I Worked Through 500+ Data Science Interview Questions

Here's what I learned along the way

Megan Dibble
Jun 14 · 5 min read ★

Thanks to a group of dedicated data science boot camp grads, Zoom, and the World Wide Web, I have listened to and answered many data science interview practice questions to count. These past months of practice have taught me a lot, not only about data science, but also about self-confidence, hard work, and the power of building relationships via networking.

Photo by Green Chameleon on Unsplash

I'll start with the data science stuff because this is a data science blog and that's what you all are likely here for.

I'll take "What is Data Science?" for 200, please.

While going through a data science bootcamp, I felt like I was grasping most of the topics pretty well. But, when answering conceptual questions about key topics and models, I realized that some of my knowledge was shallow.

Answering conceptual questions on the ear in front of knowledgable colleagues was incredibly valuable. I knew (from college) that studying with others is a good idea because it is easy to convince yourself you know topics when you are on your own, but answering questions from friends will test and solidify your knowledge.

I recommend finding someone else already in the field or someone who also wants to learn data science and working through a bunch of conceptual questions to prepare yourself for interviews and strengthen your data science knowledge and comprehension. Here is a list of *some* questions that I would consider "basic" conceptual questions, based on working through 10+ collections of questions on different websites.

- Define machine learning.
- What is an ensemble model?
- Explain how the decision tree model works.
- What is the bias/variance trade-off in data science?
- Explain cross-validation. Why is it used?
- What are Type I and Type II errors? Give examples.
- What is PCA? When would you use it?
- How do you handle missing values in your data?
- Walk me through the general steps of exploratory data analysis. What are you looking for in this stage of a project?
- Explain discrete vs. continuous variables.
- Explain the difference between descriptive, predictive, and prescriptive models.
- What is the difference between supervised and unsupervised machine learning?
- Explain overfitting and underfitting. How can you reduce each of these?

Do you feel shaky on any of your answers to these? Time to find a friend! And perhaps a textbook, or more likely Google. I am not going to provide the answers I would give — maybe I will in a later article — because I have found that it is a valuable learning process to seek these answers out, to have discussions and debates with others, and arrive at an answer yourself.

Photo by JESHOOTS.COM on Unsplash

There are enough coding questions and challenges on the internet to keep you occupied for a long time. And it's important for a data scientist to be proficient in programming. But, in my opinion, what makes data scientists valuable is mostly their knowledge of the field and their ability to apply that knowledge. You can always take a crash course on Python, brush up on R, or google the formula for a 95% confidence interval. But, if you are not confident about fundamental data science concepts, how do you even know where to start when handed a project?

If you are looking for interview questions, I have included links at the end of this article. If you would like to hear what I found helpful from the hundreds of questions taught me, keep reading.

The importance of networking and well-roundedness

Through a family member, I was connected to the group of data science boot camp grads that I mentioned at the beginning of this article. They were committed to answering practice data science interview questions with each other for one hour, three times a week, over Zoom. (Impressive, right?)

At first, it was easy to talk myself out of joining their Zoom calls. I was not finished with my boot camp, and I felt the time I did have understand either the question or their answer. I felt like I would mess up an easy answer or just not be able to contribute at all.

Photo by Allie on Unsplash

I have had similar feelings around networking in the past. It is hard, uncomfortable, and often the benefits are not immediately (or ever) evident. But, these Zoom calls taught me the value of *sticking with it!* The other aspiring data scientists in these calls have become amazing resources for me, and most importantly, they have become friends. I think networking is much less exhausting when it is mutually beneficial and more like building bonds than listing a business card or LinkedIn connection.

I learned that I could be incredibly prepared for a data science interview, but actually getting the interview was another, even more crucial challenge that I had not prepared as much for.

All of this culminates into one big lesson that I have been learning for the last few months: the importance of being well rounded.

I have heard the term *data science unicorn* used to describe position descriptions where an employer wants a data scientist who can do literally everything. While this expectation is unrealistic, there are so many different requirements for being a data scientist. Thus, it is important to focus on building your skills in all of these areas as you are applying to jobs! I learned that I needed to spend time widening my coding knowledge, deepening my understanding of statistics and math, learning more about data science concepts, and practicing communicating these things to a non business-focused audience.

This all takes time.

I am generally an impatient person. But, you cannot expect yourself to take a deep dive into all these areas in just a couple of months! And that's ok. My boot camp experience went from a target of 5 months to a total of 7 months, and I am so glad I took the time to slow down and really understand things, as well as build a network by writing articles and becoming more active on LinkedIn.

I hope for your sake that you are already a data science unicorn. But, since I have not seen one myself out in the wild, I hope that this article was helpful to you in some way on your journey to become a data scientist.

Now for those practice question links:

101 Commonly Asked Data Science Interview...
Preparing for an interview is not easy there is significant...
www.springboard.com

100+ Data Science Interview Questions an...
In this Data Science Interview Questions blog I will...
www.edureka.co

Top 70 data science interview questions.
Data science interview questions
towardsdatascience.com

OVER 100 Data Scientist Interview Questions ...
Interview Questions from Amazon, Google, Facebook,...
www.techcareerbooster.com

Top 210+ Machine Learning Interview...
In case you are searching for Machine Learning Interview...
www.gangboard.com

p.s. I apologize, some of these have some questions or answers that are confusing and/or poorly written. However, on the whole, they are good resources :)

Good luck, and happy studying!

To read my other articles on Towards Data Science, click here.