



Democratizing Data Science in Your Organization

The Need for Business Insights

Data science has made what was previously impossible possible. The government of Rwanda studies anomalies in revenue-collection data to find tax cheats. A sociologist in Indiana is using data science to find patterns in the massive amounts of text people use each day to express their worldviews—patterns that no individual reader would be able to recognize.

Yet despite the explosion of data collected in recent years, many organizations—from financial institutions and health care firms to management consultancies and governments—are simply not equipped to learn from their data in an efficient and effective manner.

Data-driven organizations are three times more likely to report significant improvements in decision making ([PwC Global Data and Analytics Survey](#)), but less than 0.5% of data is analyzed ([MIT Tech Review](#)). The challenge is to rapidly scale a company's ability to work with data to drive better decisions.

Many companies have a centralized data science team that prioritizes and works on information requests from other departments. In this model, questions come in and answers are expected to come out. The centralized data science team is the gatekeeper of data, creating a silo of data tools that aren't easily accessible.



« *The challenge is to rapidly scale a company's ability to work with data to drive better decisions.* »

Problems with the Centralized Data Model

Negative Effects on the Data Science Team

When the data science team must always react to the demands of other teams, they cannot be forward-thinking. Much of their time is spent on routine reporting rather than innovative analyses. Other teams control their time since they are a support center to the rest of the organization. And once the data team fulfills a request, they face the challenge of communicating the results to outside stakeholders, as other groups may not speak the same language of data.

The hurdle of translating their findings and creating actionable insights is compounded by the data tools they use. Data scientists typically use tools such as SQL, Python, and R, while other teams may only be familiar with Excel. Conveying data results outside of data scientists' preferred toolkit is frustrating, time-consuming, and often done poorly. This may cause a lack of follow-through on the analyses provided by the data science team. All of that work with little to no application—really, what's the point?

Negative Effects on the Business Stakeholders

The business stakeholders requesting data from the centralized data science team aren't happy in this scenario either. There is often a tedious—and sometimes political—process in place for prioritizing data requests. This is because the data science organization drives the roadmap on how they provide value to the organization. Data science leaders **may guide priorities** to protect their teams from what they might view as egregious requests. And data scientists on the receiving end of a request may not understand the full context of what

they're trying to solve, since they aren't as familiar with the same daily business needs as the person making the request. Product owners, for example, understand every step in the thought process leading to a product decision, but the data team will not have that level of understanding. Making data accessible, relevant, and actionable may require multiple conversations and iterations. Often, this is a complicated and drawn-out process. Worse, the original questions may never be sufficiently addressed.

« When the data science team must always react to the demands of other teams, they cannot be forward-thinking. Much of their time is spent on routine reporting rather than innovative analyses. »

Negative Effects on the Organization

Of course, none of this is good for the organization as a whole. On top of complicated ad hoc requests, the data produced exists in silos. Employees are effectively blind to the proliferation of inter-departmental data that may be relevant to their role. While the business may value transparency, it isn't able to share information and resources across teams, which hinders the company's ability to learn quickly, iterate, and grow.

When data science is centralized, it also limits the company's ability to solve problems independently. Functional teams can take the easy way out and push the responsibility of critical thinking to the data team. Because the data science team is seen as the owner of data, the rest of the organization can claim less responsibility in understanding the implications of data and how to use it.

How Companies Got Stuck in this Data Rut

There are multiple reasons companies choose to centralize their data science team. Perhaps managers want to limit employee access to data because they are afraid that non-technical employees won't have the necessary skills to interpret data and apply it correctly.

Perhaps there are strict data governance policies in place and data privileges are only granted to a select group of executives, data scientists, and IT. These reasons presume that an organization can become data-fluent when its employee base is not. This is a false assumption.

You might ask...

« Why would anyone besides professional writers need to know how to write? »

« I'm not the office manager, I don't have to clean up my desk! »

« Why would non-data scientists need to learn data science? »

« Why would people who aren't in HR need management skills? »

« We're not accountants, we can't be expected to keep track of our budget. »

Data Science is Not Just for Data Scientists

Data skills are necessary to solve problems, and are just as crucial to a professional skill set as communication, teamwork, time management, and interpersonal skills. You might ask why non-data scientists would need to learn data science. That's like asking non-accountants why they need to know how to manage a budget. Every industry is swimming in data, and the organizations that succeed are those that can quickly make sense of their data and adapt to what's coming.

Data science expertise must be dispersed across an organization to enable deeper insights. Alternatives to the centralized model for data science are the embedded model and the hybrid model.

In the embedded model, data scientists are dispersed across an organization and report to the functional leaders. The embedded model maximizes the utility of the data scientists, but sometimes at the expense of their personal growth, since their leaders may not have the data expertise to help them develop more fully.

At DataCamp, we have a hybrid model where data scientists sit with functional teams and also report to the Chief Data Scientist. This model ensures that data scientists have ownership over the decisions that are being made in respect to their work, and that data science is an organizational priority.

The Hybrid Data Science Model



FINANCE



PRODUCT



DATA SCIENCE



MARKETING



ENGINEERING

Companies that believe in the democratization of data science, like Airbnb, care about empowering every employee to make data-driven decisions. A workforce of “citizen data scientists” ensures that decisions are grounded in data and that everyone is able to make decisions autonomously. This is important because the person asking the question always has the best context on the question they are trying to answer.

Companies that want to compete in the age of data need to do three things: share tools, share skills, and share responsibility.

Share Tools

In many organizations, the majority of data tools sit with the data science team. While this may seem logical, creating a silo of data tools and restricting access to a narrow group of employees places too heavy a burden on a single team. Most inquiries from other departments are relatively simple requests that anyone with basic training could fulfill. By saddling data scientists with basic gatekeeping tasks, organizations divert their attention from the larger projects that require their deep expertise.

The Learning and Development leadership team at Airbnb shares tools to continuously train its workforce through their proprietary Data University. Airbnb also encourages collaboration by allowing anyone to post an article to a knowledge repository, where the rest of the company can view analyses in a news feed.



This is where new problems that have been solved can be easily found, and anyone with further questions will know who to reach out to for more information. In addition to sharing learnings across the company, such articles give recognition to the people who post them—which incentivizes others to do the same.

At DataCamp, we've set up our environment in R and consolidated data sources under a cloud-based data warehouse product. Our data scientists have built several dashboards that every employee can access at any time to view the company's performance across key metrics. We've also built custom dashboards so each functional team can stay on top of the metrics that matter most to them.

"Companies that want to compete in the age of data need to do three things: share tools, share skills, and share responsibility."

Share Skills

As companies share data tools, they must also enable their workforce to use those tools. Not every company can create its own Data University, like Airbnb. Depending on the data tools your organization uses, a variety of educational programs, both online and in person, can get your team up to speed.

At DataCamp, we work with the best instructors and industry experts to create data science and analytics courses for every skill level. Unlike other platforms, DataCamp features a modern learning experience with bite-sized videos and hands-on coding exercises, so employees **learn by doing** and stay engaged. Our mobile app makes it easy to hone skills on-the-go with **short practice sessions** that reinforce what they've learned. And with DataCamp projects, they can tackle real-world problems in a risk-free environment and **apply** their new data skills right away.

As your team becomes more comfortable with the language of data, they'll be more comfortable bringing data to bear on important business decisions. It will become clear that some team members are more comfortable using data skills than others are. Encourage the proficient ones to mentor the others. Even at DataCamp, where data science is our business, some people don't work with data continuously. When they need help on a complex problem, they pair up with those who do.

A data-fluent team also makes better requests. Even a basic understanding of tools and resources greatly improves the quality of interaction among colleagues. When the amount of back-and-forth needed to understand and fulfill each request goes down, speed and quality go up.

Shared skills improve workplace culture and results in another way, too: They improve mutual understanding. If you know how hard it will be to get a particular data output, you'll adjust the way you interact with the people in charge of giving you that output. Such adjustments improve the workplace for everyone.

1



Learn

DataCamp features bite-sized videos and hands-on practice exercises that empower employees to learn by doing.

2



Practice

Quick coding challenges reinforce new concepts and motivate learners to stay engaged from their desktop or the DataCamp mobile app.

3



Apply

With DataCamp projects, users can tackle real-world problems in a risk-free environment so they can start applying their new data skills right away.

There are no shortcuts to writing code, but with practice, anyone can grow the skills needed to solve problems using data. These skills are useful across departments and seniority levels. An added bonus is that programming languages are open source, so people can take coding skills with them wherever they go, making data science and analytics skills useful for career mobility as well.

"If your workforce is data-fluent, your centralized data science team can shift from doing everyone else's data work to building the tools that enable everyone to do their data work faster."

Share Responsibility

Once an organization is delivering the access and education needed to democratize data among its employees, it may be time to adjust roles and responsibilities. At a minimum, teams should be able to access and understand the datasets most relevant to their own functions. But by equipping more team members with basic coding skills, organizations can also expect non-data science teams to apply this knowledge to departmental problem solving—and greatly improve outcomes.

If your workforce is data-fluent, your centralized data science team can shift from doing everyone else's data work to building the tools that enable everyone to do their data work faster. At DataCamp, our own data science team doesn't run analyses every day. Instead, it builds new tools that everyone can use so that 50 projects can move forward as quickly as one project moved before.

Data science isn't just for data scientists anymore, if it ever was. Smart companies today ensure that many of their employees can speak the language of data and use it to improve work outcomes. By empowering employees with these foundational skills, companies are realizing unprecedented levels of innovation and efficiency.

Let's take a look at some real-world examples of how different departments can share data responsibility.



Predictive Analytics for Finance

Finance teams are tasked with forecasting the future with a certain amount of accuracy. They may want a simple way to create custom graphs and analyses beyond the functional capabilities of Excel. Open-source coding languages like R and Python have data exploration and visualization tools to identify seasonal trends. By removing and discounting these trends, finance teams can better see how their sales team are performing, and improve pricing and discounting policies accordingly.

Applications of Marketing Analytics

Data science and analytics can help marketers gain insights to create successful campaigns. Marketers may perform a cluster analysis to identify natural groupings of users when analyzing a campaign to improve audience segmentation. They may use SQL and R in A/B testing to compare the performance of test and control groups for a campaign. Or predict the ROI of a planned campaign using benchmarks and multivariate regression to predict metrics such as potential reach, estimated click-through rate, and approximate cost-per-acquisition. Marketers can also improve their lead scoring and campaign performance by using code instead of Excel, which isn't designed to perform complex and reproducible analysis.

Data-Driven HR Insights

HR analytics is a growing field that can lead to valuable company insights into recruiting, hiring, and employee retention. Data can help determine the makeup of an organization's talent, what skills are represented, and what skill gaps may exist. Using analytics to identify talent gaps can help a company determine the attributes needed to compose a successful team. Talent acquisition teams can focus on hiring for diversity to position the company for success. And to maximize employee engagement, companies can measure key indicators in employee pulse surveys to determine where to focus efforts.

Shortcuts for Sales Reporting

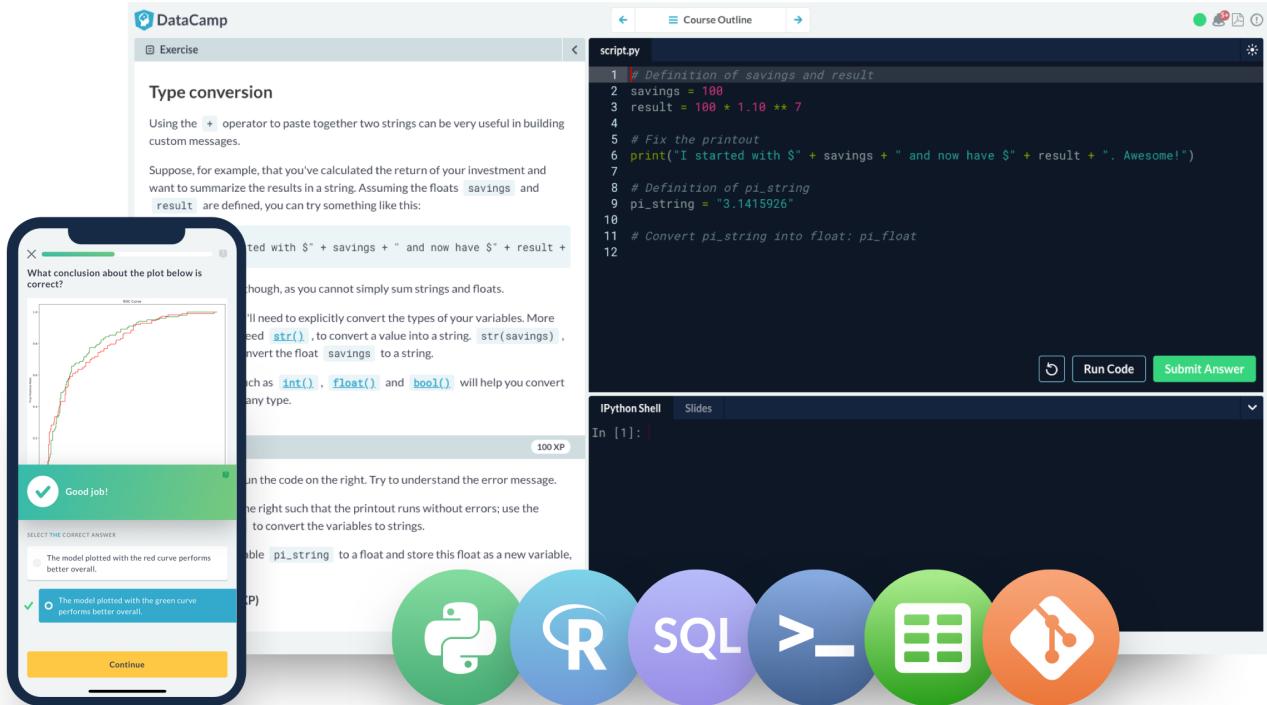
A common application of data science for sales teams is streamlining reporting for specific partners, accounts, or products. Instead of exporting results and filtering by a particular name on a spreadsheet, a simple script (a set of instructions in code) can automate the process. Data science allows sales executives to go beyond spreadsheet use to improve accessibility, efficiency, and collaboration. For more advanced users, data science tools for sales analytics can help teams build dashboards to forecast future performance, pull geographic insights, set better targets, and improve performance.

Data Science for Executives

Understanding and wielding the power of data science can help executives in a few key ways. Data can describe a company's current state, detect anomalous events, or diagnose the causes of observed events and behaviors. It can predict future events and help companies prepare for them. And of course, executives must understand how to structure their data team to best meet their organization's needs. For more examples, [McKinsey's Executive's Guide to AI](#) is a great resource.

Upskilling Your Workforce

Begin the journey of democratizing data science in your organization with [DataCamp for Business](#).



The image illustrates the DataCamp platform's integrated learning environment. On the left, a smartphone screen shows a mobile exercise titled "Type conversion" where learners can select the correct answer regarding a plot. On the right, a desktop browser window displays a Python script editor with code for type conversion and an iPython Shell. Below the browser are five circular icons representing Python, R, SQL, and other data science tools.

Our scalable learning platform provides easy onboarding and management for teams and organizations of any size, and LMS integration options. And it's easy to measure, with a dedicated dashboard to track everyone's engagement and progress over time. We are constantly expanding our curriculum to keep up with the latest technology trends and to provide the best learning experience for all skill levels.

And learners stay engaged through hands-on learning, which means our course completion rates significantly exceed those of traditional online courses. We have more than 3.7 million learners around the world at companies such as Nielsen and REI—and we're just getting started. Close the skills gap. Visit [datacamp.com](#).