

Acing your Data Science Interview



Emily Kearney

PhD, Environmental Science, Policy, and Management

Data Science Lead - NYC



Kyle Frankovich

PhD, Cognitive Neuroscience

Data Science Lead - Remote



Roadmap

1. What are my transferable skills?
 2. The Interview
 3. Common Pitfalls
 4. Questions
-

What are my transferable skills?

Who are Academic Scientists?

Job of an academic scientist:

- Plan and execute a study ~ years
- Collect and clean data
- Use programming and statistics/ML to discriminate between signal and noise
- Convey results to the scientific community

Who are Data Scientists?

Job of a data scientist in industry:

- Plan and execute a study ~ week-long sprints
- Collect and clean data
- Use programming and statistics/ML to discriminate between signal and noise
- Convey results to the team/company/investors
- Make data-informed decisions that directly impact a product and a business

The Modern Data Scientist

Math and Statistics

- Machine learning
- Statistical modeling
- Experimental design
- Bayesian inference
- Supervised learning
- Unsupervised learning
- Optimization

Programming and Database

- Computer science fundamentals
- Scripting language e.g., Python
- Statistical computing packages, e.g., R
- Databases; SQL
- Parallel databases and parallel query processing
- MapReduce concepts
- Hadoop, Hive, Spark
- Experience with AWS

Domain Knowledge and Professional Skills

- Passionate about the business
- Curious about data
- Influence without authority
- Hacker mindset
- Problem solver
- Strategic, proactive, creative, innovative, and collaborative

Communication and Visualization

- Able to engage with senior management
- Storytelling skills
- Translate data driven insights into decisions and actions
- Visual art design
- Knowledge of visualization tools

The Modern Data Scientist

Math and Statistics

- Machine learning
- Statistical modeling
- Experimental design
- Bayesian inference
- Supervised learning
- Unsupervised learning
- Optimization

Programming and Database

- Computer science fundamentals
- Scripting language e.g., Python
- Statistical computing packages, e.g., R
- Databases; SQL
- Parallel databases and parallel query processing
- MapReduce concepts
- Hadoop, Hive, Spark
- Experience with AWS

Things to highlight

Domain Knowledge and Professional Skills

- Passionate about the business
- Curious about data
- Influence without authority
- Hacker mindset
- Problem solver
- Strategic, proactive, creative, innovative, and collaborative

Communication and Visualization

- Able to engage with senior management
- Storytelling skills
- Translate data driven insights into decisions and actions
- Visual art design
- Knowledge of visualization tools

Outline of the Insight Interview

- Tell me about your background
- The code
- Steps towards data science

**Tell me about your
background**

Describing your background effectively

- **Listen** to directions and cues
- Be **concise** and clear
- **Start big** and become more specific when the interviewer shows interest
- Keep it **conversational**
- Anything in your background (and resume!) is fair game - know it!

Example: Tell me about...

...your research

...the machine learning tools you have used

...the team you were working with & how you communicated effectively

...this very specific thing from your resume

The Code

i.e. How do you problem solve?

Clear Communication

- What is the big-picture goal of your project?
- Explain technical tools in non-technical but also mathematical/technical terms
- Focus on the end user and how you provided data-driven answers
- Be able to articulate **why** you made the analytical and technical choices that you did

Collaboration

- Be willing to evaluate your code and approach
- Ask for feedback and incorporate that new perspective into your thinking
- Brainstorm trade offs, limitations, and next steps
 - Practicality? Usefulness? Validation?
 - If you took your work one step further...

Technical Proficiency

- Self-defined functions
- Good comments that communicate the why behind technical choices
- In-depth understanding of the stats/ML methods you are using

Even better:

- Self-defined classes
- Modularity

Showing your code effectively

- Start at the top (high-level)
 - Dig into most interesting aspects...
- Keep it conversational
- Mind the time
- Never go through line-by-line

Steps towards Data Science

Not just running away, but running towards

Demonstrate Your Interest

- Why DS in particular?
- Which type of data scientist do you want to be?
 - Have you evaluated the field?
 - Do you know what a data scientist does?
 - Have you looked at blogs, job postings, companies, industries...?
- What DS skills have you learned to prepare for the transition?
 - Do you know or are you learning industry standard tools? (Python/R, SQL, inferential stats, etc.)

Common Pitfalls

Interviews as
Tests

Fixation on
Novelty

Academic
Burnout

Interviews are a conversation!

- Ask questions
- Verbalize your thought process
- Check in with your interviewer
- All interviews are a culture-fit test

Most Data Science is not cutting edge

- “Interesting” is not always valuable
- Find the correct tool for the job
- All algorithms have trade-offs; be aware of them!
- Start simple, then increase in complexity as needed

Your enthusiasm is critical

- Focus on moving forward!
- Emphasize the positives, be specific
- Why data science?

Questions?



Emily Kearney

Data Science Lead - NYC

 [linkedin.com/in/emily-kearney-phd](https://www.linkedin.com/in/emily-kearney-phd)



Kyle Frankovich

Data Science Lead - Remote

 [linkedin.com/in/kylefrankovich/](https://www.linkedin.com/in/kylefrankovich/)

 INSIGHT 