Collibra Data Quality & Observability

Surface data issues in minutes instead of months

Data teams are often constrained by manual rule writing and management, with limited data connectivity and a siloed view of data quality. As a result, they lack an enterprise data quality foundation to respond to regulatory, analytics and AI demands in a scalable and continuous way. According to the MIT Sloan Management Review the cost of bad data is an astonishing 15% to 25% of revenue for most companies.

Collibra Data Quality & Observability leverages machine learning to generate adaptive checks and rules that proactively identify data quality issues across diverse data sources. Predictive, continuous, and self-service data quality helps rule writers focus on high-impact tasks such as eliminating the root causes of data errors, and helps business users access high quality data for their analytics and AI modeling work. Collibra replaces manual processes with automated data monitoring and rule management that make it easy to manage and scale data quality across the enterprise.

Customer stories

- A top 10 bank reduced 60% of their manual data quality workload leading to $1.7M+ cost savings

- A top three healthcare organization saved 2000 hours during a cloud migration requirement

- A top insurance organization completed a regulatory audit in just four weeks, instead of two years with their previous tool

> **With Collibra Data Quality & Observability, I spend now less than 30% of my day on identifying data issues and fixing broken pipelines.**
>
> A data engineer at a Fortune 2000 company

> **With Collibra Data Quality & Observability, I spend 60% less time on manual data discovery and rule management.**
>
> A data steward at a Fortune 2000 company

Benefits

Reduce manual rule management efforts
Reduce repetition and guesswork in data quality rule management by 50–70%

Gain efficiency through streamlined data operations
Continuously monitor data pipelines to reveal blindspots in your data operations

Respond to regulations in a scalable way
Automate data discovery, validation, and compliance with intelligent tags and 100s of out-of-the-box rules

Deliver data trust and transparency
Unite data engineers and data stewards to deliver trusted data for downstream users

Product Highlights

Auto-discovered and adaptive rules
Leverage machine learning to autogenerate SQL-based, non-proprietary, explainable and adaptive data quality rules

Horizontal and vertical scalability
Scan large, diverse databases and files with Spark-based parallel processing

Anomaly detection
Detect outliers, patterns, data freshness and schema change to ensure accuracy and performance of analytics models over time
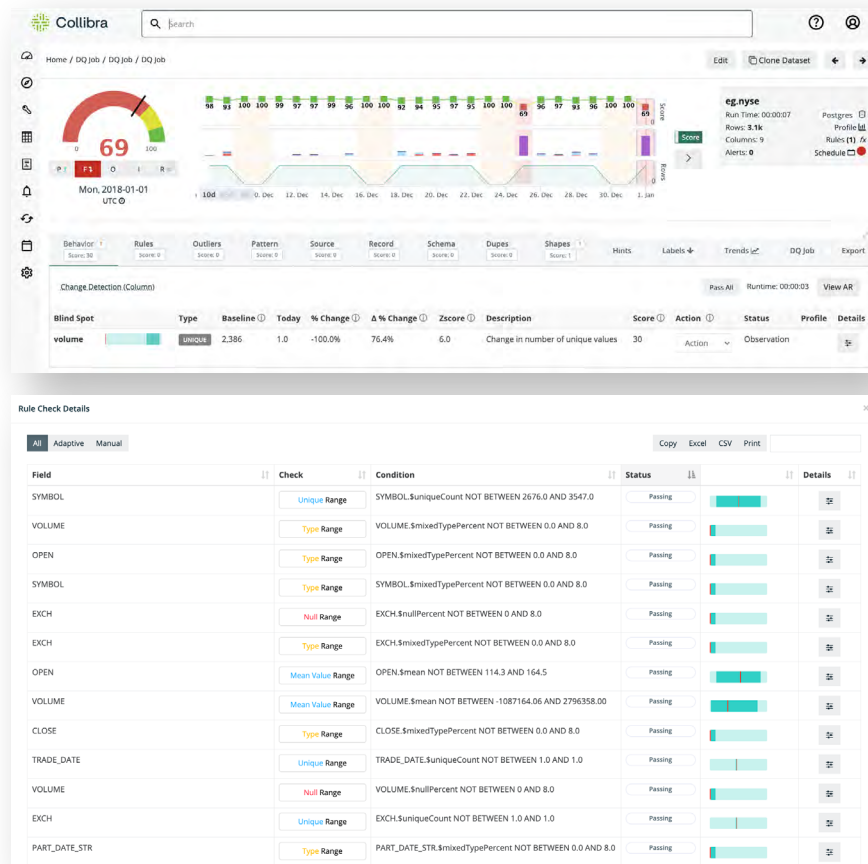
Data discovery and enforcement
Automatically discover and validate sensitive as well as non-sensitive data with pre-built, industry-specific rules

Self-service data quality
Empower technical and business users with a unified scoring and personal alert system to proactively identify and resolve the issues

Fast and flexible to meet your evolving needs
Run data quality and observability on any cloud and connect to more than 40 databases and file systems while scanning data via pushdown or pull-up processing

Replace manual processes with automated rule management