Pentaho | HITACHI Inspire the Next

Pentaho Data Quality

**Build Business Trust Through Data Quality
Unlock Hidden Value with Intelligent Automation**

Pentaho Data Quality Helps Businesses Optimize Their Data Assets

Pentaho Data Quality serves as a comprehensive data quality solution, offering a scalable framework and ecosystem for data classification, observability, quality standards compliance and governance. It enables business users to discover, observe, measure and deliver reliable and accurate data sets, facilitating innovative generative AI, improved daily operations, and better decision-making.

Unlock Hidden Value with Intelligent Automation

Modern data teams employ various architectures, ranging from Cloud Warehouses to Data Lakehouses, Data Virtualization, and Data Fabrics or Mesh, to cater to their diverse needs for speed and computational power. To harness data for improved day-to-day operations and strategic decisions, organizations must consider both the perspectives of data experts and business leaders.

Businesses, regardless of size and industry, require a scalable framework to gain a deeper understanding of data, manage data quality, and support data governance. Data collected and used by businesses often resides in multiple silos, spread across diverse repositories, databases and applications. This data fragmentation results in data quality issues, creating inconsistencies and generating duplicate, incorrect and missing data.

The lack of data quality makes it challenging to ascertain the accuracy of data assets, discover hidden relationships and extract business value for decision-making.

Implementing a modern data quality platform addresses these challenges by ensuring data accuracy, completeness, consistency, governance and privacy. As businesses increasingly seek to leverage generative artificial intelligence (AI), emerging AI technologies like GPT (Generative Pre-trained Transformer) and LLM (Large Language Model) require quality data to deliver accurate and meaningful insights. Pentaho Data Quality is essential because the review, understanding and correction of large and diverse datasets can only be effectively achieved through intelligent automation.

Data Reliability

Established centralized and reliable master data for consistent governance.

Data Modernization

Enhance accuracy and reliability during data ingestion and migration.

Mastering Transformation

Enable data discovery, observability, and valuable insights for decision-making data curation.

Data Sustainability

Delivers automated data classification for security, privacy and compliance.

Unlock Hidden Value with Intelligent Automation

Automation saves time and reduces errors.

- Identifying data anomalies, errors and outliers.
- Validate data using rules to detect problems.
- Cleanse data and conform it to standards.
- Use analytics models to predict data quality issues.

Orchestrate data pipelines throughout their lifecycles.

- Reveal hidden data patterns and relationships.
- Organize and archive data by lifecycle stage.
- Profile data to understand quality and resolve issues.
- Measure data accuracy and completeness.

Monitor quality, protect privacy and report utilization.

- Continuously tracking data quality trends.
- Visualize and report data quality in real-time.
- Anonymize personally identifiable information (PII)
- Restrict access by policy to authorized users only.

Continuous Data Quality Monitoring

Automating data profiling, standardization and monitoring underpin consistent decision-making and operational effectiveness with trustworthy data.

Optimize Data Workflows and Outcomes

Streamline processes with high-quality data, enhancing collaboration and operational effectiveness across diverse business systems.

Understand More Data for Better Insights

Establish uniform practices, reveal data patterns and identify inconsistencies to enrich operational systems and customer data insights.
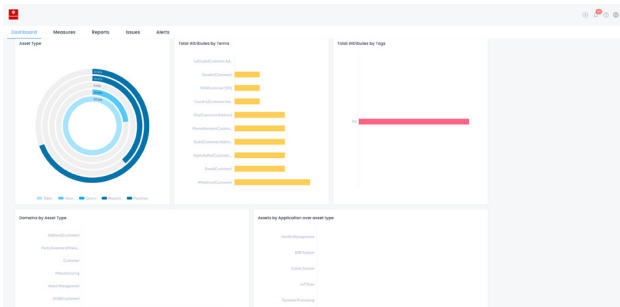

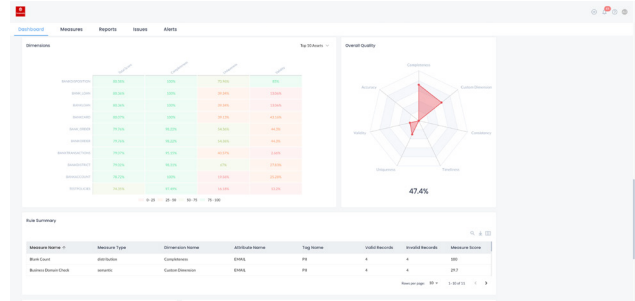Figure 1: Pentaho Data Quality metrics

Figure 2: Pentaho Data Quality metrics (continued)

Discover Pentaho Data Quality →

Pentaho