One Data Science Job Doesn’t Fit All

- Published on July 24, 2018

Elena Grewal, Ph.D.

One of the fun things about being a leader at a hyper-growth company is that you don’t just have the opportunity to change things — you *must* drive change to keep up. And working in the new and rapidly evolving field of Data Science (DS) entails another level of rapid change. We are evolving within the company and as an industry in parallel.

At Airbnb, we think of data as the voice of our users at scale. Our goal is for data scientists to maximize their impact and to look forward to coming to work. Achieving this goal is a work in progress, and we’re continually looking for ways to improve. We recently established a role-defining framework as a part of this evolution. My hope is that what we've learned along the way can help other companies be strategic in defining data science roles.

The main takeaway I will share is that companies consider three tracks of data science work to meet the needs of your business — **Analytics**, **Inference**, and **Algorithms**. Below I'll describe the evolution of how we came to these three tracks of work and how it helps.

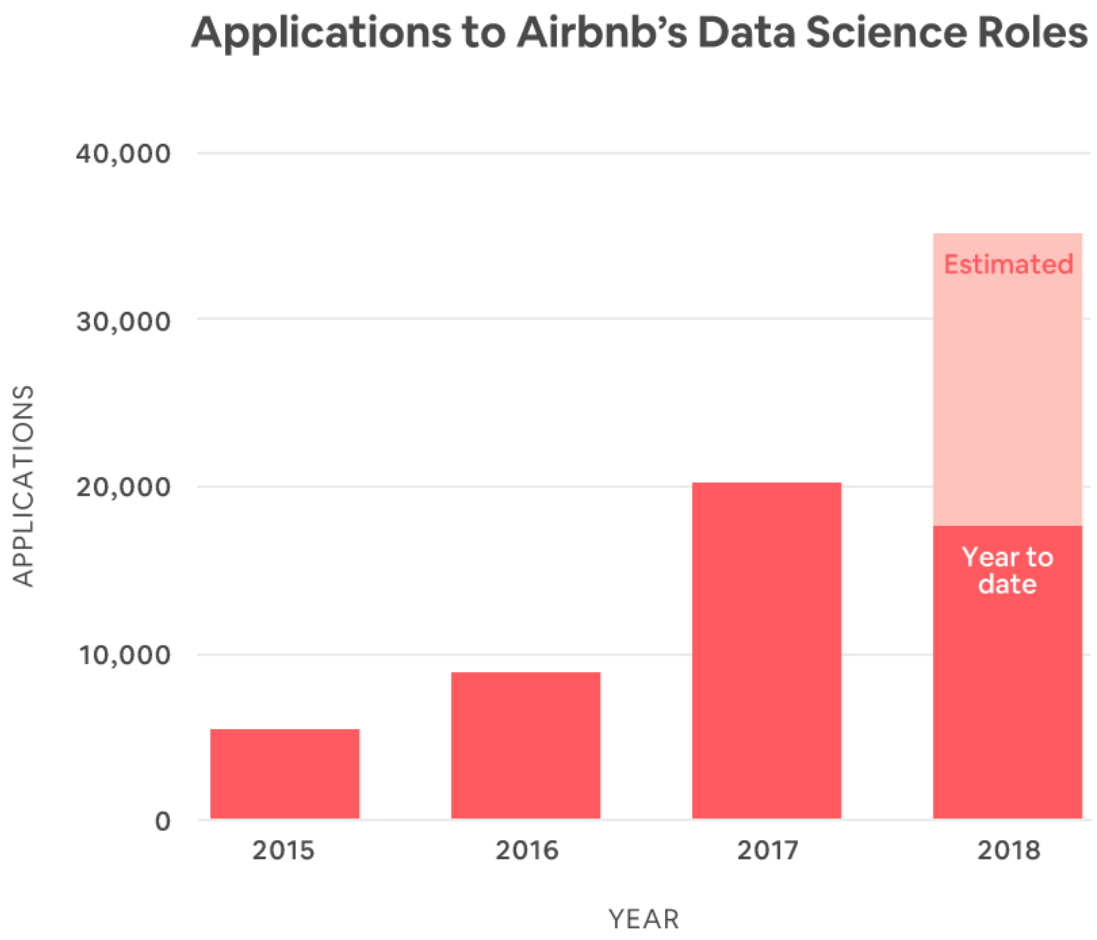
A Data Scientist by Any Other Name...

We started off as the “A-team” — **A**nalytics team — with the first hire an “Analytics Specialist.” In 2012, I was hired as a “Data Scientist.” Later, we hired a “Data Architect,” to tackle data quality, then a “Data Analytics Specialists” to help solve gaps in data access and tools. Then we saw additional needs in machine learning so we hired “Machine Learning Data Scientists.” These title evolutions were both reactions to team needs and also to the competitive landscape. We became the “Data Science” function in 2015, though we still use “A-team” because it’s fun and has a history we value.

When I took on leadership of the data science function in mid-2017, we had about 80 data scientists, embedded on various teams. Some were building dashboards, some were building NLP (natural language processing) models, others were building models for decision making and designing experiments. The landscape was incredibly varied.

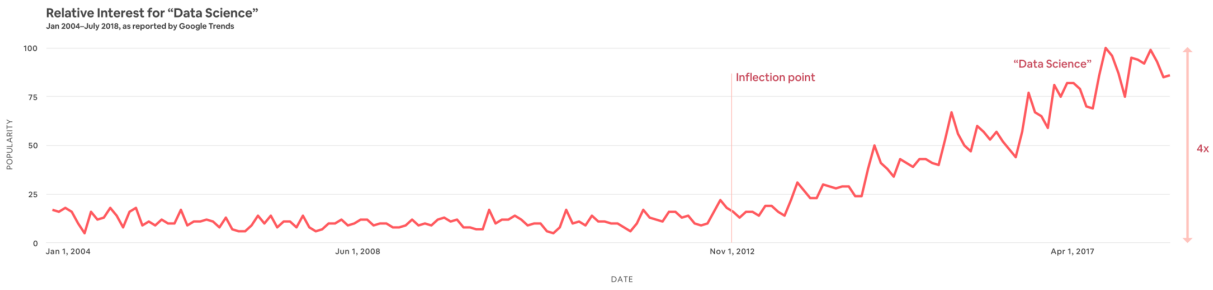
Fast Growth in an Emerging Discipline

This variety isn't totally unexpected. Data science is relatively new and growing rapidly. We see this in the data. First, looking internally, we see that applications to data science roles at Airbnb have grown 4x between 2015-2018 --



(Of course, this is also driven by interest in Airbnb and many other factors.)

And, according to Google Trends data, queries for data science have also grown (link):



The increase begins in 2012, now 4x in six years.

Not only is a new field, but what people mean when they say “data science” is also incredibly varied. Sometimes it is purely machine learning. Sometimes it’s business intelligence at a tech company. It's new, and it's evolving.

Recognizing the Diversity of Data Science Skills

We found that expectations weren’t clear about what Data Science worked on.

The downside of this variety within a given company is it can result in organizational confusion and churn, since partner teams don’t know what to expect from data scientists, and the data scientists themselves might be unclear on their role. People who come from places where DS do only modeling might not consider it a good use of data science skills to do more simple analytics. Others from places where DS only does analytics might feel that it is best for engineers to do modeling.

We also had an additional challenge: team members who were doing analytics work felt like their work was not as valued as machine learning work, and yet their work was critical for the business. Business partners craved more actionable insights to drive decisions and expanded tools to understand the data themselves. We had invested in data education through our very popular Data University, but we still needed experts. We identified one reason was that we were using the title “Data Analytics Specialist” though the team members were part of the "Data Science" function, and also that there were cues in how we talked about “data science work” that gave the impression that the analytics work was not of equal importance.

I spoke with leaders at peer companies to understand how other teams handled this— at one point I even created a shared spreadsheet with different organizations’ structures. I heard stories of new analytics teams being created from scratch, teams splitting off machine learning, tooling teams being incorporated into data science, and many more.

It was clear there was no one size fits all approach, but that being strategic and intentional about defining who we are and how we add value would be critical. We knew that our goal was to “champion the mission,” i.e., do the work the company needed most. Therefore, we needed roles that matched the current business needs, while also allowing for individuation and clarity of expectations.

The Solution: Three Flavors of Data Science Work

We decided to restructure data science along three tracks. These described what we were looking for and are areas we want to attract talent.

The Analytics track is ideal for those who are skilled at asking a great question, exploring cuts of the data in a revealing way, automating analysis through dashboards and visualizations, and driving changes in the business as a result of recommendations. The Algorithms track would be the home for those with expertise in machine learning, passionate about creating business value by infusing data in our product and processes. And the Inference track would be perfect for our

statisticians, economists, and social scientists using statistics to improve our decision making and measure the impact of our work.

Data Scientist – Analytics

Defines and monitors metrics, creates data narratives, builds tools

Data Scientist – Algorithms

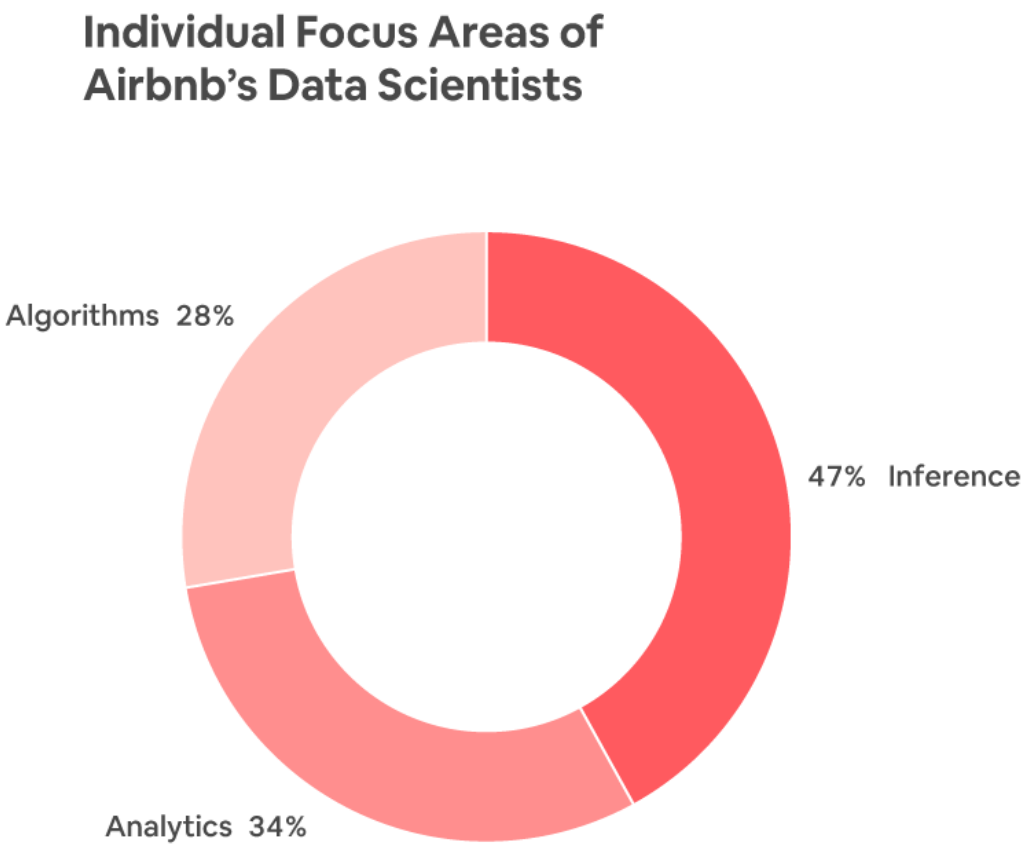
Builds and interprets algorithms that power data products

Data Scientist – Inference

Establishes causal relationships with statistics

Every data scientist on the team is expected to have expertise in one of these areas, and to acquire skills across these areas, depending on business needs and their own interests. Within each track there can be further specialization, but everyone has the title "Data Scientist" and then the descriptor that follows provides more clarity.

If we look at another discipline like Engineering, there is the helpful shorthand of "Front end" and "Back end" engineering, which helps you get a sense of someone's skills or the area of focus. I realize it's an imperfect distinction, but it gives a better sense of someone's expertise than simply "Engineering." Data science is still far from this point; this is what we are moving towards.



Clarifying Expectations

We also modified our performance evaluation criteria to reflect our new structure. We have multiple levels of data scientists and also managers. We determine success by looking at impact on the business. For those on the technical track, we modified our evaluation framework to align to these major areas --

Technical

- Analytics - Defines and monitors metrics, creates data narratives, and builds tools to drive decisions
- Algorithms - Builds and interprets algorithms that power data products
- Inference - Employs statistics to establish causal relationships
- Foundation - Demonstrates ownership and accountability for data quality and code *(expected for all tracks)*

Business (*expected for all tracks*)

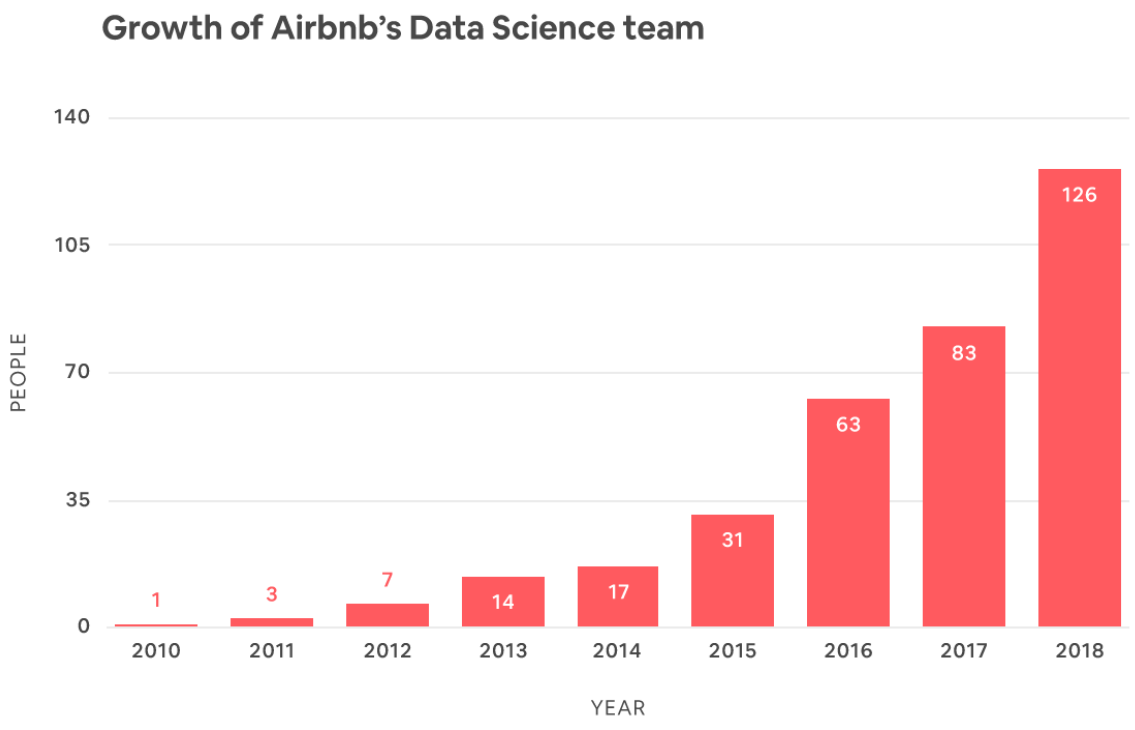
- Ownership - Able to drive projects to success, enables others, owns impact
- Influence - Communicates clearly, demonstrates teamwork, and builds relationships
- Enrichment - Contributes to team-building through mentorship, culture, recruiting, and diversity efforts

There's so much more we could write about here, but the main takeaway is that we also were explicit in changing how we evaluate performance to reflect the three tracks of work and to clarify expectations.

When to Specialize

Airbnb is big enough that having all these distinctions and nuance makes sense. When speaking with smaller companies who wonder if they should build their team with specialists, I advise them to start general. It was really helpful in the early days for us to be able to tackle whatever project was most pressing, rather than sitting inside a rigid speciality. Over time, specialization makes sense but it's best to start general, unless you see a business case for it sooner.

We didn't start to specialize until around 2015 when our team was 30 people.



We also expect to continue to change the roles of the function, as the needs of the business change.

Reaping the Benefits

Even today with our specialties, Data Scientists within each track also perform other types of work, and we encourage team members to be generalists as well (sometimes that's a point of confusion). Overall though, we have heard much less confusion after making this change. I've also started to hear partners say things like "we need someone with inference and algorithms expertise." So the language has been useful for communicating business needs.

This helps us identify gaps. I recently connected with a Product Manager who expressed concerns that the team wasn't coming up with innovative ways to run experiments in her challenging product area. I immediately diagnosed the issue: on that particular data science team, there was no one with inference expertise. That was something we could fix with our next hire or by encouraging the team members to learn from other inference experts.

We were glad to hear from team members working on analytics that they no longer felt alienated or inferior. Analytics experts understood they would have **less** impact if they tried to apply machine learning to the business problems they were working on.

Where we go from here

My hope in sharing our story is for other companies to adopt this framework! Hiring is complicated when candidates come with a vague title of "Data Science" which could mean so many different things. If all companies used a similar framework, that would make it easier for data science as a whole to communicate our value

If you like this concept, tell your data science leader or make the change yourself if you are that leader. Or if you have a better model, I'd love to hear about that too - please reach out (data-science-org-ideas@airbnb.com). Given how new and rapidly developing the data science field is, the best naming conventions will evolve over time. In data science, the more we can join forces to establish norms, the more rapidly our industry will mature and the more equipped we as individuals will be to navigate it.

It's exciting to me to share what we have done, and I'd love to share more in future posts. If there are commonly asked questions, I can respond to those as well.

If you're interested in learning more about Airbnb's Data Science work, check out our blog and follow us on Twitter at @airbnbdata. We're hiring! Apply for roles here.