

Data Science and the Art of Persuasion

Organizations struggle to communicate the insights in all the information they've amassed. Here's why, and how to fix it.

by [Scott Berinato](#)

Kristen Meyer

Summary. Despite heavy investments to acquire talented data scientists and take advantage of the analytics boom, many companies have been disappointed in the results. The problem is that those scientists are trained to ask smart questions, wrangle the relevant data, and uncover insights—but not to communicate what those insights mean for the business. To be successful, the author writes, a data science team needs six talents: project management, data wrangling, data analysis, subject expertise, design, and storytelling. He outlines four steps for achieving that success: (1) Define talents, not team members. (2) Hire to create a portfolio of necessary talents. (3) Expose team members to talents they don't have. (4) Structure projects around talents.

Data science is growing up fast. Over the past five years companies have invested billions to get the most-talented data scientists to set up shop, amass zettabytes of material, and run it through their deduction machines to find signals in the unfathomable volume of noise. It's working—to a point. Data has begun to change our relationship to fields as varied as [language translation](#), [retail](#), [health care](#), and [basketball](#).

But despite the success stories, many companies aren't getting the value they could from data science. Even well-run operations that generate strong analysis fail to capitalize on their insights. Efforts fall short in the last mile, when it comes time to explain the stuff to decision makers.

In a question on Kaggle's 2017 [survey](#) of data scientists, to which more than 7,000 people responded, four of the top seven "barriers faced at work" were related to last-mile issues, not technical ones: "lack of management/financial support," "lack of clear questions to answer," "results not used by decision makers," and "explaining data science to others." Those results are consistent with what the data scientist Hugo Bowne-Anderson found [interviewing 35 data scientists](#) for his podcast; as he wrote in a 2018 HBR.org article, "The vast majority of my guests tell [me] that the key skills for data scientists are...the abilities to learn on the fly and to communicate well in order to answer business questions, explaining complex results to nontechnical stakeholders."

In my work lecturing and consulting with large organizations on data visualization (dataviz) and persuasive presentations, I hear both data scientists and executives vent their frustration. Data teams know they're sitting on valuable insights but can't sell them. They say decision makers misunderstand or oversimplify their analysis and expect them to do magic, to provide the right answers to all their questions. Executives, meanwhile, complain about how much money they invest in data science operations that don't provide the guidance they hoped for. They don't see tangible results because the results aren't communicated in their language.

Gaps between business and technology types aren't new, but this divide runs deeper. Consider that 105 years ago, before coding and computers, Willard Brinton began his landmark book *Graphic Methods for Presenting Facts* by describing the last-mile problem: "Time after time it happens that some ignorant or presumptuous member of a committee or a board of directors will upset the carefully-thought-out plan of a man who knows the facts, simply because the man with the facts cannot present his facts readily enough to overcome the opposition....As the cathedral is to its foundation so is an effective presentation of facts to the data."

Executives complain that data science doesn't provide the guidance they hoped for.

How could this song remain the same for more than a century? Like anything else this deeply rooted, the last-mile problem's origins are multiple. For one, the tools used to do the science include visualization functionality. This encourages the notion that it's the responsibility of the data person to be the communicator. The default output of these tools can't match well-conceived, fully designed dataviz; their visualization often isn't as well developed as their data manipulation, and the people using the tools often don't want to do the communicating. Many data scientists have told me they're wary of visualization because it can dumb down their work and spur executives to draw conclusions that belie the nuance and uncertainty inherent in any scientific analysis. But in the rush to grab in-demand data scientists, organizations have been hiring the most technically oriented people they can find, ignoring their ability or desire (or lack thereof) to communicate with a lay audience.

That would be fine if those organizations also hired other people to close the gap—but they don't. They still expect data scientists to wrangle data, analyze it in the context of knowing the business and its strategy, make charts, and present them to a lay audience. That's unreasonable. That's unicorn stuff.

To begin solving the last-mile problem, companies must stop looking for unicorns and rethink what kind of talent makes up a data science operation. This article proposes a way for those that aren't getting the most out of their operations to free data scientists from unreasonable expectations and introduce new types of workers to the mix. It relies on cross-disciplinary teams composed of members with varying talents who work in close proximity. Empathy, developed through exposure to others' work, facilitates collaboration among the types of talent. Work is no longer passed between groups; it's shared among them.

A team approach—hardly new, but newly applied—can get data science operations over the last mile, delivering the value they've created for the organization.

Why Are Things Like This?

In the early 20th century, pioneers of modern management ran sophisticated operations for turning data into decisions through visual communication, and they did it with teams. It was a cross-disciplinary effort that included gang punch operators, card sorters, managers, and draftsmen (they were nearly always men). Examples of the results of this collaboration are legion in Brinton's book. Railroad companies and large manufacturers were especially adept, learning the most efficient routes to send materials

through factories, achieving targets for regional sales performances, and even optimizing vacation schedules.

How Communication Fails

I've learned in my work that most leaders recognize the value data science can deliver, and few are satisfied with how it's being delivered. Some data scientists complain that bosses don't understand what they do and underutilize them. Some managers complain that the scientists can't make their work intelligible to a lay audience.

In general, the stories I hear follow one of these scenarios. See if you recognize any of them.

The Statistician's Curse

A data scientist with vanguard algorithms and great data develops a suite of insights and presents them to decision makers in great detail. She believes that her analysis is objective and unassailable. Her charts are "click and viz" with some text added to the slides—in her view, design isn't something that serious statisticians spend time on. The language she uses in her presentation is unfamiliar to her listeners, who become confused and frustrated. Her analysis is dead-on, but her recommendation is not adopted.

The Factory and the Foreman

A business stakeholder wants to push through a pet project but has no data to back up his hypothesis. He asks the data science team to produce the analysis and charts for his presentation. The team knows that his hypothesis is ill formed, and it offers helpful ideas about a better way to approach the analysis, but he wants only charts and speaking notes. One of two things will happen: His meeting will be upended when someone asks about the data analysis and he can't provide answers, or his project will go through and then fail because the analysis was unsound.

The Convenient Truth

A top-notch information designer is inspired by some analysis from company data scientists and offers to help them create a beautiful presentation for the board, with on-brand colors and typography and engaging, easily accessible stories. But the scientists get nervous when the executives start to extract wrong ideas from the analysis. The clear, simple charts make certain relationships look like direct cause and effect when they're not, and they remove any uncertainty that's inherent in the analysis. The scientists are in a quandary: Finally, top decision makers are excited about their work, but what they're excited about isn't a good representation of it.

The team approach persisted through most of the century. In her 1969 book *Practical Charting Techniques*, Mary Eleanor Spear details the ideal team—a communicator, a graphic analyst, and a draftsman (still mostly men)—and its responsibilities. "It is advisable," Spear writes, "that [all three] collaborate."

In the 1970s things started to split. Scientists flocked to new technology that allowed them to visualize data in the same space (a computer program) where they manipulated

it. Visuals were crude but available fast and required no help from anyone else. A crack opened in the dataviz world between computer-driven visualization and the more classic design-driven visualization produced by draftspeople (finally).

Chart Wizard, Microsoft's innovation in Excel, introduced "click and viz" for the rest of us, fully cleaving the two worlds. Suddenly anyone could instantly create a chart along with overwrought variations on it that made bars three-dimensional or turned a pie into a doughnut. The profoundness of this shift can't be overstated. It helped make charts a lingua franca for business. It fueled the use of data in operations and eventually allowed data science to exist, because it overcame the low limit on how much data human designers can process into visual communication. Most crucially, it changed the structure of work. Designers—draftspeople—were devalued and eventually fell out of data analysis. Visualization became the job of those who managed data, most of whom were neither trained to visualize nor inclined to learn. The speed and convenience of pasting a Chart Wizard graphic into a presentation prevailed over slower, more resource-intensive, design-driven visuals, even if the latter were demonstrably more effective.

With the advent of data science, the expectations put on data scientists have remained the same—do the work and communicate it—even as the requisite skills have broadened to include coding, statistics, and algorithmic modeling. Indeed, in HBR's landmark 2012 [article](#) on data scientist as the sexiest job of the 21st century, the role is described in explicitly unicornish terms: "What abilities make a data scientist successful? Think of him or her as a hybrid of data hacker, analyst, communicator, and trusted adviser. The combination is extremely powerful—and rare."

Core Talents for Communicating Data

Here are the ways that various talents are involved as a data science project proceeds from gathering data to developing ...

A rare combination of skills for the most sought-after jobs means that many organizations will be unable to recruit the talent they need. They will have to look for another way to succeed. The best way is to change the skill set they expect data scientists to have and rebuild teams with a combination of talents.

Building a Better Data Science Operation

An effective data operation based on teamwork can borrow from Brinton and Spear but will account for the modern context, including the volume of data being processed, the automation of systems, and advances in visualization techniques. It will also account for a wide range of project types, from the reasonably simple reporting of standard analytics data (say, financial results) to the most sophisticated big data efforts that use cutting-edge machine learning algorithms.

Here are four steps to creating one:

1. Define talents, not team members.

It might seem natural that the first step toward dismantling unicorn thinking is to assign various people to the roles the “perfect” data scientist now fills: data manipulator, data analyst, designer, and communicator.

Not quite. Rather than assign people to roles, define the talents you need to be successful. A talent is not a person; it’s a skill that one or more people possess. One person may have several talents; three people may be able to handle five talents. It’s a subtle distinction but an important one for keeping teams nimble enough to configure and reconfigure during various stages of a project. (We’ll come back to this.)

Any company’s list of talents will vary, but a good core set includes these six:

Project management. Because your team is going to be agile and will shift according to the type of project and how far along it is, strong PM employing some scrumlike methodology will run under every facet of the operation. A good project manager will have great organizational abilities and strong diplomacy skills, helping to bridge cultural gaps by bringing disparate talents together at meetings and getting all team members to speak the same language.

Data wrangling. Skills that compose this talent include building systems; finding, cleaning, and structuring data; and creating and maintaining algorithms and other statistical engines. People with wrangling talent will look for opportunities to streamline operations—for example, by building repeatable processes for multiple projects and templates for solid, predictable visual output that will jump-start the information-design process.

Data analysis. The ability to set hypotheses and test them, find meaning in data, and apply that to a specific business context is crucial—and, surprisingly, not as well represented in many data science operations as one might think. Some organizations are heavy on wranglers and rely on them to do the analysis as well. But good data analysis is separate from coding and math. Often this talent emerges not from computer science but from the liberal arts. The software company Tableau ranked the infusion of liberal arts into data analysis as one of the [biggest trends](#) in analytics in 2018. Critical thinking, context setting, and other aspects of learning in the humanities also happen to be core skills for analysis, data or otherwise. In an online lecture about the topic, the Tableau research scientist Michael Correll explained why he thinks infusing data science with

liberal arts is crucial. “It’s impossible to consider data divorced from people,” he says. “Liberal arts is good at helping us step in and see context. It makes people visible in a way they maybe aren’t in the technology.”

Subject expertise. It’s time to retire the trope that data science teams are stuck in the basement to do their arcane work and surface only when the business needs something from them. Data science shouldn’t be thought of as a service unit; it should have management talent on the team. People with knowledge of the business and the strategy will inform project design and data analysis and keep the team focused on business outcomes, not just on building the best statistical models. Joaquin Candela, who runs applied machine learning at Facebook, has [worked hard](#) to focus his team on business outcomes and to reward decisions that favor those outcomes over improving data science.

Design. This talent is widely misunderstood. Good design isn’t just choosing colors and fonts or coming up with an aesthetic for charts. That’s styling—part of design, but by no means the most important part. Rather, people with design talent develop and execute systems for effective visual communication. In our context, they understand how to create and edit visuals to focus an audience and distill ideas. Information-design talent—which emphasizes understanding and manipulating data visualization—is ideal for a data science team.

Storytelling. Narrative is an extremely powerful human contrivance and one of the most underutilized in data science. The ability to present data insights as a story will, more than anything else, help close the communication gap between algorithms and executives. “Storytelling with data,” a tired buzz phrase, is widely misunderstood, though. It is decidedly not about turning presenters into Stephen Kings or Tom Clancys. Rather, it’s about understanding the structure and mechanics of narrative and applying them to dataviz and presentations.

2. Hire to create a portfolio of necessary talents.

Once you’ve identified the talents you need, free your recruiting from the idea that these are roles you should hire people to fill. Instead focus on making sure these talents are available on the team. Some of them naturally tend to go together: Design and storytelling, for example, or data wrangling and data analysis, may exist in one person.

Sometimes the talent will be found not in employees but in contractors. For my work, I keep a kitchen cabinet of people who have talents in areas where I’m weak. You may want to engage an information-design firm, or contract with some data wranglers to clean and structure new data streams.

Thinking of talents as separate from people will help companies address the last-mile problem, because it will free them from trying to find the person who can both do data science and communicate it. Nabbing some people who have superior design skills will free data scientists to focus on their strengths. It will also open the door to people who might previously have been ignored. An average coder who also has good design skills, for example, might be very useful.

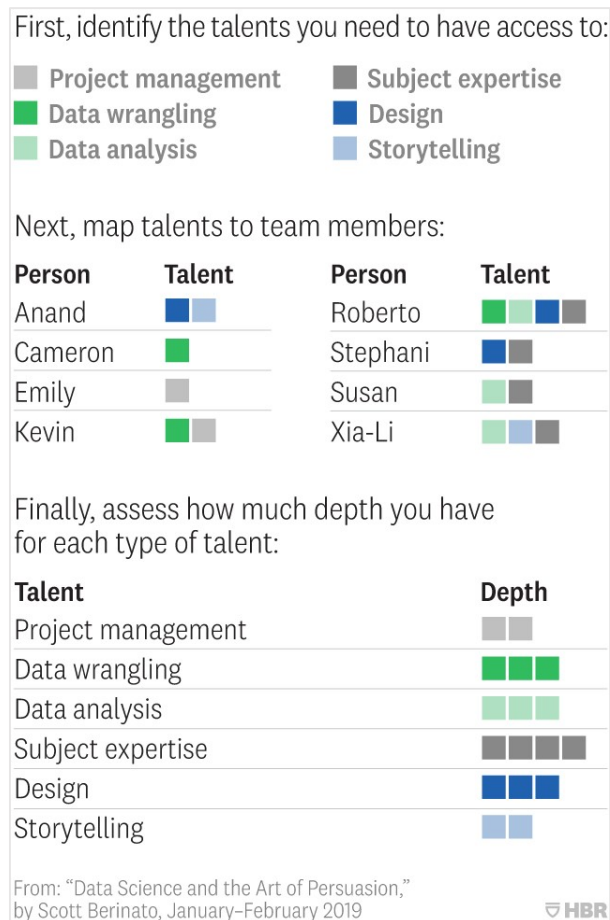
Randal Olson, the lead data scientist at Life Epigenetics and curator of the Reddit channel Data Is Beautiful (devoted to sharing and discussing good dataviz), used to focus solely on how well someone did the technical part of data science. “I know, when I started, I had zero appreciation for the communication part of it,” he says. “I think that’s common.” Now, in some cases, he has changed the hiring process. “You know, they come in and we immediately start white-boarding models and math,” he says. “It’s data scientists talking to data scientists. Now I will sometimes bring in a nontechnical person and say to the candidate, ‘Explain this model to this person.’”

3. Expose team members to talents they don’t have.

Overcoming culture clashes begins with understanding others’ experiences. Design talent often has no exposure to statistics or algorithms. Its focus is on aesthetic refinement, simplicity, clarity, and narrative. The depth and complexity of data work is hard for designers to reconcile. Hard-core data scientists, in contrast, value objectivity, statistical rigor, and comprehensiveness; the communication part is not only foreign to them but distracting. “It goes against their ethos,” says one manager of a data science operation at a large tech company. “I was the same way, working in data science for 10 years, but it was eye-opening for me when I had to build a team. I saw that if we just learned a little more about the communication part of it, we could champion so much more for the business.”

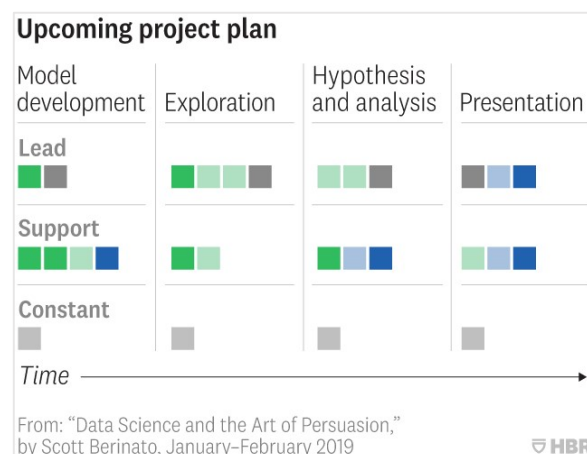
Build a Talent Dashboard

Performing a talent audit helps managers do a better job of planning for projects and configuring teams.



Put It to Use

Knowing what talents are available, managers can now assign units of talent to a project according to when it's needed. Typically, one cluster of talents will take the lead early in a project, and different clusters will do so in the later stages. Project management usually plays a big role throughout.



There are many ways to expose team members to the value of others' talents. Designers should learn some basic statistics—take an introductory course, for example—while data scientists learn basic design principles. Neither must become experts in their counterparts' field—they just need to learn enough to appreciate each other.

Stand-ups and other meetings should always include a mix of talents. A scrum stand-up geared mostly to updating on tech progress can still include a marketer who makes presentations, as happens at Olson's company. Subject-matter experts should bring data wrangling and analysis talent to strategy meetings. Special sessions at which stakeholders answer questions from the data team and vice versa also help to bridge the gap. The chief algorithms officer at Stitch Fix, Eric Colson (who is something close to a unicorn, having both statistical and communication talents at a company where data science is intrinsic), asks his team members to make one-minute presentations to nontechnical audiences, forcing them to frame problems in smart ways that everyone can understand. "To this day," Colson says, "if you say 'coconuts' here, people will know that was part of a metaphor one person used to describe a particular statistical problem he was tackling. We focus on framing it in ways everyone understands because the business won't do what it doesn't understand." Another manager of a data science team created a glossary of terms used by technical talent and design talent to help employees become familiar with one another's language.

If your organization contains some of those rare people who, like Colson, have both data talents and communication and design talents, it helps to have them mentor one another. People who express interest in developing talents that they don't have but that you need should be encouraged, even if those strengths (design skills, say) are far afield from the ones they already have (data wrangling). Indeed, in my workshops I hear from data scientists who would love to develop their design or storytelling talent but don't have time to commit to it. Others would love to see that talent added to their teams, but their project management focuses primarily on technical outcomes, not business ones.

All this exposure is meant to create empathy among team members with differing talents. Empathy in turn creates trust, a necessary basis for effective teamwork. Colson recalls a time he used storytelling talent to help explain something coming out of data analysis: "I remember doing a presentation on a merchandising problem, where I thought we were approaching it the wrong way. I had to get merchandising to buy in." Instead of explaining beta-binomial distribution and other statistical concepts to bolster his point of view, he told a story about someone pulling balls from an urn and what happened over time to the number and type of balls in the urn. "People loved it," he says. "You watched the room and how it clicked with them and gave them confidence so that at that point the math behind it wasn't even necessary to explain. They trusted us."

4. Structure projects around talents.

With a portfolio of talents in place, it's time to use it to accomplish your goals. The shifting nature of what talents are needed and when can make projects unwieldy. Strong project management skills and experience in agile methodologies will help in planning the configuration and reconfiguration of talents, marshaling resources as needed, and keeping schedules from overwhelming any part of the process.

Putting It All Together

You'll want to take other steps to make your projects successful:

Assign a single, empowered stakeholder.

It's possible, or even likely, that not all the people whose talents you need will report to the data science team manager. Design talent may report to marketing; subject-matter experts may be executives reporting to the CEO. Nevertheless, it's important to give the team as much decision-making power as possible. Stakeholders will most often be people with business expertise who are closely connected to or responsible for business goals; the aim of the work, after all, is better business outcomes. Those people can create shared goals and incentives for the team. Ideally you can avoid the responsibility-without-authority trap, in which the team is dealing with several stakeholders who may not all be aligned.

Data scientists find communicating their work not only foreign but distracting.

Assign leading talent and support talent.

Who leads and who supports will depend on what kind of project it is and what phase it's in. For example, in a deeply exploratory project, in which large volumes of data are being processed and visualized just to find patterns, data wrangling and analysis take the lead, with support from subject expertise; design talent may not participate at all, since no external communication is required. Conversely, to prepare a report for the board on evidence for a recommended strategy adjustment, storytelling and design lead with support from data talent.

Colocate.

Have all team members work in the same physical space during a project. Also set up a shared virtual space for communication and collaboration. It would be undesirable to have those with design and storytelling talent using a Slack channel while the tech team is using GitHub and the business experts are collaborating over e-mail. Use "paired analysis" techniques, whereby team members literally sit next to each other and work on one screen in a scrumlike iterative process. They may be people with data wrangling and analysis talent refining data models and testing hypotheses, or a pair with both subject expertise and storytelling ability who are working together to polish a presentation, calling in design when they have to adapt a chart.

Make it a real team.

The crucial conceit in colocation is that it's one empowered team. At Stitch Fix "our rule is no handoffs," Colson says. "We don't want to have to coordinate three people across departments." To this end he has made it a priority to ensure that his teams have all the skills they need to accomplish their goals with limited external support. He also tries to hire people many would consider generalists who cross the tech-communication gap. He augments this model with regular feedback for, say, a data person who needs

help with storytelling, or a subject expert who needs to understand some statistical principle.

Reuse and template.

Colson also created an “algo UI team.” Think of this as a group of people who combine their design talents and data wrangling talents to create reusable code sets for producing good dataviz for the project teams. Such templates are invaluable for getting a team operating efficiently. Conversations that an information designer, say, would have with a data analyst about best practices in visualization become hard-coded in the tools. Graham MacDonald, the chief data scientist at the Urban Institute, has successfully fostered this kind of cooperation on templating. His group produces data by county for many U.S. counties. By getting data wrangling and subject expertise together to understand the communication needs, the group built a reusable template that could customize the output for any particular county. Such an outcome would have been difficult without the integration of those talents on the team.

CONCLUSION

The presentation of data science to lay audiences—the last mile—hasn’t evolved as rapidly or as fully as the science’s technical part. It must catch up, and that means rethinking how data science teams are put together, how they’re managed, and who’s involved at every point in the process, from the first data stream to the final chart shown to the board. Until companies can successfully traverse that last mile, data science teams will underdeliver. They will provide, in Willard Brinton’s words, foundations without cathedrals.