

PA 446

Coding for Civic Data Applications

Will be starting at 6:05pm

Class #9

Logistics

Course Logistics

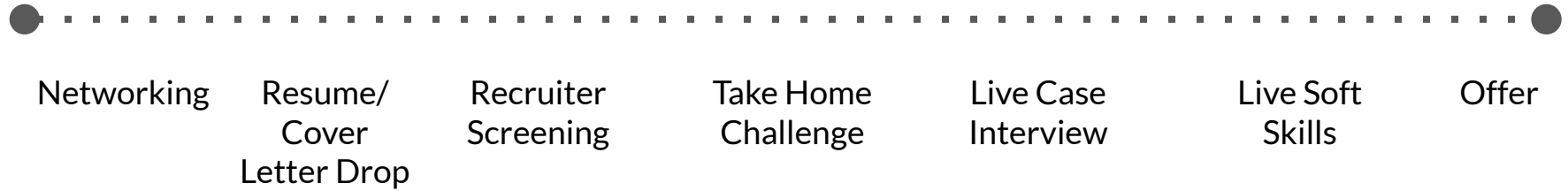
- “Midterm” next week
 - Released during class
 - I will be available for questions and clarifications
- Class next week: Shen still finalizing details
 - Recruiter from Civis
 - Mock live case interviews

Focus This Week

1. Wrapping up Data Science Hiring Process
2. HW4

Data Science Hiring Process Wrap Up

Job Search



Live Case Interview

Live Case Interview

What Is It?

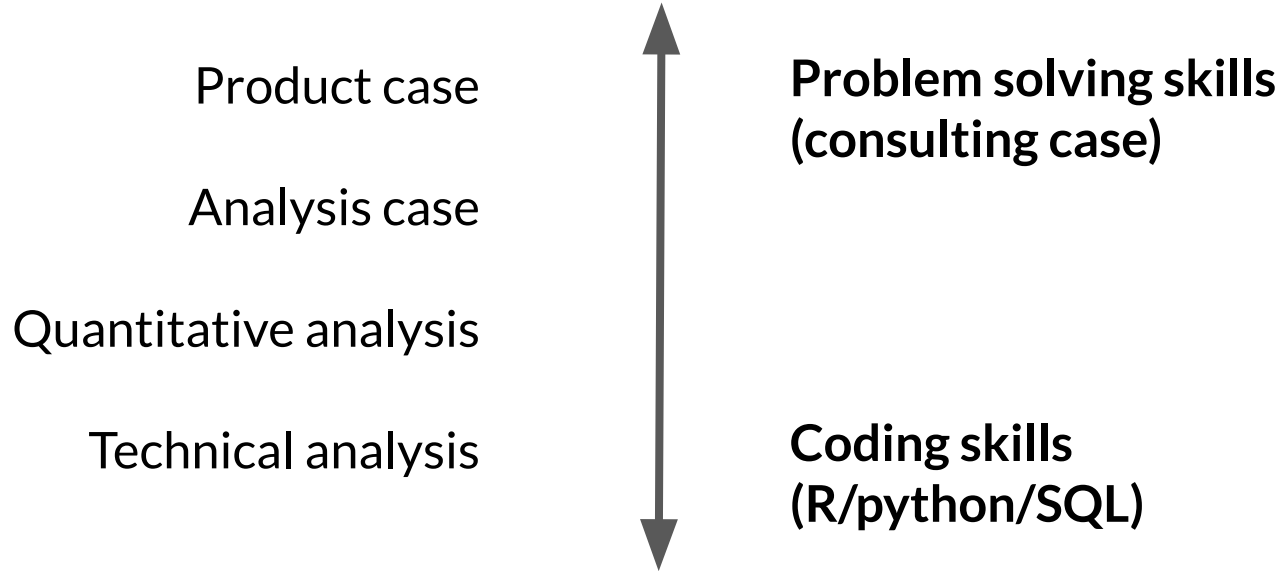
30-45 minute live interview with another data scientist(s)

Solve a problem in front of the interviewer(s)

Types of questions can significantly range

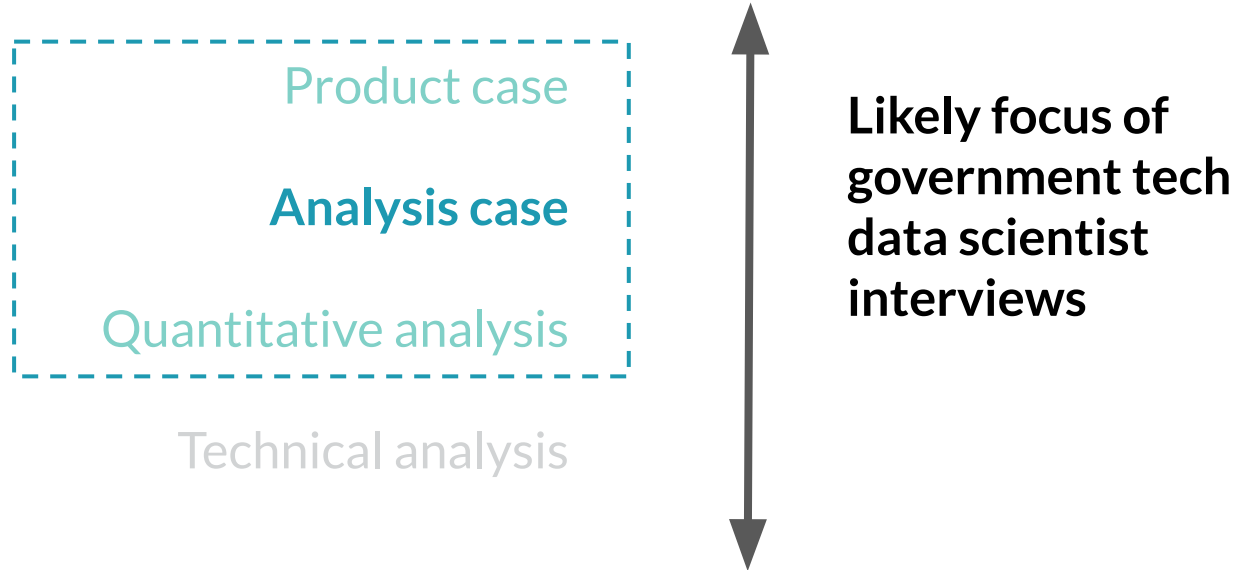
Live Case Interview

Interview Types



Live Case Interview

Interview Types



Live Case Interview

Interview Types

Interview Type	Examples
Product case	What KPIs would you use to measure the success of the city's 311 service?
Analysis case	Given a table of 311 calls, how would you create a model to detect spam?
Quantitative analysis	Explain hypothesis testing Explain the bias and variance tradeoff
Technical analysis	Given a list, search for consecutive numbers (n) whose sum is equal to a specific number (x).

Live Case Interview

What Is It for?

Assess the knowledge you already have in your head, without Google/Stackoverflow

Assess how you think and how you solve problems

Product/Analysis Case

Ambiguity

Most problems you will encounter in reality require a team of data scientists + weeks to solve

You usually have 30-45 minutes

Product/Analysis Case

What They Are Testing For

Structure: have a systematic approach to the problem.
Answers that lack structure can be viewed as red flags.

Comprehensiveness: your solution covers **all important** aspects of that problem and you don't get caught up with the lesser issues

Feasibility: approach is practical enough to be implemented.
Creativity is the cherry on top, but not at the expense of feasibility

Product/Analysis Case

Framework

Clarifying the Question: what are the product goals? What are the goals around user growth/retention?

Make Assumptions and Hypothesis: what you'll explore in your analysis, what you won't, and why

Design Approach: what analysis will you do and why

Package Findings: tie everything back to business “so whats?”
Make your work client ready: simplify and prioritize

Quantitative Case

Concepts to Know

<https://www.mastersindatascience.org/learning/statistics-data-science/>

Behavioral Interviews

Behavioral Interviews

What Is It?

30 minute live interview with a data science manager,
likely the hiring manager

Behavioral Interviews

What is it for?

Assessing soft skills

- How do you handle roadblocks and criticism
- How do you work in a team
- Can you handle clients/stakeholders

Behavioral Interviews

Intro

Tell me about yourself

- You will always get asked this, so prepare + practice
- Keep it short
 - Shoot for 1.5-2 minutes - no more than 3
- Majority is recycled; minor customization is ideal

Behavioral Interviews

Intro

Framework: the odyssey

- How you got to where you are today, professionally
 - Highlight things you have done along the way that makes you a good fit for the role
- Time permitting: why company X

Behavioral Interviews

How you got to where you are today

1st 3 years of career, I was an architect

Loved the idea of building something from nothing. But over time, I wanted to have to have a bigger social impact in my job.

Joined LC, Our mission was to provide people free education to help them pivot into a tech job.

\$2M+ in grants & all-hands-on-deck mode.

With a minor in business, I somehow had the most knowledge of HR, accounting and contracting, and setting up those departments became my job.

9 in stl to over 40 across the country

learned nearly everything about operating a mission driven organization on the job

wanted to build on knowledge with formal education and decided to pursue an MBA and MPP at Univ Michigan

I became interested in data science in grad school, and my goal coming out of grad school was to use data to improve how organizations operated.

Led me to Civis, where I am an applied data scientist, which means that I'm a client facing data science consultant. About 60% of my time is spent on soft skills, like client communication, project scoping and management. The other 40% is hands on keyboard data processing and analysis.

Behavioral Interviews

Why Wikimedia

What made me interested in this role is 2 things:

I enjoy my current role. But the one thing I really miss is being part of a team and building something for the long term. From my conversations with Marshall, it seems that the GDI Team is early in its existence and something I get to help shape.

Lastly, I believe that the way to make an impact at an organization with data isn't with advanced machine learning models or the most cutting edge data analytics, instead it is with persistent stakeholder engagement and tracking simple KPIs. I believe this role does that.

Behavioral Interviews

Additional Questions

Give me an example of a time you faced a conflict while working on a team. How did you handle that?

Describe a time when it was especially important to make a good impression on a client. How did you go about doing so?

<https://www.themuse.com/advice/30-behavioral-interview-questions-you-should-be-ready-to-answer>

Behavioral Interviews

Additional Questions Framework

- Situation Threat Action Results (STAR) format
- Prepare and ideally practice
- Keep it short
 - 3 minutes at most (not including follow up questions)
- All can be recycled

Behavioral Interviews

STAR format

- Situation: context for interviewer
- Threat: context for interviewer
- Action: how you approach and solve soft skills problems
- Results: your impact/why you are awesome

Behavioral Interviews

Tell me a time you had to learn something new

S	T	A	R
Data science consulting for Ford Motors. They had a basic random forest and logistic regression model built on customer car ownership data. They got a set of time-stamped events data for the same set of customers, and wanted us to see what we can do with that additional dataset. I led a team of 4 other MBAs and did some basic aggregate feature engineering - how many events per customer. Number of unique events etc.	The new data, which took about 2 weeks to wrangle, improved the model performance by around 1%. Demoralizing result and we only had another 4 weeks left	Went online and reached out to professors. Taught myself and implemented an LSTM model	LSTM was the best at finding people not likely to purchase. Ensemble, LSTM and random forest, increased accuracy by 2%

Behavioral Interviews

Tell me a time you had a difficult teammate

S	T	A	R
I was co-leading the development of a data dashboarding app with another applied data scientist. We were supported by an engineering team.	This was one of the first time the other data scientist has lead a project. He wanted to be the primary lead with the engineering team and the client, which were the two most demanding roles. He had a CS degree, so working with engineers made sense. But, I was worried things were going to fall through the cracks with the client	<ol style="list-style-type: none">1. Wanted to support his professional development and make sure we successfully deliver the project2. Had a frank conversation with my co-lead, and we put some narrower boundaries around his client responsibilities: he was the client lead, but I led support, doing app QC, and managing timelines and project	<ol style="list-style-type: none">1. we delivered a successful product2. even though I did some of the less sexy roles, because of my lead role in support, I ended up ensuring client facing tasks went smoothly. My co-lead felt like he got the experience he needed, and ended up giving me the client lead role about a month into the project

Behavioral Interviews

Additional Questions Framework

Endless amount of questions. Some basic to prepare:

- A time you made a mistake/failed (and recovered)
- A time you worked with a difficult teammate/client
- A time you had to take initiative/ownership

Offer

Understand What You Are Getting Yourself into

- People most useful to talk to: people who used to be in the role you are going into and have moved onto other companies - use LinkedIn
- People less useful to talk to: your future manager and the recruiter
- You are wanted! You are just asking for people's opinions, so most people will not mind

Ask for More

- “Perks”: vacation, salaries, pushed back start dates
 - You are not being a hassle
 - Most recruiters expect it
 - For salary, rule of thumb is 5%
 - For salary, ideal if you have another offer to benchmark, BUT NOT NECESSARY
- Have a promotion conversation - set expectations with your boss

Might not apply for certain government roles, but still doesn't hurt to ask

Ask for More

Thanks for taking the time to speak with me earlier. Per our chat, below are my two asks. I feel comfortable committing to Civis if the two points below are ok with you, ____ and the rest of the Civis team:

[point 1 redacted]

2. I am really excited about the team and the type of work I get to do as an ADS 2. I also feel that an MBA will add something unique to the commercial ADS team - to help me offset some of the cost of the MBA, is it possible for Civis to increase my salary from ____ to ____?

HW 4 Walk Through

15 - minute break

Be back at 7:45pm

First Pass

First Pass

Answering the Question Simply and Quickly

- Data cleaning
- Minimal feature selection
- Impute/remove missing values
- Create a modeling pipeline / analysis code

First Pass

Answering the Question Simply and Quickly

- Data cleaning
- Minimal feature selection
- Impute/remove missing values
- Create a modeling pipeline / analysis code

What did people find?

First Pass

Answering the Question Simply and Quickly

[coding]

First Pass

Reflect + Discuss

- What did we find?
- What should we tweak?

First Pass

Reflect + Discuss

- What did we find?
- What might we have missed?
 - Confounding factors - what other variables impacts close time?

First Pass

Reflect + Discuss

- What did we find?
- What might we have missed?
 - Confounding factors
 - Average - what might it miss?
 - What exactly does the different columns mean?

First Pass

Part 2

[coding]

Story Telling

We Covered Extensively Last Week, So I will Not Harp on This Week

Take Home Challenge

Storytelling Framework

Who: understand your audience

Goal: what are you trying to convince your audience

How: what are you going to show your audience to convince them

Who

Mayor, again

What findings will interest her?
What will not?

Goal

Assumptions

Make them and move on

- department + queue VS subject, reason, type
- Only including case_status: closed
- Ignoring ontime status
- Including all sources: constituent and city workers

Goal

Key Findings

Certain tickets take a lot longer to resolve than others. Thus, looking at the department level will not suffice

[to coding]

In Summary

- Case Interview + Behavioral Interview + Offer
- HW4