Title: SAS (software)

URL: https://en.wikipedia.org/wiki/SAS_(software)

PageID: 990677

Categories: Category:Business intelligence software, Category:Data mining and machine learning software, Category:Data warehousing, Category:Extract, transform, load tools, Category:Fourth-generation programming languages, Category:Mathematical optimization software, Category:Numerical software, Category:Proprietary commercial software for Linux, Category:Proprietary software programmed in C, Category:SAS Institute, Category:Science software for Linux

Source: Wikipedia (CC BY-SA 4.0).

SAS (previously " Statistical Analysis System ") [1] is data and artificial intelligence software developed by SAS Institute for data management , advanced analytics, multivariate analysis , business intelligence , and predictive analytics .

SAS was developed at North Carolina State University from 1966 until 1976, when SAS Institute was incorporated. SAS was further developed in the 1980s and 1990s with the addition of new statistical procedures, additional components and the introduction of JMP . A point-and-click interface was added in version 9 in 2004. A social media analytics product was added in 2010. SAS Viya , a suite of analytics and artificial intelligence software, was introduced in 2016.

Technical overview and terminology

SAS is a data analysis and artificial intelligence software suite that can mine , alter, manage and retrieve data from a variety of sources and perform statistical analysis on it. [2] SAS provides a graphical point-and-click user interface for non-technical users and more through the SAS language . [2]

SAS programs have DATA steps, which retrieve and manipulate data, PROC (procedures) which analyze the data, and may also have functions. [3] Each step consists of a series of statements. [4]

The DATA step has executable statements that result in the software taking an action, and declarative statements that provide instructions to read a data set or alter the data's appearance. [3] The DATA step has two phases: compilation and execution. In the compilation phase, declarative statements are processed and syntax errors are identified. Afterwards, the execution phase processes each executable statement sequentially. [5] Data sets are organized into tables with rows called "observations" and columns called "variables". Additionally, each piece of data has a descriptor and a value. [3] [6]

PROC statements call upon named procedures. Procedures perform analysis and reporting on data sets to produce statistics, analyses, and graphics. There are more than 300 named procedures and each one performs a substantial body of statistical work. [3] PROC statements can also display results, sort data or perform other operations. [4]

SAS macros are pieces of code or variables that are coded once and referenced to perform repetitive tasks. [7]

SAS data can be published in HTML, PDF, Excel, RTF and other formats using the Output Delivery System, which was first introduced in 2007. [8] SAS Enterprise Guide is SAS's point-and-click interface. It generates code to manipulate data or perform analysis without the use of the SAS programming language. [9]

The SAS software suite has more than 200 add-on packages, sometimes called components [10] [11] [12] Some of these SAS components, i.e. add on packages to Base SAS include: [2] [13]

SAS/STAT – Statistical analysis

SAS/GRAPH – Graphics and presentation

SAS/OR – Operations research

SAS/ETS – Econometrics and Time Series Analysis

SAS/IML – Interactive matrix language

SAS/AF – Applications facility

SAS/QC – Quality control

SAS/INSIGHT – Data mining

SAS/PH – Clinical trial analysis

SAS Enterprise Miner – data mining

SAS Enterprise Guide – GUI-based code editor & project manager

SAS Enterprise BI – Suite of business intelligence applications

SAS Grid Manager – Manager of SAS grid computing environment

SAS Customer Intelligence 360 – Customer intelligence

History

Origins

The development of SAS started in 1966 after North Carolina State University re-hired Anthony Barr [14] to program his analysis of variance and regression software so that it would run on IBM System/360 computers. [15] The project was funded by the National Institutes of Health . [16] and was originally intended to analyze agricultural data [11] [17] to improve crop yields. [18] Barr was joined by student James Goodnight , who developed the software's statistical routines, and the two became project leaders. [14] [15] [19] In 1968, Barr and Goodnight integrated new multiple regression and analysis of variance routines. [20] [21] In 1972, after issuing the first release of SAS, the project lost its funding. [16] According to Goodnight, this was because NIH only wanted to fund projects with medical applications. [22] Goodnight continued teaching at the university for a salary of $1 and access to mainframe computers for use with the project, [16] until it was funded by the University Statisticians of the Southern Experiment Stations the following year. [15] [22] John Sall joined the project in 1973 and contributed to the software's econometrics, time series, and matrix algebra. Another early participant, Caroll G. Perkins, contributed to SAS' early programming. Jolayne W. Service and Jane T. Helwig created SAS's first documentation. [20]

The first versions of SAS, from SAS 71 to SAS 82, were named after the year in which they were released. [23] In 1971, SAS 71 was published as a limited release. [2] [24] It was used only on IBM mainframes and had the main elements of SAS programming, such as the DATA step and the most common procedures, i.e. PROCs. [23] The following year a full version was released as SAS 72, which introduced the MERGE statement and added features for handling missing data or combining data sets. [25] The development of SAS was described by the CNBC website as an " inflection point " in the history of artificial intelligence . [26] In 1976, Barr, Goodnight, Sall, and Helwig removed the project from North Carolina State and incorporated it as the SAS Institute, Inc. [27]

Development

SAS was redesigned in SAS 76. The INPUT and INFILE statements were improved so they could read most data formats used by IBM mainframes. Generating reports was also added through the PUT and FILE statements. The ability to analyze general linear models was also added [28] as was the FORMAT procedure, which allowed developers to customize the appearance of data. [23] In 1979, SAS 79 added support for the IBM VM/ CMS operating system and introduced the DATASETS procedure. Three years later, SAS 82 introduced an early macro language and the APPEND procedure. [23]

Beginning with SAS 4, released in 1984, SAS releases have followed a sequential naming convention not based on year of release. [29] SAS version 4 had limited features, but made SAS more accessible. Version 5 introduced a complete macro language, array subscripts, and a full-screen interactive user interface called Display Manager. [23] In 1985, SAS was rewritten in the C programming language . This enabled the SAS' MultiVendor Architecture which allows the software to run on UNIX , MS-DOS , and Windows . It was previously written in PL/I , Fortran , and assembly language . [19] [23]

In the 1980s and 1990s, SAS released a number of components to complement Base SAS. SAS/GRAPH, which produces graphics, was released in 1980, as well as the SAS/ETS component, which supports econometric and time series analysis. A component intended for pharmaceutical users, SAS/PH-Clinical, was released in the 1990s. The Food and Drug Administration standardized on using SAS/PH-Clinical for new drug applications in 2002. [19] Vertical products like SAS Financial Management and SAS Human Capital Management (then called CFO Vision and HR Vision respectively) were also introduced. [30]

JMP was developed by SAS co-founder John Sall and a team of developers, in order to take advantage of the graphical user interface introduced in the 1984 Apple Macintosh . [31] JMP's name originally stood for "John's Macintosh Project". [32] JMP was shipped for the first time in 1989. Updated versions of JMP were released continuously after 2002 with the most recent release in 2016. [33] [34] [35] [36] In January 2022, JMP became a wholly owned subsidiary of SAS Institute, having previously been a business unit of the company. [37]

SAS 6 was used throughout the 1990s and was available on a wider range of operating systems, including Macintosh , OS/2 , Silicon Graphics , and PRIMOS . SAS introduced new features through dot-releases. From 6.06 to 6.09, a user interface based on the Windows paradigm was introduced and support for SQL was added. [38] Version 7 introduced the Output Delivery System (ODS) and an improved text editor. Subsequent releases improved upon the ODS. For example, more output options were added in version 8. The number of operating systems that were supported was reduced to UNIX , Windows and z/OS , and Linux was added. [39] SAS 8 and SAS Enterprise Miner were released in 1999. [19]

Recent history

In 2002, SAS Text Miner software was introduced. Text Miner analyzes text data like emails for patterns in business intelligence applications. [40] In 2004, SAS Version 9.0 was released, referred to as "Project Mercury" internally, and was designed to make SAS accessible to a broader range of business users. [41] [42] SAS 9.0 added custom user interfaces based on the user's role and established the point-and-click user interface of SAS Enterprise Guide as the software's primary graphical user interface (GUI). [41] The Customer Relationship Management (CRM) features were improved in 2004 with SAS Interaction Management. [43] In 2008, SAS announced Project Unity, designed to integrate data quality, data integration, and master data management . [44]

SAS Institute Inc v World Programming Ltd was a lawsuit with developers of a competing implementation, World Programming System , alleging that they had infringed SAS's copyright in part by implementing the same functionality. The case was referred by the United Kingdom's High Court of Justice to the European Court of Justice on 11 August 2010. [45] In May 2012, the European Court of Justice ruled in favor of World Programming, finding that "the functionality of a computer program and the programming language cannot be protected by copyright." [46]

A free version of SAS was introduced for students in 2010. [47] SAS Social Media Analytics, a tool for social media monitoring, engagement and sentiment analysis , was also released that year. [48] SAS Rapid Predictive Modeler (RPM), which creates basic analytical models using Microsoft Excel , was introduced the same year. [49] [50] In 2010, JMP 9 included a new interface for using the R programming language and an add-in for MS Excel. [51] [52] The following year, a High Performance Computing platform was made available in a partnership with Teradata and EMC Greenplum . [53] [54] In 2011, the company released SAS Enterprise Miner 7.1. [55] The company introduced 27 data management products from October 2013 to October 2014 and updates to 160 others. [56] At the SAS Global Forum 2015, SAS announced several new products

that were specialized for different industries, as well as new training software. [57]

The company has invested in the development of artificial general intelligence, or "strong AI", with the goal of advancing deep learning and natural language processing to the point of achieving cognitive computing. [58]

In 2019, SAS announced that it would be investing $1 billion into the development of advanced artificial intelligence, deep learning, natural language processing and machine learning . [59] It announced an additional $1 billion investment into these areas in 2023, particularly for industries such as finance, insurance, government, health care and energy. [60] In September 2023, the company announced its expansion of research into the applications of generative AI in analytics, data management and modeling. [61]

Software products

As of 2011, SAS's largest set of products was its line for customer intelligence . SAS modules for web, social media and marketing analytics may be used to profile customers and prospects, attempt to predict their behaviors, and manage communications. [62]

SAS also provides the SAS Fraud Framework, which is designed to monitor transactions across different networks and use analytics to identify anomalies that are indicative of fraud. [63] [64] [65] This software uses artificial intelligence to monitor income and assets. [66] [67]

SAS has various analytical tools related to risk management . [68] [69]

SAS has products for specific industries, such as government, retail, telecommunications, aerospace, marketing optimization, and high-performance computing . [70] The company also has a suite of analytical products related to health care and life sciences. [71]

Comparison to other products

In a 2005 article for the Journal of Marriage and Family comparing statistical packages from SAS and its competitors Stata and SPSS , Alan C. Acock wrote that SAS programs provide "extraordinary range of data analysis and data management tasks," but were difficult to learn and use. [72] SPSS and Stata, meanwhile, were both easier to learn but had less capable analytic abilities, though these could be expanded with paid (in SPSS) or free (in Stata) add-ons. Acock concluded that SAS was best for power users , while occasional users would benefit most from SPSS and Stata. [72] A 2014 comparison by the University of California, Los Angeles , gave similar results. [73]

Competitors such as Revolution Analytics and Alpine Data Labs advertise their products as considerably cheaper than SAS's. In a 2011 comparison, Doug Henschen of InformationWeek found that start-up fees for the three are similar, though he admitted that the starting fees were not necessarily the best basis for comparison. [74] SAS's business model is not weighted as heavily on initial fees for its programs, instead focusing on revenue from annual subscription fees. [75] [76]

SAS Viya

In 2016, SAS Viya, an artificial intelligence, machine learning, analytics and data management platform, [77] [78] was introduced with a new architecture optimized for running SAS software in public clouds. Viya also increased interoperability with open source software, allowing models to be built in tools such as R , Python and Jupyter , and then executed on SAS's Cloud Analytics Services (CAS) engine. [79] In 2020, a further architectural revamp in Viya 4 containerized the software. [80] SAS sells Viya alongside SAS 9.4, and has not positioned it as a replacement for SAS 9.4. [81]

In 2023, two new software as a service (SaaS) modules for SAS Viya were released as a private preview: Workbench, for use in creating AI models, and App Factory, for use in creating AI applications. Both modules support multiple programming languages and are expected to become generally available in 2024. [82] SAS Viya also became available on Microsoft Azure Marketplace under a pay-as-you-use model in 2023. [83]

In 2023, the company introduced SAS Health, a common health data model built on the SAS Viya platform. [84]

Adoption

According to IDC, SAS is the largest market-share holder in "advanced analytics" with 35.4 percent of the market as of 2013. [85] It is the fifth largest market-share holder for business intelligence (BI) software with a 6.9% share [86] and the largest independent vendor. It competes in the BI market against SAP BusinessObjects , IBM Cognos , SPSS Modeler , Oracle Hyperion , and Microsoft Power BI . [87] SAS has been named in the Gartner Leader's Quadrant for Data Integration Tools and for Business Intelligence and Analytical Platforms. [88] A study published in 2011 in BMC Health Services Research found that SAS was used in 42.6 percent of data analyses in health service research, based on a sample of 1,139 articles drawn from three journals. [89]

Uses and applications

Education

SAS' analytical software is used in education to measure and visualize student outcomes and growth trends. [90] [91] [92] Several states, including Virginia, North Carolina, Mississippi, Missouri, and North Dakota use its software to measure and analyze learning loss and learning recovery in students. [93] [94] [95]

Environmental science

SAS and the International Institute for Applied Systems Analysis launched an app that crowdsources image data related to deforestation to train AI algorithms that can identify human impact on the environment. [96] The University of Florida 's Center for Coastal Solutions partners with SAS to develop research, training programs and analytical tools related to environmental issues affecting coastal communities. [97]

The UNC Center for Galapagos Studies partnered with SAS in 2023 to create a model that can track the health and migratory patterns of species such as sea turtles and hammerhead sharks , as well as the health of the phytoplankton population. [98]

Finance and insurance

SAS's fraud detection and prevention software is used by the tax agencies of various countries, such as the United States, [99] and Malta . [100]

Healthcare and life sciences

SAS has been a partner of the Cleveland Clinic since 1982. [101] During the COVID-19 pandemic , the clinic used predictive models developed by SAS to forecast factors such as patient volume, availability of medical equipment and bed capacity in various scenarios. SAS joined UNC Chapel Hill's Rapidly Emerging Antiviral Drug Development Initiative (READDI) in 2021. [102]

See also

Comparison of numerical-analysis software

Comparison of OLAP servers

JMP (statistical software) , a subsidiary of SAS Institute Inc.

SAS language

R (programming language)

References

Further reading

Greenberg, Bernard G.; Cox, Gertrude M. ; Mason, David D.; Grizzle, James E.; Johnson, Norman L.; Jones, Lyle V.; Monroe, John; Simmons, Gordon D. Jr. (1978). Nourse, E. Shepley (ed.). "Statistical Training and Research: The University of North Carolina System". International Statistical Review . 46 (2): 171– 207. doi : 10.2307/1402812 . JSTOR 1402812 .

Wikiversity:Data Analysis using the SAS Language

External links

Official website

SAS OnDemand for Academics No-cost access for learners (free SAS Profile required)

A Glossary of SAS terminology

SAS for Developers

SAS community forums

v

t

e

James Goodnight

John Sall

Anthony James Barr

Jane Helwig

SAS (software)

JMP (statistical software)

SAS Viya

SAS language

Artificial intelligence

Data management

Machine learning

Generative artificial intelligence

Statistical software

Category

v

t

e

Dataplot

Epi Info

CSPro

X-12-ARIMA

ADMB

DAP

gretl

jamovi

JASP

JAGS

JMulTi

Julia

Jupyter (Ju lia, Py thon, R)

GNU Octave

OpenBUGS

Orange

PSPP

Python (statsmodels, PyMC , IPython , IDLE)

R (RStudio)

SageMath

SimFiT

SOFA Statistics

Stan

XLispStat

BV4.1

CumFreq

SegReg

XploRe

WinBUGS

Data Desk

GAUSS

GraphPad InStat

GraphPad Prism

IBM SPSS Statistics

IBM SPSS Modeler

JMP

Maple

Mathcad

Mathematica

MATLAB

OxMetrics

RATS

Revolution Analytics

SAS (SAS Viya)

SmartPLS

Stata

StatView

SUDAAN

S-PLUS

TSP

World Programming System (WPS)

BMDP

EViews

GenStat

LIMDEP

LISREL

MedCalc

Microfit

Minitab

MLwiN

NCSS

Shazam

SigmaStat

Statistica

StatsDirect

StatXact

SYSTAT

The Unscrambler

Unistat

Analyse-it

Unistat for Excel

XLfit

RExcel

v

t

e

Advanced Simulation Library

ADMB

Chapel

Euler Mathematical Toolbox

FreeFem++

FreeMat

Genius

Gmsh

GNU Octave

gretl

Julia

Jupyter (Ju lia, Pyt hon, R ; IPython)

MFEM

OpenFOAM

Python

R

SageMath

Salome

ScicosLab

Scilab

X10

Weka

Fortress

DADiSP

FEATool Multiphysics

GAUSS

LabVIEW

Maple

Mathcad

Mathematica

MATLAB

MWorks

SAS (SAS Viya)

Speakeasy

VisSim

Comparison

Czech Republic