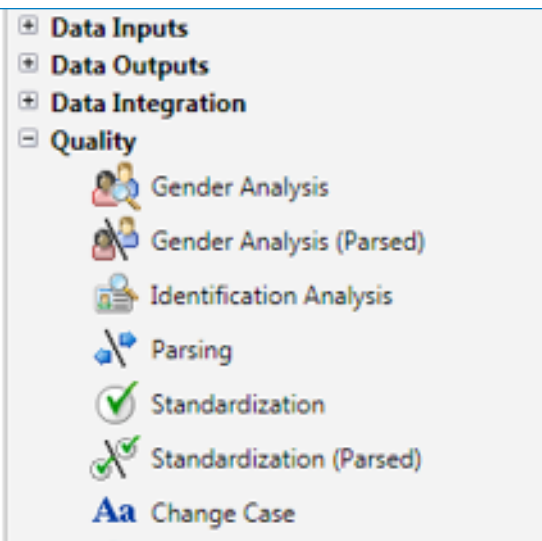


SAS® Data Quality

Make decisions you can trust to build a data-driven business.



Ask any thriving organization the secret to success and you'll get many answers: smart strategies, calculated risks, careful budgeting. They're all good business practices. And they come from the same place: a solid foundation of high-quality data. When you have accurate, up-to-date information driving your business, you're not just breaking even. You're breaking records.

With SAS Data Quality, you can manage the entire data quality life cycle: profiling, standardizing, matching and monitoring – as well as establish business rules to make sure that good data stays that way. And you can apply data quality to your traditional IT infrastructure as well as Hadoop environments.

SAS Data Quality is a fast, easy way to improve your data lifespan. And in the end, you can depend on that data to get your business where it needs to be.

What does SAS® Data Quality do?

SAS Data Quality helps you make data-driven decisions you can trust. Use it to improve and monitor the health and value of your data so you can confidently fuel operations, compliance and analytics initiatives.

Why is SAS® Data Quality important?

SAS Data Quality enables you to standardize and improve new and existing data so you can trust the information that serves as the basis for your corporate initiatives and systems.

For whom is SAS® Data Quality designed?

Managers and executives can make better decisions using trusted data. Marketing teams can reduce mailing costs and improve customer experience. And data governance teams can create business rules to monitor data health and document data assets, owners and relationships in the business glossary.

Key Benefits

- **Build a foundation of trust for all business initiatives.** Data serves as a foundation for customer relations, sales, production, forecasting and other initiatives. So it needs to be flawless. Outdated or unreliable data can harm your business decisions and your organization.
- **Get a more accurate view of operations.** Reliable, up-to-date and consistent data gives you a more vivid picture of your operations; you can see what's working – and what isn't. As a result, it's easier to instill confidence in upstream and downstream applications that consume that data.
- **Minimize unnecessary costs.** When the quality of your data is poor, things slip through the cracks. It's easy to lose track of spending and make other costly mistakes when you don't have the right answers.
- **Meet compliance and risk objectives.** Whether you need transparency for regulatory or other business purposes, you'll achieve unified, consistent, documented data with SAS Data Quality. The technology provides validation and integrity controls, as well as audit trails, a business glossary, security and lineage information.

Product Overview

SAS has always led the charge when it comes to data quality, which is why we invest in developing technology for all types of organizations, in every scenario – and for any infrastructure environment.

Data quality

SAS Data Quality enables you to approach data quality from every angle, including data standardization, deduplication and data correction. Interested in extending data quality to your Hadoop data sets? The same data quality functionality is available for your big data initiatives. You can also establish data hierarchies and reference data definitions for even more control over your business information.

Business glossary and lineage

The ability to visualize and manage your data – as well as its associated business terms and definitions – becomes part of an ongoing data quality process with this patented web-based interface. Create descriptions and requirements for system-to-business term translations, and access, share and collaborate on this information with other team members. With SAS Metadata Bridges, users can import third-party lineage metadata, including third-party

ETL jobs, data transformation logic and data models. This provides a view of data lineage that spans the enterprise, including SAS Data Management, SAS Analytics and third-party tools, tables, data models and other metadata objects.

Business rule validation

Develop and refine business rules to help your data meet organizational standards for quality and governance policies. You can validate data against standard statistical

measures as well as customized rules. You can also monitor and detect data that falls outside your standards – and take action before poor-quality data harms your organization.

Data profiling

With data profiling, you get an in-depth look at your data, examining its structure, completeness and suitability. Relationship discovery helps you uncover relationships across tables, databases and different

source applications, while data validation helps verify that the data in your tables matches its appropriate description. In addition, statistical analysis establishes trends and commonalities in your information.

Entity resolution

Confidently linking and consolidating entity information is critical to data management initiatives. Information about the same customer, product, location or employee may exist in multiple databases and in various forms. The challenge is to find and resolve similar records from different data sources – which is easily managed with SAS Data Quality.

Master data management foundation

MDM efforts are often part of a larger data management or data governance initiative. SAS Data Quality helps you configure a prebuilt master data model, providing the foundation for accelerated MDM deployments.

Reference data management

From a web-based interface, you can store and manage different types of lists, code tables or hierarchies used across enterprise applications.

Visual process orchestration

SAS Data Quality helps reduce the complexity of running jobs that have multiple steps through a visual method of scheduling and coordinating jobs. You can execute data integration or data quality jobs, execute them in parallel, and even trigger jobs based on external or internal factors (such as reaching a preset threshold).

Visualization and reporting

SAS Visual Analytics Administration and Reporting allows you to create reports and share information about data management initiatives, data health and remediation issues.

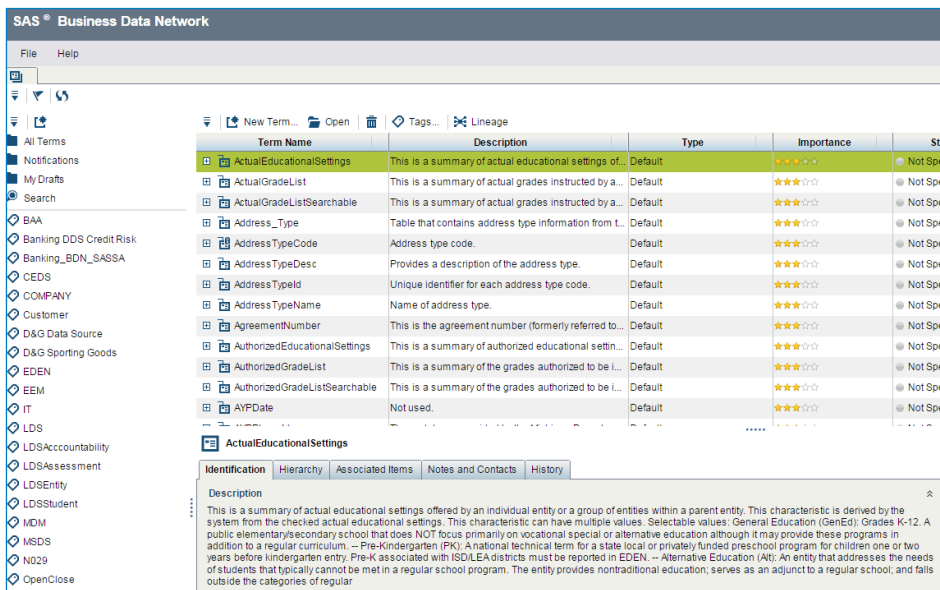


Figure 1. SAS Business Data Network allows business and IT users to collaborate on the creation of business terms, establish links to the term in a hierarchy of reference data and examine related data elements.

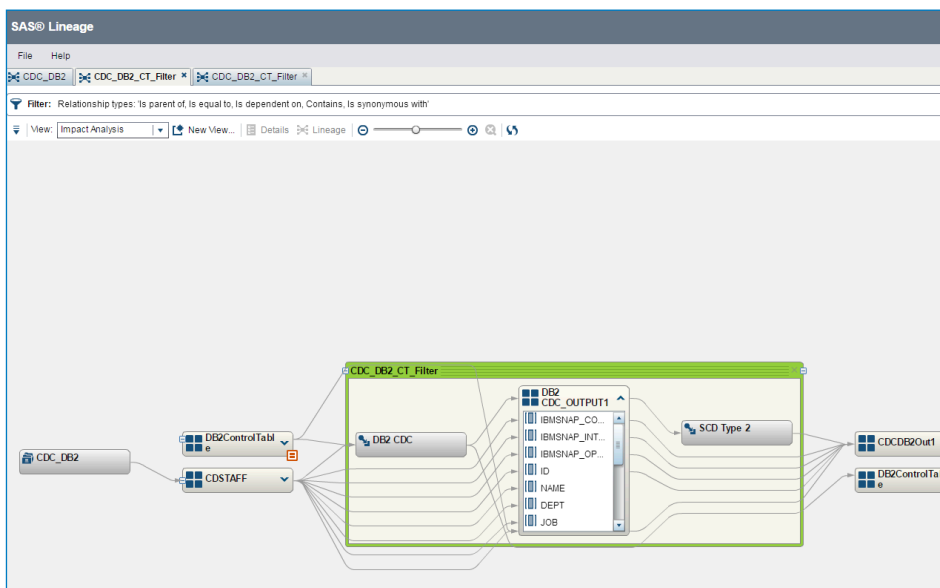


Figure 2. SAS Lineage capabilities help you visualize technical and business metadata from sources and tools that come from SAS and other software – and see how a change in one area can affect other data elements.

Key Features

Data cleansing

- Correct duplicate records, nonstandard data representations, unknown data types and other data quality issues.
- Enforce standardization rules.
- Establish data hierarchies and reference data definitions.
- Separate values using natural language parsing.
- Customize the quality knowledge base (preset business rules) to suit your unique needs with built-in data quality rules for both customer and product data.
- Apply data quality rules, standardization and natural language parsing that are specific to the context, region and language.

Data profiling

- Validate data against standard statistical measures, as well as against customized business rules.
- Uncover relationships across tables, databases and different source applications.
- Verify that data in your tables matches the appropriate description.
- Establish trends and commonalities in corporate information and assess with statistical metrics.

Business rule validation

- Verify that data meets organizational standards for data quality and business processes using validation routines assessed by statistical measures as well as customized business rules.
- Detect data that falls outside predetermined thresholds and gain insight into source data integrity.

Entity resolution

- Identify individuals across multiple data sources from incomplete and non-obvious relationships.

- Manage entity resolution routines through advanced fuzzy-matching technology, including cross-field matching.
- Create multirecord clusters, confidence scores and scatter plots to determine potential clusters.
- Recognize when slight variations suggest a connection between records.

Business glossary, lineage and metadata management

- The patented SAS Business Data Network provides a web-based interface to allow business and IT to collaborate on the creation and management of business terms in a business glossary, and relationships with technical metadata, including metadata contained in tables, reports or analytical models.
- Define and document terms, link them to governance policies and create associated requirements of a term, providing the technical team with guidance on creating business rules.
- View relationships and impact analysis across data transformation jobs from both SAS and other software, data models and other data elements with the patented SAS Lineage component.
- Use documented audit trails to know which data was changed and who changed it.
- Store and apply definitions to guide all data management processes.
- Apply role-based, granular security that manages who has access to different types of business terms and technical metadata.
- Import third-party lineage metadata including ETL jobs, data transformations or data models.

Foundation for master data management

- Create a hub of master data based on subsets of cleansed data through a phased approach.

- Integrate the creation and management of master data resources with comprehensive and consistent data management practices.
- Propagate high-quality data to an MDM hub so master data records are accurate.

Reference data management

- Identify, profile and validate reference data from multiple sources.
- Integrate accurate and reliable reference data.
- Link relationships across data elements, including parent-child, one-to-many or coded relationships.
- Maintain and manage history and data change management in a single interface.
- Create business rules across processes and manage existing business rules.
- Integrate business terms and glossaries defined in the SAS Business Data Network.
- Create enterprisewide standards for data quality at the attribute level.

Centralized enterprise activity

- Eliminate the need to write and test code to enable real-time data quality, data integration and data governance.
- Ensure controlled, managed and centralized security.

Data management console

- Monitor data quality jobs and view data issues and governance activities.
- Access all data management activity from a single, common control point.
- Secure role-based access and actions to authorize specific data quality tasks.
- Examine data lineage to trace business terms to source system definitions in a single interface.

Key Features (continued)

Data integration

- Embed data quality into extract, transform and load (ETL) and extract, load and transform (ELT) activities from multiple sources using both traditional batch processing and in-database methods.
- Transfer data to new or different locations while improving the accuracy and consistency of data through data migration.
- Match information within or across data sources, standardizing formatting differences.

Visual process orchestration

- Manage both SAS and third-party technologies in a single process.
- Coordinate the execution of data integration and data quality jobs, execute jobs in parallel and trigger job execution based on events.
- Reduce the complexity of multistep execution and minimize dependence on third-party scheduling systems.

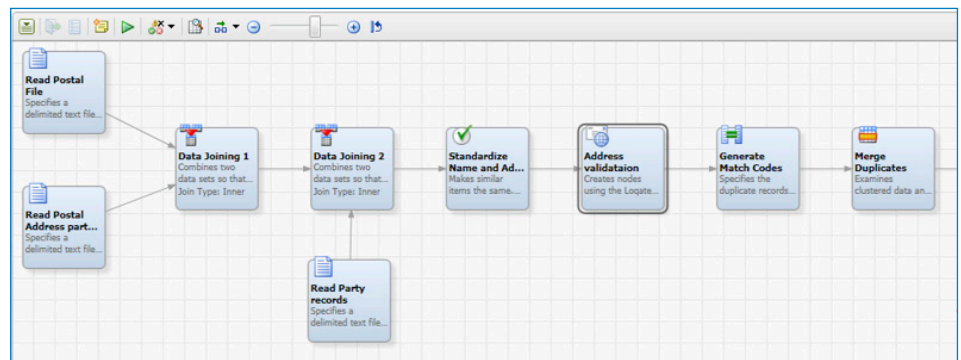


Figure 3. Join, standardize and remove duplicates from your data with the powerful and intuitive SAS Data Quality user interface.

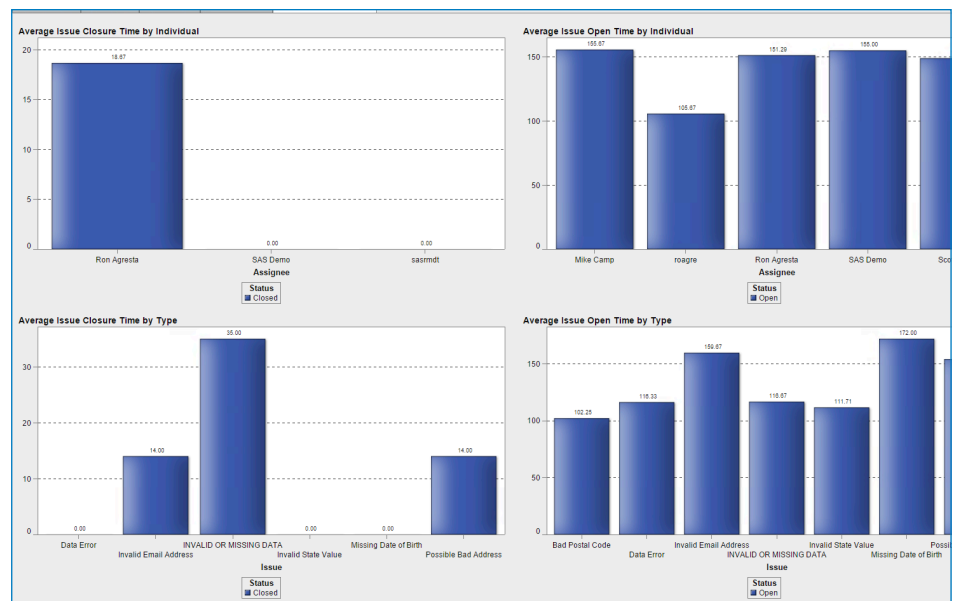


Figure 4. Users can take advantage of SAS Visual Analytics Administration and Reporting to create reports and share information about data management initiatives, data health and remediation issues.

[TO LEARN MORE »](#)

To learn more about SAS Data Quality, view screenshots and see other related material, please visit sas.com/dataquality.

To contact your local SAS office, please visit: sas.com/offices

