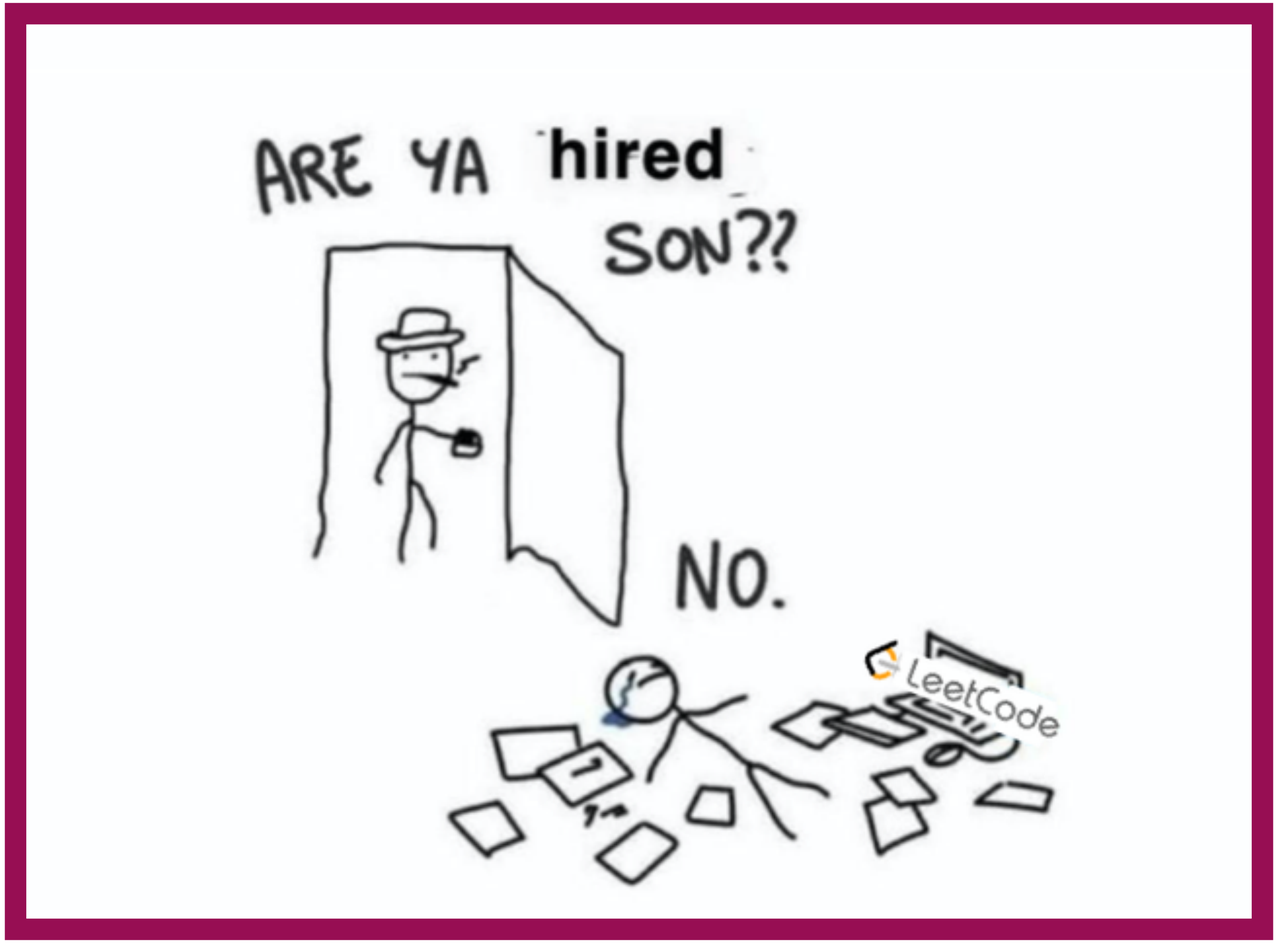
Advancing Careers: Grad School & Jobs

Week 3

DS198-003: Data Discovery Scholars Seminar
UC Berkeley - Computation, Data Science, and Society

Spring 2022

Meme of the Week


ARE YA hired
SON??
NO.
LeetCode

Announcements

Announcements

- Please complete the following
  - Mid-semester survey: tinyurl.com/ds198-mid
    - Thanks to those who've already completed it
  - Career Accelerator Program form: https://forms.gle/C1pG8YsvfGpvgPhi9
    - Complete if you are interested. Please do so by 6:30PM today.
  - Mid-semester presentations next week

Tech Recruitment

Software Engineering

Product Manager

Data Scientist

What job do I apply to? 😱

Data Analyst

Business Analyst

ML Engineering

Data Engineering

Tech Jobs

Business Analyst

- Improves business processes between the business and IT
- MS Office, SQL, Data Visualization, business intelligence, understanding, modeling

Software Engineer

- Building Products, Real Time API, Servers, Cloud Computing, Automation, Mobile/ Web Applications
- Java, Ruby,Python, Swift, JS, Go, C, C++, C#,

Product Manager

- The CEO of a product
- Asana, JIRA, Tableau, General Data Analysis,
- Product Analysis, project management, interpersonal skills, elementary technical skills to step in

Tech Jobs

ML Engineer

- Developing universal algorithms to extract meaningful information from large amounts of data
- C++, Python, Java or R
- Deep Learning, Heavy Probability, Statistics, Linear Algebra

Data Analyst

- Collects, process, and performs statistical data analyses
- Excel, SQL, Tableau, Power BI
- Math, Stats, Machine Learning

Data Engineer

- Develops, constructs, tests and maintains architecture
- SQL, Hive, Pig, R, Hadoop, Spark, Matlab, Python, Java, Ruby, C++
- Database systems, modeling, APIs

Data Scientist

- Cleans and organizes big data from the POV of expertise
- R, SAS, Python, SQL, Data Science Toolkit
- Predictive modeling, story-telling
- Math, Stats, Machine Learning

Check the job description!

Application Styles

Spray & Pray

- Same resume
- Boring to no cover letter
- No networking
- No company research
- **High** number of applications

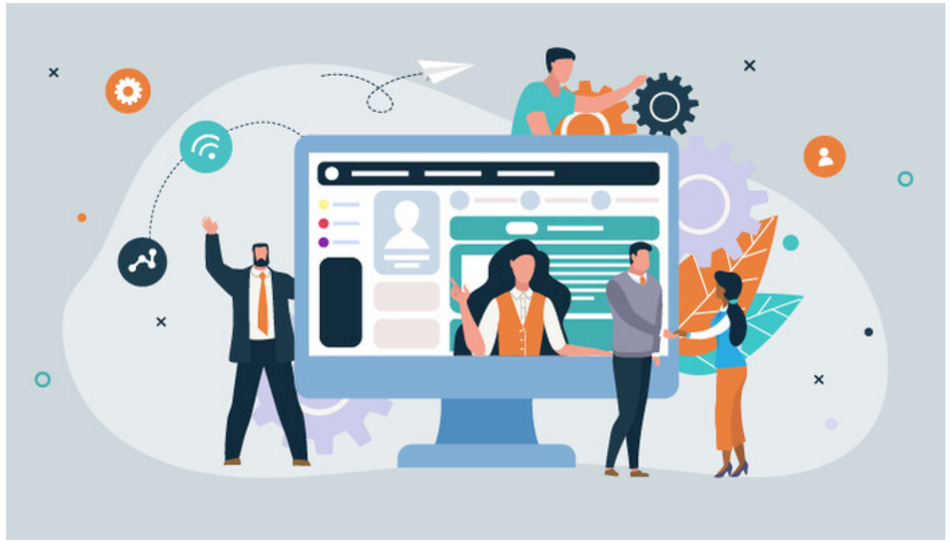
Tailored Approach

- Targeted resume
- Individual cover letter
- Crazy networking
- **Low** number of applications

Hybrid

- Resume for different jobs
- Cover letter template
- Networking when possible
- **High** number of applicants

Networking



- **Linkedin**
- **Phone Chats/Zoom Calls**
  - Acquaintances, alumni, co-workers, bosses
- Question to ask
  - Do you have advice for someone who is in this field?
  - What do you do at your company?
- You are not bothering that person! Networking is beneficial for both.

Resources: Applying to Jobs

- Tech Jobs
  - https://skillcrush.com/blog/41-tech-job-titles/
  - https://towardsdatascience.com/data-scientist-vs-machine-learning-engineer-skills-heres-the-difference-93eb2f4f6f98
- Resume
  - https://career.berkeley.edu/Tools/Resume
  - https://www.themuse.com/advice/how-to-make-sure-your-resume-is-as-current-as-your-skills
- Coffee Chat/Networking
  - https://www.indeed.com/career-advice/finding-a-job/how-to-network-for-a-job

The Interview

- **Phase 1: Coding/Take Home Challenge**
  - Goal: Assessing general programming knowledge as a way to filter out unfit candidates
  - *Time:* 30 minutes to 3 hours
- **Phase 2:** Technical Phone Screen
  - Goal: Screening a combination of technical, interpersonal skills and history
  - *Time:* 30 minutes to 1 hour
- **Phase 3:** Onsite Interview & Whiteboarding
  - Goal: Looking for coding skills, problem-solving ability, creativity, culture fit, analytical thinking, handling feedback, understanding the bigger picture
  - *Time:* 1 hour to a day

The Interview: Differences

- Though the previous slide is a general overview, there are wide variations in interview setups
- **Startups:** Tend to have shorter interviews that exist of behavioral, SQL coding tasks and resume questions.
- **Data Science:** Varies widely. Take-home exercises, SWE interview, statistics, behavioral/resume questions.
- **Machine Learning:** Combination of Data Science & SWE interview with heavy emphasis on job-related skills (think: CV, NLP, RL)
- **Data Engineering:** Combination of SWE & Systems Design questions.
- **Business Analyst:** Light SWE interview and focus on business topics and drawing conclusions.

The Interview: Prepare

- CS61B & DATA100 are your best friend
    - Data Structures, Big O, Trees, Sorting, Heaps, Hashing
  - Domain Specific (e.g. DL for ML Engineering)
  - System Design
  - Condensed forms can be found online (Cracking the Coding)
- Practice your riddles: LeetCode, HackerRank
  - Remember, there is only a finite number of problems they can give you
- Review your projects
  - Do you remember to explain all your decisions?

Tips

- Think out loud while solving problems
  - The interviewer cannot pick your brain
  - Explicitly mention each step and the motivation behind it
- **Know your audience**
  - ❌ Bragging about your AUC score in front of a recruiter
  - ❌ Covering up a big flaw in your project in front of an engineer
- Don’t lie, don’t pretend, focus on what you do know
- Do your research

The Interview: Resources

- https://github.com/yangshun/tech-interview-handbook
- https://github.com/datastacktv/data-engineer-roadmap
- https://github.com/khangich/machine-learning-interview
- https://www.crackingthecodinginterview.com
- https://www.kdnuggets.com/2021/08/common-data-science-interview-questions-answers.html

Graduate School

Do I need a Graduate Degree?

- Ask yourself the following **questions**:
  - What do you want to get out of the program?
  - Is it worth the money and labor?
  - Do I have the right background and preparations?

  For instance, you want to:
    - become a professor
    - get a higher job prospect
    - are hunting for specific niche jobs
    - become more employable
    - get more depth

Master's vs. PhD

Master's	Doctoral
• 1-2 years • Can do research and teaching, but mostly course based • Not always funded • Breadth • Diverse population • Normally enough if you don't want to go into academia	• 4-5 years • Teaching & Research-based • Requires pre-lims and quals • Funded • Depth & Expertise • Normally needed for teaching or certain specific jobs

Applying to schools

- Factors to take into considerations
  - **Research Field**
  - **Course**
  - **Program Length**
  - **Cost**
  - **Prerequisites & Competitiveness**

The Essay: SOP vs. PS

- The **Statement of Purpose** is mostly an academic document about:
  - What you did, you want to do during graduate school and afterwards **research/course**-wise
- The **Personal Statement** is more like the college essay. It asks about your personal story and how that ties in with your ambition/passion.
- **Strategy: Applying to multiple programs**
  - Start with the school with the most specific prompt. Then match and edit it accordingly for other schools.

Other Tips

- **Letter of Recommendation**
  - Ask a letter from a professor, teacher or manager that knows you well and has worked with you rather than a distant letter from a famous professor you have seen once.
    - A good letter from an unknown person is much better than a bad letter from a famous person
- **GRE**
  - There are free resources to prepare for the GRE online (Magoosh, ETS)
  - Many schools have gotten rid of the GRE
- **Extra Help**
  - Many schools provide resources for underrepresented students to help with their applications
  - Ask your friends to go over your essays!

Resources

- **Choosing a program**
  - https://www.gograd.org/resources/choosing-graduate-program/
  - https://www.cs.cmu.edu/~pavlo/blog/2015/10/how-to-write-a-bad-statement-for-a-computer-science-phd-admissions-application.html
- **GRE**
  - www.reddit.com%2Fr%2FGradSchool%2Fcomments%2F1kvk0f%2Fbest_way_to_prep_for_gre%2F&usg=AOvVaw2tctrj6YeF-alL6n72qsN_
- **Essay**
  - https://www.cs.cmu.edu/~pavlo/blog/2015/10/how-to-write-a-bad-statement-for-a-computer-science-phd-admissions-application.html
  - https://career.berkeley.edu/Grad/GradStatement

Resources (cont’ed)

- **Programs for URM**
  - https://github.com/gwisk/gradguide
  - https://github.com/chinasaokolo/csGraduateApps

Data Science Interview

Data Science Interview

Form pairs. One person will be the interviewer, the other person the interviewee. The interviewee has 5 minutes to elaborate their thought process on the first question/case. After 5 minutes, switch roles and proceed to the second question/ case.

Case/Questions

- (1) You work as a data scientist for a ride-sharing company. An executive asks how you would evaluate whether a 50% rider discount promotion is a good or bad idea. How would you implement it? What metrics would you track?
- (2) We're working on a new feature for LinkedIn chat, and we want to implement a green dot to show an “active user”. Given engineering constraints, we can't AB test it before release. How would you analyze the effectiveness of this new feature?

Check-In

Check-In

- No class next week! Instead, we will be conducting 15-20 min one-on-one check-ins via Zoom. This meeting serves to:
  - get to know you better
  - discuss how you are feeling overall
  - talk about how the project is going
  - plan ahead your project
  - answer any questions you have related to your project

Work Session & Office Hours