

# Collibra Data Quality & Observability with Pushdown for Databricks

Quickly and securely surface data quality issues at the source

Poor data quality translates to unreliable insights, which can ultimately impact business growth. A lack of visibility into the quality of data residing in data warehouses, data lakes, and data pipelines, makes it a tedious job to quickly identify and fix data quality issues. When that data resides in a data cloud, it becomes even more difficult to remediate the problem because you need to pull the data out to a testing platform, which can be time-consuming and costly. As a result, delivering trusted data for regulatory compliance, analytics, and AI in a scalable and continuous way is a challenge.

Collibra Data Quality & Observability enables you to monitor the quality of your data across your Databricks environment, and automatically generate alerts for data issues found. It leverages machine learning (ML) to generate adaptive checks and rules that automate anomaly detection without the need for additional code. This reduces manual processes, enabling rule writers to focus on higher-impact tasks. Collibra makes it easy to manage and scale data quality across the enterprise, delivering high-quality data for operations, analytics, and AI.

---

**“By 2024, 90% of data quality technology buying decisions will focus on ease of use, automation, operational efficiency and interoperability as the critical decision factors.”**

2022 Gartner® Magic Quadrant™ for Data Quality Solutions

---

With the addition of Pushdown for Databricks you can now address data quality issues by doing all of the data quality processing at the source. By pushing auto-generated SQL queries to the Databricks Data Cloud to perform the required processing, you eliminate reliance on an intermediary compute plane. And because the data never leaves your Databricks environment, processing is faster, more efficient and more secure.

## Benefits



Reduce data vulnerability with faster and more secure data quality processing in Databricks



Increased efficiency by tapping into the built-in scalability, flexibility, and elasticity of Databricks



Reduce latency and remove Spark dependency by leveraging your existing Databricks investment



Identify and fix data quality issues faster with complete visibility into the quality of data in your Databricks environment



Improve data governance with data discovery and enforcement for compliance

## High-quality Databricks data



Drive trusted business decisions



Streamline data operations



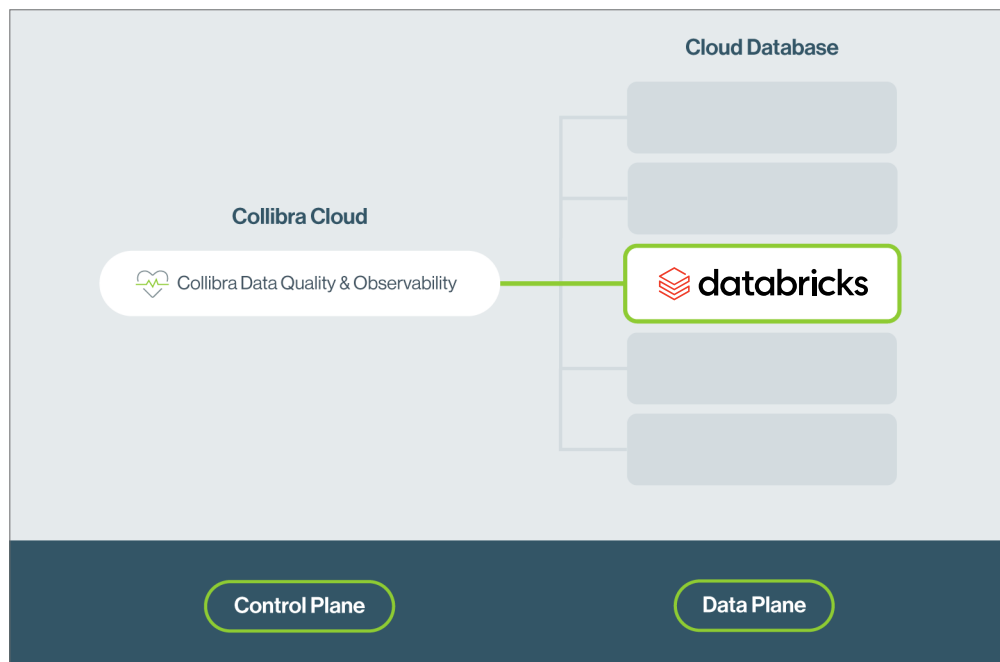
Improve productivity



Comply with regulations in a scalable way



Achieve faster results and ROI while lowering costs



Ready to identify and fix data quality issues faster?  
Get started by [requesting a demo](#) or [starting a free trial](#) today.