

RESEARCH REPORT

The Business Value of Data Engineering



Introduction

Today's mandate is clear: become data-driven or fall behind.

Across industries, organizations must harness the immense potential of data to drive decision-making and innovation or lose ground to their data-savvy competitors. For the last decade-plus, most enterprises have sunk significant time and resources into digital transformation initiatives chasing this goal.

Infrastructure has been moved to the cloud, processes have become agile, and organizational structures, roles, and goals have changed. Where data leaders once focused on protecting data at all costs, they now face increasing pressure to drive tangible value from data.

To fully realize this mission, data leaders must enlist and empower their data engineering teams. Over the past decade, StreamSets has seen first-hand the incredible impact data engineers have on deriving value from data quickly and at scale. But it's also clear that plenty of data engineers out there — and their business counterparts — don't yet understand the full potential of their work.

As a catch-all data role in many organizations, data engineers are used to rolling up their sleeves and getting stuff done. They build the underlying infrastructure and pipelines to deliver data to downstream users smoothly and lay the groundwork for analytics. Though not flashy, their work enabling data science and analysis is crucial. And **the more data engineers understand, the better business outcomes are.**

To start the conversation, we decided to get data straight from the source. We surveyed 523 data engineers and 778 business data consumers worldwide to get a baseline view of how data engineers and business people see data engineering's impact on business results. In the process, we uncovered several insights that can help organizations set their data engineering teams up for success in a value-driven world.

What's Inside?

Meet the surveyees

Perception matters

Understanding the data problem: business context and impact

The state of collaboration and communication

You can't improve what you don't measure

Learnings from completely satisfied data engineers

Meet our Surveyees

It's always nice to know who's on the other end of a survey. For this study, we surveyed two audiences: data engineers (DEs) and the line of business (LOB) data consumers who work with them. Data consumers could include anyone from sales and marketing operations to business and finance analysts to Fred in HR, who has to migrate an old system to the cloud.

In a nutshell, most data consumers and data engineers work at enterprise companies in managerial, senior professional, or individual contributor roles. In other words, these are the people doing the day-to-day work who are in a position to know the answers to these questions.

They come from a wide variety of industries and, on the data consumer side, various departments across the enterprise. Though most data engineers report to IT, Data and Analytics, or Engineering, 64% are aligned with a business unit as well.

The data engineers who responded were an experienced bunch, with the most common tenure reported at 6–10 years. Finally, most survey respondents were from the US, UK, Canada, Australia, and the EU.



Business Data Consumers

COMPANY SIZE

80% Over 2500 employees

JOB LEVEL

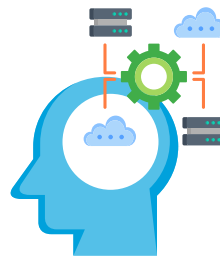
74% Manager/ individual contributor

TOP 3 DEPARTMENTS

31% Operations

18% Customer Success

15% Finance / Accounting



Data Engineers

COMPANY SIZE

68% Over 2500 employees

JOB LEVEL

70% Manager / individual contributor

TOP 3 DEPARTMENTS

49% IT

33% Data and Analytics

12% Engineering

Perception Matters

As we all know, data engineers are often behind the scenes, architecting data systems, building pipelines, and keeping the data “plumbing” running. It’s critical work, but it doesn’t usually get a lot of internal publicity.

While it’s a good sign that most data consumer respondents think data engineering is important to their analytics and decision-making, only 32% believe it’s critical. That’s unfortunate. Those who do say it’s critical seem to have good strategic partnerships that lead to better data. They’re statistically *much* more likely to:

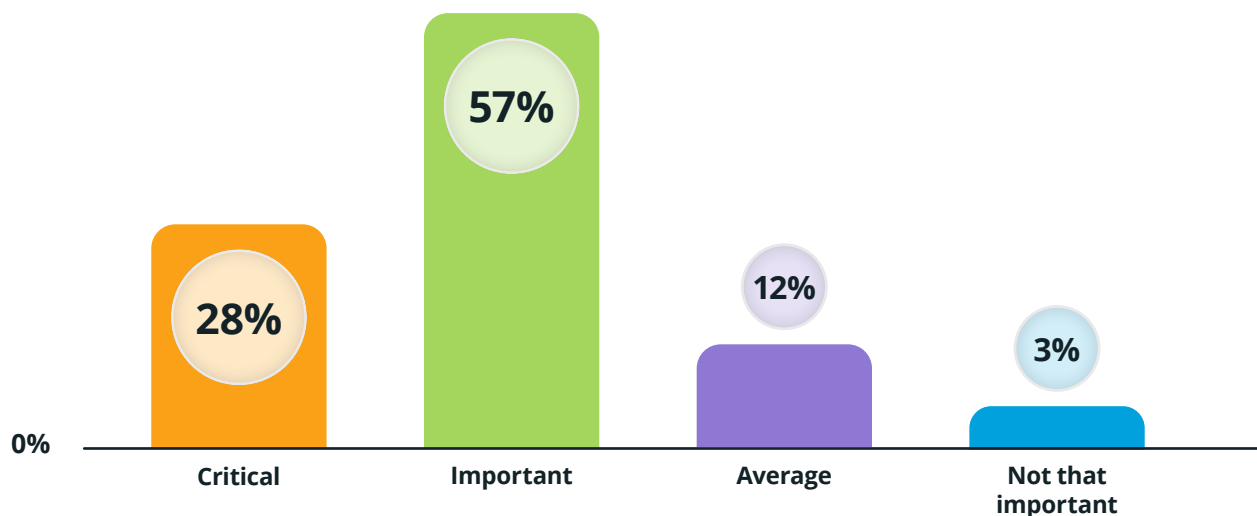
- Say they get their data within 48 hours
- Rate the quality of the data they get as excellent
- Think their data engineers are problem solvers, critical thinkers, clear communicators, customer service-oriented, and good listeners
- Claim they always provide business context and business impact (we’ll get into this in the next section)

These findings present a chicken-and-egg conundrum. Are these things happening because data consumers believe data engineering is critical? Or are their data engineers rockstars who drive these kinds of results and that leads to an understanding of just how essential data engineering really is? It could be a bit of both. But, either way, results are better when the vital role data engineering plays is recognized.

RECOMMENDATIONS:

- Raise the profile of your data engineering team.
- Help your business users understand how they’ll benefit from collaborating more closely with data engineering.

Q: How important a role does data engineering play in your analytics and decision-making?



Understanding the Data Problem: Business Context and Impact

Every data engineering challenge stems from a business problem. Let's say a SaaS company recently launched a new feature, and a data analyst at the company wants to gauge its adoption and impact.

When the analyst puts in a request to data engineering, it simply asks how to query all user activity for X feature from the system logs. And a data engineer could certainly tell them how to do exactly that. But what could the DE do if they understood that the goal was to see if this new feature was driving more engagement with the platform?

Focusing on available datasets, the data engineer could steer the analyst toward a multi-dimensional approach to evaluate the feature's impact. For example, they might suggest looking at API call data and cross-referencing with client support tickets to see if there are usage challenges. By doing so, the analyst can gather a more nuanced understanding of not just feature adoption but also its broader implications, challenges, and the value it brings to clients.

At least theoretically, most data engineers (73%) and their data consumer counterparts (63%) agree with this. Yet, when it comes to aligning actions with their words, the survey tells a different story. Most data consumers admit they don't always provide business context — and a paltry 17% of data engineers report getting that information all the time.

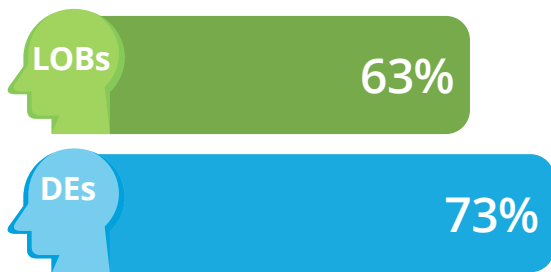


Knowing the bigger picture is essential for data engineers to pin down requirements and design optimized data solutions accordingly." ¹

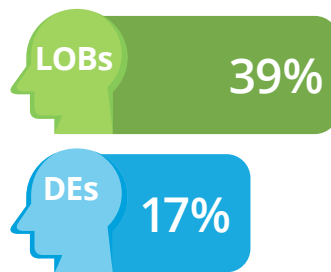
Xinran Waibel
Data Engineer, Netflix

¹ [The Question that Every Data Engineer Should Ask](#), Data Engineering Things, Xinran Waibel

Say they consider it **VERY IMPORTANT** for a DE to understand business context



ALWAYS provide/get business context



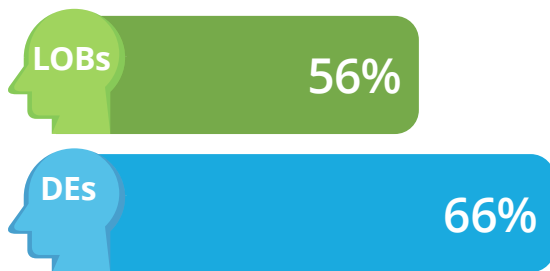
■ LOB Data Consumers ■ Data Engineers

The same is true of business impact, which is the desired result of the data requested (e.g., improving time to market, making the customer experience better, or increasing revenue). Understanding the desired business impact may seem like a bonus, but we'd argue that tying it to strategic corporate goals is essential. It shows data consumers just how important the data engineer is to their success and can boost data engineer morale, too.

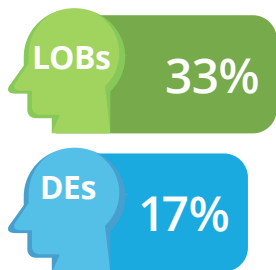
The numbers speak volumes; when data engineers always receive the business context with a data

request, they are significantly more likely to be completely satisfied with their career. Give them both data context and data impact, and they're much more likely to believe their organization leverages data engineers to their fullest potential. They're also more likely to report that their work impacts cost reduction, increasing productivity, identifying new business opportunities, better serving customers, enabling data-driven decision-making, and improving sales and marketing

Say they consider it **VERY IMPORTANT** for a DE to understand business impact



ALWAYS provide business impact



■ LOB Data Consumers ■ Data Engineers



Treating data engineers as a core part of the strategic mission rather than back-end service providers gives this critical role the respect it deserves and leads to better results.

RECOMMENDATIONS:

- Emphasize the importance of business context and impact in developing data solutions.
- Put processes in place to ensure data engineers can gather this information, either with the initial request or through collaborative work.
- Ensure both sides understand what question(s) the data will answer. When LOB communicates those questions to DE, the DE can potentially integrate additional data sets or perform data transformations for a richer data set, thereby providing more accurate answers — or even help uncover deeper questions that need to be answered.
- Align data engineers with business units or consider a decentralized data model like a data mesh, with some data engineers embedded in business units as data product owners.

The State of Collaboration and Communication

To get to that all-important business context and impact, collaboration and top-notch communication processes are essential. When data engineers have a clear understanding of the big picture, including desired outcomes, they can actively consult on how to optimize data utility. But do they?

How Data Engineers and Business Data Consumers Work Together

Currently, it's a mixed bag, and there are discrepancies between how data engineers say they work with data consumers and vice versa. We asked whether data engineers are more likely to take a request and do precisely what it tells them (order-taker) or get an understanding of the desired outcome and make strategic recommendations (strategic partner).

Both sides say they work both ways for the most part, but data engineers reported they more often function as strategic partners. In contrast, data consumers said the data engineers they work with lean toward the order-taking side.

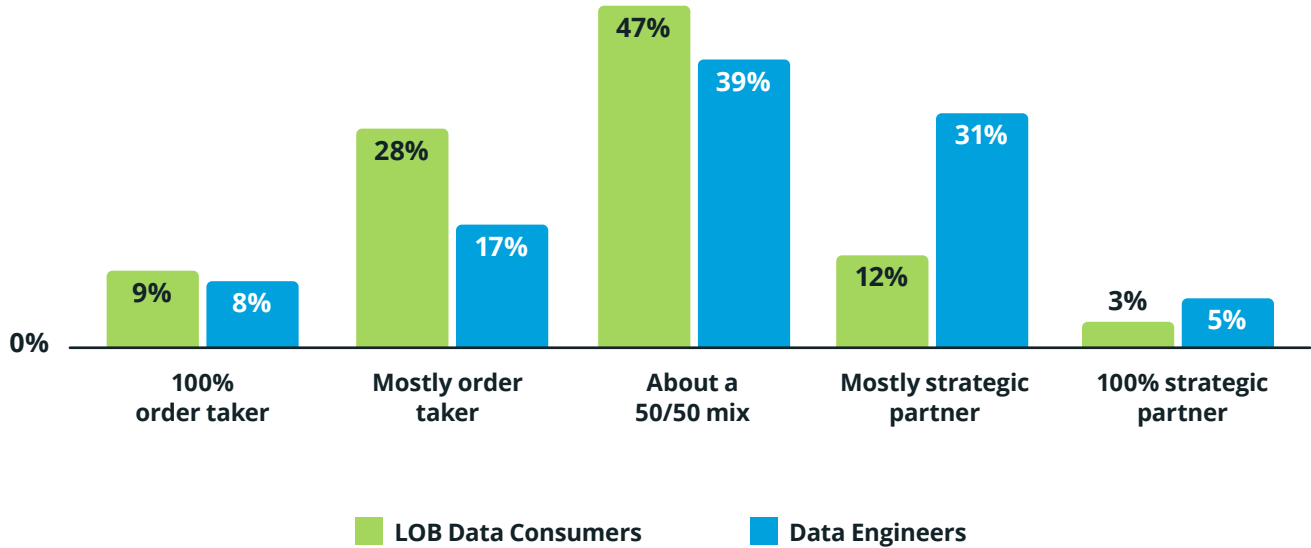
When asked how they'd prefer to work, it becomes clear that data engineers, on the whole, are looking to collaborate. Over half (55%) want to be strategic partners, period, while 30% are open to a mix of work.

Data consumers aren't entirely on board with a full partnership model yet, though, at only 32%. The majority prefer a mixed model where they can put in orders or get a consult when they feel they need one.

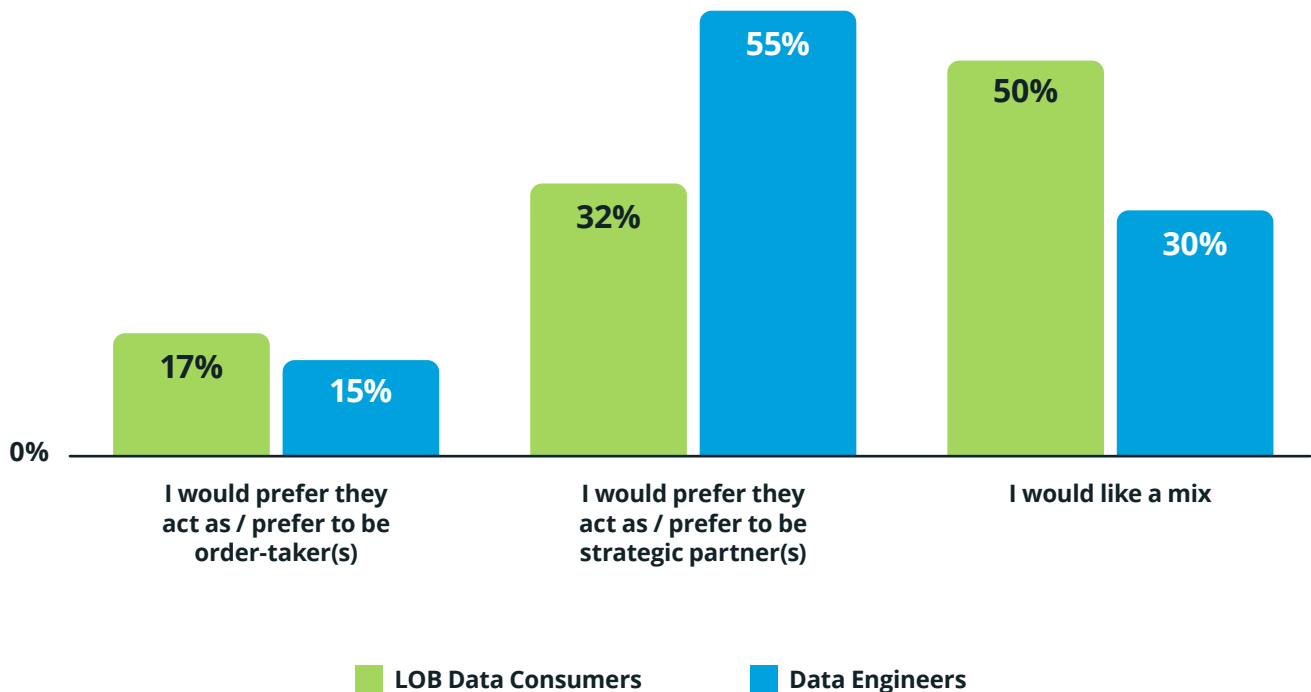
RECOMMENDATIONS:

- Educate lines of business and data engineering on the importance of business context and impact.
- Teach data consumers what kind of information to provide and data engineers what kinds of questions to ask to get this information.

How LOB data consumers say they **CURRENTLY WORK** with data engineers vs. how DEs say they work with LOBs



How LOB would **PREFER** to work with data engineers and vice versa



How Data Engineers Get Data Requests

Data engineers typically get data requests either through a standard tool or form, on an ad hoc basis (by email, Slack or Teams, phone, or in-person), or through a combination of the two. The results were pretty evenly distributed across all three options, though the combo method had the majority with both data consumers and data engineers.

That means that almost a third (31%) of data consumers don't have a standard form or tool to submit data requests, and over a third (38%) are submitting requests on an ad hoc basis at least some of the time. This lack of consistency is a breeding ground for insufficient and inconsistent information.

Data engineers reported that they were more likely to get enough information when requesters used a standard form or tool. And when they received enough information, the survey showed that:

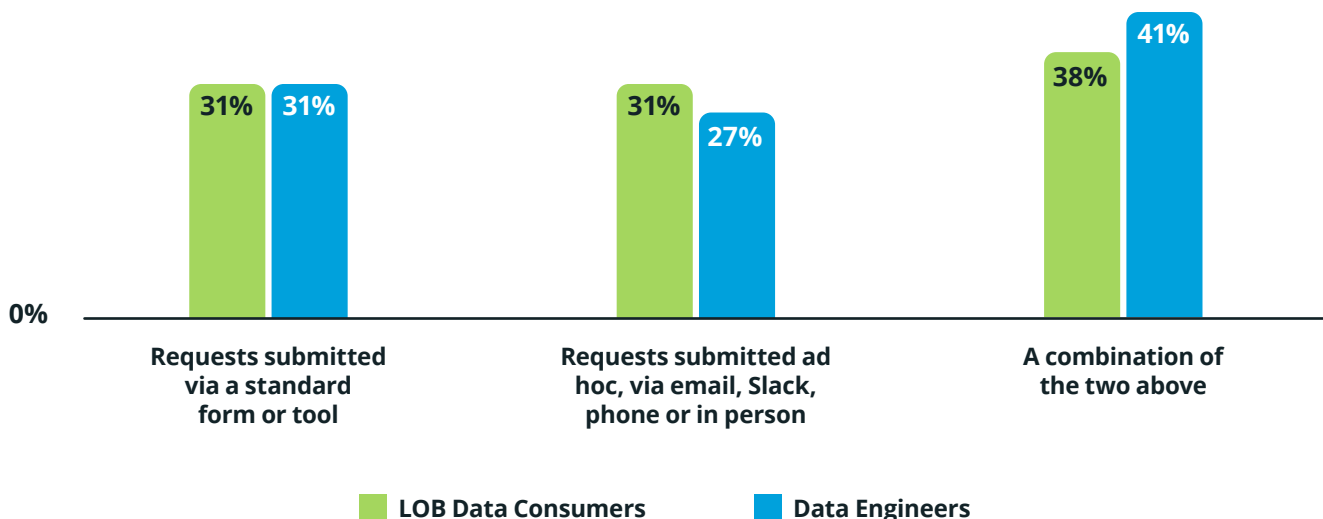
- Information was more likely to be consistent
- They were *much* more likely to get business context and business impact
- They were more likely to track external strategic goals (we'll get to this shortly)

But there's another intriguing impact this single step has on data engineers. According to the survey, data engineers who always get enough information with data requests are also significantly more likely to believe that their organization leverages them to their fullest potential. Additionally, when they receive business context with their requests, they're more likely to be completely satisfied with their career choice.

RECOMMENDATIONS:

- Align data engineers with business teams. This can lead to better collaboration.
- Implement a standardized way for data consumers to make data requests; this will help data engineers get enough and consistent information.

Q: How do you make / receive data requests?



You Can't Improve What You Don't Measure

Since the goal is to get data engineers focused on delivering business value, we wondered what type of measurements they're using to track success. We gave examples of each type to keep everyone on the same page.

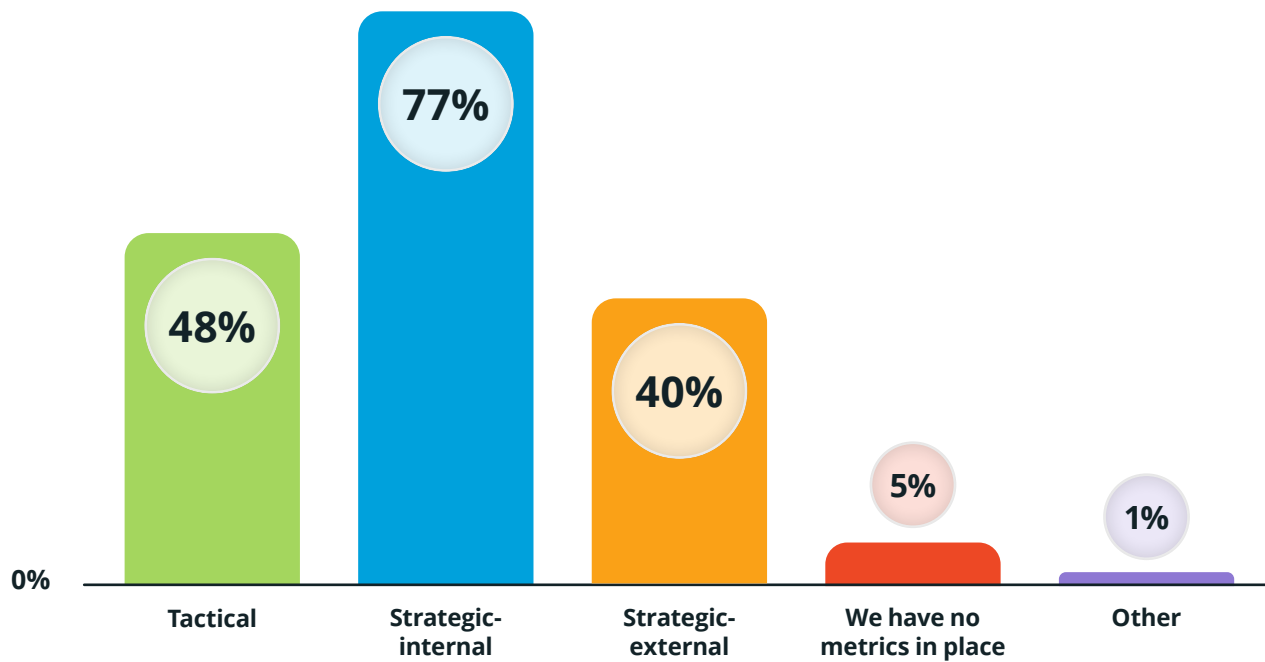
- **Tactical:** metrics like number of requests/tickets closed and time to close requests/tickets.
- **Strategic internal:** impact on organizational goals like increasing productivity and operational efficiency.
- **Strategic external:** impact on business outcomes like getting to market faster.

While it's a small percentage, it was surprising that 5% have no metrics in place. But the good news is that most data engineers are beginning to measure business value. However, since only 40% are measuring their impact on business outcomes — better serving customers, getting to market faster, and the like — there's plenty of room for improvement here.

RECOMMENDATIONS:

- Pick 1 strategic external metric and start tracking
- Add a new one quarterly or bi-yearly

Q: What type of measurements do you use to track your success?



Respondents were instructed to select all that applied.

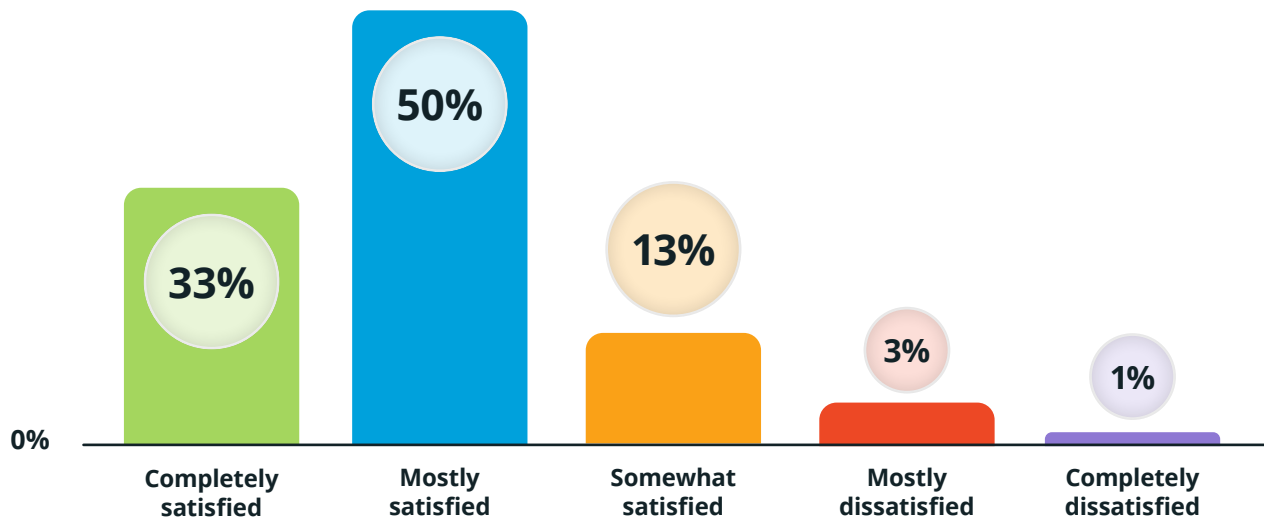
Data Engineers Are Overwhelmingly Satisfied

Surprised? You're not alone. We've seen more than one survey find that data engineers are overworked, overwhelmed, and burned out. And that may be true.

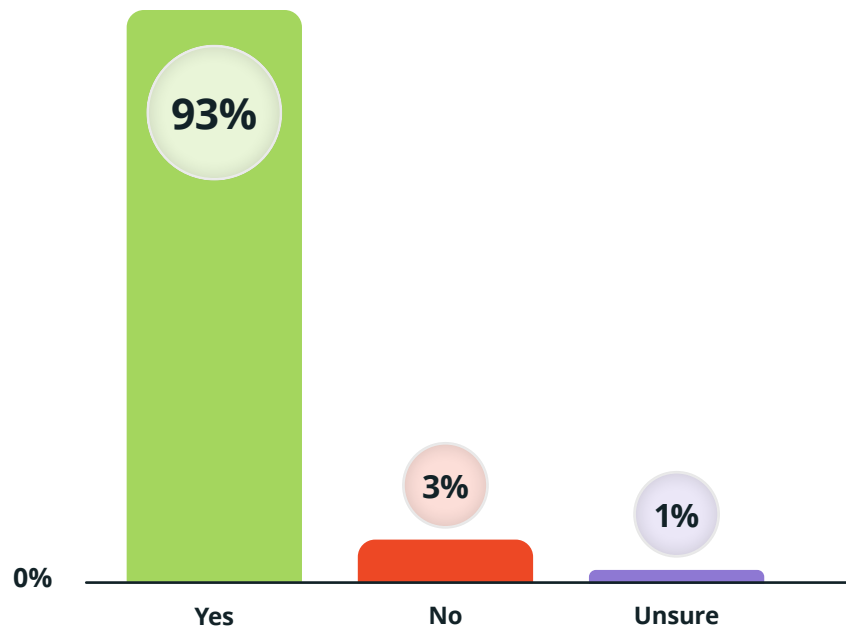
But the 523 data engineers we surveyed, which included 150 from our own community, are also

overwhelmingly satisfied with their chosen field. 88% are at least mostly satisfied with their career, and nearly all of them (93%) would recommend data engineering to others.

Q: How satisfied are you with your career as a data engineer?



Q: Would you recommend data engineering as a career?



Here's What You Can Learn From Them

There's a lot to be learned from the 33% of data engineers who are completely satisfied. Those who reported feeling this way were much more likely to:

- Be aligned with a specific line of business
- Receive data requests through standard form or tool only
- Always get enough information with data requests
- Get consistent input from different data requesters
- Believe it's very important to understand the business context and impact of data requests
- Get business context and impact for data requests
- Think their organization uses data engineers to their full potential

When we asked data engineers if their organization leverages them to their fullest potential, the positive numbers are still strong. However, a non-trivial 29% disagree. What do they think is standing in the way?

- Communication between data engineers and business users needs work (63%)
- Not enough data engineering staff (52%)
- The organization doesn't understand the potential data engineering has to impact business outcomes directly (44%)
- There's no formal process for data requests (39%)

Aside from staffing, which is a persistent challenge, three of their top four reasons speak directly to the issues we've been discussing throughout this paper.

Other Interesting Findings

Self-Service Isn't the Be-All and End-All (Yet)

We're huge proponents of enabling self-service data, and StreamSets has helped many enterprise organizations do just that. But when we asked business data consumers about self-service, 67% said they still want to work with data engineers most or all of the time.

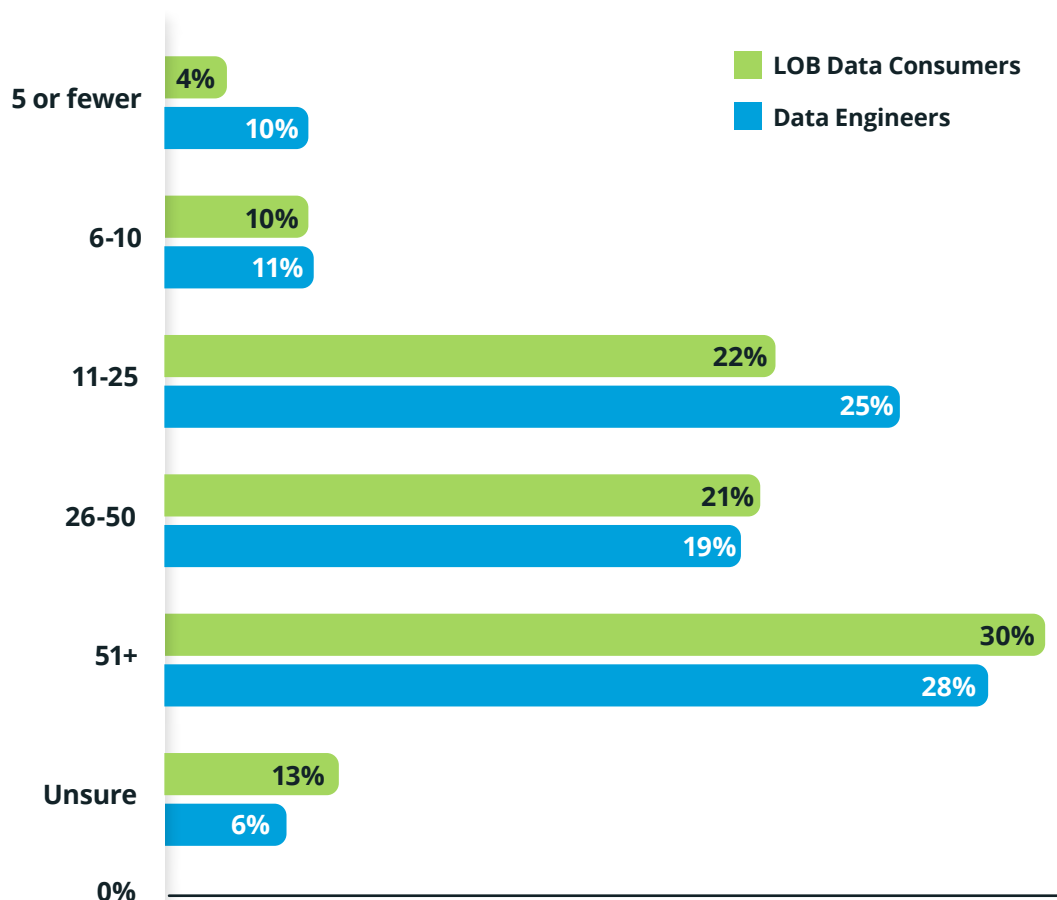
Size Matters

Data engineering teams of all sizes were represented in this study. Those with the largest data engineering teams (51+) are likelier to think their organizations leverage data engineering to

its fullest potential. On the flip side, those with the smallest teams (5 or fewer) are much less likely to feel the same. Additionally, data engineers on these small teams are significantly less likely to be completely satisfied with their careers.

This is likely due to resources. Larger data engineering teams are sitting in big organizations with the resources to establish processes, hire enough staff, invest in the best infrastructure and tools, etc., while the smallest teams are stretched beyond their limits.

Q: About how many data engineers are in your organization?



What's Next?

It's time to start acting on the recommendations throughout this report. Data engineers are ready to be the strategic partners your business needs to drive business value from data initiatives.

As it stands, many data engineers function more as order-takers than strategic advisors and lack the full business context behind requests. However, the data shows that when data engineers get the big picture and work hand-in-hand with business teams, outcomes improve markedly. Data consumers get higher-quality data faster and data engineers feel more empowered and satisfied.

Tellingly, the most fulfilled data engineers tend to be tightly integrated with business units and objectives. The report suggests there is significant room for improvement in processes, communication, metrics, and mindset to unlock data engineering's full potential. With a focused effort to involve data engineers in strategic decisions and goals, enterprises can transform this technical function into a true driver of business value.

How StreamSets Can Help

If you're looking to help your data engineers focus on delivering business value rather than building pipelines — or you struggle with data engineering staffing — consider a data integration tool that lets you do more data engineering with fewer data engineers.

StreamSets can 10x your data engineers' productivity.

- Learn once to create many different data integration pipelines — and build and deploy CDC, streaming, ETL, and ELT pipelines for any platform, on-premises or in the cloud, from one interface. This reduces the required tools and skill sets, maintenance, and other related technology overhead, driving out complexity.
- Simplify all transformations with 50 pre-defined extensible drag-and-drop processors. Meet 99% of your analytics requirements out

of the box and give your “pro-level” users the ability to include custom code and deliver it as a new element that can be easily reused.

- Pipeline fragments that easily capture, reuse, and refine business logic, so you can encapsulate expert knowledge in portable, shareable elements and keep them up to date no matter where they are used. Common transformation logic and processing elements can be independently reused across multiple pipelines without specialized knowledge, too.
- Create repeatable patterns and easily build hundreds of pipelines with just a few lines of code with StreamSets Python SDK.

Try StreamSets [free for 30 days](#) to see how easy it is!



About StreamSets

StreamSets, a Software AG company, eliminates data integration friction in complex hybrid and multi-cloud environments to keep pace with need-it-now business data demands. Our platform lets data teams unlock data—without ceding control—to enable a data-driven enterprise. Resilient and repeatable pipelines deliver analytics-ready data that improve real-time decision-making and reduce the costs and risks associated with data flow across an organization. That's why the largest companies in the world trust StreamSets to power millions of data pipelines for modern analytics, smart applications, and hybrid integration.

To learn more, visit www.streamsets.com and follow us on [LinkedIn](#).