Pentaho

Pentago Data Quality

Ensure Trusted Data Throughout the Pipeline for all Your Business Initiatives

The Business Risk You Might Be Unaware Of: Poor Data Quality

Poor data quality is a hidden threat that can derail even the most strategic business initiatives and AI projects. You simply can't deliver trusted insights or drive successful outcomes if the data underneath is incomplete, inconsistent, or inaccurate.
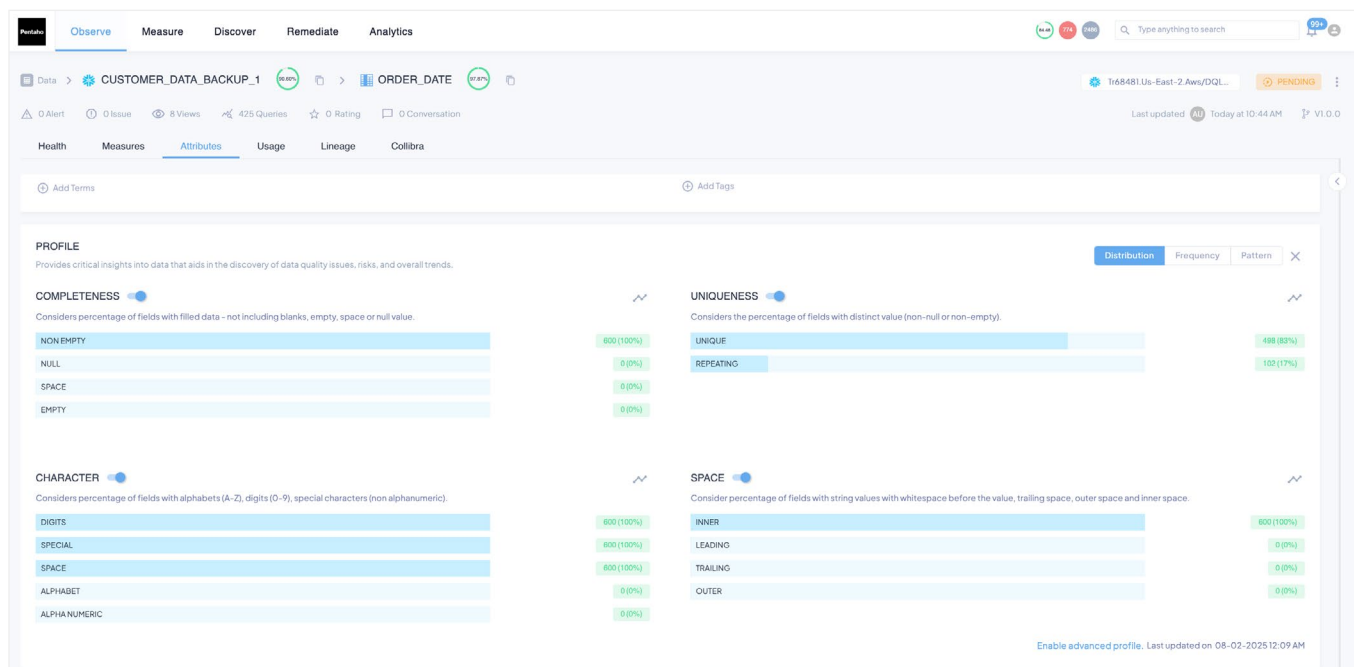
Pentaho Data Quality helps eliminate a lot of the data challenges that impede businesses today. It is built into our smart and simple data management platform—so you can become truly data-fit for AI. At Pentaho, we believe that simplifying complexity is what unlocks business value, which is why we've made it extremely easy to manage data quality whether it's on-prem, hybrid, or in the cloud.

Secure by Design

Pentaho keeps your data safe. Only metadata is pulled into the data quality platform, and your actual data remains secure at the source. There's no data duplication or storage within the platform. Instead, quality checks are executed directly in the source system. With over 20 prebuilt connectors for popular data warehouses, data lakes, and BI tools, Pentaho Data Quality fits seamlessly into your environment.
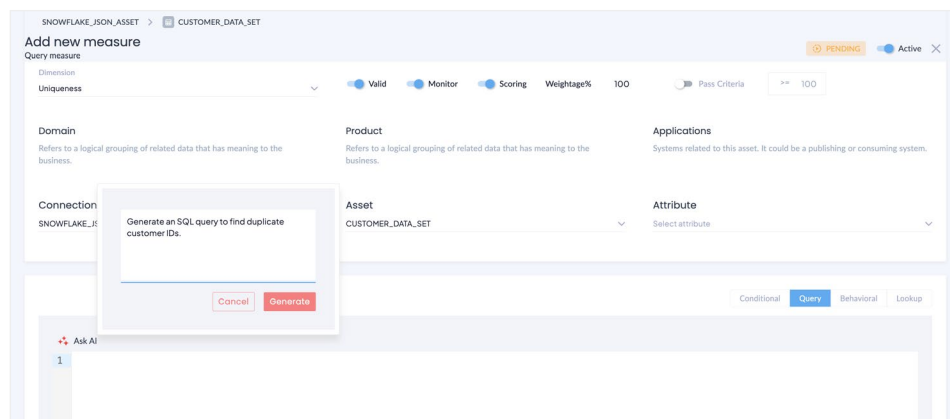
Data Profiling with Built-In Intelligence

Pentaho makes it simple to assess your data for completeness and uniqueness. Using AI/ML, the platform automatically flags outliers and patterns that don't align. When anomalies are detected, alerts are generated instantly. The system can even auto-tag these issues and generate reusable rules to prevent similar errors in the future.

Easily Define and Apply Rules

Choose from over 250 prebuilt data quality rules or create your own using a no-code/low-code interface. Business users can set up powerful conditional rules without writing a single line of code. In addition, AI can recommend rules based on your data profiling results.

Stay Ahead with Real-Time Alerts

Alerts can be set up by you or automatically by Pentaho Data Quality. Set a manual threshold for the check or let the platform decide, and it will automatically create high and low priority alerts. These can also be routed to a third-party service management tool for remediation.

Built-In Data Observability

The data observability part of the platform checks the health of the data based on three key things:

- Is the data fresh?
- Are there any changes in the schema?
- Are there any changes in the volume?

Data is monitored for anomalies on a continuous basis and visualized in charts that inform DataOps teams about issues in the pipeline that need to be fixed.

The Foundation of Smarter Business Decisions

With Pentaho Data Quality you can ensure trusted data for every initiative, whether that is regulatory compliance, reporting, analytics, or AI. Pentaho gives you the power to improve data accuracy and reliability, observe data as it moves through the pipeline, and enhance data transparency while reducing risk. At Pentaho, we know that making anything simple takes hard work. That's why smart simplicity is embedded in every aspect of our platform.

Discover Pentaho Data Quality →

Pentaho Data Quality Technical Specifications and Compatible Components Learn More

About Pentaho

Trusted by over 70% of the Fortune 100, Pentaho's end-to-end data integration and intelligence platform enables the foundational and AI-grade data fitness that drives faster and more confident decision-making.

Pentaho

P-BTD-DS-PDQ-22July25-A