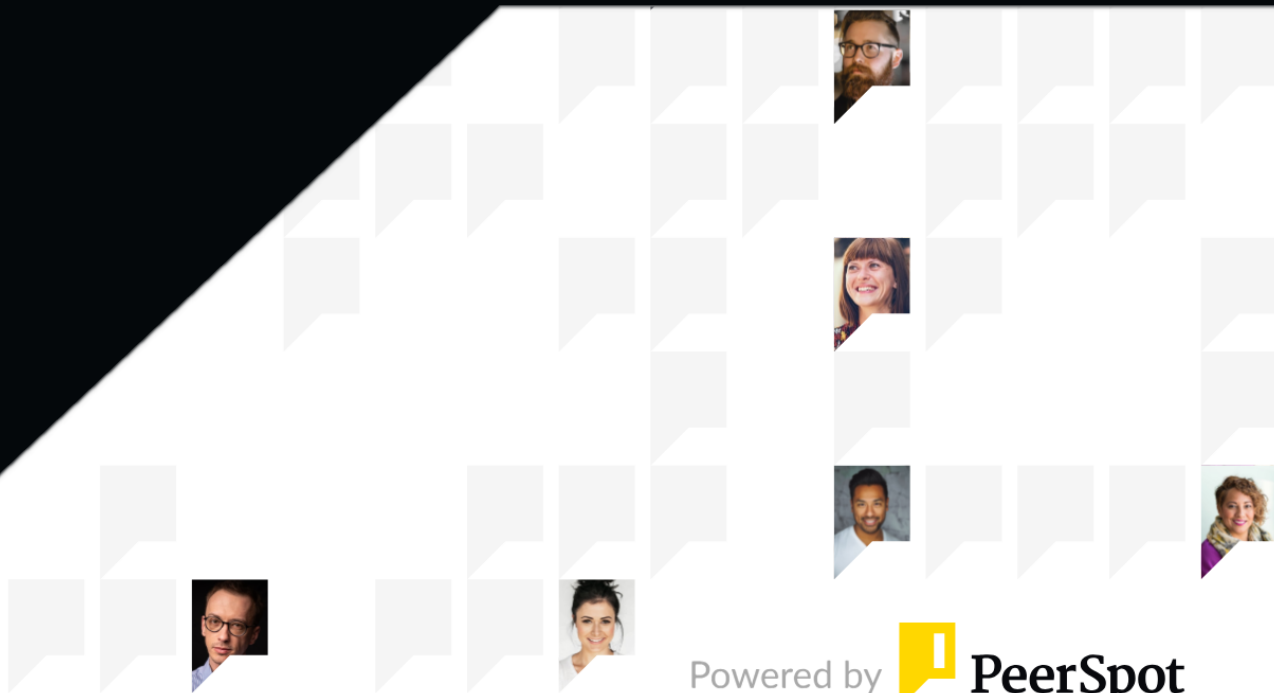


aws marketplace

Datafold

Reviews, tips, and advice from real users



Powered by  PeerSpot

Contents

Product Recap..... 3 - 4

Valuable Features..... 5 - 9

Other Solutions Considered..... 10 - 11

ROI..... 12

Use Case..... 13 - 15

Setup..... 16

Customer Service and Support..... 17

Other Advice..... 18 - 20

Trends..... 21 - 22

About PeerSpot..... 23 - 24

Product Recap



Datafold

Datafold Recap

Datafold enhances data engineering by streamlining data quality and security processes, offering robust insights and automation for faster and more accurate analytics outcomes.

Datafold provides a comprehensive set of tools for data engineers to manage and validate data pipelines while ensuring data accuracy. By automating data quality checks and offering in-depth analytics, it supports a seamless transition from data collection to actionable insights. Datafold reduces risks associated with data anomalies and accelerates the development of reliable data-driven applications.

What are the key features of Datafold?

- **Data Quality Assurance:** Automated checks to ensure data consistency and prevent errors.
- **Visualization Tools:** Intuitive interfaces for exploring patterns and discrepancies in data pipelines.
- **Data Lineage Tracking:** Trace the flow and evolution of data through systems for better transparency.
- **Scalability:** Designed to handle increasing data volumes without compromising performance.

What benefits and ROI should users expect?

- **Improved Data Accuracy:** Enhances trust in data operations and decision-making.
- **Increased Efficiency:** Reduces time spent on manual data checks, allowing focus on development.
- **Risk Mitigation:** Identifies potential issues before they impact operations.
- **Cost Savings:** Lower operational costs by minimizing data-related issues.

In the finance industry, Datafold helps institutions automate compliance checks and ensure transactional data integrity. E-commerce businesses rely on it to optimize inventory management by providing accurate sales forecasts. Healthcare organizations use Datafold to maintain patient data integrity, facilitating better outcomes and operational efficiency.

Valuable Features

Excerpts from real customer reviews on PeerSpot:

- ✓ “Datafold has proven to be a game changer for organizations, and its pricing is quite reasonable, making it easy on our budget and encouraging annual subscription renewals.”



Shrinkhala Singh

Deputy General Manager at Agriculture Skill Council of India

- ✓ “Datafold has positively impacted my organization by helping us catch data regressions before they reach production.”



JohnBosco Obi

Data Analyst at Africa Initiative

- ✓ “Datafold helped significantly because it has great resources for customizing queries and defining primary keys for comparison.”



Caio Gabriel Guimarães

Data Engineer at Qubika

What users had to say about valuable features:

“I used Datafold for data quality checks while working on a data migration project where we moved data and applied business rules and transformations. After that movement, we had to check if the data loaded was correct, and Datafold helped significantly because it has great resources for customizing queries and defining primary keys for comparison. We could set the source and target tables, specify rules, and then press play, which provided excellent results regarding data matching percentages.

“Sometimes the mismatches were minor, but other times they required more attention. Being able to use Datafold for these quality checks while working on other processes was really helpful..”

Caio Gabriel Guimarães

Data Engineer at Qubika

[Read full review](#) 

“The data diff stands out the most for me as the best feature because it gives me that column-level and row-level comparison between any two datasets, highlighting what changed, even down to characters and whitespaces.

“It also tells me when something has failed, and it makes debugging faster and more precise.

“Column-level lineage derived from SQL static analysis helps me trace how columns flow through transformations across the entire pipeline.

“Datafold has positively impacted my organization by helping us catch data regressions before they reach production.

“Even while we are still building the pipeline and mapping out how the data will flow, we catch any data regressions and alterations before they reach production. This helps us avoid a lot of debugging when the migration has happened or when we reach production level and data goes live.

“Additionally, it significantly reduces manual testing time, saving us time that we can put into useful work. Furthermore, it builds our team's confidence when pushing code changes, as the data reviewers, data annotators, and engineers can actually see the data impact of every change, not just the code change..”

JohnBosco Obi

Data Analyst at Africa Initiative

[Read full review](#) 

“Regarding the best features Datafold offers, tracking and monitoring of our data pipeline has become easier than ever, especially in the field of agriculture where collecting vast amounts of data is critical. India, being an agrarian country, has a majority of its population dependent on agriculture, making the tracking of complex data essential. The user interface is very useful and easy to navigate, allowing newcomers to train for just a week before being able to work independently with Datafold. It simplifies the entire process of collecting data from source to sync and addresses concerns related to secret and confidential data by supporting integration effectively. Datafold has consistently proven itself with its real-time alerts and visualizations when analyzing data for insights, enabling users to check on real-time anomalies and resolve them quickly. From our perspective, the platform has improved our testing environment significantly, ensuring the quality and consistency of captured data while saving us time and effort from manual intervention.

“I find Datafold to be superior, given its strong positive feedback from numerous users on platforms such as Google. My team has validated that the data testing capabilities are impressive, helping users validate data quality and identify any issues or lag, thus enabling us to address root causes before they become significant problems. Datafold has proven to be a game changer for organizations, and its pricing is quite reasonable, making it easy on our budget and encouraging annual subscription renewals.

“Datafold has positively impacted my organization as the user interface is extremely easy to navigate, allowing my team to efficiently engage with the environment and receive real-time updates regarding data capturing, quality, migration, and flow. It alerts us to any bugs, enabling us to reduce or eliminate errors almost entirely. Datafold provides an environment for creating tests, allowing users to independently ensure data matches benchmark quality and consistency, which saves valuable manpower and time. Thanks to Datafold, our team is now focused on more critical tasks rather than debugging and monitoring data flow. Overall, the feedback regarding data handling is consistently very positive.

“In terms of measurable impact, when we used traditional methods, our team

consumed approximately four to five man hours daily for data transfers, but with Datafold, we have reduced that to only one hour per day, which is a substantial time-saving and a game changer for our organization..”

Shrinkhala Singh

Deputy General Manager at Agriculture Skill Council of India

[Read full review](#) 

Other Solutions Considered

“We have not used any other solutions long-term as the homegrown options developed by our tech team were not effective, leading us to switch to Datafold immediately..”

Shrinkhala Singh

Deputy General Manager at Agriculture Skill Council of India

[Read full review](#) 

“We did not evaluate many options before choosing Datafold as our experience during the trial version was so positive that we decided to stick with it instead of exploring others..”

Shrinkhala Singh

Deputy General Manager at Agriculture Skill Council of India

[Read full review](#) 

“Before working with Datafold, I have never used similar tools. I know there are some options, but I haven't worked with them. My experience required writing customized queries or automation scripts for tasks that Datafold automates.

“I didn't work with a similar tool before or after Datafold. Whenever I needed to do data quality checks, I always had to write a customized query or script..”

Caio Gabriel Guimarães

Data Engineer at Qubika

[Read full review](#) 

ROI

Real user quotes about their ROI:

“There has been a clear return on investment since the first year because we have been able to generate value from Datafold's implementation. Previously, eight people worked on data processes, but now only two are required, significantly reducing time, costs, and human resources, all while enhancing our image in front of clients. Overall, we have no issues with data errors or migration failures, showcasing all the positives that contribute to our ROI..”

Shrinkhala Singh

Deputy General Manager at Agriculture Skill Council of India

[Read full review](#) 

Use Case

“I used Datafold for data quality checks while working on a data migration project where we moved data and applied business rules and transformations. After that movement, we had to check if the data loaded was correct, and Datafold helped significantly because it has great resources for customizing queries and defining primary keys for comparison. We could set the source and target tables, specify rules, and then press play, which provided excellent results regarding data matching percentages..”

Caio Gabriel Guimarães

Data Engineer at Qubika

[Read full review](#) 

“My main use case for Datafold is as a data observability platform that we leverage for many of our data products, particularly because our data is very susceptible to any kind of breakdown or catastrophe due to its confidential nature. Datafold helps us identify, prioritize, and deep dive into data quality issues which can occur anytime. The platform finds these issues and investigates them in a proactive manner, and appries us to rectify them before they hit production. Currently, we are using it to improve the quality of our organizational assessments and tests. It helps us test our various products and services, validate various parameters, and monitor whether the production cycle is happening in a synchronous manner. It is assisting us significantly in terms of migrating data from less secure platforms to more secure platforms as per clientele need. Datafold is continuously working 24/7 contextualizing the data ensuring that migrations happen accurately, as well as reconciling data to ensure there are no gaps, providing complete monitoring from start to end.

“Our organization started deploying Datafold primarily due to outstanding peer feedback we received regarding its user-friendly interface and the quality of data monitoring, which is amazing. Since we deal in agriculture, where personal farmer data must not leak, Datafold helps us simplify the complicated process of collecting diverse data in a streamlined manner and track it from start to end, ensuring a high level of accuracy is maintained during data collection, transferring, migrating, and ultimately monitoring. This has significantly improved our operations, and we are continuing with Datafold..”

Shrinkhala Singh

Deputy General Manager at Agriculture Skill Council of India

[Read full review](#) 

“My main use case for Datafold is to automatically compare data differences through data diffing, where I can compare datasets row-by-row, column-by-column, to surface unexpected changes.

“I also use it for CI/CD integration for data pipelines, where we embed data quality checks into GitHub and GitLab pull requests. Another interesting feature I use Datafold for is column-level lineage to trace how specific columns flow through SQL transformations across the data stack we have.

“Recently, for data diffing, there was one time we wanted to conduct a lot of comparisons on a very large dataset to be sure that all was in check before we migrated to Snowflake, which is a complex and resource-heavy process.

“We used Datafold to do the automatic checks across rows and columns to ensure that when we completed the migration, no data was lost and the migration into Snowflake was very smooth. That was in March, and it was very useful.

“For a friend who uses Datafold in their enterprise, an engineer made a change from an SQL model and did not realize it would cascade and alter downstream dashboard metrics.

“However, with the help of Datafold integrated into the pull request workflow, the data diff actually surfaced exactly which columns changed, which rows were affected, and by how much..”

JohnBosco Obi

Data Analyst at Africa Initiative

[Read full review](#) 

Setup

The setup process involves configuring and preparing the product or service for use, which may include tasks such as installation, account creation, initial configuration, and troubleshooting any issues that may arise. Below you can find real user quotes about the setup process.

“The initial deployment was really easy. Once we understood how it worked and set it up, connecting to databases like Redshift, SQL Server, and Databricks was straightforward..”

Caio Gabriel Guimarães

Data Engineer at Qubika

[Read full review](#) 

“In terms of pricing, setup costs, and licensing, I find everything to be reasonable, quite affordable, and stable over the years as the prices have not increased more than ten percent in two years. The initial licensing and setup were also straightforward, and while I cannot disclose specifics about the quote, it was definitely negotiable. Implementation and integration costs were minimal, allowing small and marginal companies to consider Datafold..”

Shrinkhala Singh

Deputy General Manager at Agriculture Skill Council of India

[Read full review](#) 

Customer Service and Support

“Customer support is excellent, available 24/7 through various channels, including phones, emails, and a ticketing system. I would rate customer support as a four out of five as there are two types of customer support: one for general feedback and another technical team addressing integration issues. Overall, it is a great support system..”

Shrinkhala Singh

Deputy General Manager at Agriculture Skill Council of India

[Read full review](#) 

Other Advice

“For others considering Datafold, I advise them to assess their own needs before committing to the platform. Datafold can be a game changer for handling vast data, data migration, and monitoring. It is important for organizations to clearly define their ROI metrics relevant to their context and establish clear, understandable SLAs that their team can implement. My final thoughts about Datafold are that any size organization would benefit from its ability to manage complex data infrastructure and architecture effortlessly, providing a seamless data engineering and automation experience. I would rate this solution a nine out of ten..”

Shrinkhala Singh

Deputy General Manager at Agriculture Skill Council of India

[Read full review](#) 

“I talked with someone on LinkedIn about [PeerSpot](#), and he mentioned something about a gift card to do the review, which is why I asked for this meeting because I'm not comfortable using the company email and I'm not sure if I'm authorized to do that. He explained that I should do a review about some tool that I have experience with to be eligible for a possible gift card, but I didn't know about the company before, to be honest.

“I have been working in my current field for 11 years overall, dealing with data-related projects.

“I have no questions and will be waiting for the next steps in the process. My review rating for Datafold is 8..”

Caio Gabriel Guimarães
Data Engineer at Qubika

[Read full review](#) 

“I give Datafold a seven out of ten because it is excellent at its core specialization, which is data diffing and CI/CD integration, and its migration automation capabilities are strong and very AI-powered. However, I lose three points due to the weak reporting.

“Even though it provides reports when there is a breakage in your push or migration, I sometimes cannot get the full scope of what I want when producing an actual report. Additionally, the lack of a free trial is a downside.

“Regarding Datafold's governance and security, I rate it high because its migration agent uses LLMs for SQL translation and validation, which is a strong point.

“Additionally, there is a self-hosted deployment option available, so organizations with strict data residency or compliance requirements can run Datafold entirely within their own cloud environment, which could be either [AWS](#), [GCP](#), or [Azure](#), ensuring that data never leaves the perimeter of the organization.

“Datafold's output is highly accurate and very reliable. The migration agent's accuracy, which utilizes LLMs to convert SQL dialects, is excellent.

“Furthermore, the data diff, which is the most reliable AI-adjacent feature, is deterministic, not generative, and it compares actual data values mathematically rather than using inference.

“Because it does the comparison mathematically, the outputs are highly accurate and consistently reliable.

“Datafold is deployed in my organization as a cloud, specifically as a SaaS, which is fully managed by Datafold. This means that hosting, maintenance, and automatic updates are all managed by Datafold, making it simpler for us and easier to get started.

“The advice I would give others looking to use Datafold is that whoever is handling it, perhaps the head of IT, should have a sit-down with the analysts to ensure it fits into the stack that the organization is already conversant with.

“Datafold is purpose-built for SQL and warehouse-based analytic pipelines with DBT, so if the current stack does not include a data warehouse and DBT, I would advise them to evaluate alternatives first. I also recommend using it for CI/CD quality, not just for general observability, because Datafold excels at pre-merge testing and data diffs.

“Therefore, if the primary need is broad production or observability, the organization should also check out other options.

“I rate Datafold a seven out of ten overall..”

JohnBosco Obi
Data Analyst at Africa Initiative

[Read full review](#) 

Top Industries

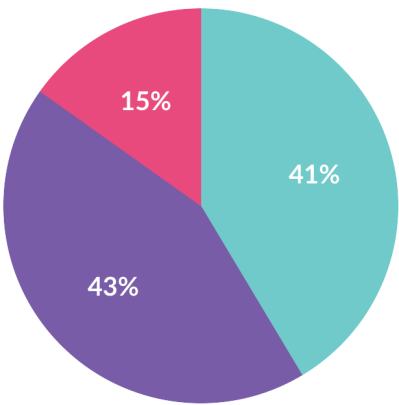
by visitors reading reviews



Company Size

by reviewers

by visitors reading reviews



Large Enterprise Midsized Enterprise Small Business

About this buyer's guide

Thanks for downloading this PeerSpot report.

The summaries, overviews and recaps in this report are all based on real user feedback and reviews collected by PeerSpot's team. Every reviewer on PeerSpot has been authenticated with our triple authentication process. This is done to ensure that every review provided is an unbiased review from a real user.

Get a custom version of this report... Personalized for you!

Please note that this is a generic report based on reviews and opinions from the collective PeerSpot community. We offer a [customized report](#) of solutions recommended for you based on:

- Your industry
- Company size
- Which solutions you're already considering

The customized report will include recommendations for you based on what other people like you are using and researching.

Answer a few questions in our short wizard to get your customized report.

[Get your personalized report here](#)

About PeerSpot

PeerSpot is the leading review site for cloud, AI, and business software. We created PeerSpot to provide a trusted platform to share information about software, applications, and services. Since 2012, over 22 million people have used PeerSpot to choose the right software for their business.

PeerSpot helps tech professionals by providing:

- A list of products recommended by real users
- In-depth reviews, including pros and cons
- Specific information to help you choose the best vendor for your needs

Use PeerSpot to:

- Read and post reviews of products
- Access over 30,000 buyer's guides and comparison reports
- Request or share information about functionality, quality, and pricing

Join PeerSpot to connect with peers to help you:

- Get immediate answers to questions
- Validate vendor claims
- Exchange tips for getting the best deals with vendor

Visit PeerSpot: www.peerspot.com

PeerSpot

244 5th Avenue, Suite R-230 • New York, NY 10001

reports@peerspot.com

+1 646.328.1944