

Your data scientist hiring guide

Ask the right interview questions and compare your candidates to our data scientists

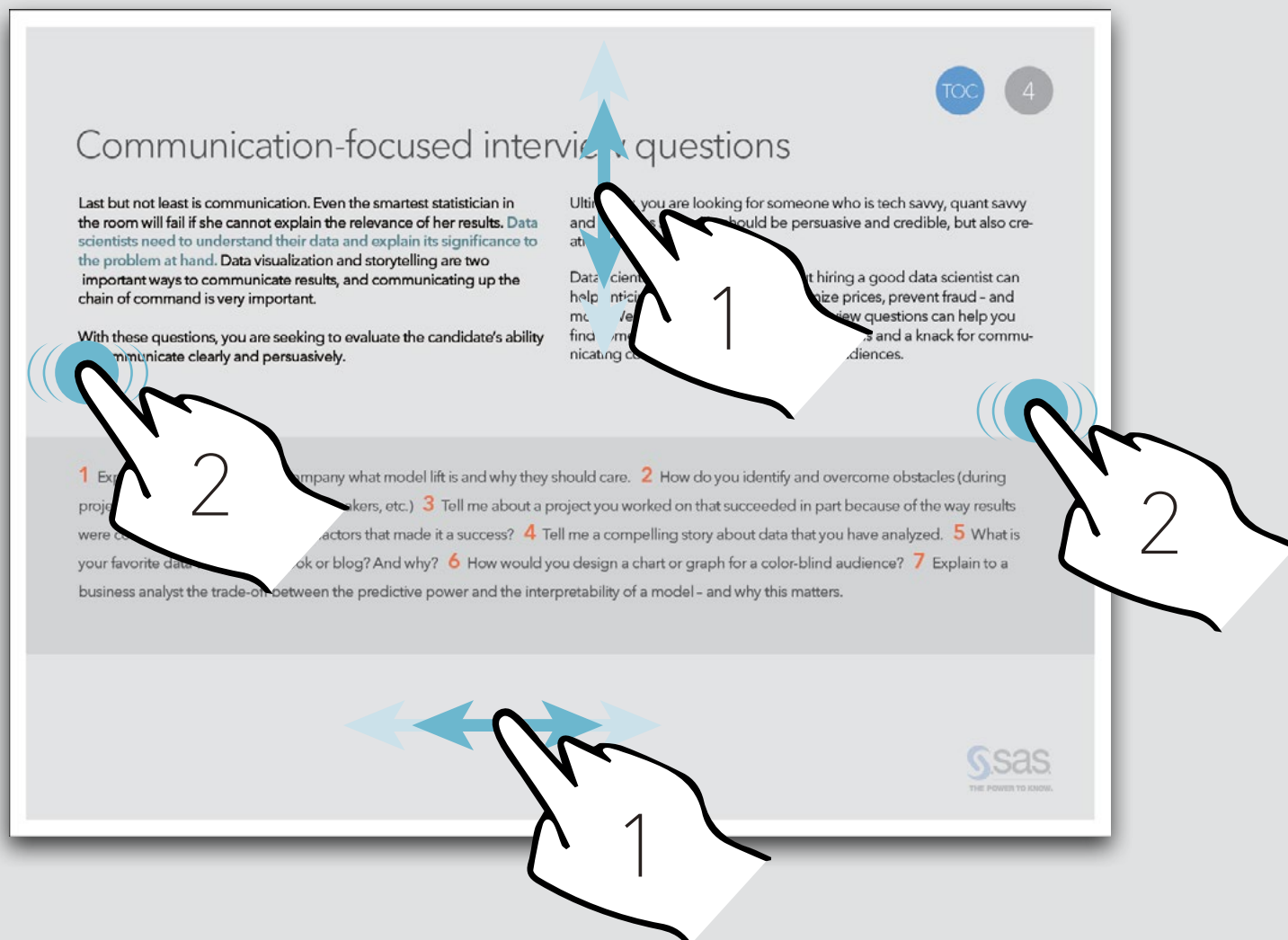
By 2018, the United States alone could face a shortage of 140,000 to 190,000 people with deep analytical skills, according to a McKinsey study. Likewise, the study anticipates a deficiency of 1.5 million managers and analysts with the know-how to use big data and analytics to make effective decisions. How can you get ahead of the trend and hire this rare breed of analytical thinker who combines business know-how with computational smarts? Use our interview questions and mold your data scientist role after our three sample data scientists.



How to navigate this e-book (mobile)

How you move from page to page depends on your PDF viewer's settings

1. A) Swipe left to proceed to the next page, or swipe right to return to the previous one.
B) Swipe up to proceed to the next page, or down to return to the previous one.
2. Tap the right edge to proceed to the next page, or tap the left edge to return to the previous one.



Icons and links



Indicates an active link



Return to table of contents



Links to sas.com

* All links open externally from the e-book. To return, navigate to your open files and tap the e-book.

Table of contents

20 questions for recruiting
well-rounded data scientists

1

Technical interview questions

2

Practical experience
interview questions

3

Communication-focused
interview questions

4

Meet the data scientists

5

6 Meet Kristin Carney

8 Meet Daymond Ling

11 Meet Colin Nugteren

14 Final tips on hiring
data scientists

20 questions for recruiting well-rounded data scientists

Data scientists have a unique blend of skills. Not only do they solve problems by squeezing information out of data, they also communicate results and persuade others to apply that information to their decisions.

When hiring data scientists, how can you evaluate the full range of required skills?

Focus the interview on more than technical skills. Ask questions that evaluate logical reasoning agility, problem solving skills and the ability to collaborate and communicate with business analysts and decision makers. You also want to evaluate practical experience, but keep in mind that university projects and volunteer work can provide that background in addition to on-the-job experience.

To help you evaluate this full range of skills, we've developed a list of 20 questions, but we've broken the list down into three key categories:

1. Technical questions.
2. Practical experience questions.
3. Communication questions.

Consider our questions to be a starting point for your data scientist interviews. They are designed to evaluate the three sets of skills above, but you'll want to add your own questions to make sure the candidate is a good fit with your organization's culture and requirements.



Technical interview questions

Statistics and machine learning are important technical skills for data scientists. These questions help measure knowledge, plus the ability to explain complex topics. Some of the questions are also designed to bring out the art and science of data science.

1

What is the curse of dimensionality and how should one deal with it when building machine-learning models?

2

Why is a comma a bad record separator/delimiter?

3

Explain the difference between a compiled computer language and an interpreted computer language.

4

How do you determine "k" for k-means clustering? Or, how do you determine the number of clusters in a data set?

5

What's more important: predictive power or interpretability of a model?

6

Explain finite precision. Why is finite precision a problem in machine learning?

7

Explain the "bias-variance trade-off" and why it is fundamental to machine learning.



Practical experience interview questions

Technical skills are important, but they must be applied to solve problems. Your data science candidates should be able to describe projects they have worked on, and how they turned out. They also should be able to articulate what aspects of their technical training have been important in their day-to-day data scientist tasks, and how they can apply their skills to your business.

Remember, the candidate's practical experience does not need to be within your industry. Some hiring managers say their best hires have been people outside their industries who can look at problems from a fresh perspective. Evaluate practical skills by asking some of these questions:

- 8 Describe a recent use of logistic regression.
- 9 Describe an analysis you have recently completed, including strategies and findings. How were the findings used by the business? (This can be from a student research project or thesis if the candidate is a recent graduate.)
- 10 Give examples of data cleaning techniques you have used in the past.
- 11 What subjects would you include in a one-day data science crash course? And why?
- 12 Describe a situation where you had to decide between two different types of analyses – and why you chose the one you did.
- 13 Explain the benefits of test-driven software development; or explain the benefits of unit testing.



Communication-focused interview questions

Last but not least is communication. Even the smartest statistician in the room will fail if she cannot explain the relevance of her results. Data scientists need to understand their data and explain its significance to the problem at hand.

Data visualization and storytelling are two important ways to communicate results, and communicating up the chain of command is very important. With these questions, you are seeking to evaluate the candidate's ability to communicate clearly and persuasively.

14 Explain to the leaders of this company what model lift is and why they should care.

15 How do you identify and overcome obstacles (during projects, with customers, with decision makers, etc.)

16 Tell me about a project you worked on that succeeded in part because of the way results were communicated. What were the factors that made it a success?

17 Tell me a compelling story about data that you have analyzed.

18 What is your favorite data visualization book or blog? And why?

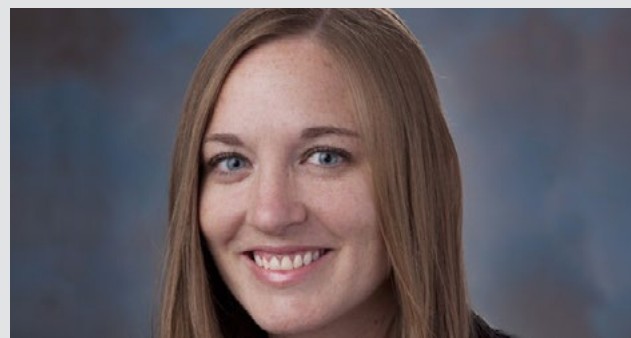
19 How would you design a chart or graph for a color-blind audience?

20 Explain to a business analyst the trade-off between the predictive power and the interpretability of a model - and why this matters.

Ultimately, you are looking for someone who is tech savvy, quant savvy and business savvy. He should be persuasive and credible, but also creative and passionate.

Data scientists are in short supply, but hiring a good data scientist can help anticipate customer needs, optimize prices, prevent fraud - and more. We hope these data scientist interview questions can help you find someone with a range of technology skills and a knack for communicating complex subjects to a variety of audiences.

Meet the data scientists



Data scientists come from a variety of backgrounds, and they wear many different hats. Where do they fit in the organization? And what does a typical day look like for a data scientist? There is no right answer and no consistent description.

The only commonalities among data scientists are the basic skill sets: curiosity, communication skills and mathematical aptitudes. Beyond those essential skills, the roles and daily activities vary from company to company.

Here, we introduce you to three data scientists from different backgrounds and different industries. As you read their stories, think about the needs of your organization, and use their stories to further refine and develop the data scientist role you are building to help solve complex problems in your organization.

Meet Kristin Carney

When Kristin Carney graduated with a BS in mathematics, she wasn't sure what she wanted to do with her degree. That's when she began researching data science. Here, we interviewed Carney, who works in Research & Analysis for World's Foremost Bank.



Kristin's background

I have a BS in mathematics and an MS in statistics from the University of Nebraska-Lincoln. I was the 2013 vice president of the American Statistical Association Nebraska Chapter and the 2014 president of the American Statistical Association Nebraska Chapter.

Kristin on data scientist skills

I like to solve puzzles, and as a data scientist for a credit card portfolio, those puzzles often look like questions about how to mitigate risk and increase profit. Fortunately, in this industry there's a lot of data.

In my opinion, logical reasoning is the most important skill for data scientists, closely followed by communication skills, because you have to explain your analysis. You also need to be detail oriented. I believe I've been more successful due to my curiosity. I want to know everything. My family pokes fun at me since my favorite question is "why?"

Kristin's role in the organization

I work in Research & Analysis for World's Foremost Bank (owned by Cabela's Inc). I report to the vice president of portfolio management. I started at Cabela's in July 2010 as a research analyst. I transitioned to a statistician role in August 2010 and started my current position as a risk manager in April 2013.

Kristin's team

Our team is made up of all data scientists and their managers.

I manage a team of between four and eight individuals. They are all analytical problem solvers. Some of them have statistical modeling experience and some do not. Their education is typically in mathematics, statistics or economics.

Kristin's typical day

On a typical day, I brainstorm and answer business questions from my team, I review analysis and recommendations completed by my staff, and I attend a variety of meetings.

Kristin's biggest challenge

My biggest challenge has been in my management role, helping my staff obtain resources and budget to execute changes based on analytical findings.

Kristin's biggest accomplishment

My biggest accomplishments are seeing my team succeed, gaining experience in decision making, and acquiring the ability to say no professionally. Sometimes a decision, even if it is the wrong one, is better than no decision at all - and there are only so many hours in a day and so many resources.

Meet Daymond Ling

Daymond Ling, a data scientist at CIBC for the past 20 years, believes the right personal traits are more important than technical skills when it comes to being a successful data scientist. He says the latter can be learned; the former is much harder to shape and cultivate.



Daymond's background

I've always had a fascination with how the world and universe works; things like looking for order, patterns, purity and truth. That led me to a BS degree in honors physics. I then obtained a Master of Science in business administration with a focus on statistics and operations research. I liked learning about unpredictability, chaos and a statistical approach to systems. Operations research involves learning how a complex system's behavior can be described mathematically followed by the analysis of optimal decisions of desired states and outcomes. And actually, the learning never stops.

Daymond on data scientist skills

Advanced analytics requires deep technical competencies. You need to have a good grasp of mathematics and statistics. Knowledge of operations research further extends one's ability into optimal decision making in complex trade-off problems, a very useful skill. You need

these skills to be in the advanced analytics game. They are entrance requirements, but they are not enough – I always tell people that quantitative skills are necessary but insufficient conditions for success.

You need to be curious about the business processes and how they interconnect. You need clear, structured thinking to identify well-defined problems. Then, take a common-sense approach to the solution, balancing sophistication and complexity against resource and speed of delivery.

You need the ability to shape and manipulate very large amounts of data at will. You need to be able to manipulate the data every which way you want with absolute ease. The ability to express your analytical ideas and thoughts via crisp, clean, modular code with lightning speed is a big boost to productivity.

And when you have the solution to the problem, you need great storytelling abilities to help business people understand what you saw, what you did, what you found out and what they should do – all in a clear, persuasive, captivating manner using their language, not your analytical language.

In many ways, I believe personal traits are more important than technical skills, as the latter can be taught and learned, whereas the former are much harder to shape and cultivate.

Daymond's role in the organization

I am always attached to the business unit where the problems reside. Advanced analytics is not an IT function, and being distant from the problem creates, in my opinion, unnecessary and detrimental barriers. You cannot solve a problem well when you are distant from it. At American Express, I worked in card operations and risk management. I analyzed client behavior across multiple departments that look after the card product life cycle, from application, authorization, credit, collections and fraud. Currently at CIBC, I am attached to marketing and report to the vice president of customer intelligence and competitive insights.

The shift from risk management to marketing is interesting. Whereas risk management is about controlling operational processes and mitigating losses, marketing brought a fresh new perspective of using data and analytics to identify business opportunities and create new possibilities. I joke with my peers in risk management that it is now my turn to create headaches for them as we look for new ways to grow the business.

Daymond's team

Over the years, I've had the pleasure to work with both experienced people as well as talent straight out of school. They are physicists, mathematicians, econometricians, statisticians, biostatisticians – all



people with strong quantitative skill sets and disciplined scientific minds. Many started knowing little about formal statistics or SAS®, but their strong quantitative foundation and bright mind enabled them to pick up skills on the job. I run a training school of sorts.

Daymond's typical day

My days are a mix of many things - working with business partners to understand their business problems and concerns, educating people on what we can do and how it can benefit them, proposing frameworks and methods to solve problems, presenting findings and driving for change. Even politely declining engagement where analytics is not what is needed.

Specific examples of projects we work on include customer acquisition for the bank, predicting individual behaviors, determining risks and responding to changing market conditions and business strategies.

We have weekly full-team meetings where all new discoveries are put on the table and probed to make them better. We cross-check what we just discovered with what we already know from before. I try to challenge people. I want them to think about how problems can be approached differently and what possibilities we have for improving on what we've already done. We also keep an idea log of what would be interesting to investigate - what we call our "skunk-work projects."

Daymond's biggest accomplishment

My biggest joy is introducing advanced analytics to many bright people and instilling analytics passion within them. It's wonderful when you see that "they get it." The light bulb comes on, their torch is lit and it will continue to spread.

As far as technical accomplishments, I have a long list, but the one that stands out is identifying a billion-dollar funds impact, a research project that hit gold. I did it in three months' spare time. Just an idea, gobs of data and SAS - that's all it took. It's not every day that you can find a prize of that size.

Meet Colin Nugteren

DirectPay Chief Analytics Officer and Data Scientist Colin Nugteren says while every day is different, one thing remains the same. He ends each day with a hands-on look at important customers using SAS® Visual Analytics.



Colin's background

I studied econometrics and management science at the Erasmus University School of Economics in Rotterdam. I graduated with a Master of Science in the field of marketing econometrics based on research I did for DirectPay about consumer credit risk in a factoring environment.

Colin on data scientist skills

It's quite an open door, but in essence it requires an analytical mind. Next to that, it is very useful to understand computers and data and the way they work together, because all major analytical advances are highly dependent on these systems.

Several other skills like autodidactic learning and an organizational mindset are very important. When you work with huge amounts of data in several systems, the competence of fast learning and a high-quality structure of the analysis, models and business rules are very important.

Colin's role in the organization

I have worked as an analyst for DirectPay for four years, and I reported to the IT manager for the systems that we build. At the same time, I reported to the CEO for the analysis of our clients and portfolios. I learned a lot from that combination of IT and business.

I was simply hired to analyze the database with payments and I thought I would be done within a year or two. I learned that this process is ongoing and new possible improvements come up daily. In 2015 I was promoted to CAO, and now I am responsible for both the analysis department and the analytical decisions underlying our operational processes that actually try to collect money. Within these processes the use of analytics was subordinated for a long time, but we have achieved great results with analytics in the last few years.

Colin's team

In our management, I work closely with my colleagues who are responsible for IT and sales. I work with my team of data scientists and our team works closely with the development team, the DBA, R&D and the business team. We realized that there are huge possibilities in analyzing more and higher-quality data. I like to work with intelligent people that challenge me to think out of the box.

Colin's typical day

At this moment I am responsible for the implementation of the SAS Decision Management system. So I actually pay a lot of attention to the different SAS systems that I used to work with daily, like SAS Visual Analytics, SAS/STAT® and SAS® Enterprise Miner™. We work hard to bring these models and analyses to SAS Decision Manager and SAS Data Integration Studio.

I continue to use SAS Visual Analytics because I like to do the analysis of important clients myself. I understand our business and clients best this way. In my current role, every day is different, but I end each day with SAS Visual Analytics and a look at the production reports. Our main goal is to implement as much intelligence into our systems and decisions as possible, and I am very busy discovering opportunities to increase the use of intelligent systems.

Colin's biggest challenge

It is a serious challenge to use time efficiently. My work is never finished, but I try to set goals every day and achieve them. But this can be a real challenge some days.

Colin's biggest accomplishment

Building the credit rating and collections systems that empower our organization.

Final tips on hiring data scientists

Once you've hired your first data scientist, use this plan from MIT Sloan Management Review to learn how to build a data science program that will work for your business.

- 1 Appoint and empower a data and analytics leader.
- 2 Point data scientists to your biggest problems and then get out of the way.
- 3 Cultivate support for your data scientists among decision makers.
- 4 Connect data scientists within the organization but locate them near decision makers.
- 5 Encourage data scientists to get their hands dirty.
- 6 Build analyst teams with diverse skills.
- 7 Reward data scientists in ways they care about - with recognition and intellectual challenges.



Need more? Click for full data scientist job descriptions and longer plans for building a data science program



What about training for your data scientists? Send them to the SAS[®] Academy for Data Science.

Get more Insights from SAS.

