EBOOK

Databricks and Tableau: The Best Data Warehouse is a Lakehouse

Bringing together business intelligence, real-time analytics, and ML for data analysts.

databricks ✕ +tableau from Salesforce

Are you getting the most out of your organization's data?

Siloed data is a challenge faced by organizations of all sizes, from the largest multinationals to small businesses. Data from multiple sources can cause inconsistent results across the organization. That's because organizations often store different types of data in separate data platforms that are designed for different purposes. Data warehouses store structured tables of historical data that have already been processed and can be used for traditional analytics (reporting), such business intelligence (BI). Data lakes store massive unstructured, detailed data that can be used for advanced analytics (predictive), such as artificial intelligence (AI) and machine learning (ML). Many organizations lack access to ML models that business users can leverage in existing tools.

Everyone who handles data, from data architects and data engineers to data analysts and data scientists, struggle with this chasm between incompatible—yet essential—platforms. In a recent survey, 80% of business leaders see data as critical to their organization's decision-making. But 41% are challenged because they think their data is too complex or inaccessible. However, continuing to store data in different places is expensive and fails to take full advantage of valuable insights that could be gained through your organization's data. Furthermore, attempting to implement data streaming at scale while manually building complex operational tooling is a daunting task for any data team.

> In this ebook, discover how the Databricks Data Intelligence Platform, through a simple, native connection from Tableau, can unify your organization's data in one place for self-service analytics.

¹Salesforce, "73% of Business Leaders Believe Data Reduces Uncertainty and Drives Better Decisions—So Why Aren't They Using It?," February 22, 2023
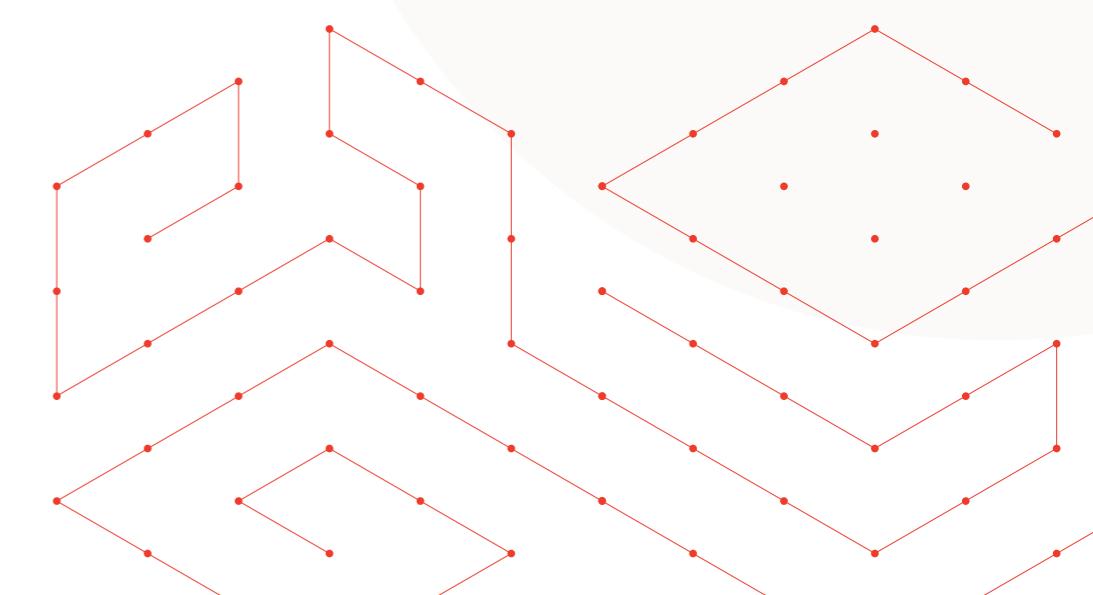
The simulator we've built on Databricks and Tableau enables us to test every business hypothesis before we take action. We could never gain this level of insight if our data were still scattered among data silos.

Marcia Freitas
Group Product Manager, iFood
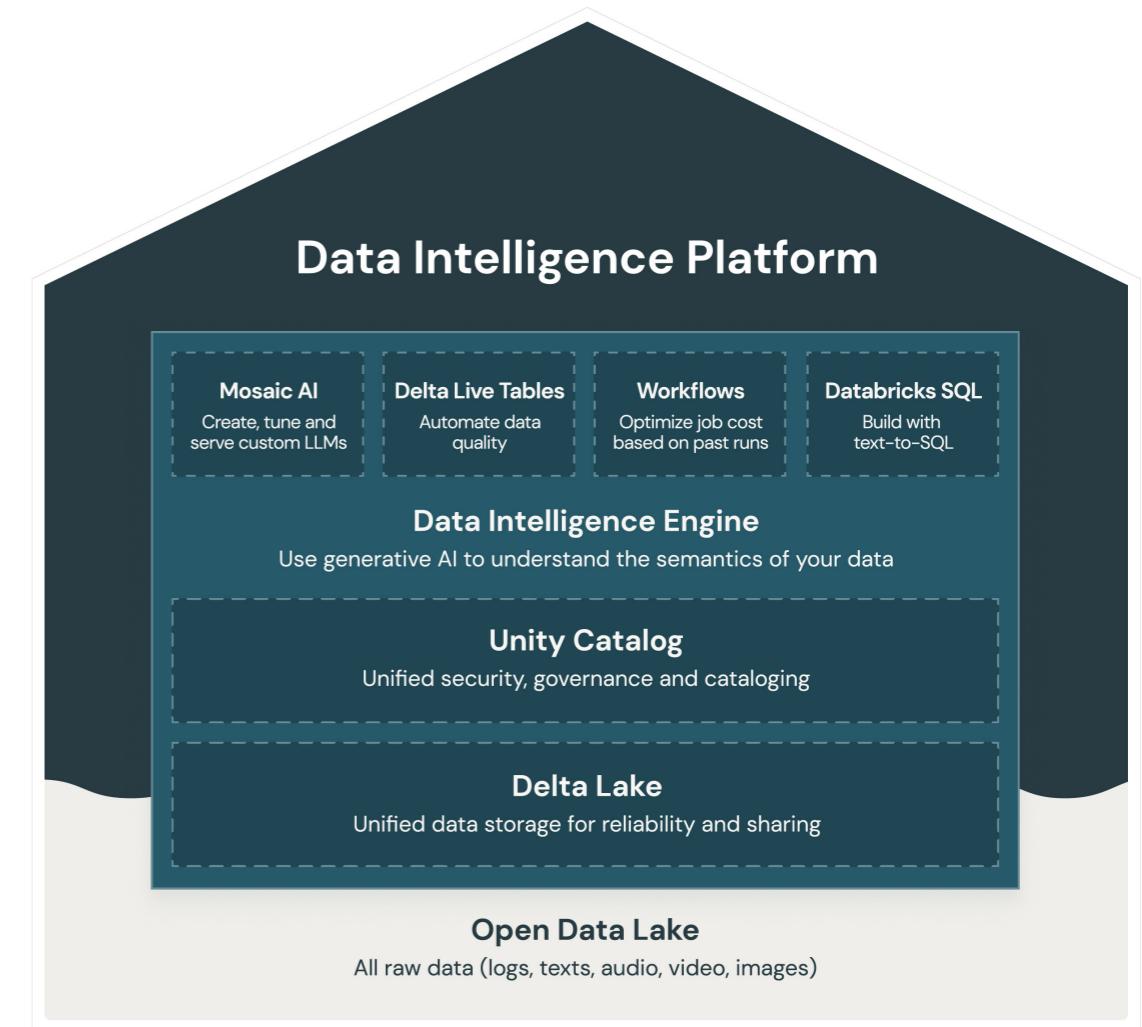
LEARN MORE ›

Get the best of data lakes and data warehouses with a lakehouse

Clearly, both data lakes and data warehouses play critical roles as data users across your organization attempt to learn more about its business and customers. The Databricks Data Intelligence Platform combines the best elements of data lakes and data warehouses, delivering consistent results and enabling your team to deliver on data and AI initiatives faster. A lakehouse cuts costs, too, because it eliminates the need to store and pay for multiple copies of data.

Databricks SQL (DB SQL) is the intelligent data warehouse. Built with DatabricksIQ, the Data Intelligence Engine that understands the uniqueness of your data, Databricks SQL democratizes analytics for technical and business users alike. Businesses can innovate faster with an intelligent and auto-optimizing platform that provides the best price/performance in the market. As part of the Databricks Data Intelligence Platform, Databricks SQL benefits from the simplicity, unified governance, and openness of lakehouse architecture.

With Databricks SQL, there's just one toolset for Python and SQL, notebooks and integrated development environments (IDEs), batch and streaming, and all major cloud providers. Costs are reduced as your team only needs to be trained to manage one system. And because the Databricks Data Intelligence Platform is built on open source and open standards, your data estate is simplified, eliminating those silos that complicate data and AI. Importantly, your data is always under your control, without proprietary formats or a closed ecosystem, with reliability and extremely fast query performance.

Data Intelligence Platform

Mosaic AI	Delta Live Tables	Workflows	Databricks SQL
Create, tune and serve custom LLMs	Automate data quality	Optimize job cost based on past runs	Build with text-to-SQL

Data Intelligence Engine
Use generative AI to understand the semantics of your data

Unity Catalog
Unified security, governance and cataloging

Delta Lake
Unified data storage for reliability and sharing

Open Data Lake
All raw data (logs, texts, audio, video, images)

Databricks SQL set a new world record for TPC-DS, the gold-standard performance benchmark for data warehousing. Databricks SQL outperformed the previous record by 2.2x.

LEARN MORE ›
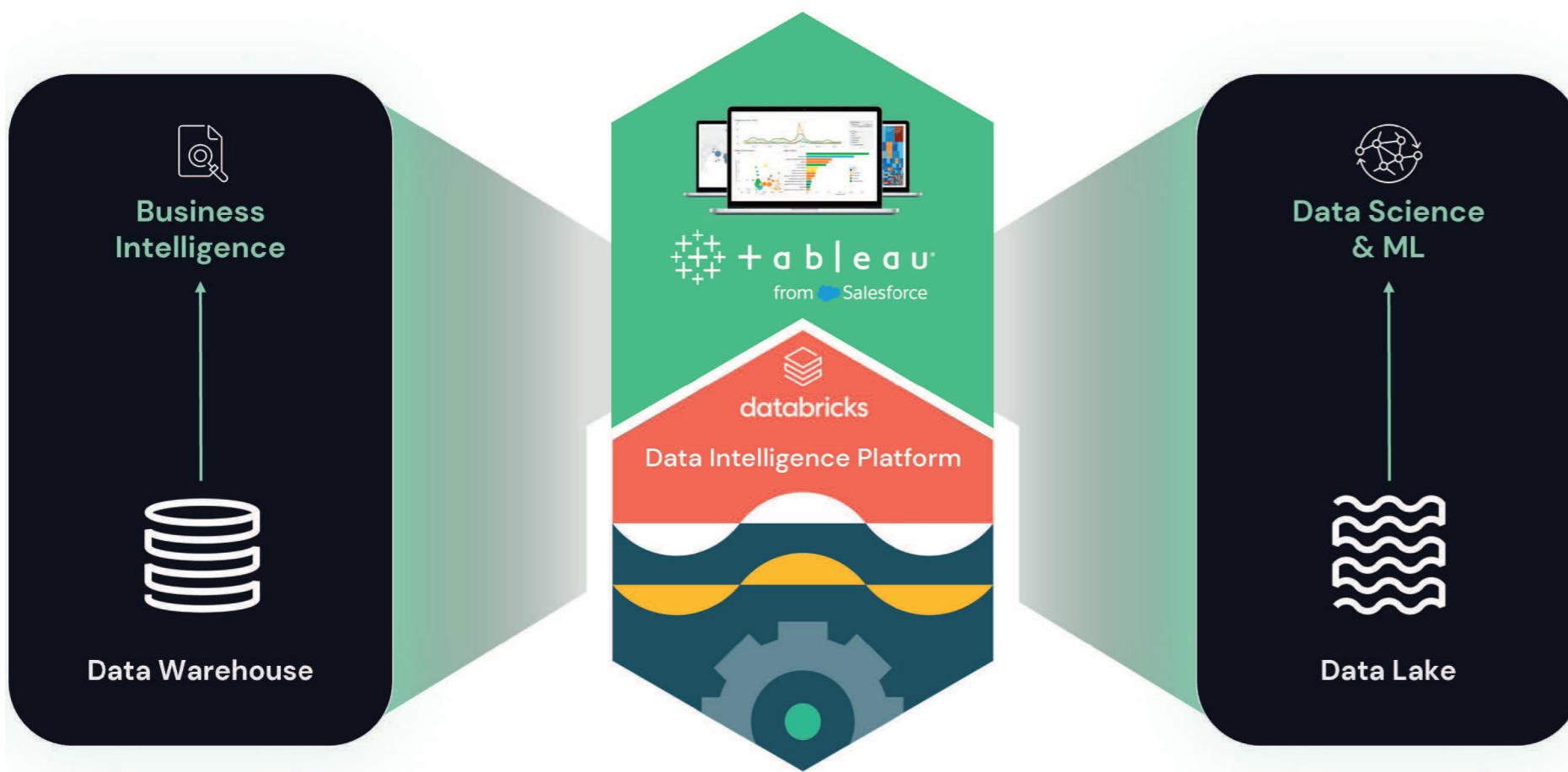
databricks ✕ +ableau from Salesforce

See and understand all your data with Databricks and Tableau

Now all that unified data is great, but it's only truly valuable when you can see and understand it. That's where Tableau, the visual analytics platform, comes in with an optimized integration with Databricks.

Users across your organization can connect directly to the Databricks Data Intelligence Platform from Tableau so they can explore validated, real-time data and then quickly share the insights.

Conveniently, there are likely data users throughout your organization already comfortable using Tableau to analyze and visualize data. Coupled with Databricks as a unified single source of truth, you can establish a data culture based on trusted data sources and governed self-service analytics.

Unlock real-time cross-platform collaboration with Delta Sharing Tableau connector

You may also need to share data outside of your organization—with partners, customers, or suppliers. Databricks and the Linux Foundation developed Delta Sharing—the first open—source agnostic approach to data sharing across data, analytics, and AI. With the Delta Sharing Tableau connector, organizations can easily share live data with their external partners and consume it directly in Tableau to facilitate real-time, data-driven decision-making. This expands the open ecosystem for business intelligence use cases.

The Delta Sharing Tableau connector's capabilities extends to all Tableau customers. Users can connect directly to the shared data from Delta Sharing within Tableau Desktop and Server to explore the newest data. The connector is fully integrated with Tableau's data management features, including Tableau Prep and Tableau Catalog.
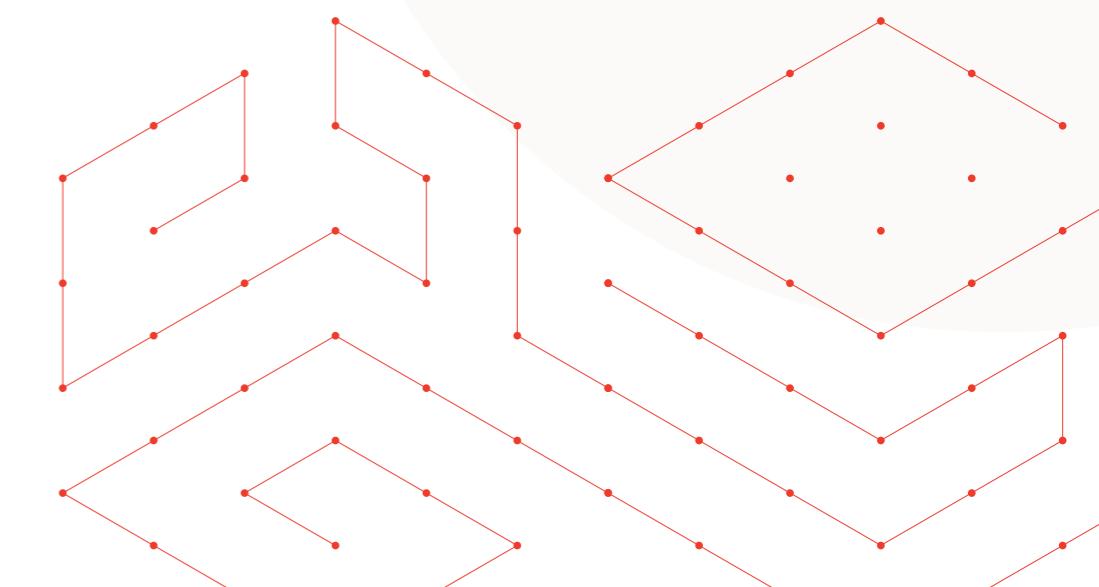
For data providers, the connector expands the reach of your data to end users on Tableau without sacrificing security while leveraging the built-in data governance from Unity Catalog. For data consumers using Tableau, the connector provides easy access to the rich Delta Sharing data ecosystem, including data from the Databricks Marketplace. This makes it effortless to tap into a vast array of shared data sources.

DOWNLOAD THE CONNECTOR ❯

FOLLOW THE TECHNICAL DOCUMENTATION ❯

Our experience has been really smooth … The Delta Sharing Tableau connector means our customer can simply select their data tables, download the credential, and get to work immediately. It is a game changer."

Caleb Benningfield
Principal Architect, Amperity
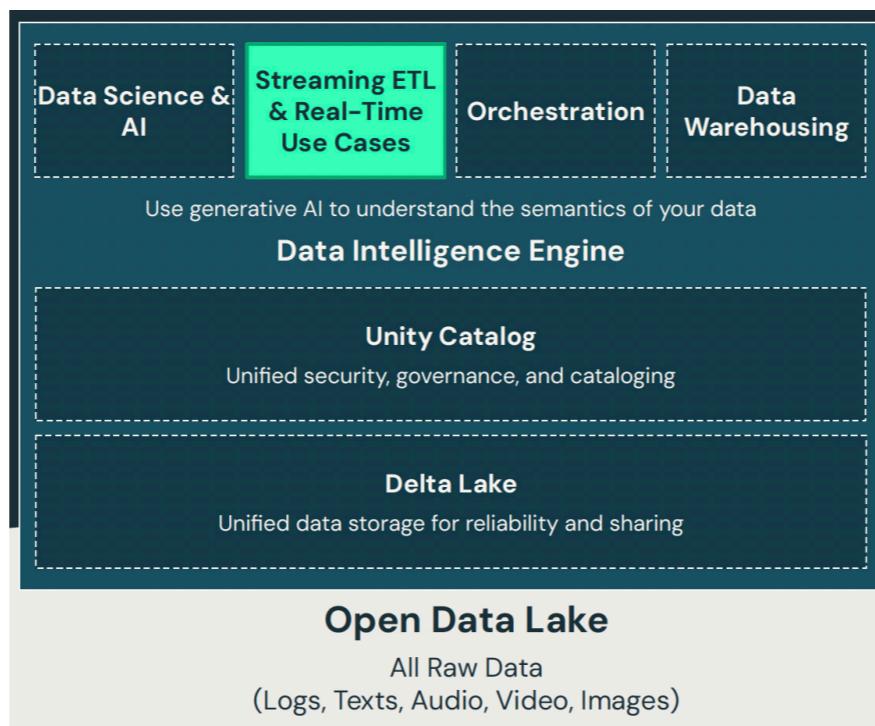
Take real-time analytics with streaming data to new heights

Does your organization have difficulty giving all of its employees access to streaming data for making decisions in real time? It's an expensive proposition, often with the need to train teams on new languages and tooling.

Databricks and Tableau have taken the heavy lifting out of real-time analytics with streaming data. Databricks lands streaming data directly in the lakehouse, and makes it instantly available to Tableau, enabling real-time dashboards for manufacturing, retail, supply chain, fleet management, and other use cases. Because you are always connecting to the latest data, your team has faster—and more reliable—time to insight.

With Tableau connected to Databricks, you can:

- **Easily deploy** and manage your real-time pipelines and applications in production.

- **Leverage built-in streaming functionality** like automatic checkpointing, fault tolerance and recovery, and performance optimization with powerful serverless compute.

- **Leverage Databricks Unity Catalog** to give your lakehouse one consistent governance model for all your streaming and batch data, simplifying how you discover, access, and share real-time data.

Data Science & AI | **Streaming ETL & Real-Time Use Cases** | Orchestration | Data Warehousing

Use generative AI to understand the semantics of your data

Data Intelligence Engine

Unity Catalog
Unified security, governance, and cataloging

Delta Lake
Unified data storage for reliability and sharing

Open Data Lake
All Raw Data
(Logs, Texts, Audio, Video, Images)

incite

Incite's Rapid OEM Automotive Data (ROAD) platform enables the world's largest automakers to build safer autonomous vehicles. Incite chose the Databricks Data Intelligence Platform as a core technology in its ROAD platform and then Tableau to provide the BI and analytics functionality of its platform.

THE RESULTS

75%+
Improvement in data acquisition and ingestion cycle time

80%+
Reduction in advanced driver assistance systems (ADAS) feature development cycles

50%
Reduction in cost per file processed

READ MORE >

databricks ✕ +ableau from Salesforce

Machine learning made easy for data analysts

When CVS Health was looking to use machine learning to analyze its customer data to implement different experiences, experiments, and a variety of segments for personalization at scale, the company turned to Databricks and Tableau.

"We're not dealing with a typical grocery store customer," says Raghu Nakka, Sr. Analytics Advisor at CVS Health. "… the predictor of the behavior could be unpredictable, which is why our data dimensionality leads to overfitting issues for any kind of machine learning model."
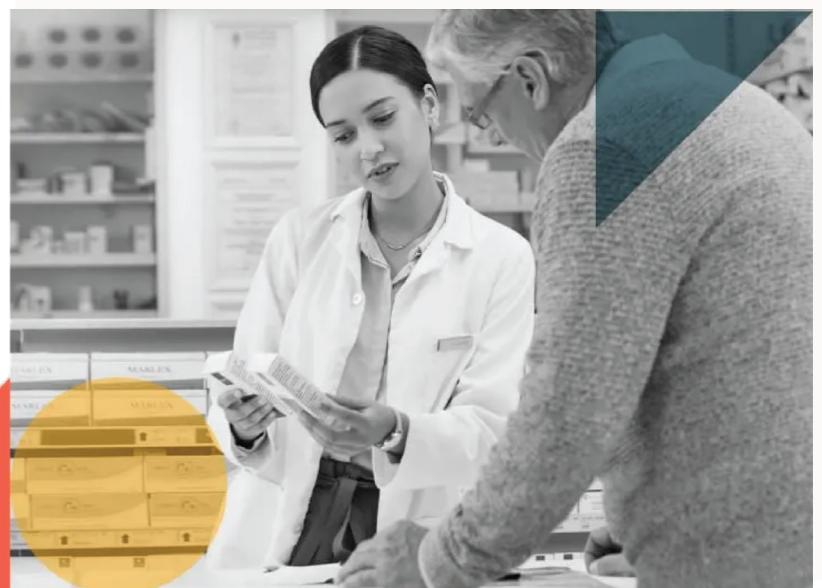
Data analysts can leverage machine learning models from Databricks directly in Tableau for advanced analytics use cases. This simplified approach to machine learning gives teams the power to prepare and process data, streamline cross-team collaboration, and standardize the full ML lifecycle from experimentation to production. A simple connection to the Databricks Data Intelligence Platform al–lows users to stream Internet of Things (IoT) data into a lakehouse while providing zero latency dashboards in Tableau.

With a new, collaborative environment, CVS Health teams were finally able to work together: data engineering on building faster data pipelines, data scientists on training ML models more efficiently, and data analysts on visualizing financial and operational metrics for smarter internal decision making through Tableau. Says Nakka: "Databricks helps us to predict customer behavior and then make our offers more relevant and increasingly personal."

> **Databricks helps us to predict customer behavior and then make our offers more relevant and increasingly personal.**
>
> **Raghu Nakka**
> Sr. Analytics Advisor, CVS Health

Start using one unified, open, and scalable platform for all your data

Ready to connect Tableau to the Databricks Data Intelligence Platform to accelerate insights? It's easy to quickly get started so you can see and understand all the data in your lakehouse.

LEARN MORE ❯

DOWNLOAD STARTER KIT ❯

databricks ✕ +ableau from Salesforce