

Paper SAS 3194-2019

Amazing End-to-End Demo in No Time: But Wait, There's More...

Gert Nissen, Vegard Hansen

SAS Institute Inc.

ABSTRACT

Amazing End-to-End Demo in No Time: But Wait, There's More...

The time is now, and with the help of this unique product, you can achieve your ultimate goal — all tools available in one solution. In this infomercial, you will see how amazing a product this is and how it's changing the whole game. But wait, there's more! This super-demo covers the entire analytical lifecycle end-to-end on one integrated platform in no time—from SAS® Add-in for Microsoft Office through SAS® Data Integration Studio, SAS® Data Preparation on SAS® Viya®, Model Studio, SAS® Visual Analytics, SAS® Lineage, and SAS® Enterprise Guide® in one demo. You will not ever need anything else, not sold in stores anywhere.

INTRODUCTION

Are you using the right SAS® tools and are you aware of the many tools (ref appendix in the end of the paper – *Table 1: Overview of [the SAS® Platform tools](#)*) that are available? With SAS it's possible to cover the entire Analytics Life Cycle with the SAS Platform.

This paper has the purpose to try to unfold both SAS® 9.4 and SAS® Viya and introduce you to as many as 14 tools. Perhaps you are unsure if it's SAS® 9.4 or SAS® Viya or perhaps how both are used in a seamless combination?

This will hopefully at the end of the paper leave you a little exhausted but also enlightened but with a better understanding that you as a SAS programmer, Analyst or Data Scientist can leverage the full potential of the entire SAS Platform – And if you're an Administrator, Project Manager or Decision maker this should also give a good understanding of the many benefits you could transfer to your business.

Whether you are a very experienced SAS professional or have limited in-dept understanding of SAS, it is our hope that this broad audience and everyone that sees themselves somewhere in-between will gain new knowledge and understanding of the entire SAS Platform covering both SAS 9.4 and SAS Viya – It is a lot to promise in only a few pages, let us know if we succeeded or not – It's almost sounds too amazing to be true 😊.

This paper doesn't try to cover or mention every possible tool in the SAS Platform, but try to put a broad selection of pieces (tools) together till one platform. So regardless of your role in your organization, you can see how it all fits together.

In next section we will describe what a platform are and how it relates to the analytics life cycle. In the appendix you will find a table with overview of the SAS tools that are part of this paper, with short description and links to more information.

WHAT IS A PLATFORM AND HOW DOES IT RELATE TO THE ANALYTICS LIFE CYCLE?

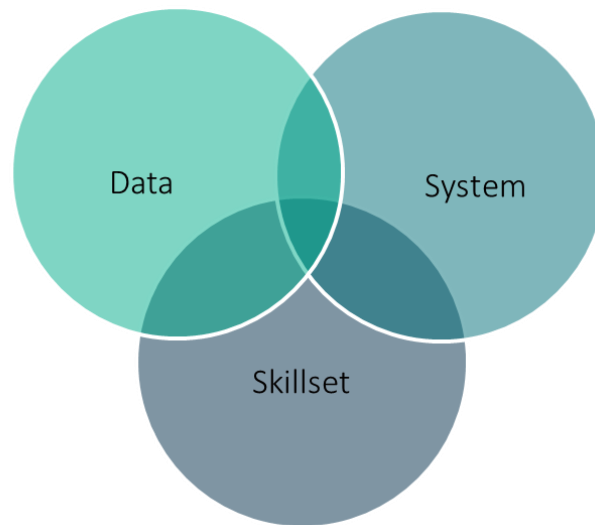


Figure 1. SAS Platform unifies your analytics eco-system

The SAS® Platform unifies your analytics eco-system, allows you to create value from your data, system and skillsets (Figure 1. SAS Platform unifies your analytics eco-system). Helping you to orchestrate the right balance between choice and control that gives organizations the ability to extend and accelerate their analytics life cycle. By doing this, IT can focus on transforming the business with technologies like machine learning, cloud, open source and tons of data and more. Enablement of users to use the tools they need to perform their jobs in a way they need. Governance – that is, rules and policies that prescribe how organizations protect and manage their data and analytics – support the secured assets that are available, depending on every organizational need for analytics (SAS Institute, 2019). In other words, good governance will help organizations become more transparent, ensuring better control.

The «heart» of the analytics life cycle is the SAS® Platform. In its simplest form, the analytics life cycle is a series of activities with the goal to extract value from raw data. There are many ways of presenting the analytics lifecycle, but the very simplified way is to represent it with its three core components; Data, Discovery and Deployment (Figure 2. The Analytics Life Cycle).



Figure 2. The Analytics Life Cycle

Business teams, from marketing and finance to operations and HR, need self-service tools to speed and simplify data preparation and analytics tasks. Such tools may include built-in, advanced techniques like machine learning, and many of them work across the analytics life cycle – from data collection and profiling to monitoring analytical models in production. These “smart” tools feature three capabilities; automation, reusability and explainability. As organizations mature analytically, it’s important for their platform to support multiple roles in a common interface with a unified data infrastructure. This strengthens collaboration and makes it easier for people to do their jobs (SAS Institute, 2018).

Governance is critical in learning to trust data and become data-driven. For instance, decisions based on poor data, or models that have degraded, can have a negative effect on the business. As more people across an organization access data and build models, and as new types of data and technologies emerge (big data, cloud, stream mining), data governance practices need to evolve (SAS Institute, 2018).

Put simple – It’s easy to pick a good tool that will solve the one task at hand, and if you’re alone solving the task and the task is simple, that tool might be just as good as any tool in the toolbox. But when you add more people, more problems to solve, regulations, governance and limited time to execute – You need a robust and flexible platform where everything is integrated, documented and not least secure!

HOW DO YOU WORK?

Sometimes a formal requirement specification is the basis for the work, but perhaps more often its more in the format of being informal and ad-hoc – but nevertheless the starting point there are some similarities – getting data, exploring and understanding the data, cleansing the data before you can finalize your analysis and visualize the result. This is often an iterative process with more trial-and-error efforts.

Even though this sounds like wild-west ruling, you need a platform that can support you in both getting your work done, while maintaining security and governance. And if you can reuse your skills from one tool to the next because of ease of use and an intuitive fast adoption, then you are well off.

In next section we will show all the tools that are part of this Platform Overview from end-to-end by going through the steps in the analytics life cycle.

”ALL” TOOLS AVAILABLE IN ONE SOLUTION – THE SAS® PLATFORM

We are slowly getting into the details about how we work as SAS analysts - How to go from Data to Decisions as quickly as possible and with as few bumps on the way.

TOOLS YOU GET IN THE PLATFORM – DATA

Data is the first step in the lifecycle. We need data – if possible beautiful data (without data quality issues), that we can *prepare* and *explore* on our way to decision. SAS can capture data of any source, complexity, speed and size, whether it is traditional structured data or a new format such as streaming sensor data. We have a robust suite of data management tools, like SAS® Data Integration Studio on SAS9 and SAS® Data Studio on SAS® Viya (Figure 3. SAS Platform tools - for data). The SAS tools work also seamlessly with other data management products on the market.

Data preparation is estimated to take up to 80 percent of the time spent on an analytics project. That is time that can be better spent building models. The SAS platform streamlines data preparation with native access engines to your many data sources, integrated data quality, and data preparation tools that leverage artificial intelligence to automate time consuming tasks. This makes it easy to create and monitor high quality ongoing comparing with new data, not only with its data quality issues, but also against your profiling baseline, to ensure that data are ready for analytics and reporting.

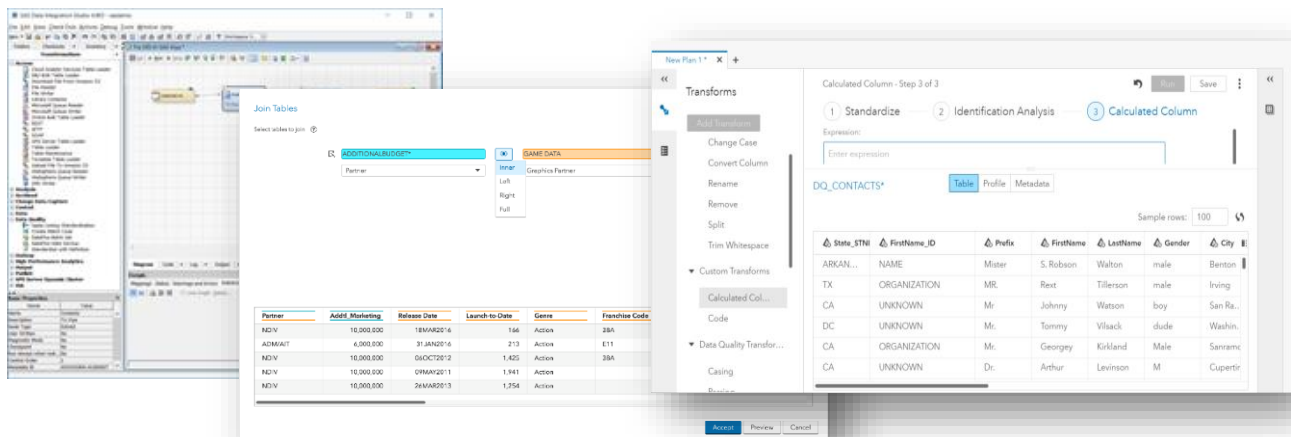


Figure 3. SAS Platform tools - for data

BUT, THERE IS MORE - DISCOVERY

The second step in the analytics life cycle – Discovery – is all about *exploration*, *visualization* and *model building*. With the platform organizations have access to an incredible breadth of analytic techniques and languages. The analytic techniques may be SAS® Analytics or they may be from Open Source languages. It means analysts can choose between several languages to explore and to build reports and models.

Regardless of the skillset, a wide variety of users can collaborate through the SAS platform and tackle analytic challenges. Business users and non-programmers can for instance use,

among many other SAS tools, the visual point and click interface of SAS® Visual Analytics, SAS® Visual Statistics and SAS® Visual Data Mining and Machine Learning, to take advantage of analytics and visualizations. Advanced analysts, statisticians and Data Scientists can write code in the language of their choice, such as Python, R or Lua, or they can code in SAS using SAS® Studio or SAS® Enterprise Guide – with the same shared data access setup and secured centrally by the administrator. They all can benefit using SAS® Model Studio which provides a suite of analytic data mining tools to facilitate end-to-end data mining analysis. Additionally, analytics and (BI) reports can be accessed through mobile devices or through SAS® Add-in for Microsoft Office. Some examples of SAS tools for Discovery are displayed below in Figure 4. SAS Platform tools - for discovery.

Other tools for discovery are for instance SAS® Times Series Forecasting, SAS® Optimization, SAS® Econometrics, SAS® Data & Text Mining, etc.

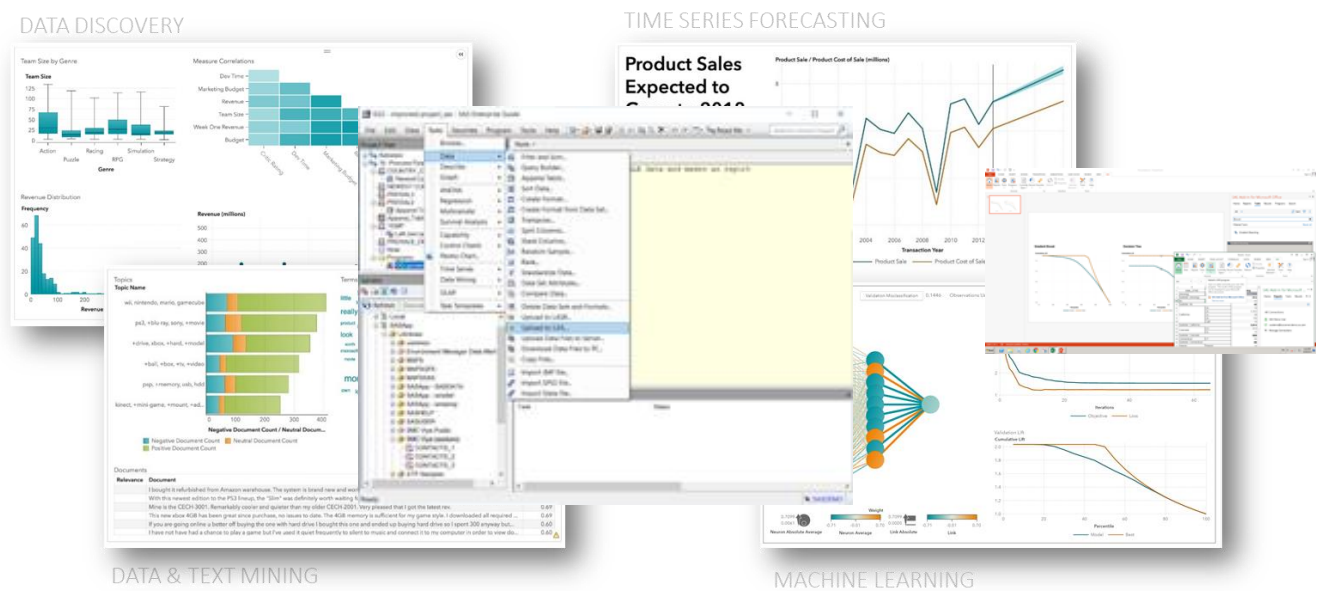


Figure 4. SAS Platform tools - for discovery

....BUT WAIT, THERE IS EVEN MORE – DEPLOYMENT

Deployment is the third step in the analytics life cycle, also called the *implementation* part, where all the efforts to get and analyze data pay off. This is also where most of the organizations struggle most.

Model management capabilities (tools) in SAS make sure that organizations can quickly move from creating and selecting a model to deploy. This is regardless of whether you are building a one-off single model or are building thousands of models. It's critical that organizations can evaluate all the models for a specific business problem, regardless of the language the model was written in. They need this to figure out which model should be deployed because it performs best, i.e. which is the champion, and which models are the challengers. This needs to happen quickly and efficiently and repeatedly. Tools and capabilities for deployment in SAS® 9 and SAS® Viya are SAS® Model Manager and SAS®

Decision Manager.¹ (Figure 5. SAS Platform - for deployment). For instance, in Model Manager you can test and administrate your models (and data), having control over model performance, version control and track of history.

When the champion model is selected then it can be deployed. With SAS tools for deployment, organizations no longer need to recode models into various languages for deployment and then test the results to insure they perform as expected. A cycle that must be repeated next time the model needs to be refreshed. The SAS Platform for Deployment therefore helps organizations to reduce their time-to-market and supports an accelerating of the analytics life cycle.

A dashboard/report in SAS® Visual Analytics or to a mobile device is an example of monitoring of the analytic deployment.

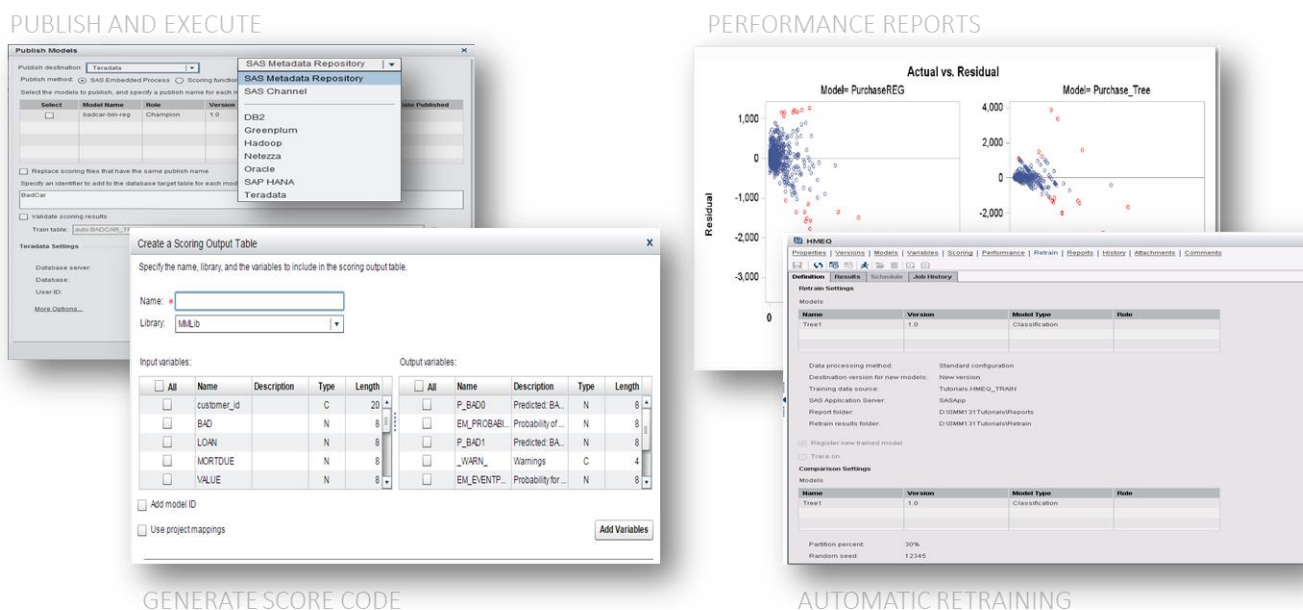


Figure 5. SAS Platform - for deployment

FINALLY – AND NOT TO FORGET – THERE IS EVEN MORE – ORCHESTRATION

A unique feature of SAS’ platform is its ability to orchestrate the entire analytic eco-system at an organization. There is no longer enough to provide data, discovery and deployment capabilities. You also need the platform and tools to connect and accelerate the activities of the analytic life cycle. This part is essential for finding the right balance between business need of choice and the organizations need for control, and some of these features include SAS tools like

SAS® Data Lineage (Figure 6. SAS® Data Lineage), which gives an upstream and downstream understanding of the relationships between analytic objects such as data,

¹ SAS® Model Manager and SAS® Decision Manager are also an important part of the orchestration part of the platform, described in next section - ‘FINALLY – and not to forget – there is even more – Orchestration’.

transformations, models and reports to help keep them in synch. This is critical for efficient modelling activities and model deployment.



Figure 6. SAS® Data Lineage

SAS® Model Manager. You see model management in both the deployment section as well as the orchestration section. For Deployment it's about which model to deploy, in Orchestration it's about the on-going monitoring and automated re-training etc. The Platform automatically monitors model performance over time to make sure they continue to perform as expected, regardless of the language it was created in. If model performance is not maintained, you are notified and can act to improve, replace or remove the model from production.

SAS® Environment Manager (Figure 7. SAS® Environment Manager). The SAS Platform provides additionally versatility in where processing occurs, allowing for efficient analytics execution closer to the data. Whether data is stored in a database, distributed in Hadoop, loaded into memory, saved in the cloud, or streaming in from the Internet of Things, a corresponding SAS Platform runtime engine is available to accelerate results by avoiding unnecessary data transfer or duplication, content, folders and/or user.

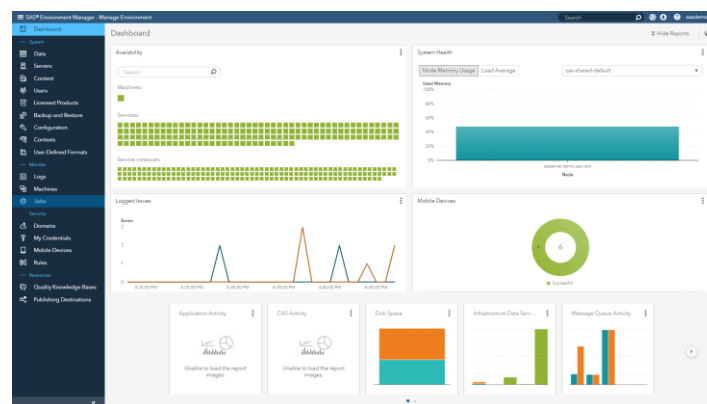


Figure 7. SAS® Environment Manager

Orchestration doesn't slow down analytic processes, it instead allows organizations to be agile because staff don't have to spend time manually performing tasks.

CONCLUSION

In this paper we have tried to present an example of end-to-end platform or life cycle in SAS. An end-to-end life cycle depends on the problem, skills and time and every business and organization have its own goals and objectives for creating values and solving high value tasks. This paper gives an overview of the SAS Platform in one context, not trying to cover every possibility. There are many other products that could be added to the toolbox, like e.g. SAS® Text Mining, SAS® Optimization, SAS® Forecasting and SAS® Econometrics and tools for real-time decisioning like SAS® Event Stream Processing (ESP) and SAS® Real-Time Decision Manager (RTDM). Tools that are covering other business objectives than in this paper. These are not stand-alone tools but are also integrated with the platform and are part of the analytics life cycle.

TV-shop tools can be a nice affordable shortcut for a business, but knowledge needs to be added to the tool and prototyping will only bring you some way in the analytics life cycle. You will also need to bridge your results from this kind of tools to deployment and governance, in need for control, securing and orchestrating your platform.

Working with data is never linear – there are always unknowns, fallbacks and luck/experience involved in the process and with more data, more complex data, more analytics applied and faster response time to even more end users you need good tools and a robust platform to support these demands.

When you see the picture below (Figure 8. The Analytical Life Cycle) and read the different sections in this paper you might think – I don't need this many tools and why make it so complicated. Your right, one analyst will probably not need many of the tools, but the entire life cycle needs different competences and therefore tools – But it's not impossible to master several of them.

Realize that many problems perhaps are not discovered/explored because the tools/solutions are not known.

Nevertheless, the basic ideas around a platform overview in this paper should be a valid trial, and hopefully invite you to find out where your business stands regarding having the right balance between choice and control to orchestrate your analytics journey. And, regardless of your role or use of tool, there's no need for migration – you can continue to work with the tools you prefer.



Figure 8. The Analytical Life Cycle

REFERENCES

SAS Institute. 2019. The SAS® Platform - Orchestrate your entire analytics journey – from data to tangible results. Accessed March 2019. Available at:

https://www.sas.com/en_us/software/platform.html

SAS Institute. 2018. Managing the Analytics Life Cycle for Decisions at Scale. Accessed March 2019. Available at:

https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/manage-analytical-life-cycle-continuous-innovation-106179.pdf

SAS Institute. 2016. Data Mining From A to Z, How to Discover Insights and Drive Better Opportunities. Accessed March 2019. Available at:

https://www.sas.com/en_us/whitepapers/data-mining-from-a-z-104937.html

SAS Institute. 2018. Building your data and analytics strategy. E-book, available at:

<https://www.sas.com/en/white-papers/building-data-analytics-strategy-110139.html>

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Vegard Hansen

SAS Institute Inc.

Phone: +47 977 44 445

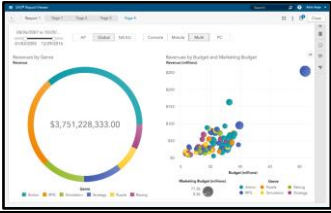
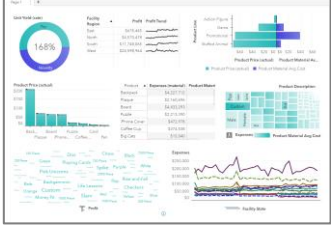
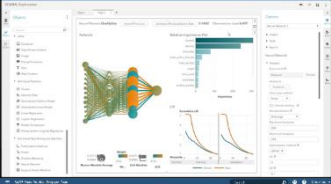
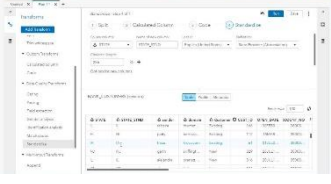
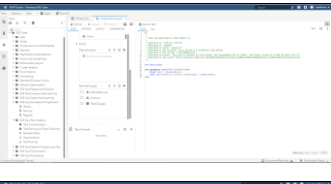
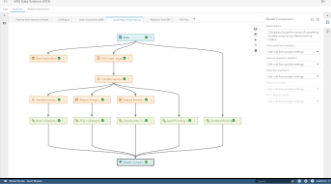
E-mail: vegard.hansen@sas.com

<http://www.sas.com/norway>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

APPENDIX: OVERVIEW OF TOOLS

Table 1: Overview of [the SAS Platform](#) tools

Tool	Doc	Role	Desc	Print Screen	Free Trials
SAS Visual Analytics	Doc	Report consumer / Exploration Report Developer / Deployment	By using SAS Visual Analytics, users can enhance the analytic power of their data, explore new data sources, uncover relevant patterns, and create reports.		Trial
SAS Visual Statistics	Doc	Exploration / Model developer	SAS Visual Statistics extends SAS VA capabilities by creating, testing, and comparing models based on the patterns discovered		Trial
SAS Visual Data Mining and Machine Learning	Doc	Exploration / Model developer	SAS Visual Data Mining and Machine Learning enables you to rapidly create powerful data mining and machine learning models in an easy-to-use, web-based interface.		Trial
SAS Data Studio	Doc	Data Steward	SAS Data Studio offers an easy way for you to prepare data with data transforms such as joining tables, appending data to a table, transposing columns, creating calculated columns, and applying data quality transforms like parsing, standardizing, match operations etc. using SAS Quality Knowledge Base (QKB)		Trial
SAS Studio	Doc	SAS Programmer	SAS Studio is a development application that you access through your web browser. With SAS Studio, you can access your data files, libraries, and existing programs, and you can write new SAS programs. You can also use the predefined tasks and snippets to generate SAS code		Trial
SAS Model Studio	Doc	Model Developer	Model Studio is a visual environment that provides a suite of analytic data mining tools to facilitate end-to-end data mining analysis.		
SAS Model Manager	Doc	Model development/maintenance	Using SAS Model Manager, you can store models in a common model repository, and organize them within projects and folders. You can also evaluate models for champion model selection, monitor performance of models, and publish models.	