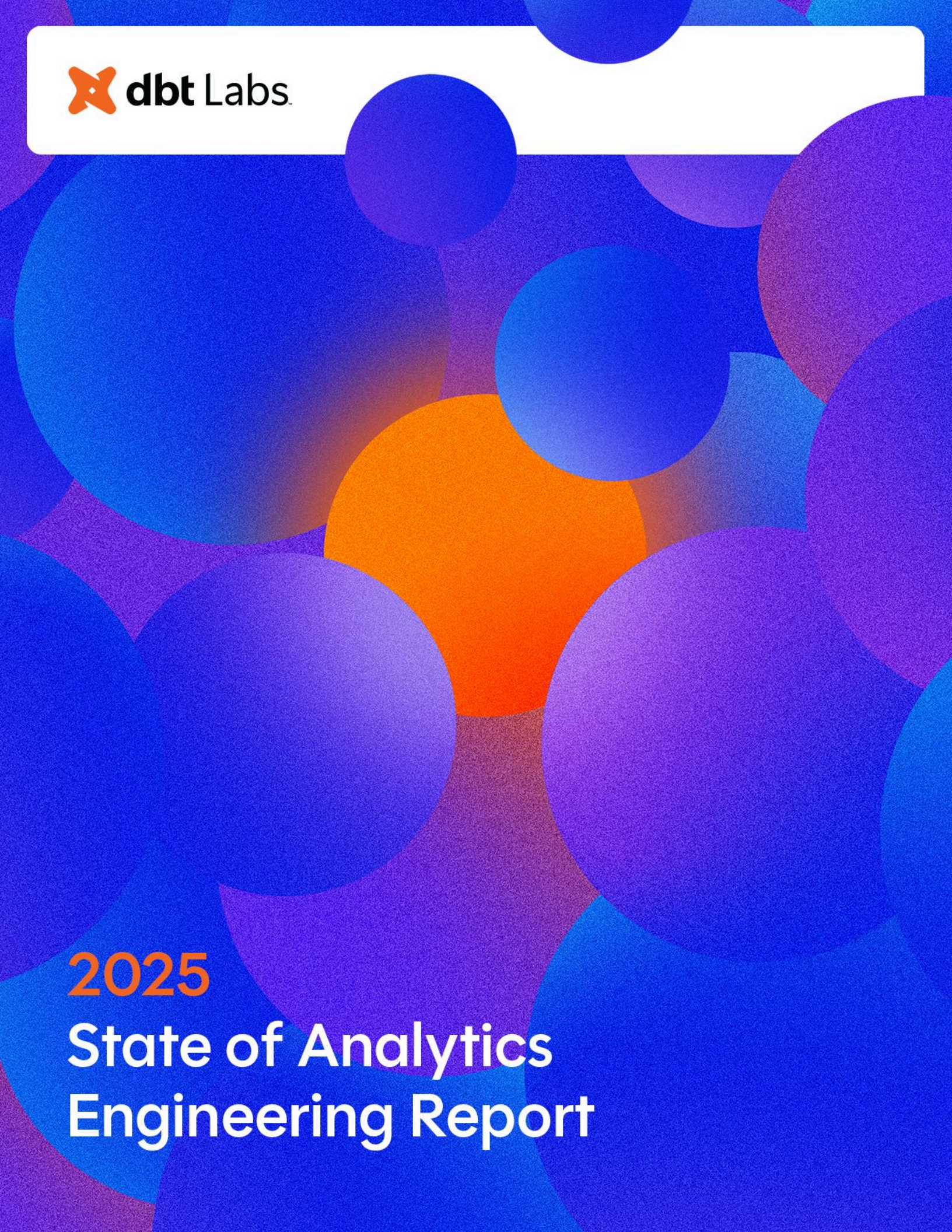

dbt Labs
2025
State of Analytics Engineering Report

Introduction

The rise of analytics engineering was inextricably linked to the most important computing story of the past decade—Cloud. The next phase of analytics engineering will see a transformation on an even larger scale powered by the most important computing story of this decade—artificial intelligence. We're starting to see glimmers of what that world looks like.

AI is a driving force behind renewed investment and restructured workflows for data teams.

The 2025 State of Analytics Engineering Report highlights a sector that is evolving at breakneck speed, where thus far AI is the catalyst to reshape—not replace—the role of data professionals, and where trust in data remains the foundation of success.

Here's what we found

- **AI is augmenting—not replacing—data teams**. Despite early fears of job displacement, data team sizes are actually increasing. 70% of analytics professionals already use AI to assist in code development, and 50% use AI for documentation. AI adoption is surging, but rather than replacing human expertise, AI is changing what jobs look like.
- **Investment is back—big time**. After a period of economic caution, data budgets are growing again, and AI is leading the charge. AI tooling is, by far, the biggest area of investment for data teams in the past year.
- **Building trust in data is still the top priority**. Even as organizations embrace AI, data teams recognize that unreliable data means unreliable outputs. Data quality remains the most critical challenge for data teams to solve.

The 2025 State of Analytics Engineering Report, crafted by dbt Labs, examines how data teams are navigating this new AI era. This report isn't just a snapshot of where we are today. It's a blueprint for where analytics engineering is headed next.

Let's dive in.

Key Insights

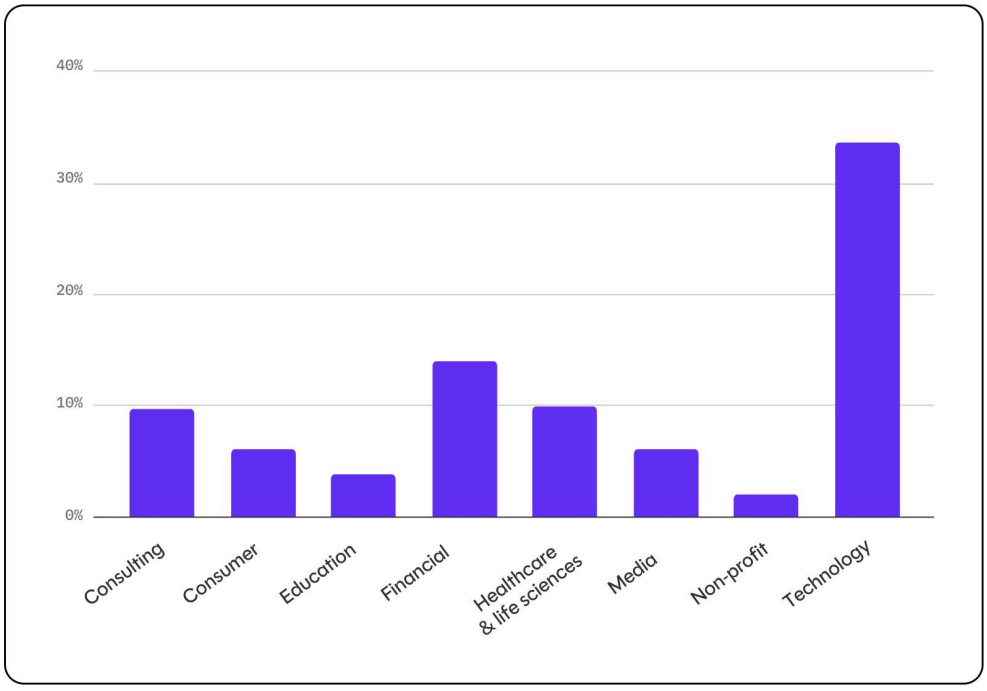
Organizations

Analytics engineering is proving its value in new and varied domains—expanding far beyond the initial "tech-forward early adopter" userbase that powered its initial base.

While the tech industry is still the most heavily represented by a substantial margin (at 34%), its share of the total has declined by 3% year over year.

Some of the most heavily regulated industries—financial services (15%) and healthcare and life sciences (10%)—represent a significant portion of the analytics engineering community. Analytics engineering is proving to be a critical capability even for organizations dealing with complex, compliance-heavy data ecosystems.

What best describes your organization's industry?



Data practitioners

Salary

The view this year is one of growing investment in data teams, and this extends to financial compensation. Salary distributions show strong growth, although this has been driven by and concentrated in North America and Europe.

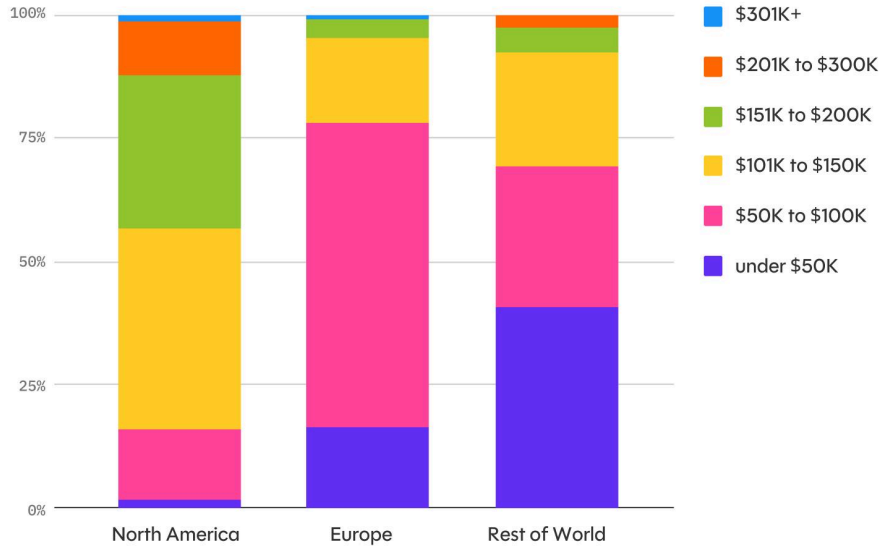
North America

- In last year's survey, 69% of individual contributors in data roles earned over $100K. This year, it's over 80%.
- Similarly, 32% of respondents in manager roles reported earning over $200K last year; this year, it's 49% of manager respondents.

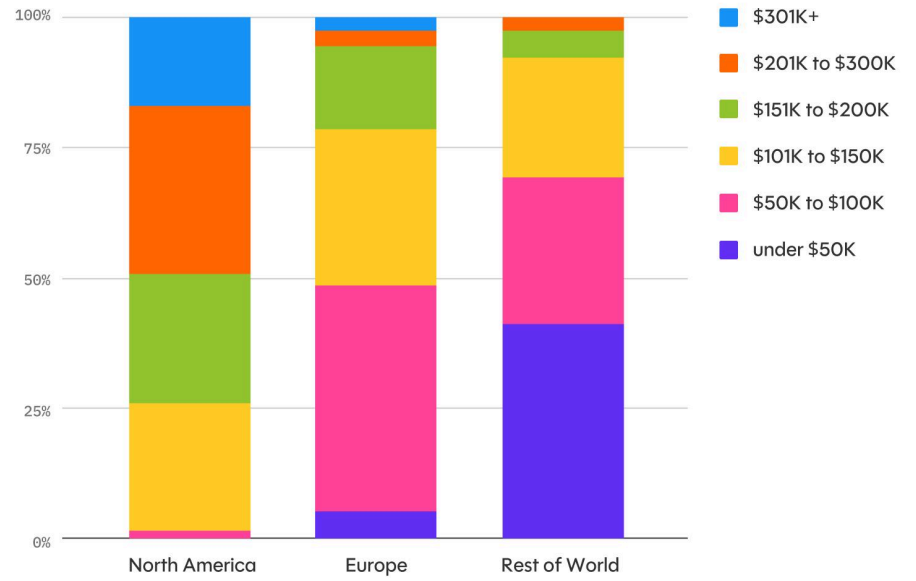
Europe

- Last year, 40% of IC respondents reported earning over $50K. This year, it's up to 51%.
- A potentially interesting thread to track is that respondents reported a decline in manager salaries in Europe, however. Last year 79% of respondents reported earning over $100K; this year it's 51%. The large number of ICs relative to managers means that this does not significantly alter the picture in aggregate.

Individual contributor salary by region (in USD)



Manager / Director / VP Salary by Region (in USD)



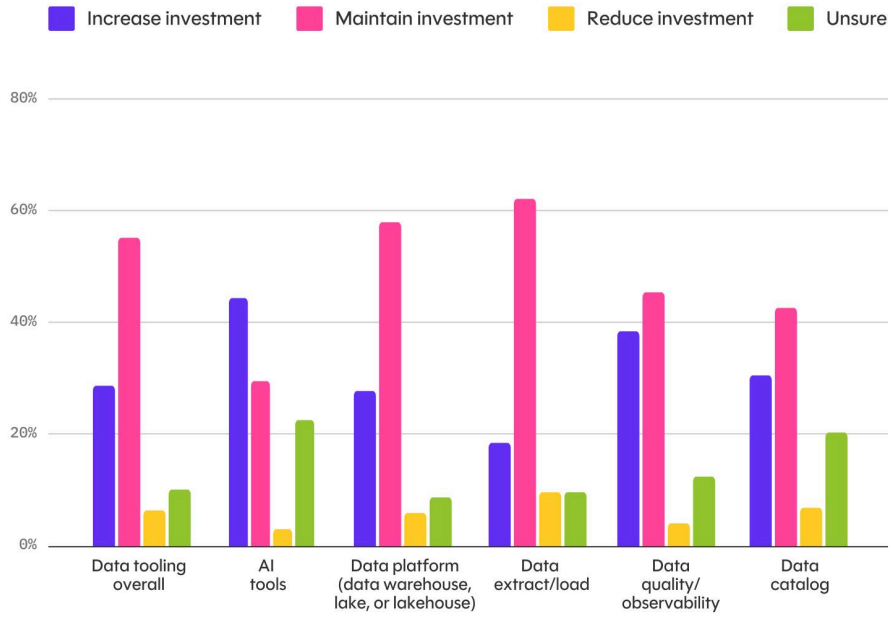
Investment areas

We asked respondents how their data teams are thinking about investment in data platforms and various data tools. AI tooling is by far the largest area of tool investment, and the only category in which the plurality of respondents (45%) plan to increase investment in the next 12 months.

Data quality/observability was the second-largest area for increased investment; 38% of respondents are planning on increasing investment in the next 12 months. The interest in increased investment in data quality/observability highlights the urgency with which organizations are trying to solve problems related to data quality. This is a theme that permeates throughout the findings of this report.

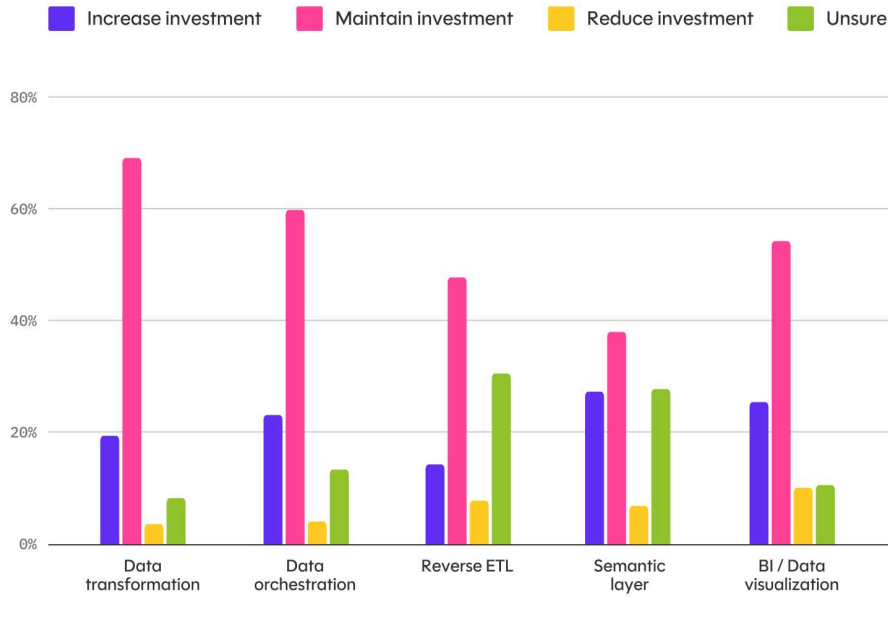
1/2

How is your team thinking about investment in the following areas over the next 12 months?



2/2

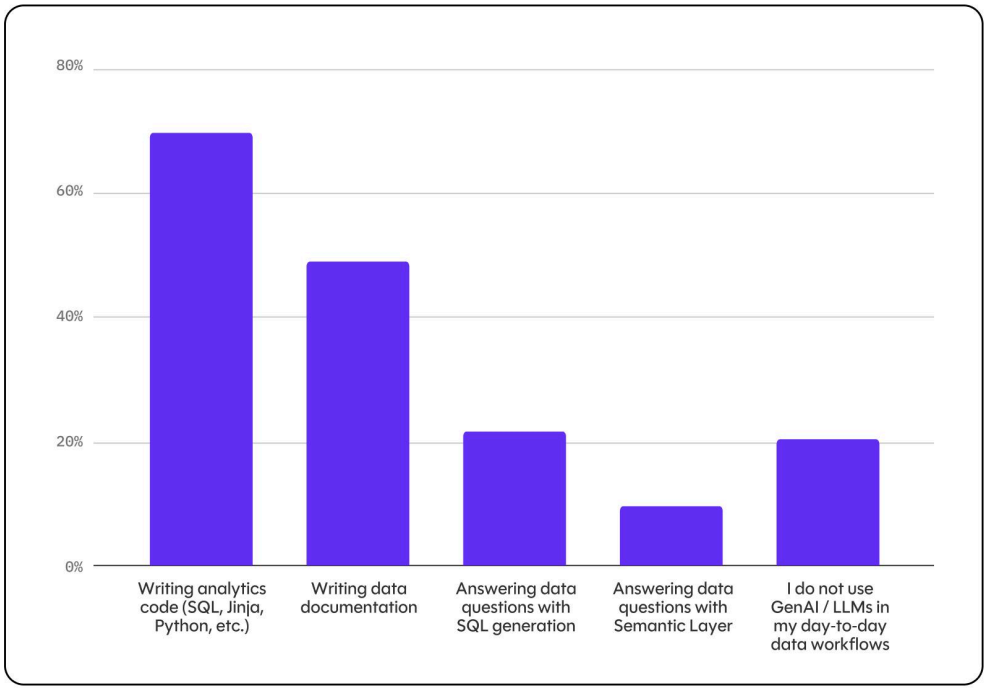
How is your team thinking about investment in the following areas over the next 12 months?



AI Investment

We are clearly in the AI era. An overwhelming majority—80% of respondents—are already using AI in their day-to-day workflow. This represents an enormous jump from just one year prior, when only 30% of respondents were doing so.

In which of the following ways do you use generative AI / LLMs in your day-to-day workflow? Select all that apply.



How are respondents using AI? Code development, followed by docs / metadata development represent the most frequently reported use cases for data teams. That is, today AI is much more commonly being used to create or describe data assets, than to consume them in order to help answer data questions.

Overall, 70% of respondents use AI for analytics development in some form. How are they going about this? Data teams are primarily using general-purpose, widely accessible LLMs like ChatGPT, Claude, and Gemini to handle code development.

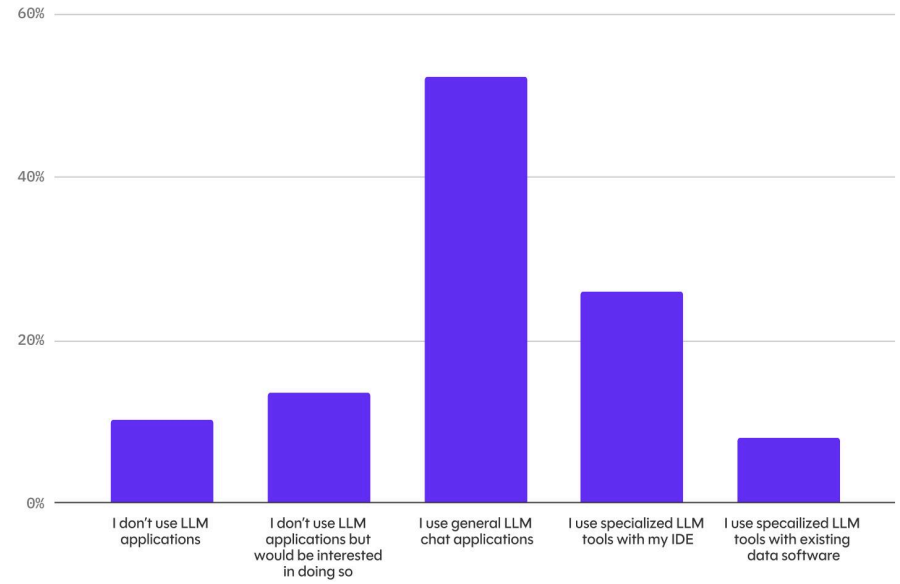
However, these general purpose tools can often run into limitations due to constraints on their ability to view important context in an analytics project, such as the rest of a codebase or metadata. It's no surprise we're already seeing ~25% of respondents use specialized gen AI solutions built into their development tooling.

We expect two things to happen here:

1 The context from your data will make its way into general purpose systems via the development of open standards for managing context to LLMs.

2 Specialized tooling that is deeply integrated with data development use cases will become more central in the workflows of data practitioners.

Which of the following best describes how you use AI tools in your analytics engineering development workflows?

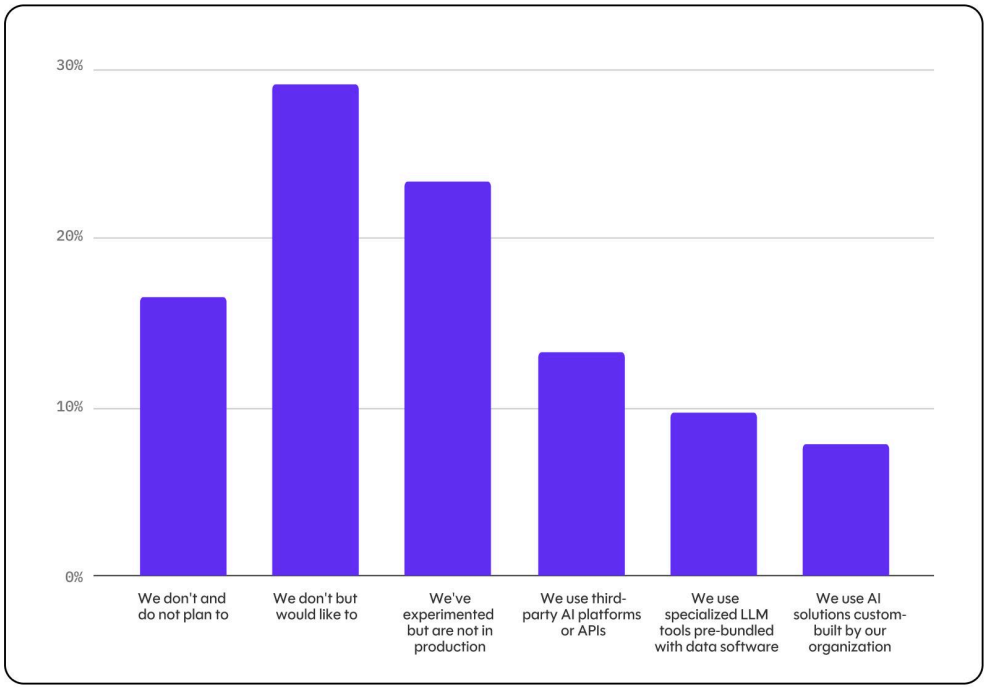


Meanwhile, the appetite for using AI to query data in natural language is there. More than half of respondents reported being interested in doing so, but can't (or won't) yet. 29% of respondents said their organizations don't use AI tools for natural language interaction with data, but would like to, while a further 23% have already experimented with natural language interaction.

The difficulty of generating consistently accurate data reporting with AI is a likely factor here. Of the 30% using AI to consume data assets by answering questions in natural language, two-thirds are doing so today using vanilla SQL generation, compared to one-third using a semantic layer. Research has in the past indicated that using a semantic layer tends to lead to significantly higher accuracy in natural language AI queries of data. As noted earlier in the study, 27% of respondents intend to increase investment in semantic layer tooling in the next 12 months.

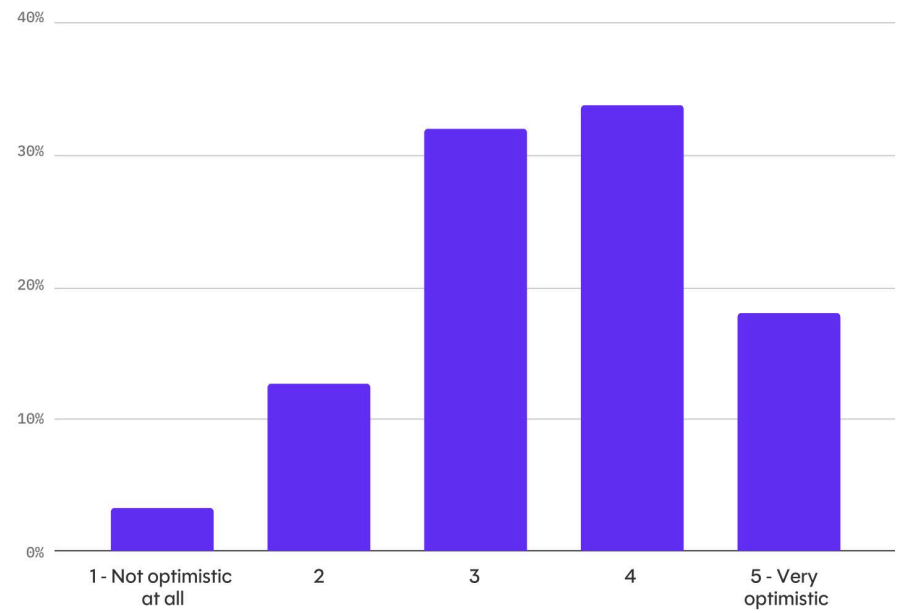
Meanwhile, nearly a third of respondents said they are already using AI tools to interact with their data using pre-bundled tools, third-party platforms, and homegrown tools. The level of interest indicated by organizations in going down this route means this number will likely continue to climb.

Which of the following best describes your organization's use of AI tools for natural language interaction with data?



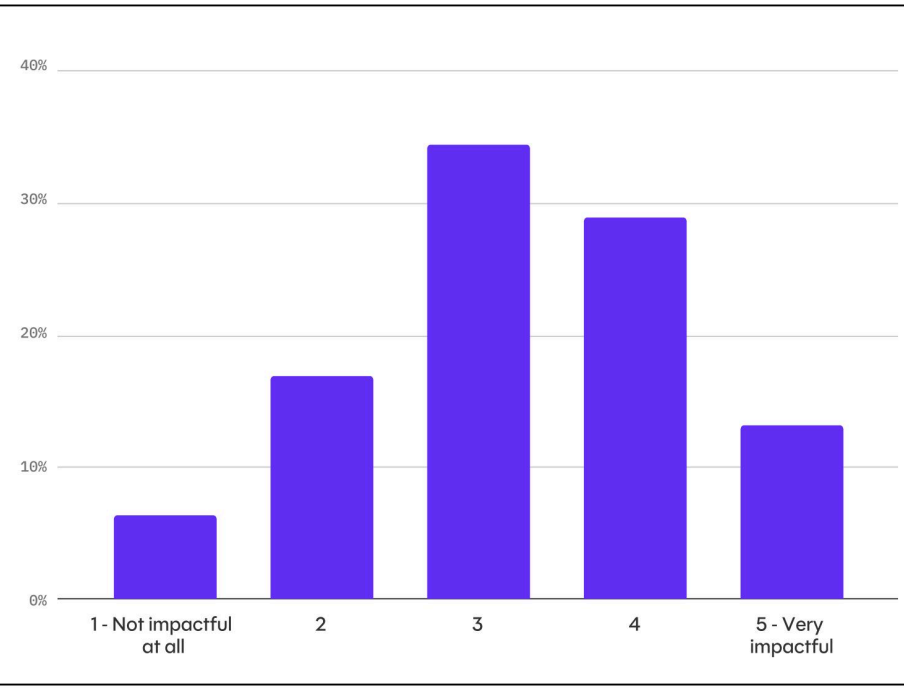
When it comes to optimism around the impact of AI, while a small percentage of respondents are skeptical, the majority express a positive outlook. The vast majority of analytics professionals see potential value in AI tools, but are still somewhat cautious in their optimism.

How optimistic are you about the potential value and impact of AI/LLMs in the data analytics workflow?



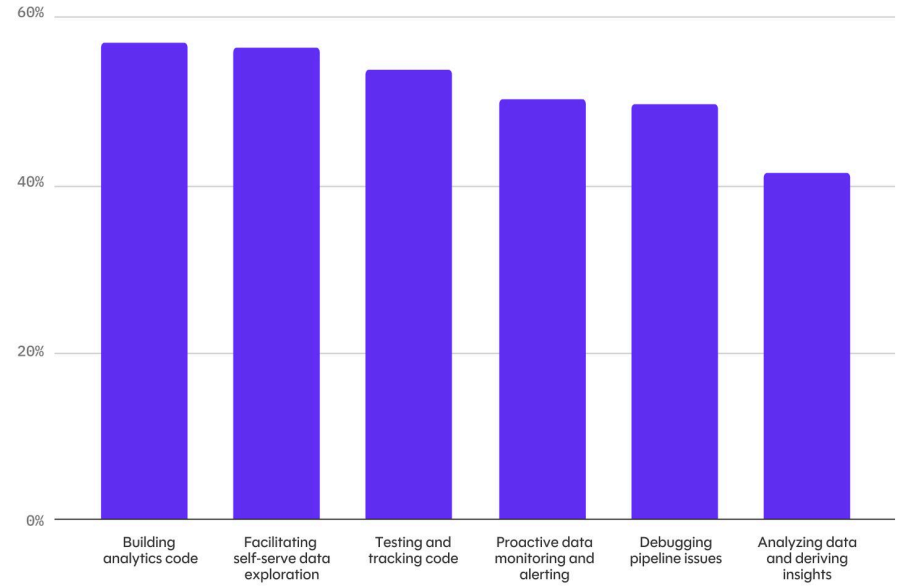
AI/LLMs hold promise for data teams, but practical impact is still evolving. The small but notable gap between optimism and real-world impact suggests that while AI-driven tooling is improving workflows, it hasn't yet been quite as transformative as hoped for all of its users. Based on the pace of adoption we're already seeing, we will watch closely how this impact changes in next year's report.

If you use AI tools as part of your day-to-day data workflow, how impactful are these tools?



Survey respondents expect a wide range of areas where AI will significantly benefit the data workflow in the near future. In particular, a great deal of optimism (cited by 56% of respondents) is expressed for the ability of AI to help with self-serve data exploration for more users, and to help bridge the data quality gap with proactive data monitoring (cited by 50% of respondents) and pipeline debugging.

In what areas do you expect AI to significantly benefit the data workflow in the near future? Select all that apply.



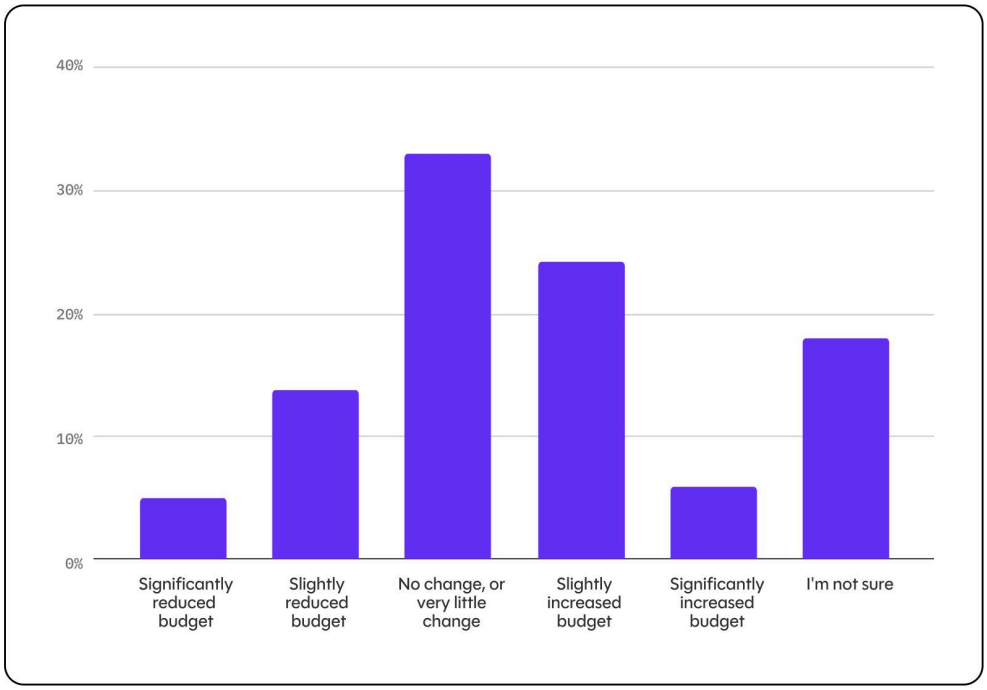
We see substantially more optimism for each individual use case than we did last year, suggesting growing conviction in the impact of AI. For instance, last year 33% of respondents expected AI would make it easier to debug pipeline issues; this year it's 50%. Similarly, the number that said they expected AI would significantly benefit the process of analytics code increased to 57% this year from 50% last year, and 54% of respondents expect AI to benefit testing/tracking code, up from 47% last year.

Data team budgets and headcount

With the increases in AI investment, data budgets and data teams are growing significantly, according to survey respondents.

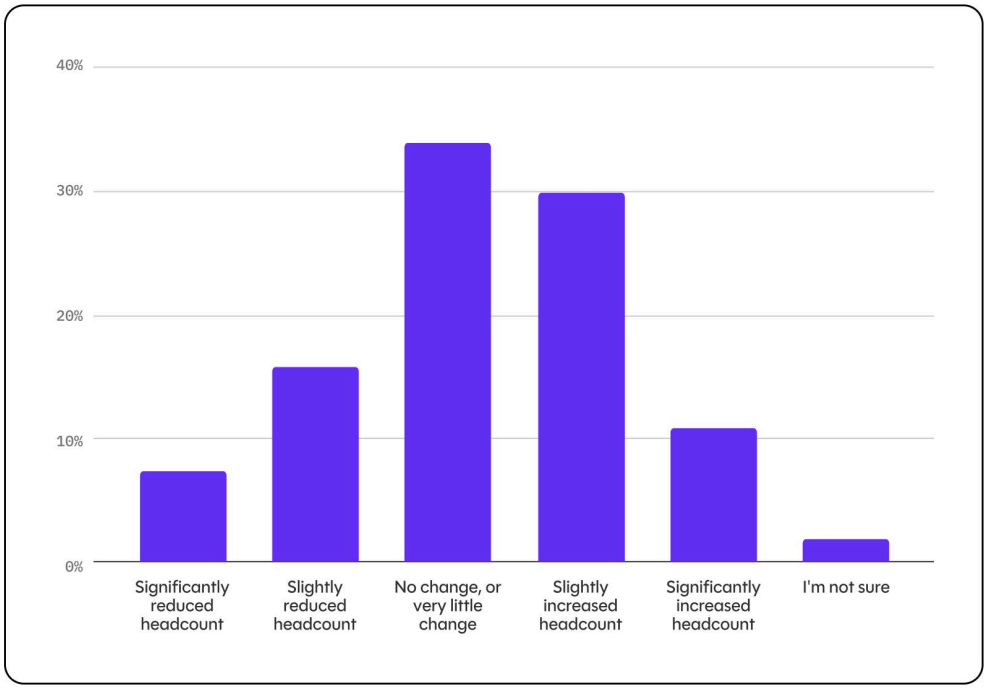
Last year, just 9% of respondents reported budget increases for their data teams; this year 30% of respondents reported budget increases from the previous year.

Is your data team's budget bigger or smaller than it was this time last year?



AI and data tooling investments aren't coming at the expense of headcount; data teams are growing too. 40% of respondents reported headcount increases in the year, compared to 14% of respondents in last year's report. Data teams are an important part of an organization's success, and investments are demonstrating that in this year's survey.

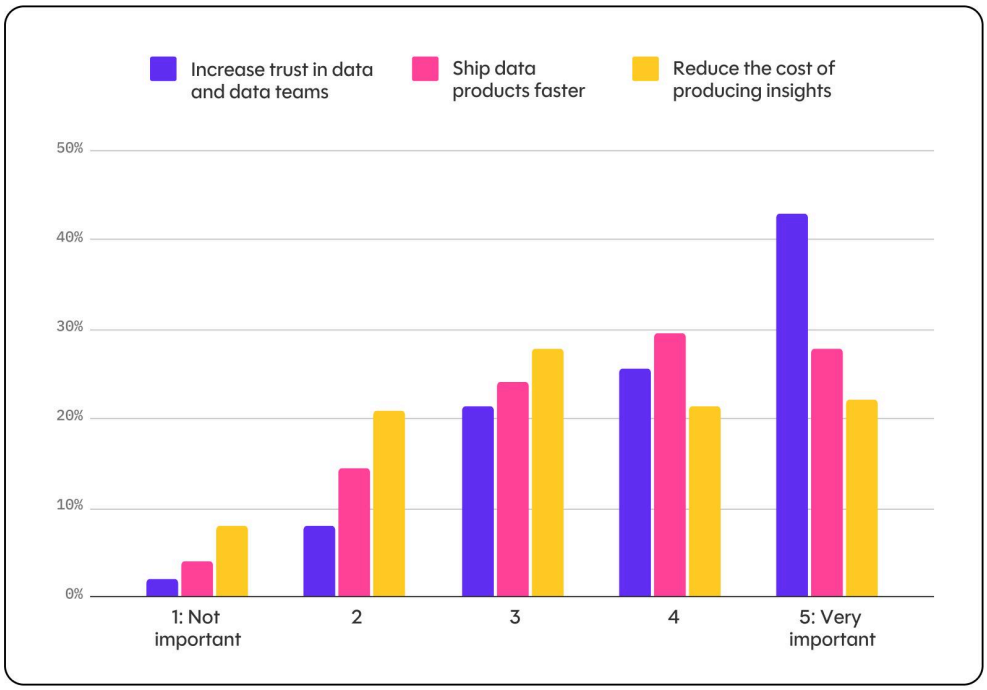
Is your data team bigger or smaller than it was this time last year?



Challenges for data teams

Data teams solve a variety of challenges for their teams and organizations. The survey results underscore a reality for data teams: increasing trust in data is the most important objective for data teams. This has not changed in the AI era. The ability to ensure accuracy, transparency, and governance in the data workflow is more than a technical necessity—it's a business imperative.

Please rate the importance of the following objectives for your organization in 2025

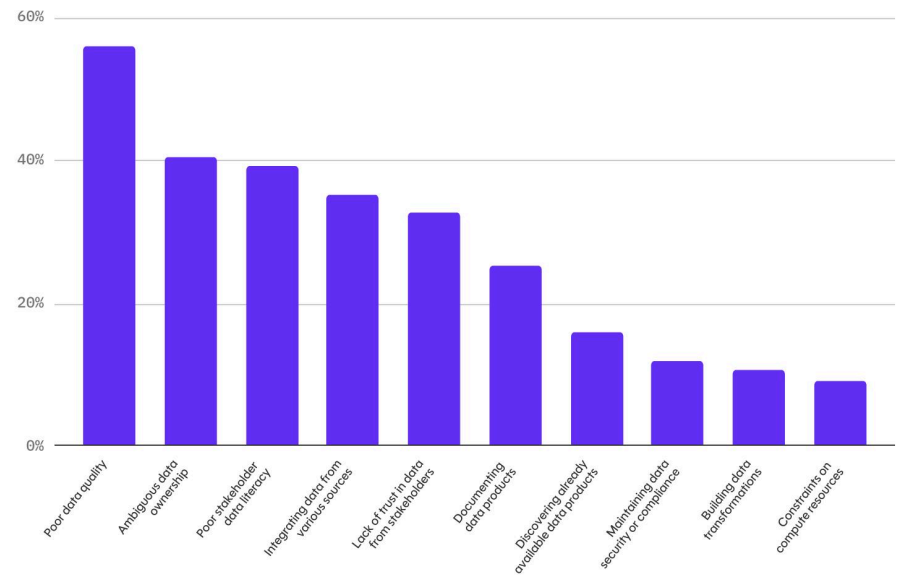


Poor data quality continues to be the challenge most frequently reported by data teams, cited by over 56% of respondents. When data is incomplete, inconsistent, or outdated, it introduces errors that ripple across dashboards, reports, AI models, and operational systems. Poor data quality compromises trust and decision quality, and severely hinders the effectiveness of data teams.

While documenting (25%) and maintaining data sets (12%), as well as constraints on compute resources (9%)—all once major hurdles—are now significantly alleviated challenges, the problem of poor data quality stubbornly persists.

Meanwhile, the prevalence of ambiguous data and poor stakeholder data literacy as significant pain points suggests that a major challenge isn't just technical but also organizational. Data quality is irrelevant when organizations are unable to make use of that data to drive meaningful insights.

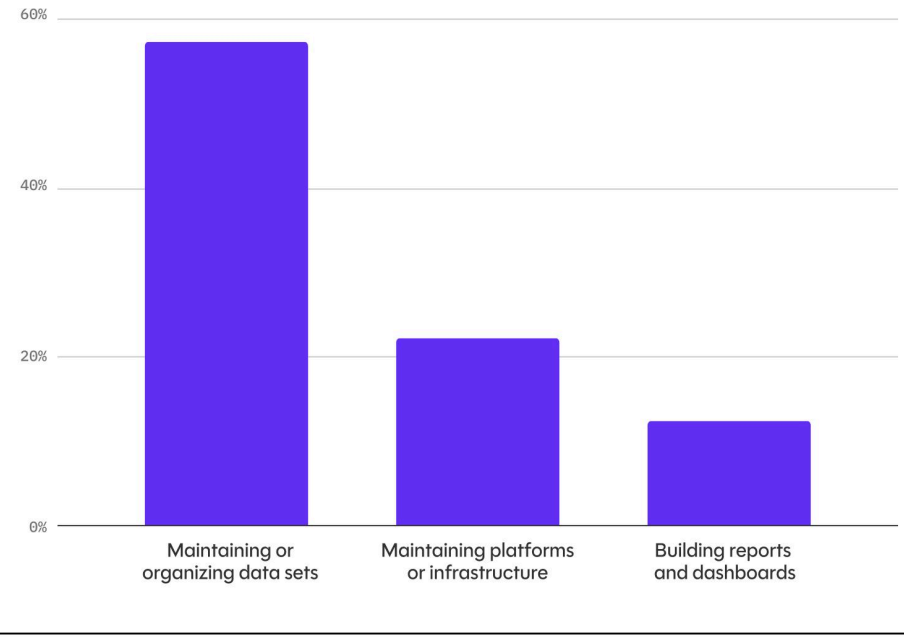
What do you find most challenging while preparing data for analysis? Select up to three.



Data practitioners' daily experience

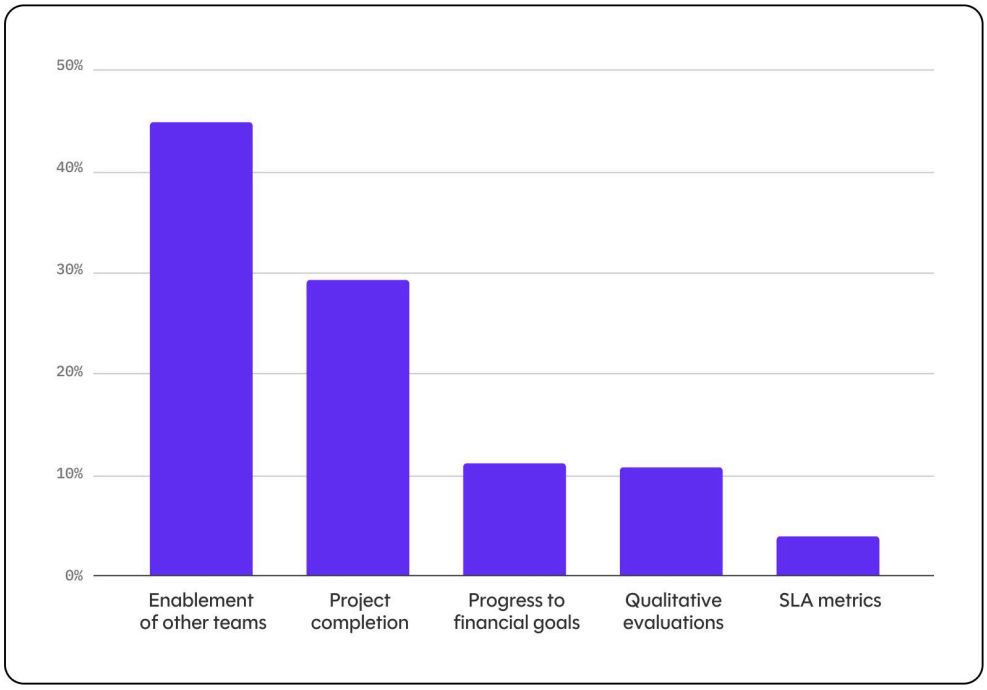
Even as more and more data teams adopt AI in their workflow, this has not fundamentally changed how contributors report spending most of their time. 57% of respondents spend most of their workdays maintaining or organizing data sets, nearly identical to last year's report.

Which of the following best describes how you spend most of your time?



Similar to last year's report, enabling other teams was identified as the primary measure of success for data teams.

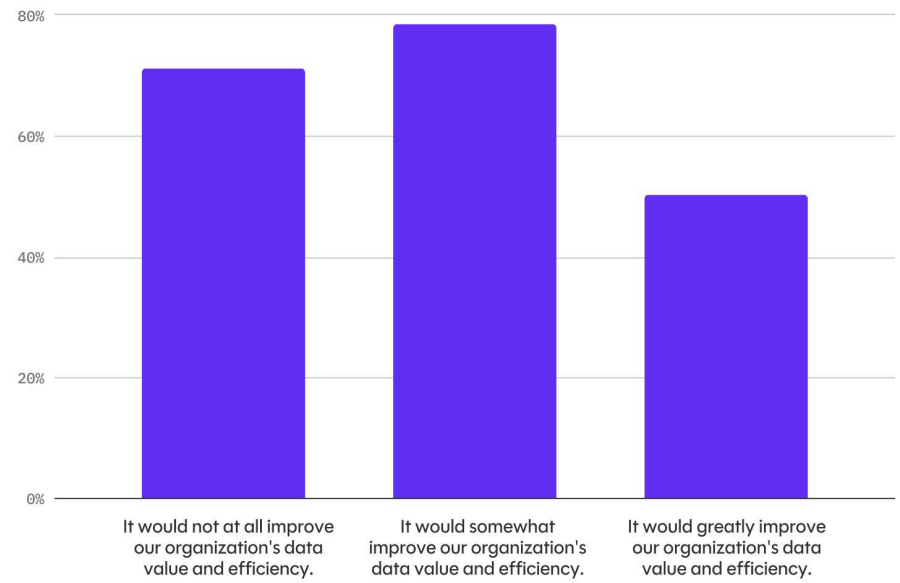
How does your team primarily measure success?



The creation of analytics engineers was substantially driven by data analysts looking to draw best practices from software engineering. Today, we continue to see the pull for nontechnical business users towards engaging in data transformation.

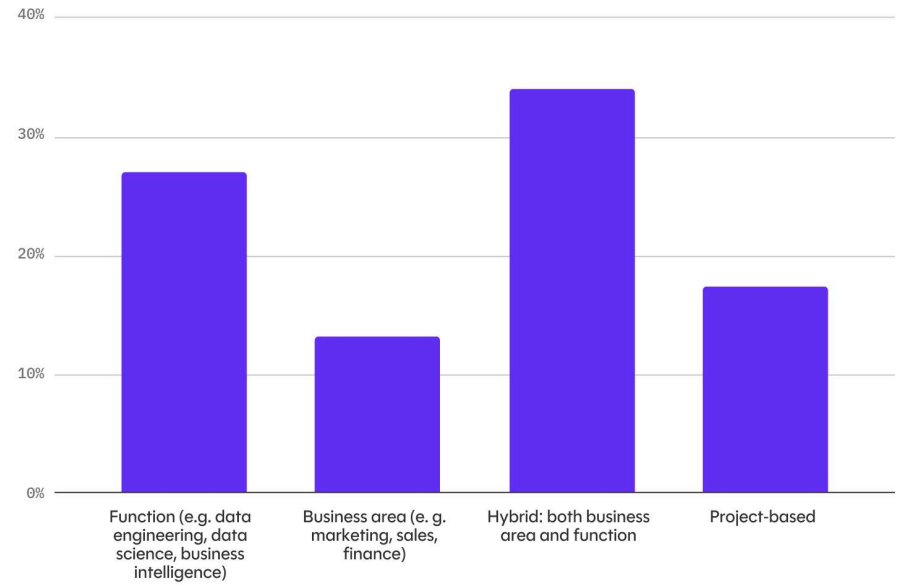
Nearly 65% of respondents said that enabling nontechnical business users—most of a data team's stakeholders—to create transformed and governed data sets would somewhat or greatly improve their organization's data value and efficiency.

How helpful would it be for your organization to enable nontechnical business users to create transformed and governed datasets?



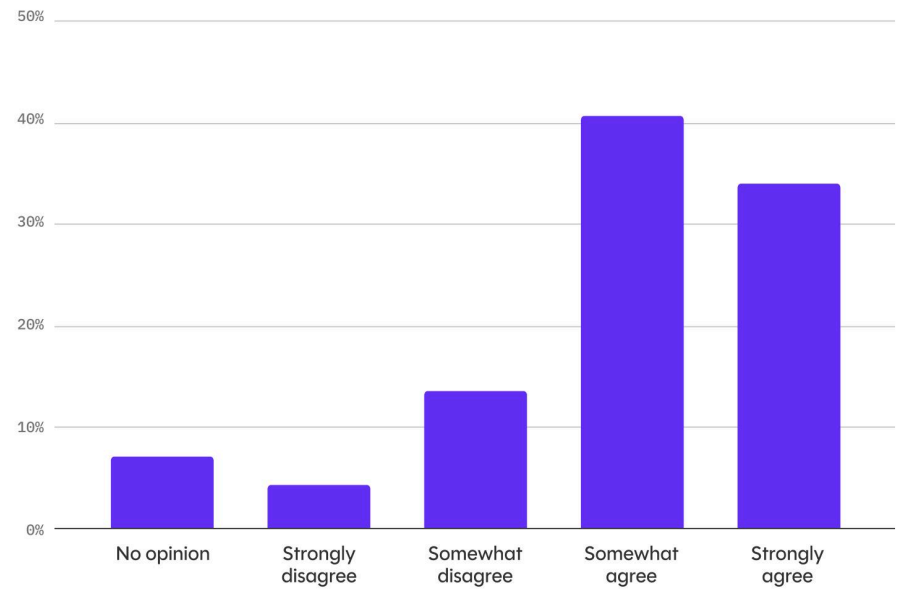
Work is distributed among data teams in a variety of ways—by function (data engineering, data science, etc.), by business area (marketing, sales, finance, etc.), by project, or by a hybrid approach. Last year we noted an increase in the hybrid model in which work is distributed by both business area and function. This signaled that data team members of all specialties are embedding more deeply within their organizations. This has held true in this year's survey as well.

What determines how work is distributed within your data team?

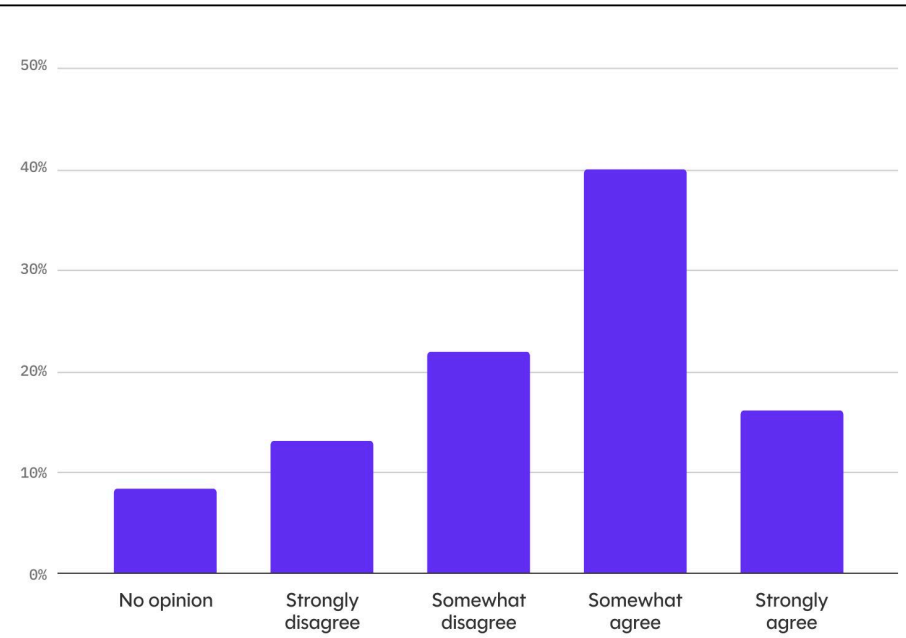


Respondents overwhelmingly agree their organization values the data team. Respondents feel less strongly that their organizations set clear goals for the data teams. Data teams feel valued, but ambiguity remains in how data teams are expected to impact their organizations.

My organization values the data team. We are respected and included in decisions that impact our work.



My organization sets clear goals for the data team. We have a roadmap on how to positively impact the organization.



What's Next?

The 2025 State of Analytics Engineering makes one thing clear: Data teams are not standing still. With AI reshaping workflows, investment flowing back into data, and maintaining data quality still a key challenge, organizations must be proactive in defining their data strategy. This report is not just a reflection of current trends—it's a call to action.

Data leaders should use these insights to guide their investments and to focus on AI-driven efficiencies while reinforcing governance to build data trust.

The path forward is one where AI and human expertise work in tandem, data reliability is non-negotiable, and the role of analytics engineering continues to expand in impact and influence. Now is the time to take these learnings and build a data strategy that drives smarter decisions, faster innovation, and a more resilient data organization.

There's more:

There's so much to explore within the analytics workflow. Explore Tristan Handy's whitepaper on The Analytics Development Lifecycle on applying software engineering best practices to all layers of the analytics stack. And check out Tristan's post How AI will disrupt data engineering as we know it.

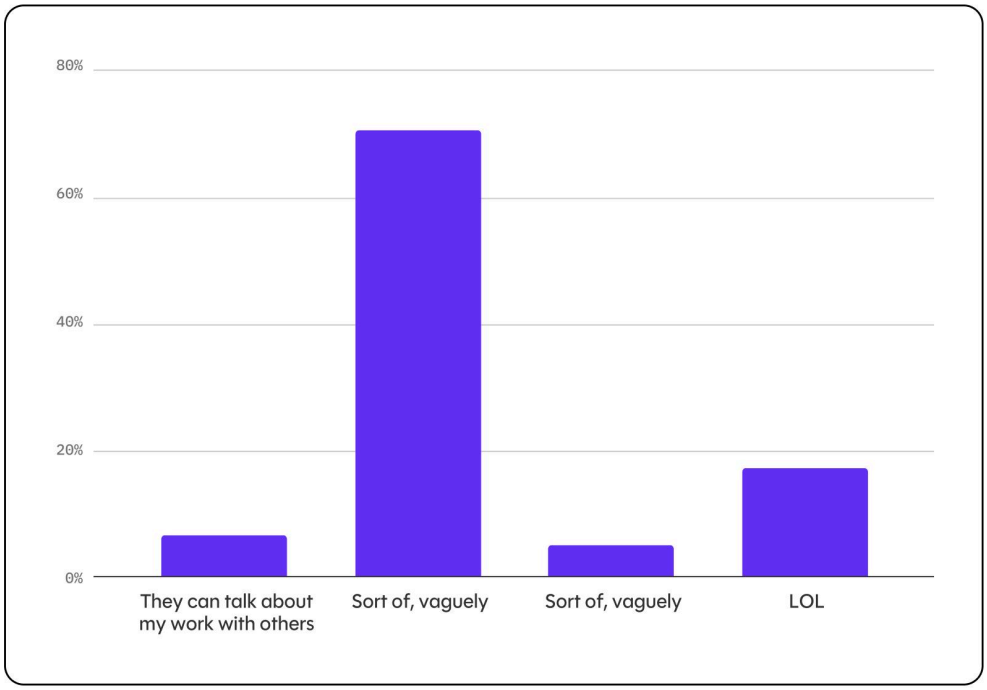
Join us!

If you want to learn more about the world of analytics engineering and connect with data practitioners and data leaders, join the dbt Community.

And one more thing...

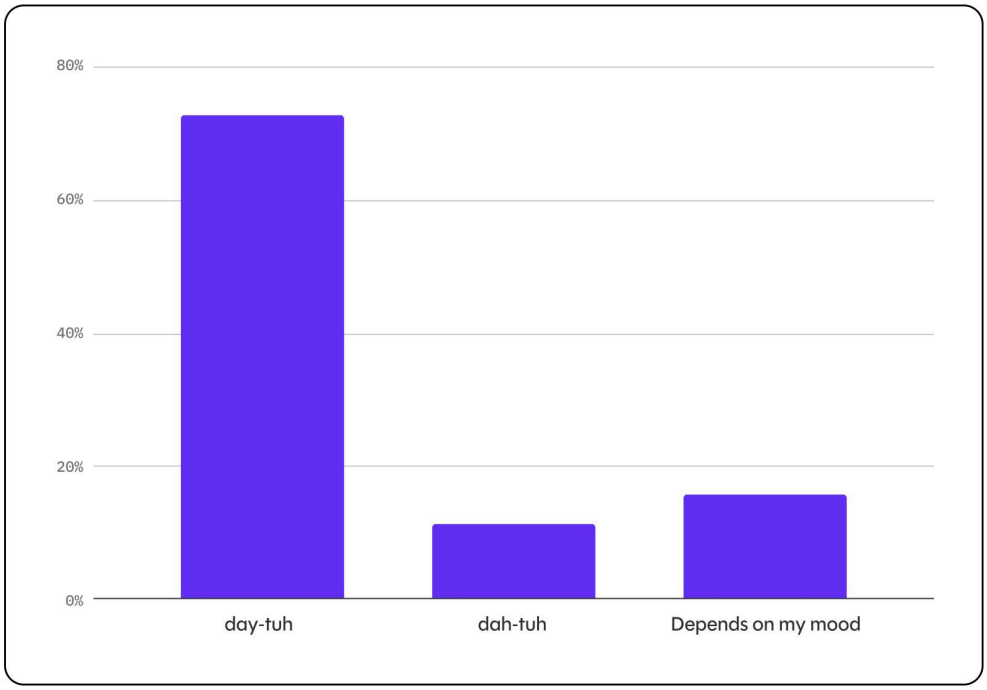
Frankly, your families and friends are doing pretty well, based on our personal experience.

Do your family and friends know what you do for a living?



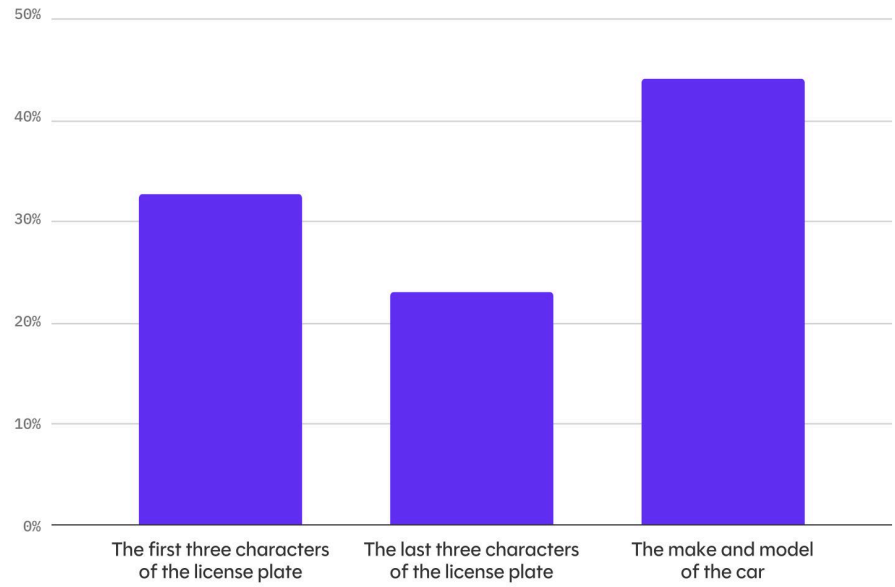
A strong showing for “mood.”

How do you pronounce "data"?



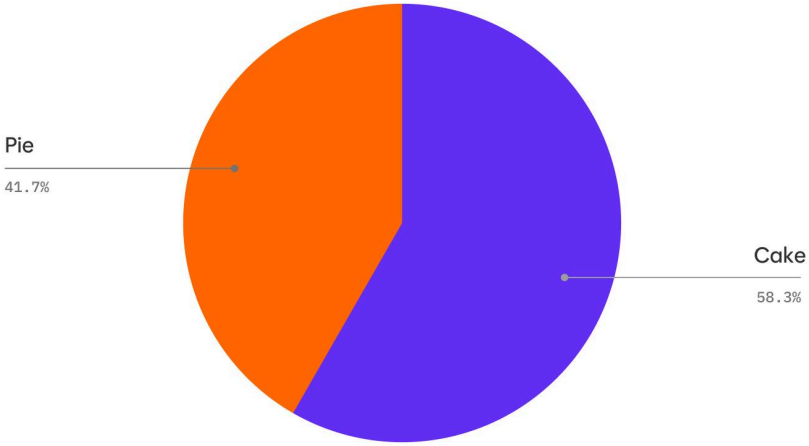
There's a right answer here.

When you're looking for your Lyft/Uber car, what do you look for?



There's a right answer here, too.

Cake or Pie?



Methodology

dbt Labs collected survey responses October 8, 2024 - December 27, 2024 from 459 data practitioners and leaders—70% of survey respondents are individual contributors (IC) and 30% are managers. Analytics engineers made up 48% of IC respondents, 36% of IC respondents are data engineers, and 16% of IC respondents are data analysts. These splits are virtually identical to the 2024 State of Analytics Engineering.