Agenda

1. Introductions
2. What are Data Science jobs?
3. How should you evaluate companies?
4. The Interview from Both Sides of the Desk:
   a. How should you prepare?
   b. What is your interviewer looking for?
What are Data Science Jobs?

Archetype 1:
What is happening in the company?
Focused on current and historical.

Archetype 2:
Creating data products.
Focused on real time and future.

Archetype 3:
Infrastructure work.
What are Data Science Jobs?

Archetype 1:
What is happening in the company?
Focused on current and historical.

Product efficacy, internal efficiency, business analytics, research

Archetype 2:
Creating data products.
Focused on real time and future.

Archetype 3:
Infrastructure work.
What are Data Science Jobs?

**Archetype 1:**
What is happening in the company?
Focused on current and historical.

*Product efficacy, internal efficiency, business analytics, research*

**Archetype 2:**
Creating data products.
Focused on real time and future.

*Product discovery. Optimization.*

**Archetype 3:**
Infrastructure work.
What are Data Science Jobs?

**Archetype 1:**
What is happening in the company?
Focused on current and historical.
*Product efficacy, internal efficiency, business analytics, research*

**Archetype 2:**
Creating data products.
Focused on real time and future.
*Product discovery, Optimization*

**Archetype 3:**
Infrastructure work.
*Data engineering*
How to Categorize a Job

- What’s their business?
  - This shapes the data available

- How mature is their pipeline/platform?
  - Companies with established data science standards and practices have different tradeoffs than startups
    - E.g. do you get to merge master in prod?

- Where does data science offer them a competitive advantage?
  - Really, really, not everyone needs RNN projection models
How to Evaluate Companies

1. Company Deep Dive
   A. Read their website.
   B. Read recent news articles.
   C. Understand the stage of the company. (Crunchbase)
   D. Try to understand their business model, and specifically how the job description fits into that business model.
   E. Generate questions to understand what it would be like to work there.
How to Evaluate Companies

2. Ask Questions:

A. Typical day of work.
B. How do they have their data housed? Who takes care of the data quality, cleaning?
C. How are the data pipelines taken care of?
D. What are the some current ongoing projects that excite them?
E. What types of interview questions do they typically include? It is totally okay to ask recruiters this!
Key Considerations for your First Job

What kind of work do you really enjoy?
Key Considerations for your First Job

What kind of work do you really enjoy?

What parts of your education and current experience can you leverage?
Key Considerations for your First Job

What kind of work do you really enjoy?

What parts of your education and current experience can you leverage? What about your skillset?
Interviews from Both Sides of the Desk

What Interviewers are Looking For
Specific Skills
How to pass

- **Technical competence**
  - How’s the code?
  - Know your (basic) stats

- **Product sense**
  - Does the tool fit the question?
  - What’s the question actually asking you to do?

- **Communication**
  - Technical communication
  - Non-technical communication
How to stand out

- **Adaptability**
  - You got more information---now what?
  - You don’t have context, but you need it?

- **Growth mindset / use of all resources**
  - ASK GOOD QUESTIONS (tech and non-tech)
  - Generalizability/flexibility

- **Simplicity**
  - Do simple things simply
  - For bonus points, do complicated things simply

- **Communication**
  - The bedrock against which most data science rests
Core Skills

SQL, Communication, Business Sense

Statistics

Computer Science

Machine Learning
Specific Must Knows: The Basics

**SQL:**
All of Mode Analytics Tutorial (https://mode.com/sql-tutorial/introduction-to-sql/)

**Behavioral:**
30s Elevator Pitch about You
Why are you leaving Academia?
3 Core Stories that cover the ground of:
1. Teamwork
2. Adaptability
3. Communication
4. Motivation and Values

**Business Sense**
Company deep dive with good business questions.
McKinley Case Studies
Churn, LTV, KPI
Understand Business Models!
Specific Must Knows: Statistics

Basics:
- EDA, Summary statistics
- Combinatorics, Probability, Bayes Theorem
- Basic Modeling (Linear, Logistic Regression)
- Central Limit Theorem, Hypothesis Testing

Resources:
- Khan Academy
- AP Statistics

Specializing in Archetype 1 or 2:
- Advanced Experimental Design
- Specialized Modeling Methods (Time Series, Bayesian Statistics, Mixed Models)
- Advanced Visualizations

Resources:
- Elements of Statistical Learning
Specific Must Knows: Machine Learning

**Basics:**
- Data manipulation, imputation, parameter tuning
- kmeans, PCA, Random Forest, SVM
- Cross Validation, Validation Metrics (F1, ROC)

**Specializing in Archetype 2:**
- Computational optimization
- XGBoost, tSNE, DBSCAN
- Recommender Systems, Dynamic Experimentation, NLP, Deep Learning

**Resources:**
- Andrew Ng Intro to Machine Learning Coursera
- Stanford CS229
- Elements of Statistical Learning

**Resources:**
- Andrew Ng Deep Learning Specialization Coursera
- Introduction to Statistical Learning
Specific Must Knows: Computer Science

**Basics:**
Data structures (up to hash, stack, linked list)
List/Dict Comprehension, RegEx
Complexity, $O(n)$

**Specializing in Archetype 3:**
Recursion, Dynamic Programming (Knapsack problem)
Binary Trees, Sorting Algorithms
Functional Programming

**Resources:**
- LeetCode (Easy)
- Hacker Rank (Easy)

**Resources:**
- LeetCode (Medium+)
- Hacker Rank (Medium+)
- Cracking the Coding Interview
Case Study Interviews

- Open-ended / brainstorming / multiple solutions
- Multi-staged ("twist" or more information added later)
- Require balancing objectives / making trade-offs

[Our company] is currently facing problem X. Let’s brainstorm some ways to solve it. How would we determine that we solved the problem?
Final Tips

In many cases, the interviewer has a specific checklist for the interview. Let them steer you.

- Answer the question the interviewer asked, especially if they appear to restate/repeat it
  - Ask clarifying questions and check assumptions
  - You may have missed something, or they may need a different level of detail
- Provide concrete examples when asked
- Offer alternatives
- Use the whiteboard, especially when prompted
  - The interviewer might not be following where you are
- Draw graphs, especially when prompted
  - The interviewer might not be following where you are
  - The interviewer might need to assess how well you dashboard/communicate to stakeholders
  - The interviewer might need to assess how well you can build data structures
  - You might be missing something that a graph would reveal
What’s a data science job really like?
Medicare beneficiaries are generally 65 and older, have multiple chronic conditions, and engage the healthcare system more than the average population.
The Business Model: Alignment of Profit and Better Member Outcomes

Identify care gaps —> Modify risk probabilities —> Member benefits —> Profit
We’re not just using technology to pay out claims. We’re building a learning machine, and we’re using it for health. By operating at the nexus of care and cost, we think we can fix how care is given and for how much.

Vivek Garipalli
Clover founder, CEO
Engagement Platform

Data Warehouse (Data Science)

Applications Database (Engineering)

The Router (Data Science, Engineering)

Eligibility logic, Prioritization logic, Member Preferences.

Calls

SMS

Structured Data Capture

Nurse Practitioners

"I stopped taking this medication"
2018 Data Science Projects

In Home Physician Program:
- High risk of hospitalizations

High Risk Chronic Care Management:
- Hypertension
- Medication Adherence
2018 Data Science Projects

In Home Physician Program:
- High risk of hospitalizations

High Risk Chronic Care Management:
- Hypertension
- Medication Adherence

Mission Control:
- Determining our likelihood of reaching a member.
- Medication Adherence

Provider App:
- Identifying suspected diagnoses.
2018 Data Science Projects

In Home Physician Program:
- High risk of hospitalizations
  93% reduction in clinician time

Mission Control:
- Determining our likelihood of reaching a member.

20% improvement in reach rate
Questions?

Contact:

jenny.yang@cloverhealth.com
andrew.ward@cloverhealth.com