



Databricks

Reviews, tips, and advice from real users



Powered by  **PeerSpot**

Contents

Product Recap..... 3 - 4

Valuable Features..... 5 - 10

Other Solutions Considered..... 11 - 13

ROI..... 14 - 16

Use Case..... 17 - 19

Setup..... 20 - 22

Customer Service and Support..... 23 - 25

Other Advice..... 26 - 28

Trends..... 29 - 30

About PeerSpot..... 31 - 32

Product Recap



Databricks

Databricks Recap

Databricks is utilized for advanced analytics, big data processing, machine learning models, ETL operations, data engineering, streaming analytics, and integrating multiple data sources.

Organizations leverage Databricks for predictive analysis, data pipelines, data science, and unifying data architectures. It is also used for consulting projects, financial reporting, and creating APIs. Industries like insurance, retail, manufacturing, and pharmaceuticals use Databricks for data management and analytics due to its user-friendly interface, built-in machine learning libraries, support for multiple programming languages, scalability, and fast processing.

What are the key features of Databricks?

- User-friendly interface: Simplifies operations and usability.
- Built-in machine learning libraries: Facilitates machine learning tasks.
- Support for multiple programming languages: Enhances flexibility.
- Scalability: Efficiently handles growing data needs.
- Fast processing: Improves performance and speed.
- Automated optimization: Reduces manual efforts.
- Data visualization: Provides insightful visuals.
- Collaborative features: Enhances teamwork.
- Delta Lake performance: Boosts data management.
- Seamless cluster management: Simplifies system operations.

What are the benefits or ROI to look for in Databricks reviews?

- Efficiency in handling large datasets: Ensures smooth processing.
- Interactive workspace environment: Improves user collaboration.
- Integration with platforms: Provides connectivity benefits.
- Performance optimization: Enhances overall system performance.
- Support for data governance and security: Ensures data integrity and protection.

Databricks is implemented in insurance for risk analysis and claims processing; in retail for customer analytics and inventory management; in manufacturing for predictive maintenance and supply chain optimization; and in pharmaceuticals for drug discovery and patient data analysis. Users value its scalability, machine learning support, collaboration tools, and Delta Lake performance but seek improvements in visualization, pricing, and integration with BI tools.

Valuable Features

Excerpts from real customer reviews on PeerSpot:



“Databricks offers various courses that I can use, whether it's PySpark, Scala, or R.”



ShubhamSharma7

Data Engineer at a engineering company with 1,001-5,000 employees



“Databricks has a Unified Catalog that assists with secured access and governance.”



Prabhakar Bonam

Data Platform Architect at KELLANOVA



“Having one solution for everything, from data engineering to machine learning, is beneficial since everything comes under one hood.”



Parag Bhosale

Senior Data Engineer at a logistics company with 51-200 employees

- ✓ “Databricks serves as a single platform for conducting the entire end-to-end lifecycle of machine learning models or AI ops.”



IshwarSukheja

Head CEO at bizmetric

- ✓ “The notebooks and the ability to share them with collaborators are valuable, as multiple developers can use a single cluster.”



Rama Subba Reddy Thavva

Solution Architect at Mercedes-Benz AG

- ✓ “I think Databricks is very good at facilitating AI and machine learning projects; they implement AI and machine learning models very well, and clients can run their models on Databricks.”



Lax Kas

Data Engineer at CRAFT Tech

- ✓ “It offers AI functionalities that assist with code management and machine learning processes.”



Monalisha Nayak

Senior Data Engineer at Shell

What users had to say about valuable features:

Databricks has a Unified Catalog that assists with secured access and governance. Additionally, serverless computing is crucial for our computing needs. Its collaboration features, such as data sharing capabilities, are also outstanding. Overall, the platform allows us to leverage cloud advantages effectively, enhancing our AI and ML projects.

Prabhakar Bonam

Data Platform Architect at KELLANOVA

[Read full review](#) 

“The most valuable features of Databricks are especially the Delta Lake and the Unity Catalog; those are the main features. The Unity Catalog is for data governance, and the Delta Lake is to build the lakehouse. Currently, they're coming up with workflow jobs, along with other supporting elements to create an end-to-end solution..”

Lax Kas

Data Engineer at CRAFT Tech

[Read full review](#) 

“I think it is difficult to determine which feature of Databricks I enjoy the most since there are many valuable features.

What's valuable about Databricks to my organization is that it is more cost-effective and provides better performance than the current AWS tools and services they offer..”

AvivCohen

Data Engineer at a tech vendor with 1,001-5,000 employees


[Read full review](#) 

“Unity Catalog is a feature I am currently using extensively. I am migrating many projects to Unity Catalog. MLflow, which I use for model registering and creating the lineage of models, is also valuable.

Additionally, Databricks serves as a single platform for conducting the entire end-to-end lifecycle of machine learning models or AI ops. I don't need to switch between various tools, making it an all-encompassing solution for development and research. I use the lake house and utilize features effectively..”

IshwarSukheja

Head CEO at bizmetric

[Read full review](#) 

“Databricks offers various courses that I can use, whether it's PySpark, Scala, or R. I can leverage all these courses in a single notebook, which is beneficial for clients as they can access various tools in one place whenever needed. This is quite significant.

I usually work with PySpark based on client requirements. After coding, I feed the Databricks notebooks into the ADF pipeline for updates. Databricks' capability to process data in parallel enhances data processing speed. Furthermore, I can connect our Databricks notebook directly with Power BI and other visualization tools like Qlik. Once we develop code, it allows us to transform raw data into visualizations for clients using analysis diagrams, which is very helpful..”

ShubhamSharma7

[Read full review](#) 

Data Engineer at a engineering company with 1,001-5,000 employees

“The notebooks and the ability to share them with collaborators are valuable, as multiple developers can use a single cluster. This reduces costs. The scheduling part is managed by Databricks itself, for example, when it is idle, it will automatically turn off. All these features are handled by Databricks, reducing costs. We do not need to schedule separately.

For example, on AWS EC2, we have to create a Lambda function or use System Manager templates to schedule EC2 and EMRs. Here, it is taken care of, saving significant resources.

Additionally, notebooks can be shared within the development team which saves effort. Developers can share their notebooks. Git and Azure DevOps integration on the Databricks side is also very helpful..”

Rama Subba Reddy Thavva

Solution Architect at Mercedes-Benz AG

[Read full review](#) 

Other Solutions Considered

Before Databricks, we were using the R Studio platform for our advanced analytics requirements. We switched to Databricks because it is an advanced platform compared to R Studio.

Prabhakar Bonam

Data Platform Architect at KELLANOVA

[Read full review](#) 

“We used Informatica before, which is perfect for data management solutions. We started using Databricks for its capabilities in data science and machine learning..”

Heba Ismail

Senior Data Engineer at a computer software company with 1,001-5,000 employees

[Read full review](#) 

“For companies focused solely on data transformation, transferring data between databases, and not tackling machine learning or deep learning problems, I recommend ADF. It would be sufficient and cost-saving compared to a full-fledged solution like Databricks. However, for data analytics and solving ETL problems, one should consider Databricks..”

Rama Subba Reddy Thavva

Solution Architect at Mercedes-Benz AG

[Read full review](#) 

“Before Databricks, I used Batch processing, Fargate, and possibly Kubernetes.

I switched from my previous solutions because they were either too expensive or too difficult to configure..”

AvivCohen

[Read full review](#) 

Data Engineer at a tech vendor with 1,001-5,000 employees

“I have considered other solutions besides Databricks, such as Snowflake, but we haven't explored it extensively yet.

We are still early in our Snowflake experience, so we don't know the pros and cons compared to Databricks..”

AvivCohen

[Read full review](#) 

Data Engineer at a tech vendor with 1,001-5,000 employees


“We looked into Azure Synapse as an alternative, as well as Azure ML and Vertex on GCP. Vertex AI would be the main alternative.

Some people consider Snowflake a competitor; however, we can't deploy Snowflake ourselves just like we deploy Databricks ourselves. We use that as an advantage when we sell Databricks to our clients. We say, "If you go with us, we are going to deploy Databricks in your environment in 15 minutes," and they really like it.

Lately Fabric was released and can offer quite a similar product as Databricks. Yet, the user experience, the CI/CD capabilities and the frequent release cycle of Databricks remains a strong advantage..”

Axel Richier

Tech Lead Consultant | Manager Data Engineering at Ekimetrics

[Read full review](#) 

ROI

Real user quotes about their ROI:

“I can't say if there's seen an ROI from the solution because I do not have exposure in that area, although I think the people who decided to implement Databricks might have done all this analysis and POCs..”

Lax Kas

Data Engineer at CRAFT Tech

[Read full review](#) 

“With the benefits we receive, the price is reasonable. However, it's important to have good use cases. If it's just for data ingestion, it might not be the best solution price-wise. For a lot of different tasks, including machine learning, it is a nice solution..”

Parag Bhosale

Senior Data Engineer at a logistics company with 51-200 employees

[Read full review](#) 

“I believe it is competitive across clouds. When it comes to big data processing, I prefer Databricks over other solutions. Cost-wise, it is very competitive. The setup process is straightforward, thanks to the use of Spark clusters. This allows for faster turnaround times with Databricks..”

IshwarSukheja

Head CEO at bizmetric

[Read full review](#) 

“It's a good return on investment for Databricks from a delivery perspective. Delivered multiple dashboards. So, it's quite a good return on investment. And being a small organization, everyone can use Databricks, and cost-wise, it's also good for small organizations..”

Rupal Sharma

Data Architect at Three Ireland (Hutchison) - Infrastructure

[Read full review](#) 

“Databricks stands out among most data platforms mainly because of its ease of use. The learning curve is not as steep, making it accessible for anyone to handle large-scale data processing on Databricks. This ease of use contributes positively to our return on investment. However, in our line of work, converting this efficiency into direct monetary gains can be challenging, given our nonprofit nature. .”

Dunstan Matekenya

Data Scientist at a financial services firm with 10,001+ employees

[Read full review](#) 

“Our ROI was of the order of USD \$75k per year for one deployment. We were able to switch our workloads from an onsite Hadoop cluster, billed to our department for more than USD \$100k per year, to a Databricks workspace in the cloud for a quarter of that expenditure.

Further, we were able to transparently and efficiently scale our queries to run under fifteen minutes per major analytics use case, while being subject to unstable queries and highly brittle data flow use cases from the in-house Hadoop cluster.

We are further reducing spending on our traditional RDBMS solution by offloading reporting workloads to the Databricks PySpark notebooks, which is reducing our expensive datacenter resources and freeing up RDBMS resources for OLTP loads. .”

Tristan Bergh

Data Scientist at a computer software company with 501-1,000 employees

[Read full review](#) 

Use Case

“A typical use case for the solution is to build the data lakehouse for the client because they have a variety of source systems, and they want to unify that data into the lakehouse platform, where they want to use the data for analytical purposes and insights..”

Lax Kas

Data Engineer at CRAFT Tech

[Read full review](#) 

“I use Databricks for various purposes, including data engineering, MLOps, machine learning training and deployment, the entire ML cycle, and dashboards. It serves different purposes for different projects..”

IshwarSukheja

Head CEO at bizmetric

[Read full review](#) 

“Databricks is used for transformations and streaming data processing. We utilize it primarily for data analytics, including the use of Delta Lake and Delta Live tables for ETL processes, dashboards for analysis, and the Unity catalog for role management..”

Monalisha Nayak

Senior Data Engineer at Shell

[Read full review](#) 

I am currently working as an IT architect. We have an AWS analytics platform, a cloud-based platform. We use Databricks for our AI/ML requirements and also the Databricks platform. For the past year, we have been using Databricks for our data scientist community to build their apps.

Prabhakar Bonam

Data Platform Architect at KELLANOVA

[Read full review](#) 

“I am working as a data engineer at Fractal. On a daily basis, I work on Azure Cloud, and I use Databricks frequently. We have EDF pipelines and utilize Synapse for our daily tasks..”

ShubhamSharma7

Data Engineer at a engineering company with 1,001-5,000 employees

[Read full review](#) 

“We work on three platforms. Databricks is hosted on Azure for us, so we work with ADFS, Azure Data Factory, and also the AWS Cloud. We work for some customers..”

Rama Subba Reddy Thavva

Solution Architect at Mercedes-Benz AG

[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 setup of Databricks could be complex. We set it up with AWS as the backend. It's crucial to have both cloud-specific knowledge and how to configure its features in the platform.

Prabhakar Bonam

Data Platform Architect at KELLANOVA

[Read full review](#) 

“The initial setup of the Databricks solution is reasonably fair enough. It doesn't give any trouble to implement the solution, and I think it's fairly easy to set up and work on Databricks..”

Lax Kas

Data Engineer at CRAFT Tech

[Read full review](#) 

“I use infrastructure as code on the cloud to deploy the infrastructure. I have all the Git repositories and code repositories for deploying the code and models in the workspace. My setup includes a shared workspace, shared clusters, and integration with Unity Catalog..”

IshwarSukheja

Head CEO at bizmetric

[Read full review](#) 

“The initial setup is not that difficult. I rate the ease of setup a seven out of ten. The solution is cloud-based. We use native services like Data Factory for orchestration. Sometimes, the customers require us to use Amazon as the cloud provider instead of Azure..”

PraveenS

Design Engineer at Cyient Limited

[Read full review](#) 

“When we have administration experience, the solution is not difficult to deploy. Technically, however, it's difficult because governance is more complex. For example, I have two warehouses on Databricks, which are clusters in this workspace, and we have to switch from workspace to workspace to have all this information. There is a system table that has all this, but I don't know if everyone can use these tables..”

Nabil Fegaier1

Chief Executive Officer at dotFIT, LLC

[Read full review](#) 

“The implementation is very simple to set up. That's why we choose it over many other tools. Its Terraform provider is our way-to-go for the initial setup as we are reusing templates to get a functional workspace in minutes.

Usually, we have two to five data engineers handling the maintenance and running of our solutions..”

Axel Richier

Tech Lead Consultant | Manager Data Engineering at Ekimetrics

[Read full review](#) 

Customer Service and Support

“I rate the technical support as fine because they have levels of technical support available, especially partners who get really good support from Databricks on new features. For us, it's so far so good with no problems, and I would rate the support quality as eight out of ten..”

Lax Kas

Data Engineer at CRAFT Tech

[Read full review](#) 

“We mostly had limited data support required from Databricks. Whenever we did need support, within two or three days the problem was solved. I would rate them ten out of ten..”

Rama Subba Reddy Thavva

Solution Architect at Mercedes-Benz AG

[Read full review](#) 

“There was a training session from Databricks where they explained how to use it. We never had to contact them because they had already given us proper training on the platform..”

Jithin James

[Read full review](#) 

Financial Analyst 4 (Supply Chain & Financial Analytics) at Juniper Networks

“Databricks' customer support is very good. They have a lot of ways in which they interact with vendors and service partners across the globe. They have periodic touch-up sessions with vendors, where their engineers answer your questions..”

Avadhut Sawant

[Read full review](#) 

Consulting Architect at a computer software company with 10,001+ employees

“I've had very good experiences with technical support where they answer me in a couple of hours. Sometimes it takes a bit longer. It's usually a matter of days, so it's very good overall.

Even if it took a bit of time, I got my answer. They never left me without an answer or a solution..”

Axel Richier

[Read full review](#) 


Tech Lead Consultant | Manager Data Engineering at Ekimetrics

“Databrick's support is great. If we need any support, they are very quick with it. And they genuinely want you to use Databricks. So, whatever we ask them, they come up with multiple solutions to problem statements. That's really good.

Overall, the customer service and support are very good..”

Rupal Sharma

Data Architect at Three Ireland (Hutchison) - Infrastructure

[Read full review](#) 

Other Advice

I would recommend Databricks as an advanced analytics platform. It offers features to quickly pull data from various sources required for report generation. Overall, I rate this solution as seven or eight out of ten.

Prabhakar Bonam

Data Platform Architect at KELLANOVA

[Read full review](#) 

“I would rate the solution seven out of ten. That rating also depends on how we have the contract with Databricks.

It's still a solid and good rating. I work as a data engineer and Databricks engineer. .”

Parag Bhosale

Senior Data Engineer at a logistics company with 51-200 employees

[Read full review](#) 

“I would rate the overall product eight out of ten.

Everything is probably good as far as I have used it, but there's room for improvement in cluster integration. Enhancing cluster capabilities while keeping costs lower would be beneficial..”

ShubhamSharma7

Data Engineer at a engineering company with 1,001-5,000 employees

[Read full review](#) 

“The product rating is nine out of ten.

Databricks serves as a single platform that can handle numerous end-to-end machine learning tasks. The configuration is simple, scalability is excellent, and monitoring cluster utilization facilitates informed business decisions.

It's easy to schedule jobs, pipelines, and handle multiple use cases in parallel, providing countless benefits..”

IshwarSukheja

Head CEO at bizmetric

[Read full review](#) 

“My deployment model for Databricks is limited as I'm not a heavy user.

I am not the person who purchased Databricks, but it was possibly acquired through the [AWS Marketplace](#).

I may not have utilized Databricks machine learning capabilities.

My experience with the pricing and licensing model is that it remains relatively expensive. Though it's less expensive than [AWS](#), we still need a more cost-effective solution.

I would rate Databricks overall a nine out of ten..”

AvivCohen

Data Engineer at a tech vendor with 1,001-5,000 employees

[Read full review](#) 

“My relationship with the vendor is that I'm not a partner of Databricks; I work for a client where we use the Databricks software for implementing the solutions.

“My clients are usually enterprise-level organizations, but the area where they're implementing is medium level here, although it might go into enterprise level in the future.

“Regarding the price of Databricks, I don't involve myself in those decisions.

“I think Databricks is very good at facilitating AI and machine learning projects; they implement AI and machine learning models very well, and clients can run their models on Databricks. I believe they are in a better place compared to competitors such as [Snowflake](#), and they are tying up with important companies such as SAP and Palantir.

“Based on my experience, I would recommend Databricks to other people. Overall, I would rate this solution as one of the best, about eight out of ten, although I might not know some of the pitfalls; it's based on use case to use case, but for us, it's working well..”

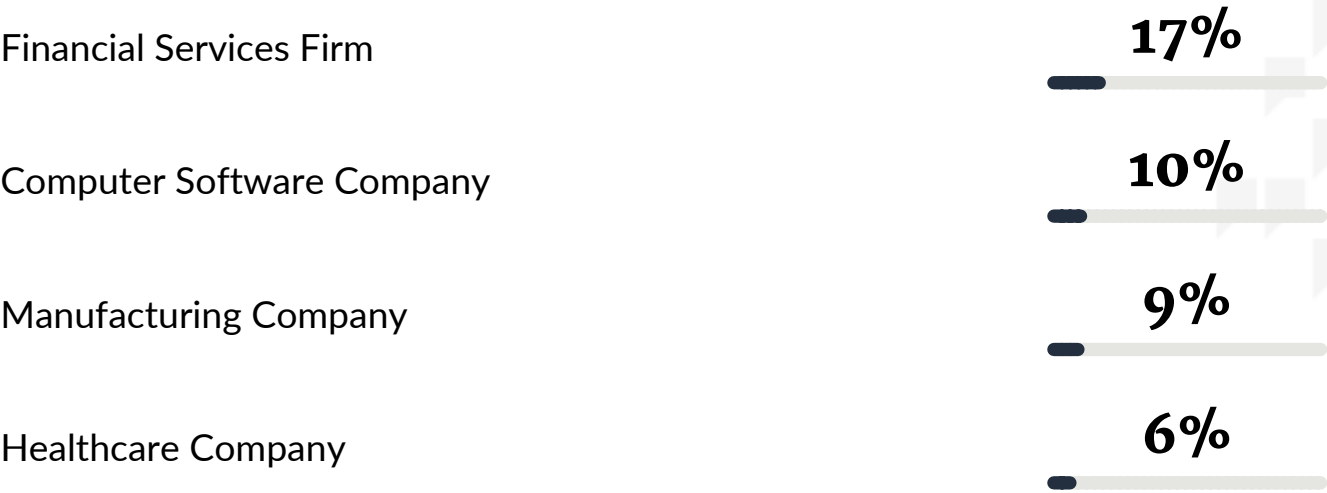
Lax Kas

Data Engineer at CRAFT Tech

[Read full review](#) 

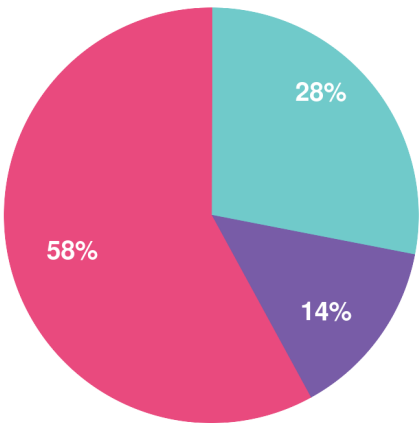
Top Industries

by visitors reading reviews

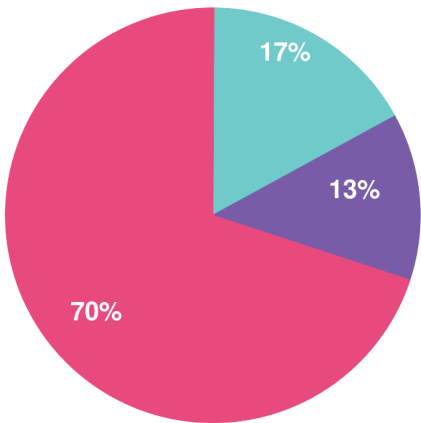


Company Size

by reviewers



by visitors reading reviews



 Large Enterprise  Midsize 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 software running on AWS and other platforms. 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