



Base SAS® Software

Flexible and extensible fourth-generation programming language designed for data access, transformation and reporting

What does Base SAS® software do?

Base SAS provides a scalable, integrated software environment specially designed for data access, transformation and reporting. It includes a fourth-generation programming language; ready-to-use programs for data manipulation, information storage and retrieval, descriptive statistics, and report writing; and a powerful macro facility that reduces programming time and maintenance headaches.

Why is Base SAS® software important?

Because it runs on all major computing platforms, Base SAS provides a powerful, versatile software foundation. It significantly reduces programming and maintenance time, while enabling your IT organization to produce the analyses and reports that decision makers need in the format they prefer.

For whom is Base SAS® software designed?

Base SAS is used by domain experts, power users and programmers who need to manipulate data, produce and distribute ad hoc queries and reports, and/or interpret the results of data analysis.

Most IT organizations struggle with problems arising from complex and distributed data. Synchronizing and reformatting data for various applications takes time. Producing accurate and visually appealing reports requires excessive programming resources. Compounding that, IT departments often manage a plethora of software packages because no single tool seems able to provide the power and flexibility needed to support the whole enterprise.

IT organizations must find a way to streamline and speed up programming, reduce costs and produce the analyses and reports that decision makers need. Also imperative is the flexibility to support new business requirements and make effective use of existing hardware resources.

Base SAS software provides an integrated and scalable software environment for transforming data and creating differentiating intelligence. It includes a highly flexible and extensible fourth-generation programming language (4GL) specially designed for data access, transformation and reporting; a rich library of encapsulated procedures for data manipulation, information storage and retrieval, statistical analysis, and report writing; and a powerful macro facility that reduces applications development and maintenance time.

With a highly scalable cross-platform architecture, Base SAS software enables optimized use of all hardware resources. Numerous components are available for additional functionality, including direct access to standardized data sources and advanced statistical analysis.

Key Benefits

- **Integrate data across environments.** Based on an open, cross-platform architecture, Base SAS integrates into any computing environment infrastructure, enabling you to unify your computing efforts and get a single view of your data.
- **Read, format and analyze any data.** From small data issues to large complex data problems, programmers can read, format, analyze and report on data in any format quickly.
- **Make programming fast and easy.** With an intuitive, easy-to-learn and easy-to-use programming language, Base SAS significantly reduces the amount of code required to deliver information, increasing your programmers' productivity.
- **Simplify reporting.** Base SAS provides maximum reporting flexibility. You can easily create reports in standard office formats such as Rich Text Format (RTF) and Portable Document Format (PDF), produce output in a variety of markup languages, and produce graphs automatically from statistical procedures.
- **Maximize use of all computing resources.** With multithreaded capabilities, Base SAS software takes advantage of parallel processing power to maximize use of computing resources and produce the timely answers your organization demands.
- **Comprehensive analytic support for Hadoop.** Using Base SAS, you can not only incorporate Hadoop capabilities such as the Pig and Hive languages and MapReduce framework, you can also apply them across all SAS products and solutions.



Product Overview

Base SAS software is a highly flexible and integrated software environment that can be used in virtually any setting to access, manipulate, manage, store, analyze and report on data.

Key Capabilities

Intuitive fourth-generation programming language

Base SAS provides a highly flexible and extensible 4GL with an easy-to-learn syntax and hundreds of language elements and functions that support programming everything from data extraction, formatting and cleansing to data analysis, reporting and information delivery.

A rich library of prewritten, ready-to-use integrated procedures

Prewritten SAS procedures handle many common tasks, including data manipulation and management, information storage and retrieval, statisti-

cal analysis and report writing. They encapsulate and deliver significant functionality that can be executed with just a few simple commands, enabling programmers to be more efficient and productive.

Powerful data analysis tools

Base SAS can be used for a wide variety of analyses ranging from simple descriptive statistics that include mean, sum, variance and standard deviation to more advanced data correlation and cross-correlation, frequency analysis and detailed data distribution analysis.

SAS® macro facility

Included as part of the SAS language, the SAS macro facility lets you extend and customize your SAS programs by reducing the amount of coding required for common tasks. With the macro facility, you encapsulate small or large amounts of text into units with names and then work with the names rather than the longer text. Your development staff can modularize its work for easy reuse and maintenance.

SAS® Code Analyzer

The SAS Code Analyzer runs an existing SAS program, executing the program as usual, while generating metadata about the SAS job that is emitted via comments. The SAS Code Analyzer can be used to capture information such as file dependencies and macro symbol usage to assist in analysis and potential modification of legacy SAS programs.

In-database capabilities

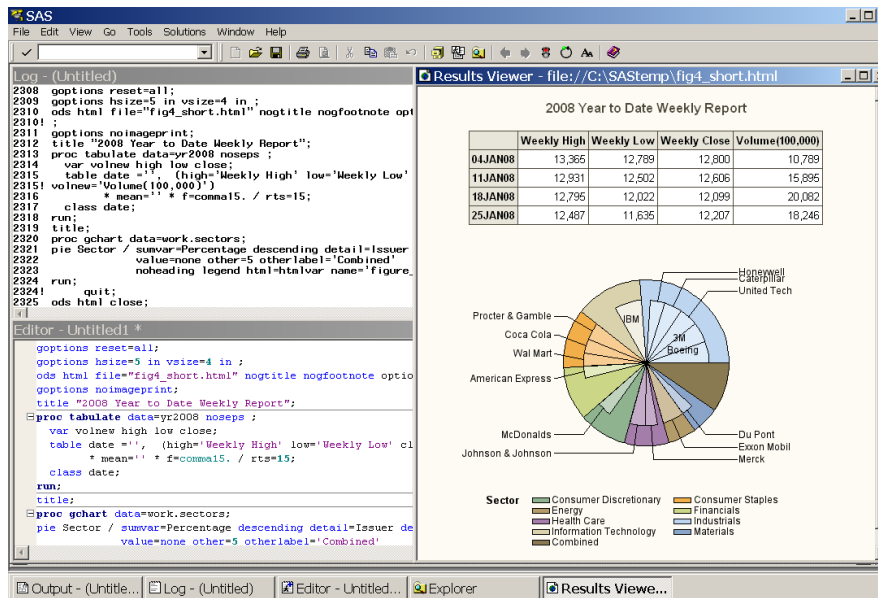
Some of the most widely used Base SAS procedures now support SQL push-down capabilities to the following third-party databases: Aster Data, Greenplum, IBM DB2, Netezza, Oracle and Teradata. SAS users can continue to program as usual, but SAS converts their code into database SQL-specific code optimized to the receiving database. SAS procedures currently supporting SQL push-down optimization are FREQ, MEANS, RANK, REPORT, SORT, SUMMARY and TABULATE. In addition, SAS format catalogs can be published and compiled inside those databases. This allows formats to be applied to the actual data values during query execution, reducing data movement.

Cross-platform and multiple-platform support

With MultiVendor Architecture™, Base SAS runs on all major computing platforms, can access virtually any data source and easily integrates into any organization's computing environment. SAS code developed on one platform runs easily on others.

Support for symmetric multiprocessing (SMP) hardware

Key SAS procedures for sorting, querying and summarizing data are threaded, enabling them to take advantage of multi-CPU computing environments. The SAS Scalable Performance



Base SAS capabilities can be extended with additional components such as SAS/GRAPH® software for advanced graphic capabilities.

Data Engine speeds throughput and processing of very large data sets by employing parallel read and indexing capabilities on partitioned SAS data sets.

SAS® Open Metadata Architecture

Base SAS delivers the SAS Open Metadata Architecture, a general purpose metadata management facility that provides common metadata services to SAS and other applications. The SAS Metadata Server allows business metadata to be shared across the enterprise from a centralized repository, ensuring that metadata is consistent, accurate and not held in isolated, hard-to-maintain silos.

SAS® Management Console

Managing a variety of applications across different platforms can be time-consuming. The SAS Management Console is a Java framework that provides a single interface for easier management of many SAS administrative tasks. Whether your SAS deployment is across multiple platforms, departments or the entire enterprise, the easy-to-use GUI enables IT professionals to more effectively manage and administer all aspects of SAS. Site-specific tasks are supported with customizable plug-ins.

SAS® XML Mapper

The XML engine delivered with Base SAS imports and exports a wide variety of XML documents. For complex XML documents, a new Java client simplifies the handling of nonstandard tags. The SAS XML Mapper provides a drag-and-drop interface for creating reusable XML maps, which describe an XML document's content or type and tell the XML engine how to convert the XML document into SAS data sets, variables (columns) and observations (rows).

Key Features

Flexible 4GL

- Intuitive 4GL with easy-to-learn syntax.
- SAS Macro Facility reduces coding for common tasks and lets you modularize work for easy reuse and maintenance.
- Runs interactively or in batch mode.

Powerful data analysis capabilities

- Analysis capabilities ranging from simple descriptive statistics to advanced data correlations.
- Library of prewritten programming procedures for managing, analyzing and presenting data.

Support for wide range of data formats

- Ability to read data in any format, from any kind of file, including variable-length records, binary files, free-formatted data and even files with messy or missing data.
- Support for Structured Query Language (SQL).
- Globalization with full support for the most widely used character encodings (such as Latin1, Latin2, and multiple-byte character sets for Hebrew, Arabic and Asian languages).
- Unicode support so that SAS works with data in multiple languages on all platforms.

Support for Hadoop

- Support for Pig, MapReduce and HDFS commands from the SAS execution environment.
- Supports external file references from within any SAS component. Delimiters are externalized, making it easy to work with Hadoop files.

Performance and scalability

- Parallel I/O optimization lets you work with huge data volumes in a timely manner. Data can be partitioned across devices for faster access but referenced as a single data set.
- Parallel index creation reduces time needed to create large data sets with multiple indexes or to append data to existing data sets.
- Key SAS procedures are multithreaded for faster execution of standard tasks such as sorting and data summarization.
- Base SAS scalability can be enhanced to span multiple machines and networks using SAS/CONNECT® software, which is available separately.
- Several Base SAS procedures support SQL push-down optimization to Aster Data, EMC Greenplum, IBM DB2, Netezza, Oracle and Teradata.
- SAS format catalogs can be published and compiled inside databases (Aster Data, EMC Greenplum, IBM DB2, Netezza, Oracle and Teradata) so that formats can be applied to the actual data values during query execution.

Interoperability and multiple-platform deployment

- MultiVendor Architecture allows programs to be written once and run anywhere, regardless of hardware or operating system.
- SAS Open Metadata Architecture enables different applications to exchange metadata.
- The GROOVY procedure enables SAS code to execute on the Java Virtual Machine.
- Cross-environment data access provides easy-to-access files across a network.

Manageability

- SAS Management Console provides an extensible Java GUI for administering SAS tasks.
- XML engine imports and exports a wide variety of XML documents. A drag-and-drop interface creates XML maps.
- Checkpoint and restart capability allows users to resubmit a failed program in restart mode to complete execution, resuming with the step that did not complete when the failure occurred.
- Application Response Measurement (ARM) interface monitors the availability and performance of transactions within and across diverse applications.

Continued on next page

Output Delivery System (ODS)

The Output Delivery System for data capture, report formatting and information delivery provides an almost limitless number of choices for reporting and displaying analytical results with a wide variety of output formats and destinations. While providing an array of options for customizing the output, ODS takes care of arranging the output in the form most appropriate for the chosen format. Programmers can create and deliver accurate and visually appealing reports in less time with reduced effort. With ODS statistical graphics, a statistical procedure generates appropriate graphical displays for its analyses by default.

Base SAS® System Requirements

To learn more about Base SAS system requirements, download white papers, view screenshots and see other related material, please visit sas.com/basesas.

Key Features (continued)

Data presentation

- ODS provides choices for reporting and displaying analytical results.
- High-quality graphics are now included in Base SAS (ODS Statistical Graphics, the SG family of procedures, the Graphical Template Language, the ODS Graphics Designer and the ODS Graphics Editor).
- Create reports in standard formats such as RTF and PDF. All formats are available on all platforms.
- Create visually appealing graphics from analytic output by default (no additional programming).
- HTML 4 and XML are among the markup languages provided. Modify any markup language that SAS provides or create your own markup language for output. HTML is the default destination for output.
- Customize or modify output hierarchy; replay output to different destinations without rerunning the program.