)



Additional section of CogSci 1 available during Session D, Summer 19! Because our first section filled up so quickly, we have added an additional section of CogSci 1 in summer 19. Course number 16006. Sign up today!!

https://classes.berkeley.edu/content/2019-summer-cogsci-n1-003-lec-003

EPS10: Earth's Greatest Volcanic Eruptions (Session A)

Consider taking EPS 10 Earth's Greatest Volcanic Eruptions with volcanologist Prof. Stephen Self: http://www.stephenself.com/index.html!

A science-based course on the most significant eruptions Earth has produced. Most eruptions discussed will be from within historic time and will involve information from geology (volcanology), geography, archaeology, history, art, and paleoenvironmental records such as tree-rings and ice-cores.

https://classes.berkeley.edu/search/class/eps10?retain-filters=1&f%5B0%5D=im_field_term_name%3A839

Upcoming Events

Women Leaders in Tech: Shattering Ceilings and Scaling Cliffs

A panel of women in senior leadership roles in the technology sector discuss their experiences in an industry that continues to be largely male-dominated.

What are the characteristics of a successful woman leader? Why is it essential to include women at the most critical decision-making levels? Why should men care?

Join us for this hour-long discussion followed by a networking reception.

Register and learn more about each panelist HERE.

Co-sponsored by:

CITRIS AND THE BANATAO INSTITUTE
citris-uc.org

WOMEN IN TECH INITIATIVE
UNIVERSITY OF CALIFORNIA

Women Leaders in Tech: Shattering Ceilings and Scaling Cliffs

Please join us this Wednesday for *Women Leaders in Tech: Shattering Ceilings and Scaling Cliffs.* A panel of women in senior leadership roles in the technology sector will discuss their experiences in an industry that continues to be largely male-dominated. The panel will be followed by a networking reception.

Wednesday, April 24, 2019
11 am – 12 pm, with a reception to follow
Banatao Auditorium, Sutardja Dai Hall, UC Berkeley

This event will be recorded.

Panelists

- Moderator: **Pratima Rao Gluckman,** Engineering leader, VMware; Author of *Nevertheless She Persisted: True Stories of Women Leaders in Tech*
- **Amanda Blevins,** Senior Director and Chief Technologist, VMware
- **Lily Chang,** Vice President, Strategic Transformation, VMware
- **Pam Kostka,** CEO of a stealth-mode startup
- **Holly Liu,** Co-Founder, Kabam, Visiting Partner, Y Combinator
- **Telle Whitney,** Former President and CEO, Anita Borg Institute for Women in Computing

---

Workshop on Critical Timescales of Hydrologic Transport Student Opportunity

BIDS
BERKELEY INSTITUTE FOR DATA SCIENCE

We are seeking a select number of participants for a workshop at that will bring together data scientists and watershed hydrologists at the **Berkeley Institute for Data Science, May 22-24, 2019**. This workshop will be part training, part hackathon, and will be focused on applying multiple time-series analysis techniques, hydrologic modeling, and isotope tracer approaches to understand fundamental controls on the timescales over which water moves through watersheds to generate streamflow. All workshop and travel costs will be fully funded for the selected participants.

Who: If you are a student or postdoc who has strong data science and computing skills and looking to learn more about how to apply those skills to real-world hydrologic problems, this workshop is for you. Alternatively, if you are trained as a hydrologist but would like to learn new techniques for working with large hydrologic datasets and performing causal inference, this workshop is also for you!

What: We are trying to compile a database relevant to understanding how watersheds respond to precipitation and climatic factors, which will enable us to produce better forecasts of streamflow. In this workshop, we will work with data to try to understand what is often one of the greatest uncertainties in forecasting: the timescales over which watersheds respond to perturbation. We will work with data from two watersheds (HJ Andrews, Oregon, and East River, Colorado) that are currently a focus for the development of predictive models and have extensive data records from sensor networks and isotope studies. This workshop will be part training, part networking, and part hackathon, with a presentation of findings and discoveries at the end.

When: May 22-24

Where: **Berkeley Institute for Data Science**, University of California-Berkeley

Why: You are interested in gaining new skills or putting your skills toward solution of an important research problem in hydrology. For committed participants, there will be opportunities to participate in journal papers that will emerge from the work done here. As an immediate short-term incentive, there will be monetary prizes associated with the hackathon.

How: Fill out this **simple application** by April 30, 2019.

Workshop organizers: Laurel Larsen, Edom Moges (edom.moges@berkeley.edu), Zexuan Xu (zexuanxu@lbl.gov), Dino Bellugi (dino@berkeley.edu)