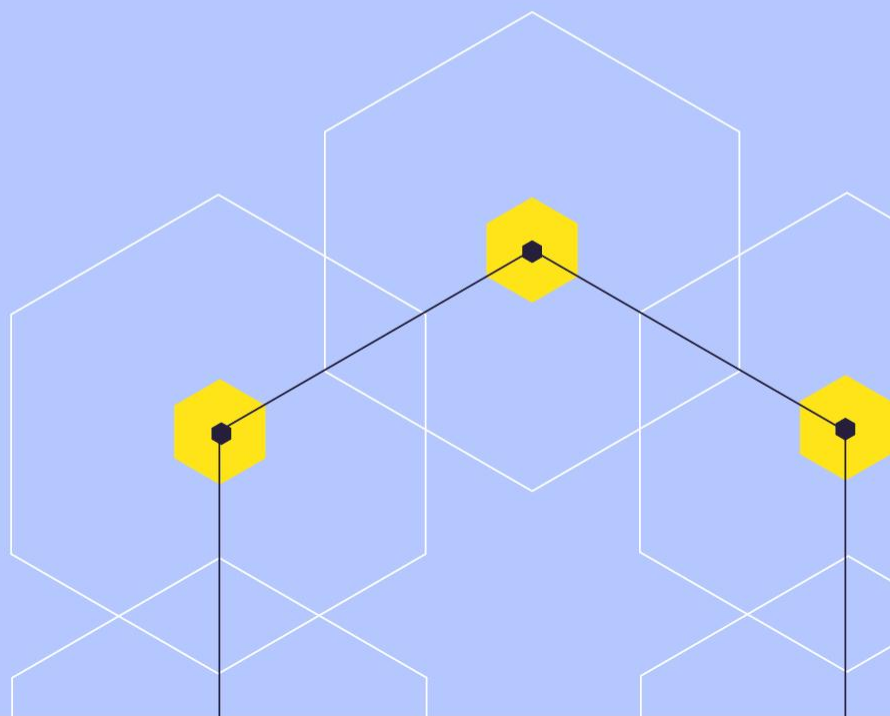
SYNQ

2025 Data Quality Benchmark Survey



Summary

In the 2025 Data Quality Benchmark Survey we've asked hundreds of data leaders and engineers about their top challenges and planned investments around data quality.

AI/ML use cases are now the most critical use cases, but data teams are struggling with how to best build reliable data pipelines, citing "insufficient knowledge of how to test well" as their top data quality challenge.

This high-stakes environment is driving nearly 40% of companies to increase investments in data quality and observability tools, which today mainly consists of built-in tooling from their data transformation tool.

This survey is made by SYNQ, the data observability platform.

SYNQ

Key takeaways

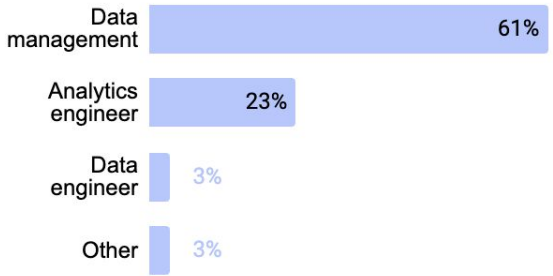
- Insufficient knowledge on how to test well is the biggest data quality challenge
- AI/ML ranks as the most important use case for data, followed by internal BI reporting
- Nearly 20% of respondents report a single data incident cost over $10,000
- The vast majority of participants rely on built-in tests from their data transformation tool
- Nearly 40% of teams plan to increase their data quality and observability investments next year
- Only 10% of respondents use AI often in their data quality workflows

Survey methodology

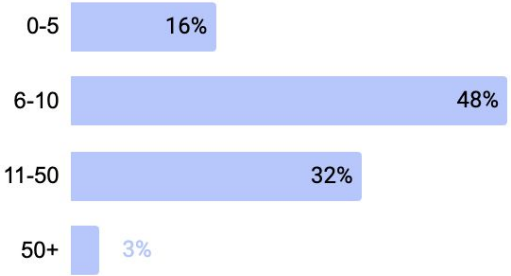
Who you'll learn from

We surveyed hundreds of data professionals, mainly data leaders (lead, head of, director, VP) and analytics engineers about their top challenges related to data quality. The survey sample contains responses from companies ranging from 100s to 10,000s of employees. The majority of respondents operate at a significant scale, with thousands of tables and dozens of people in the data team.

What's your role



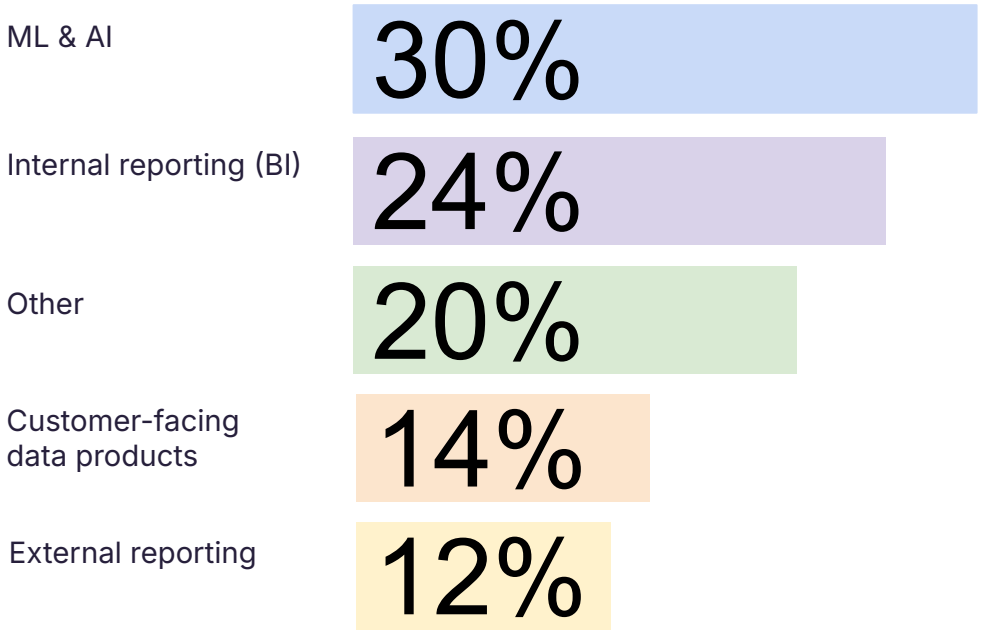
Size of data team



of tables in data warehouse

Who you'll learn from

What's the most important use case of data in your company

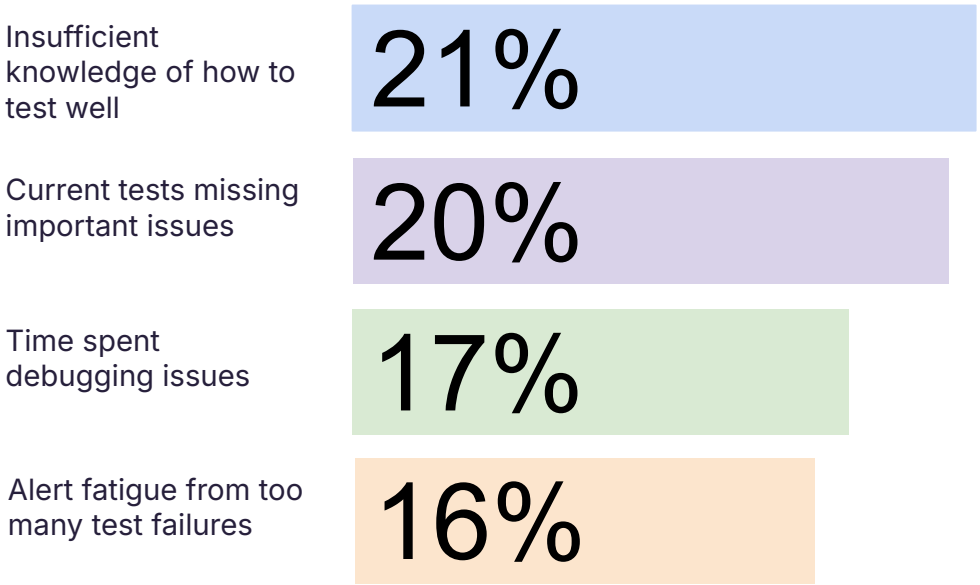


Learnings—AI/ML ranks as the top use case of data followed by internal BI reporting

Top challenges

Which of these is your biggest challenge related to data quality?

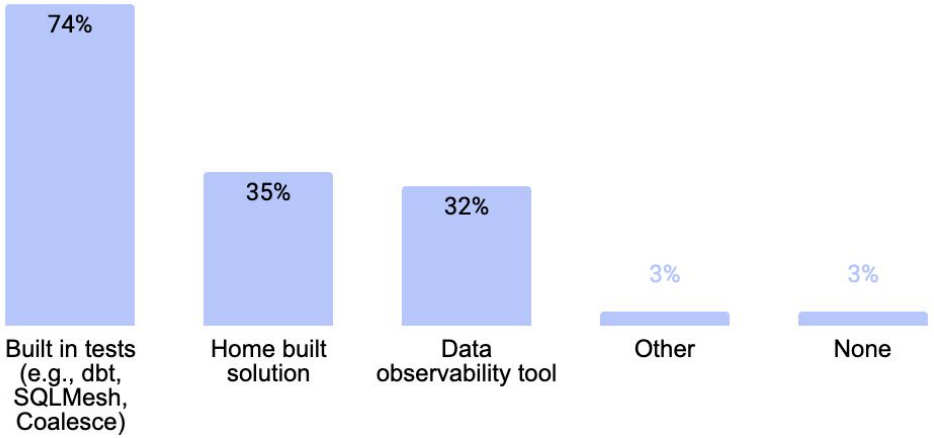
Top 4 challenges



Learnings—the top problem around data quality faced by data practitioners is lack of sufficient knowledge of how to test well. Alert fatigue and missing important issues also rank top of the list

Testing & monitoring

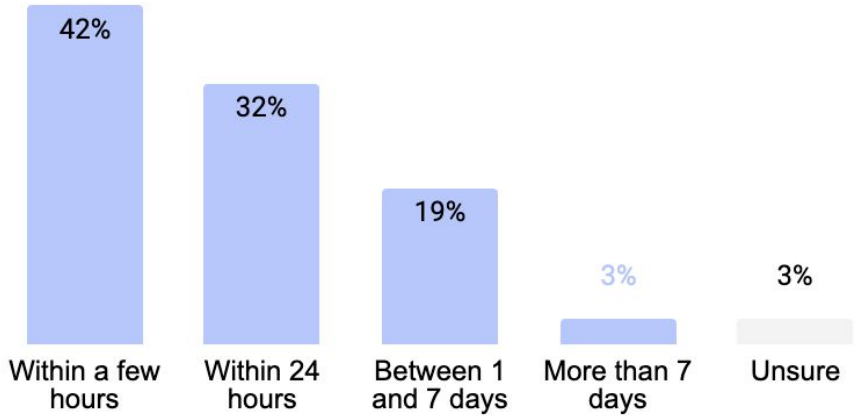
Which tools or platforms do you currently use for testing and monitoring?



Learnings—the vast majority of surveyed participants use data quality tests, mainly built in tests from their data transformation tool

Testing & monitoring

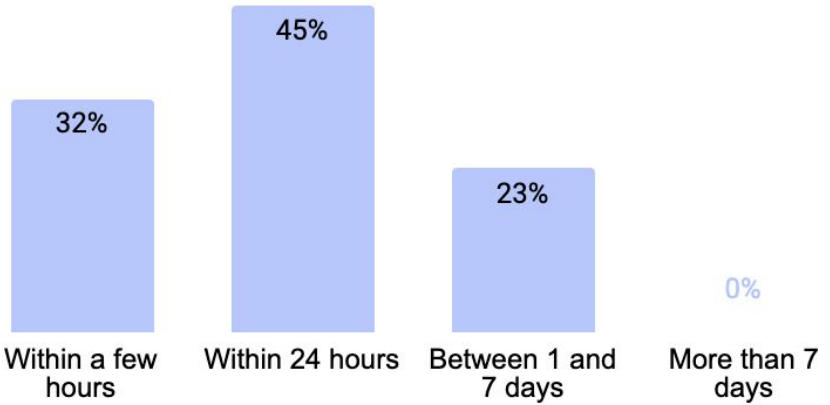
How fast do you typically detect data issues?



Learnings—the majority of respondents detect issues within 24 hours of them happening, and 97% within 7 days of them happening

Issue resolution

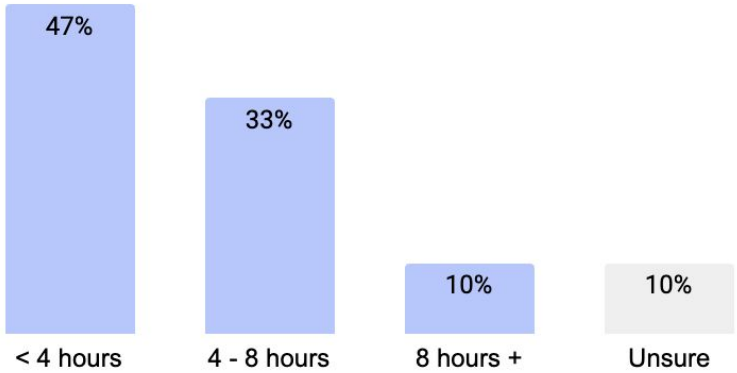
How fast do you typically resolve data issues?



Learnings—one third of respondents typically resolve issues with a few hours while most do it within the first 24 hours

Issue resolution

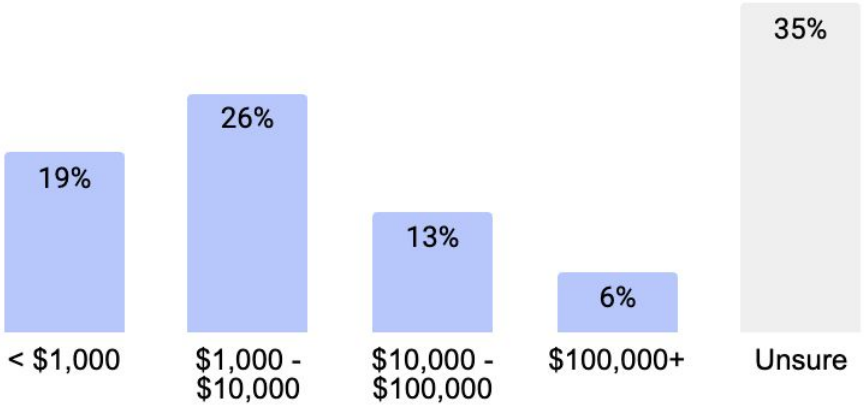
% of weekly time each data professional spends on data quality tasks (debugging, testing, etc.)



Learnings—more than ¾ of respondents report that their data team members spend between 0 and 8 hours weekly on data quality tasks

Issue resolution

What's your most costly data incident?



Learnings—19% of respondents have had a data incident cost more than $10,000. 35% are unsure what their most costly data incident is

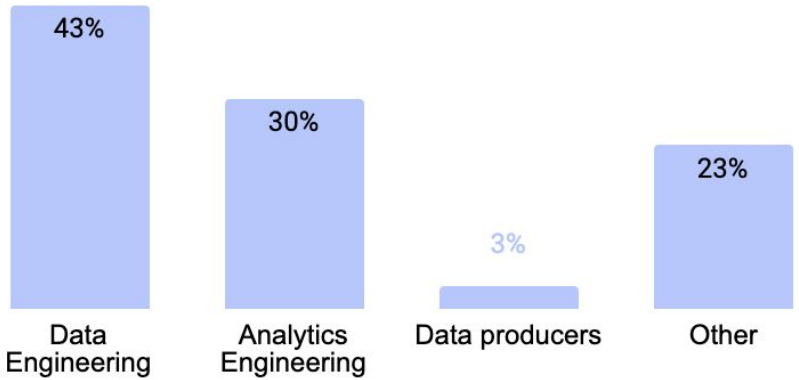
SYNQ

"There's no single way to assess the cost of data incidents, but having a shared method that everyone buys into helps create transparency and mobilize people when issues need to be addressed quickly"– *extract from [The cost of data incidents]*

Mikkel Dengsøe
SYNQ co-founder

Issue resolution

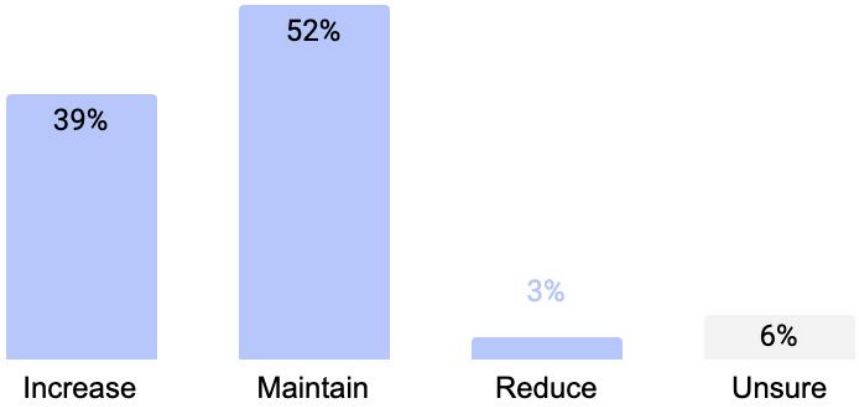
Who is primarily responsible for fixing data quality issues in your company?



Learnings—the majority of the workload around fixing data quality issues falls on data and analytics engineers

Planned investments

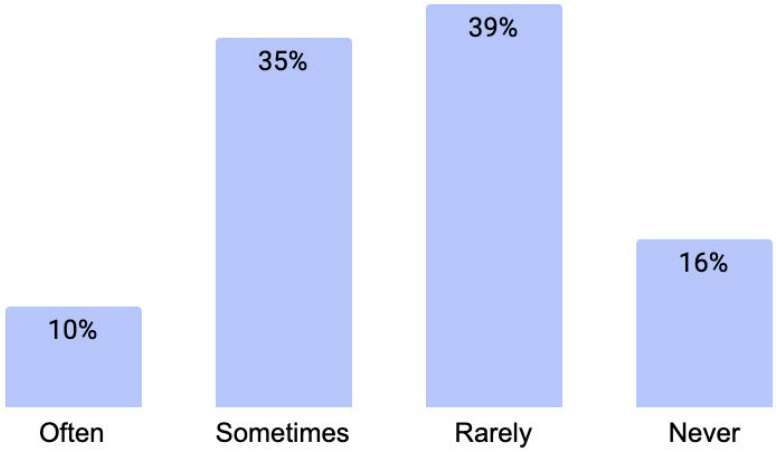
How is your team thinking about investments in data quality & observability in the next 12 months?



Learnings—39% of respondents plan on increasing their investments in data quality the next 12 months. Only 3% plan on reducing it

AI

Do you use AI today for data quality workflows (e.g., automated testing and root cause analysis)?



Learnings—only 10% use AI often in their data quality workflows while the majority use AI “sometimes” or “rarely”

AI

How are you planning on using AI in your data workflows in the next 12 months?

Based on the responses, the primary driver for AI adoption in the next 12 months is around team efficiency. Data leaders are planning to implement AI to assist with time consuming tasks, and empower the wider business through self-serve analytics.

SYNQ

Top AI focus areas

Developer efficiency (36%): This was the most common application mentioned. Teams plan to use AI for tasks like writing code, automating PR reviews, and generating documentation.

Insights & analytics (32%): The second most popular use case, with plans to use AI for process optimization, automation, generating insights, and anomaly detection for smarter testing.

Self-serve analytics (23%): Another key area is using AI for self-serve capabilities through conversational analytics, chat-based interfaces, and increasing data accessibility.

Learn more about SYNQ

This survey was run by SYNQ. SYNQ is a data reliability platform focused on precise anomaly detection, robust ownership model and fast issue resolution, and built for teams that power business-critical use cases.

Learn more on synq.io or book a demo.

Watch 5 minute platform demo