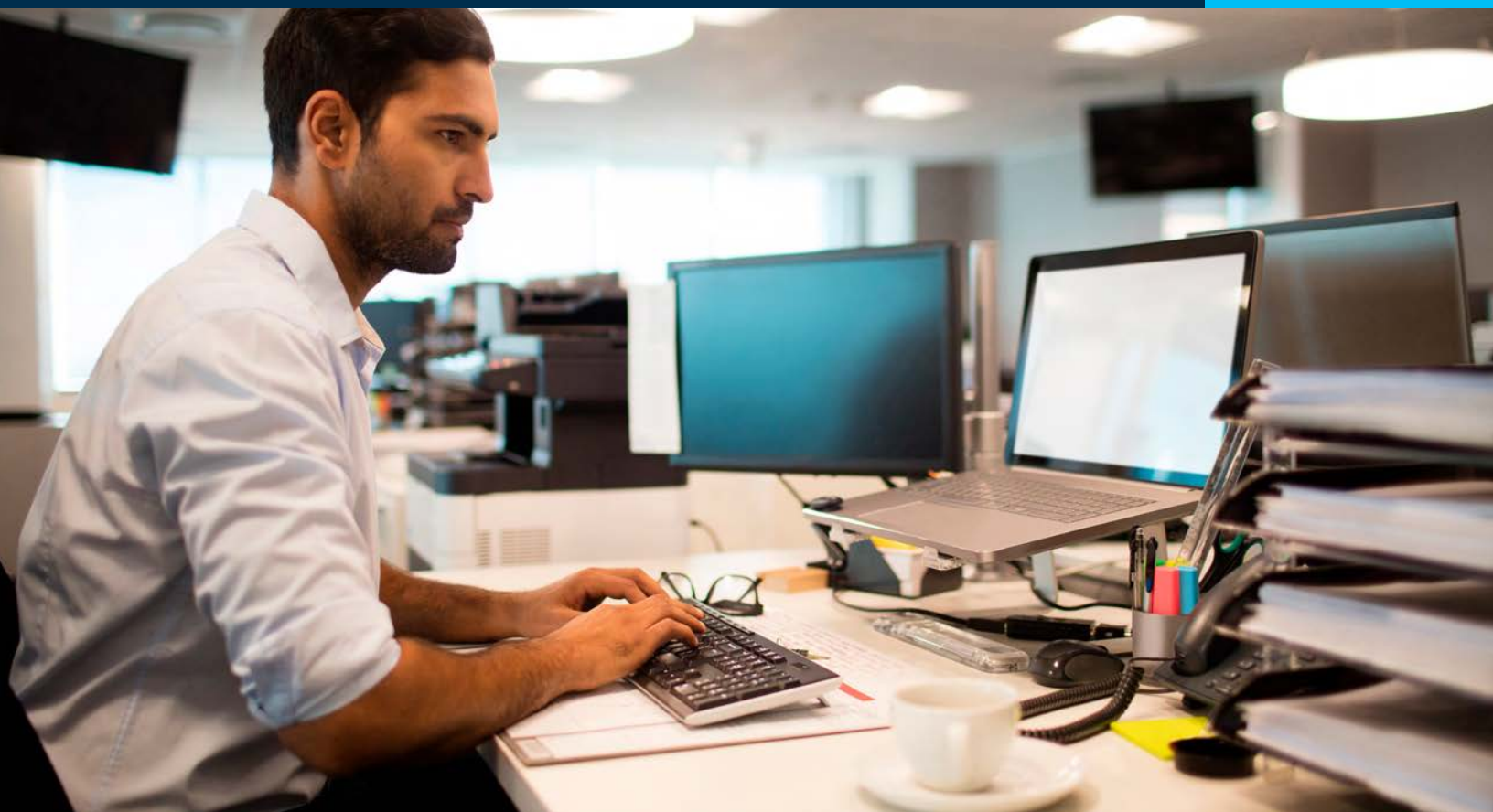


# SAS® Data Management

Better data is the foundation for better decisions



## Overview

Data is the heart and soul of every organization, regardless of size, industry or expertise. And data management underlies virtually every process organizations depend on. In fact, you will be hard-pressed to find a C-level business executive who doesn't believe data is an integral part of forming a data strategy.

But there are challenges to becoming data driven. For example:

- Accessing and integrating data continues to be a problem.
- Users often don't trust the quality of their data.
- In many cases, organizations have no enterprise data governance and security program.

Why are there so many obstacles?

A data-driven organization depends on tight collaboration between business and IT and a culture that trusts its data to inform business decisions. It demands leaders who support this mindset and a technology foundation that ensures excellent data quality and governance. Most organizations that are truly data-driven also have an integrated analytics and data management strategy.

Unfortunately, many organizations still rely on manual efforts to combine, cleanse and transform their data – processes that are hard to repeat or share. And after devoting time to search for relevant data, some find that the data is flawed.

SAS Data Management is a trusted foundation for analytics that gives you access to clean, well-governed data – at the right time and in the right format. Across the entire analytics life cycle – from data to discovery to deployment – SAS can help you make data-driven decisions for your business.

## Our Approach

To run an agile business in the face of changing markets and innovations, organizations need fast, easy access to new and variable types of data – and a software platform that promotes collaboration. Many organizations will have to adopt new technologies and modernize their existing IT infrastructures to accommodate these changes and embrace new business strategies (such as data monetization).

SAS extends data management beyond traditional boundaries to help you make better decisions, faster. With SAS, you can manage data at its source to improve productivity and performance. Our software automates and supplements manual tasks – with machine learning, natural language processing and advanced statistical techniques – to help your data direct actions for transformations, adjustments and process refinement. With data that's thoroughly prepped, high quality, cataloged, fully governed and ready to use for any purpose, you'll be positioned to stave off competitive threats and pursue new business opportunities.

### SAS® Data Management capabilities

Through an integrated, comprehensive suite of capabilities, SAS Data Management helps you take charge of all your data and treat it as a valuable business asset that guides your data-driven business decisions. Our software works consistently across data types, sources and locations to ensure you operate and make analytically sound decisions based on clean, trusted data.

SAS is an analyst-recognized leader in data integration, data quality and analytics. 92 of the top 100 companies on the 2018 Fortune Global 1000® are SAS customers.

### Data access

Give technical and business users fast, secure access to the information they need, no matter what format it's in or where it's stored. SAS/ACCESS® software provides seamless, transparent read, write and update rights to more than 60 data sources – including relational and nonrelational databases, text files, spreadsheets, Hadoop/Spark, Amazon Redshift, data warehouses and more – on-site or in the cloud. Our software honors the native security of the target data source and makes all data sources appear as though they are a native SAS data type. As a result, it facilitates data interaction across sources and with SAS procedures, SAS DATA step and SAS solutions like SAS Event Stream Processing. You can also access data through web service APIs. SAS is built on trusted integration standards, including ODBC, JDBC (on SAS® Viya®), Spark SQL (on SAS Viya) and OLE DB.

### Data integration

Break down data silos with the industry's leading integration technology, which speeds integration processes. From legacy systems like data warehouses to spreadsheets, data lakes and the cloud, SAS lets you integrate any data you need – streaming, real time or batch. An intuitive interface with a single point of control makes it easy for business users to work independently, freeing IT for other tasks. Auditing tools monitor processing and source data lineage so you know your data is ready for analytics or reporting. Data federation lets you virtually blend data from multiple underlying sources without moving or storing it – so you can keep your data secure and use the combined data whenever it's needed.

## Benefits

- **Take charge of your data wherever it lives, as fast as it comes.** Access and integrate all types of data. Apply data management techniques anywhere, whether it's in stream, in database, in memory or in cloud. Taking data management closer to the source means better speed, agility, productivity and performance.
- **Uncover analytics insights faster.** Free IT to do more by giving business users self-service data preparation tools. Automate tasks with machine learning and deep learning techniques to get a faster view into hidden problems, issues with data elements and processing, and potential new opportunities.
- **Act confidently as you make data-driven decisions based on a trusted data foundation.** With best practices embedded into every element of our technology, your data will be consistent, accurate, analytically valid and well-governed.
- **Support all users who need analytics insights.** A common platform and easy-to-use interface make analytics accessible to everyone – across all types of data and problems of any size or complexity.

## Challenges

- **Counterproductive business and IT efforts.** Business and IT don't speak the same language, which makes it nearly impossible for staff to focus on long-term goals or prioritize daily activities.
- **Fragmented systems wreak havoc on processes.** New data is often stored in silos outside of the managed environment, so it has variable structures, quality levels and process definitions.
- **Wasted time.** Manual coding that's used to reconcile different data quality definitions and structures causes bottlenecks between teams.
- **Hampered creativity.** When IT spends too much time maintaining the status quo, their efforts to improve, automate and be strategic to the business are stifled.
- **Lack of trust in the data.** Poorly governed data, or data with questionable quality, results in inconsistent reports – so no one trusts the data, and it's impossible to operate as a data-driven business.

### Data cataloging

Provide a central area for the collection of metadata, combined with data management and search tools, so your analysts and other data users find the data that they need. The data catalog serves as an inventory of available data and provides the information needed to evaluate the fitness of data for specific, intended uses. With SAS Information Governance, now you can:

- Know what data is available.
- View the health of data.
- Understand data relationships.
- Easily locate hard-to-find data and gain access it.
- Allow data users and consumers to assess differences between data sets.
- Determine data preparation requirements are needed, if any.

### Data quality

Make sure data quality is embedded into every process. There's no need to move or extract data. Analyst-validated SAS Data Quality profiles, standardizes, enriches, monitors, transforms and verifies data where it exists – in motion or at rest. You can customize, automate and reuse data quality business rules within process job flows. Establish repeatable processes to build and maintain high-quality data across the entire data life cycle. And create a master record to get a single view across multiple sources for one domain.

### Data preparation

Reduce time spent preparing data so you can spend more time on analytics. SAS Data Preparation lets you access, cleanse, manipulate and prepare data for analytics with an intuitive interface – no special coding or SQL skills required. SAS combines advanced analytics, data quality, visualization and data preparation capabilities on a single platform. As a result, you have more time to explore data, get answers and respond to situations – in near-real time. Nontechnical people can use a point-and-click interface to profile, cleanse, blend and move data on their own. Data prep tools can suggest data sets and guide users to work with the data without coding. Improve efficiency as you collaborate and share data and jobs among team members – while reusing existing plans, scheduling and automating tasks, and monitoring jobs.

### Data governance

Establish and enforce policies and get a consistent view of your data using a web-based environment and dashboards within data governance technology from SAS. You can create a repository of governed terms and sources to use across any system – including third-party systems – to enhance stewardship and enforce governance. With the SAS Visual Analytics reporting engine and data governance capabilities from SAS, you can visually and interactively flag issues, route them to the appropriate person and make sure they get fixed. An integrated business data glossary, metadata management and lineage visualization promote better collaboration.

### Metadata management

Improve trust in data by providing complete transparency. Visually trace your data's life cycle both in and out of the SAS environment using lineage capabilities – so you'll understand the impact of making changes to data rows and tables before you act. Keep a common glossary to ensure that everyone is on the same page when discussing access to new data sources and the requirements of meeting business objectives.

### Event stream processing

Analyze streaming data while it's in motion – from operations, transactions, IoT sensors and other devices. SAS Event Stream Processing analyzes millions of events per second, determining which data requires immediate attention, what you can ignore and what can be stored for later use. SAS assesses data streams using a suite of prebuilt operators, functions, routines and advanced analytics – providing instant insight into events so you can take appropriate actions. It also provides in-stream data quality and includes a single, intuitive interface to let you define patterns and address scenarios from any aspect of your business. With its machine learning capabilities, the software can teach itself what events warrant a closer look – without human intervention.

## Data protection

Meet compliance and auditing requirements as you protect data across your organization. SAS for Personal Data Protection identifies all sources that contain the data you need to protect – and includes tools like identification analysis, standardization and pattern matching. Examine data attributes, patterns and contexts as you evaluate data to minimize risk. Use data governance to establish and monitor rules and policies that link systems, processes and business owners in data flows across the organization. Safeguards like authentication and authorization control who accesses the data. Techniques like role-based data masking, pseudonymization, anonymization and encryption protect sensitive and personal data. As you log, monitor and audit data use, you'll be able to provide clear documentation and reports to auditors to prove your data is safe.

## The SAS® Difference

### Data processing at the source

Moving routine execution closer to the source of the data takes advantage of existing hardware, minimizes data movement and simplifies access. In turn, processing is faster and downstream use becomes more consistent, requiring fewer computing resources. Consider streaming data from the IoT. By embedding in-stream data management and normalization functions before data is stored, you can identify what's worth keeping – correcting it as needed before you incur unnecessary, incremental storage costs. This helps modernize your infrastructure while taking advantage of investments in legacy technology.

### Data management for multiple roles

Finding the right data source and shaping it to meet project-specific needs should be easy. SAS Data Management presents pertinent suggestions to help users fine-tune and understand what changes need to be made to the data for it to be useful. Combining legacy and modern sources, SAS sends alerts when data changes are made or need to be made, helping to streamline workflows and keep results current. Our software is designed for people with a range of skill levels: IT developers who support the data infrastructure; analysts and business users who need to prepare data for reporting and analytics; and data stewards and data governance council members who must assess data health over time and continually validate the data governance program's effectiveness.

## Built-in artificial intelligence and advanced analytics

SAS Data Management incorporates artificial intelligence capabilities (such as natural language processing) to manage data in native character sets and ensure the data is accurately tokenized, normalized and enhanced. Fuzzy matching uses scoring to reveal similarities (matches) between data. Self-tuning systems make automatic adjustments to maximize the computing infrastructure and optimize processing. Algorithms discover and curate business rules within operational data. Parallel profiling discovers and identifies new metadata insights – providing direction for data management transformation and cleansing tasks. And SAS defines groups of similar data events to be assessed for common patterns by integrating learning models in data streams.

SAS participates in the Data for Good movement, which gives us opportunities to use data – and our software – to change lives.

### Multiple cloud options

SAS Data Management can access cloud-based data from many environments, including Azure SQL Database ([learn about the new SAS Microsoft partnership](#)), as well as Amazon Aurora, Amazon Redshift, Amazon Elastic MapReduce, and the cloud versions of traditional databases. Our software can be deployed in cloud environments like Amazon Web Services, Microsoft Azure, Google Cloud Platform, OpenStack, VMware and other infrastructure-as-a-service providers. SAS also provides its own cloud hosting environment, SAS Cloud, which handles the details of installing, configuring and administering your computing environment so you can focus on running your business.

### A cohesive platform

The SAS Platform is built on more than four decades of expertise in analytics, helping customers across virtually every industry uncover insights from all types of data, in all types of environments. Our platform supports every phase of the analytics life cycle – from data to discovery to deployment. Programmers, data scientists and executives all rely on SAS to manage traditional and new types of data and ensure governance and security across the entire analytics ecosystem.

## Learn More

Find out how SAS helps organizations around the world transform their data into intelligence and build successful, data-driven businesses. Visit [sas.com/data](https://sas.com/data).

For more information about SAS® Data Management, please visit [sas.com/data](https://sas.com/data)

