## How Can I Run My DATA Step Programs in SAS Viya

#### T Winand

Senior Solutions Architect Customer Success Organization





## William "T" Winand Senior Presales Solutions Architect, SAS

T is a computer software professional and analyst with 24 years of experience in program design and development, data management, statistical analysis and reporting. A part of SAS since 1995, he helps organizations grow through analytics. He specializes in statistical analysis, descriptive and predictive data mining, text mining, forecasting and model management.



#### Viya Programming

## Agenda



- Introduction to Viya
- What is CAS?
- SAS9 and Viya
- Programming in Viya
- Considerations
- Resources for Learning More

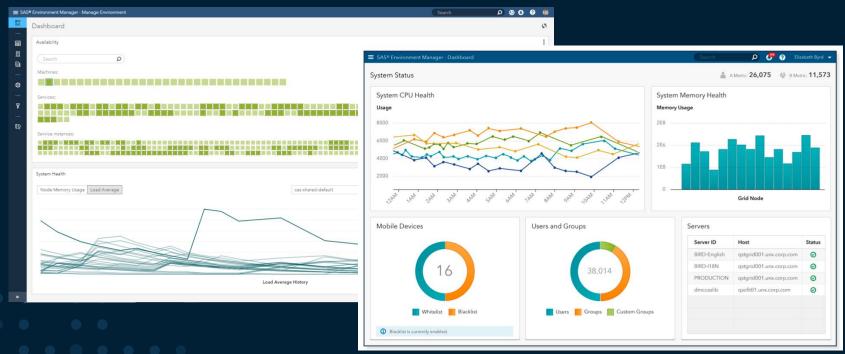


## **Introduction to Viya**



## What is SAS Viya?

SAS Viya is a cloud-enabled, in-memory analytics environment that provides quick, accurate and reliable analytical insights.





## What is CAS?



## Cloud Analytic Services (CAS)

## **In-Memory Engine**



The CAS in-memory engine is a fast, scalable, and resilient run-time environment for data management and analytics for SAS<sup>®</sup> Viya<sup>™</sup>





- Multi-threaded
- Distributed In-Memory
- Efficient Inter-node
   Communication



#### Scalable

- Single Machine to Distributed MPP
- Memory-mapping to process data larger than physical memory \* (No memory failures)



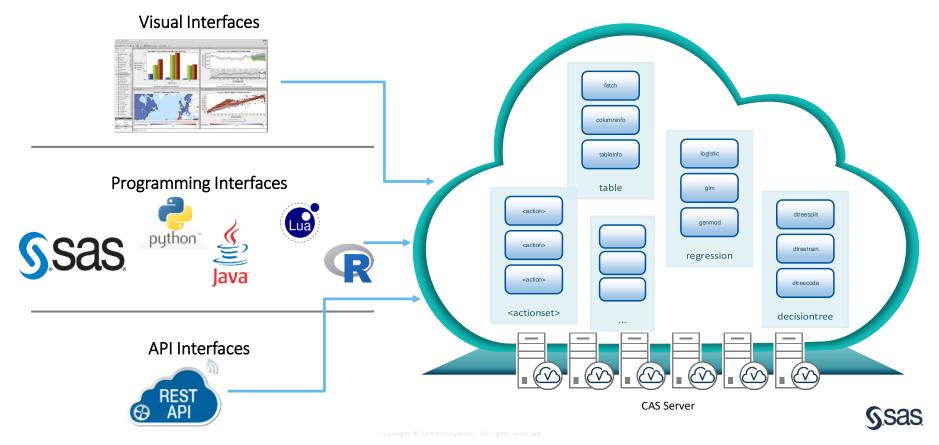
#### Resilient

- Fault-tolerant to Node and Network Failures
- Worker and Controller\* Failover
- Session Independence



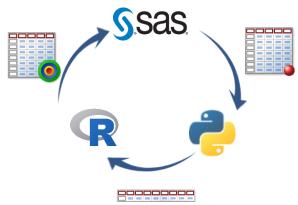
## Multiple interfaces, single code base

Clients ask CAS to run "actions" on data



## Open Architecture Advantage

- Same CAS action regardless of the interface and API.
- Results will be equivalent (given that all random elements are seeded).
- Information can be passed between different languages.
  - Promote tables to persist on the server and apply CAS actions from different APIs.





## **How APIs Work**

```
proc logselect data=mycas.getStarted;
 class c1;
 model y = c1 x1 x2;
 selection method=forward details=all plots=all;
run;
```

```
proc cas;
 regression.logistic /
   table={name='getStarted', caslib='mycas'},
   class={{vars={'c1'}}}, model={depVars={{name='y'}}},
   effects={{vars={'c1'}},{vars={'x1'}},{vars={'x2'}}}},
   selection={method='forward', details='ALL'};
run;quit;
```



```
mdl = s.logistic(table={'name':
'getStarted','caslib':'mycas'},
 classvars='c1',
 model={'depvar':'y', 'effects':['c1', 'x1', 'x2']},
 selection={'method':'forward','details':'all'})
```



```
cas.regression.logistic(s,
  table=list('name'='getStarted','caslib'='mycas'),
  classvars='c',
 model = list('depvar'='y', 'effects' =c('c1','x1','x2')),
  selection = list('method' = 'forward', 'details'= 'all'))
```



Controller





#### **Translated Action**

```
action regression.logistic /
 table={name='getStarted', caslib='mycas'},
 class={{vars={'c1'}}},
model={depVars={{name='y'}},
 effects={{vars={'c1'}},{vars={'x1'}},{vars={'x2'}}}},
```

curl -X POST http://.../cas/sessions/.../actions/regression.logistic \ -u sasdemo:XXXXXXX -H 'Content-Type: application/json' \ -d ="{\"table\":{\"caslib\":\"demodata\",\"name\":\"getStarted\"}

,\"class\":\"C\",\"model\": $\{\"depvar\":\"y\",\"effects\":[\"c1\",\"x1\",\"x2\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x2\",\"x1\",\"x2\",\"x2\",\"x1\",\"x2\",\"x2\",\"x1\",\"x2\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x2\",\"x1\",\"x2\",\"x1\",\"x2\",\"x2\",\"x2\",\"x1\",\"x2\",\"x2\",\"x1\",\"x2\",\"x2\",\"x1\",\"x2\$ 

\"selection\":{\"method\":\"forward\".\"details\":\"all\"}} "



## SAS 9.4 and SAS Viya The SAS Platform



# SAS® 9.4 and SAS® Viya® A Single Platform

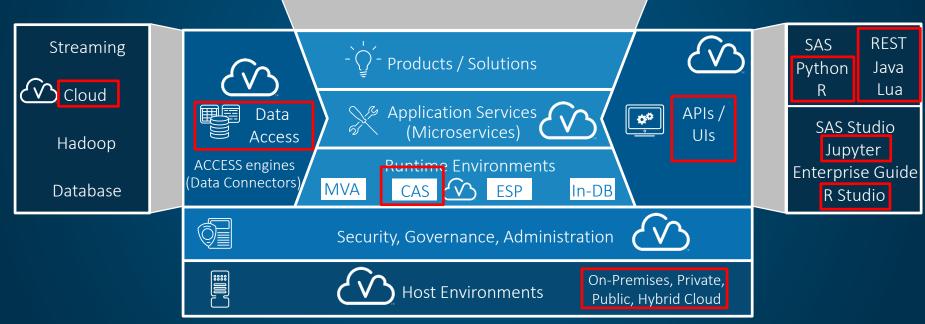
- SAS Platform
- Analytics Life Cycle
  - Data Management
  - Visualization
  - Analytics
  - Decision Management





## The SAS Platform

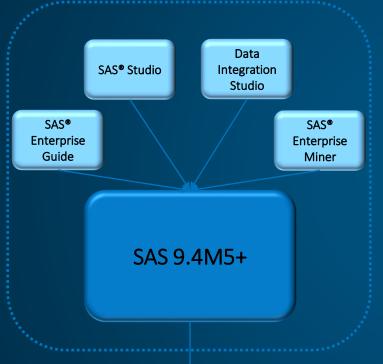
CI Visualization Risk Fraud & Security Analytics Data Management





## The SAS Platform

SAS 9.4 Language Execution SAS Viya





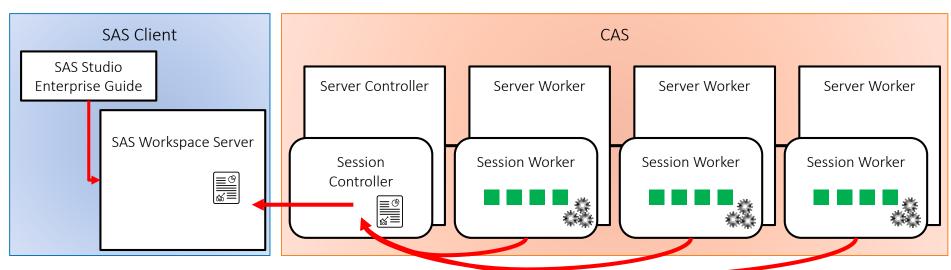


<sup>\*</sup>Many SAS Viya clients (e.g., SAS Visual Analytics) do not submit SAS code, but rather call CAS directly

## SAS Viya Data Processing

## CAS-enabled data processing

- Work is done by the CAS Session Workers
- Results passed to the CAS Session Controller for consolidation
- Passed back to the SAS 9 client

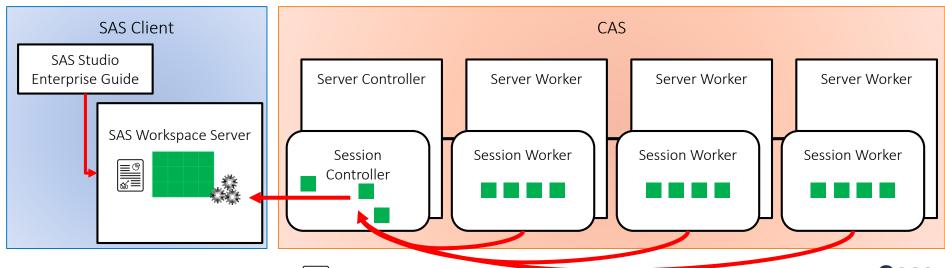




## SAS Viya Data Processing

## **NOT** CAS-enabled data processing

- Data is retrieved from CAS
- Passed to the SAS 9 Workspace Server or Compute Server
- Work is done by the Workspace Server/Compute Server single-threaded



## The SAS Platform

**Base SAS** 

- All 9.4 Base in Viya 3.3+
- SAS Viya = speed!
  - Multi-threaded DATA step
    - Rework code to leverage
    - Steven Sober's SGF Paper #1710-2018
      - SAS Viya Readiness Utility
  - Procedures That Use CAS Actions (16)



## SAS Viya Data Processing

General information

- Most analytics run in memory
  - Visual Statistics, Visual Forecasting, VDMML, Opt
- SAS Foundation PROCs
  - CAS-enabled **¬**
  - Not CAS-enabled ¬
- CAS Actions
  - PROC CAS (CASL)
  - Python
  - Lua...

Procedures That Use CAS Actions

CAS Processing of Base SAS Procedures

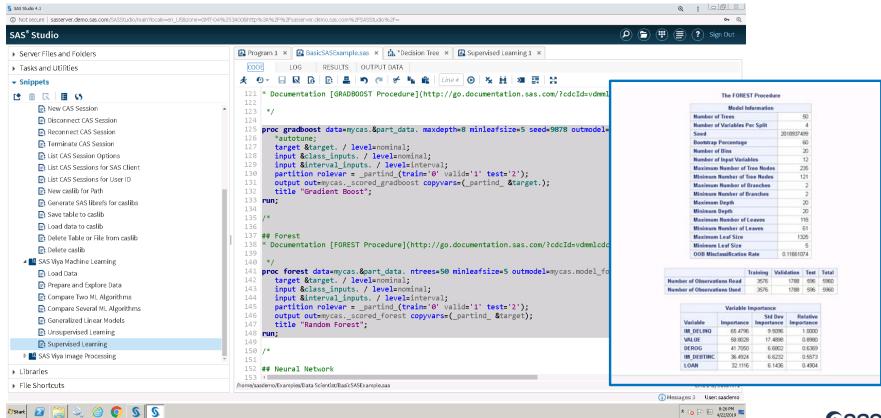
Append, Contents, Copy, Datasets, Delete, DS2, FCMP, FedSQL, Format, Lua, Means, Report, ScoreAccel, Summary, Tabulate, Transpose

Catalog, Compare, Download, DSTODS2, Export, FMTC2ITM, Hadoop, HDMD, HTTP, Import, JavaInfo, JSON, MapImport, Options, Print, PrintTo, Product\_Status, PWEncode, Registry, S3, SGPanel, SGPlot, SGRender, SGScatter, Sort, SQL, Stream, Template

## Viya Programming in four easy steps



## **SAS Studio**





### PROC versus CAS Action

```
proc cas;
    action factmac result=R / table={name="movlens"},
    outModel={name="factors_out", replace=true},
    inputs={"userid", "itemid"},
    nominals={"userid", "itemid"},
    target="rating",
    maxIter=20, nFactors=10, learnStep=0.15,
    output={casout={name="score_out", replace="TRUE"},
    copyvars={"userid","itemid","rating"}};
run;
```



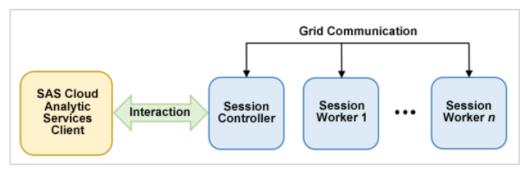
## Step 1: Establish a CAS Session

#### Purpose: Enables the client to communicate with the server

- User identification
- Fault isolation for each session

- Efficiency
- Resource tracking

#### Session Processes in a Distributed System





#### Step 1: CAS Statement

#### Purpose: Start and manage your CAS session

- Create a session
- Terminate a session
- List information about a specific CAS session or all your sessions
- List the properties of a session

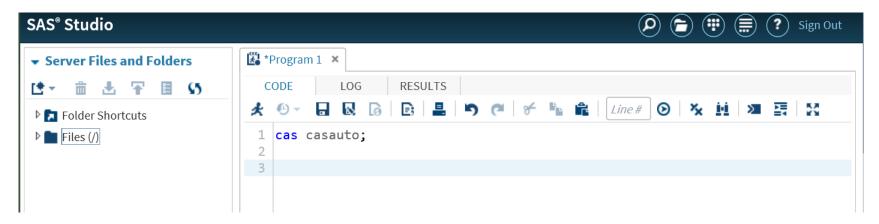
- Manage format libraries in a session
- Change one or more session properties
- Disconnect a session
- Connect to an existing session



#### Step 1 continued

#### Syntax

CAS session-name <option(s)>;



73 cas casauto;

NOTE: The session CASAUTO connected successfully to Cloud Analytic Services viya47.dept-tbt.sashq-r.xxxxxxxxxxsas.com using port

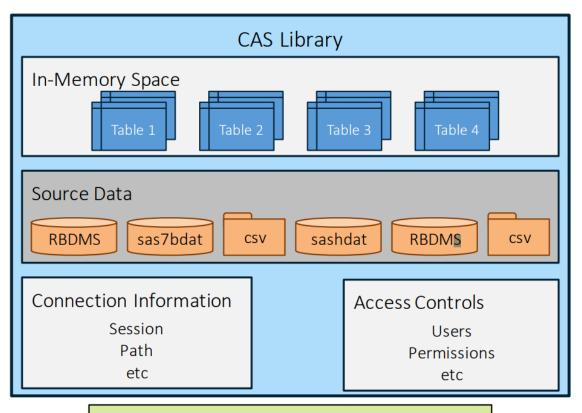
NOTE: The SAS option SESSREF was updated with the value CASAUTO.

NOTE: The SAS macro SESSREF was updated with the value CASAUTO.

NOTE: The session is using 4 workers.



## What is a CASLIB?



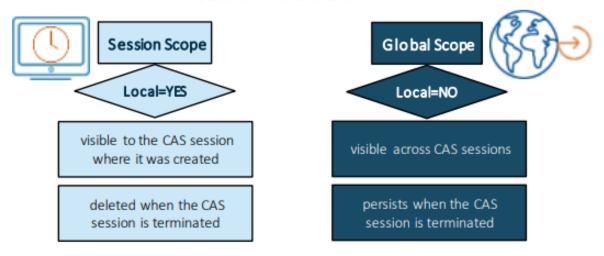
Logical CAS Library Representation



## **CASLIB SCOPE**

#### Caslib Scope

Session Versus Global





## **Step 2:** Read Data into CAS

Multiple file types, including the following:

• SAS data set

Hive

• Delimited text files

Oracle

Microsoft Excel files

Teradata

Multiple methods of reading in data, including the following:

- DATA step
- PROC CASUTIL
- PROC CAS



#### **Step 2:** Read a SAS Data Set Using the DATA Step

#### Code:

```
libname mycas cas
caslib='casuser';
```

```
data mycas.cars;
  set sashelp.cars;
run;
```

#### Log:

```
libname mycas cas caslib='casuser';
NOTE: Libref MYCAS was successfully assigned as
follows:
Engine: CAS
Physical Name: b864dbe5-75d9-9e4a-aa07-7d15ba1d33f2
76
77 data mycas.cars;
78 set sashelp.cars;
79 run:
NOTE: There were 428 observations read from the data
set SASHELP.CARS.
NOTE: The data set MYCAS.CARS has 428 observations
and 15 variables.
NOTE: DATA statement used (Total process time):
real time 0.00 seconds
cpu time 0.00 seconds
```



#### Step 2: Read in an Excel File Using PROC CASUTIL

#### Code:

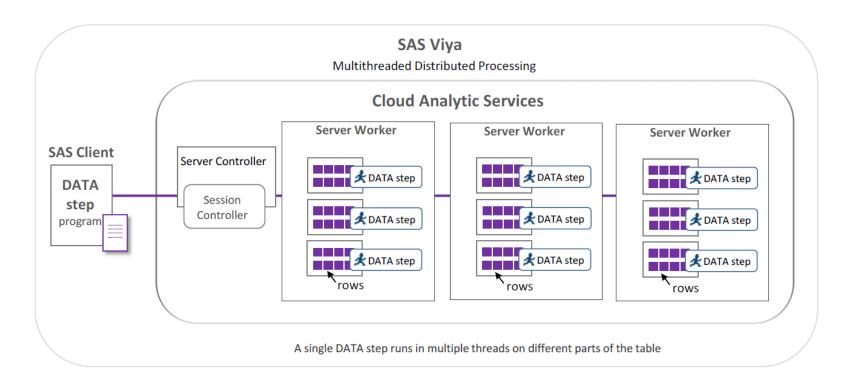
```
proc casutil;
load file='/tmp/cars.xls'
casout='cars2'
outcaslib='casuser'
importoptions=(filetype='xls'
getnames=true);
quit;
```

#### Log:

```
proc casutil;
NOTE: The UUID 'b864dbe5-75d9-9e4a-aa07-
7d15ba1d33f2' is connected using session CASAUTO.
77 load file='/tmp/cars.xls'
78 casout='cars2' outcaslib='casuser'
79 importoptions=(filetype='xls' getnames=true);
NOTE: Cloud Analytic Services made the uploaded file
available as table CARS2 in caslib CASUSER(sasdemo).
NOTE: The table CARS2 has been created in caslib
CASUSER (sasdemo)
from binary data uploaded to Cloud Analytic
Services.
80 quit;
NOTE: PROCEDURE CASUTIL used (Total process time):
real time 0.04 seconds
cpu time 0.00 seconds
```



## Step 3: Manipulate the Data Using the DATA Step





### Step 3: Manipulate the Data Using the DATA Step continued

#### Code:

```
data mycas.updated_transaction_history;
  set mycas.transaction_history;
  year=year(transaction_dt);
  fee=0;
  if transaction_status='SUCCESSFUL' then
do;
  if year=2013 then fee=1;
   else if year=2014 then fee=2;
   else if year>2014 then fee=3;
  end;
new_transaction_amt=transaction_amt+fee;
  new_act_balance=act_balance-
new_transaction_amt;
  put _threadid_=;
run;
```

#### Log:

```
NOTE: Running DATA step in Cloud Analytic
Services.
NOTE: The DATA step will run in multiple threads.
THREADID =2
 THREADID =4
THREADID =1
THREADID =3
NOTE: Duplicate messages output by DATA step:
THREADID =2 (occurred 720000 times)
 THREADID =4 (occurred 719059 times)
 THREADID =1 (occurred 721000 times)
THREADID =3 (occurred 720000 times)
NOTE: There were 2880059 observations read from
the table TRANSACTION HISTORY in caslib
CASUSER (sasdemo).
NOTE: The table updated transaction history in
caslib CASUSER(sasdemo) has 2880059 observations
and 8 variables.
NOTE: DATA statement used (Total process time):
real time 3.86 seconds
cpu time 0.00 seconds
```

#### Step 3 continued

```
88
     data updated transaction history;
     set mylib.transaction history4;
89
90
     year=year(transaction dt);
    fee=0;
     if transaction status='SUCCESSFUL' then
do;
93
     if year=2013 then fee=1;
    else if year=2014 then fee=2;
94
     else if year>2014 then fee=3;
    end;
   new transaction amt=transaction amt+fee;
    new act balance=act balance-
new transaction amt;
99
     run;
```

```
NOTE: There were 2880059 observations read from the data set
MYLIB.TRANSACTION_HISTORY4.

NOTE: The data set
WORK.UPDATED_TRANSACTION_HISTORY has 2880059 observations and 9 variables.

NOTE: DATA statement used (Total process time):

real time 47.93 seconds cpu time 1.42 seconds
```



## Step 4: Analyze the Data Using Base SAS Procedures

#### **PROC MEANS:**

```
proc casutil;
  load data=sashelp.cars outcaslib="CASUSER"
  casout="cars" replace;
quit;

proc means data=casuser.cars n mean max min;
  var weight mpg_city;
  by type;
run;
```



#### Step 4: Analyze the Data Using Viya Procedures

#### PROC MDSUMMARY:

```
data mycas.updated transaction history2;
  set mycas.updated transaction history;
 month=put(transaction dt,monname8.);
run;
proc mdsummary
data=mycas.updated transaction history2(where=(fee ne 0));
 var fee;
 groupby year month / out=mycas.summary transaction history;
run;
proc print data=mycas.summary transaction history label;
  title 'Summarized Transaction History Data';
 var year month min max nobs sum;
  label sum ='Total collected ($)' nobs_='Number of Fees
 Collected';
  format sum dollar8.;
run;
```



## Step 4: Analyze the Data Using PROC CAS

PROC CAS using the Simple Action Set:

```
proc cas;
session casauto;
simple.freq /
    inputs={'transaction_status'}
    table={caslib='casuser', name='updated_transaction_history2',
        groupby={name='year'}};
quit;
```



## Demonstration



## Keep In Mind...

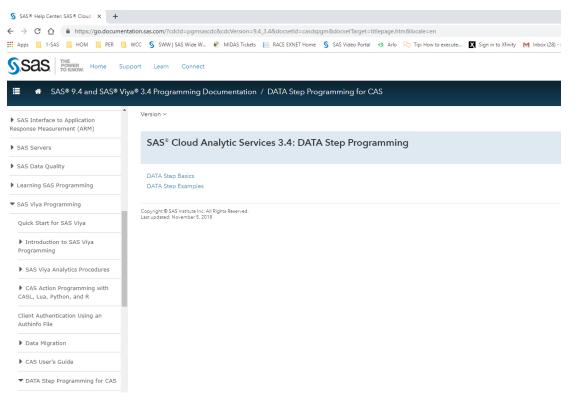
- Your SAS 9 code will work in Viya
- Code & Procedures that cannot effectively take advantage of CAS, run in SPRE
- SAS is smart (it will run process in CAS when it can no user intervention required)
- If your tables are loaded into CAS, much of your data step code will run in CAS
- Use PROC CASUTIL to load your data into CAS (versus DATA Step)
- BUT REMEMBER... You don't need to convert everything to run in CAS.
  - and you may not want to (e.g. By Variables with High Cardinality)
  - think architecture, strengths, and tradeoffs
  - Go for the big wins long running steps, large tables, analytical procedures
- If use SAS Enterprise Guide in SAS 9, you can still use it in Viya
- CAS does not care about order
- PROC SQL runs in SPRE; PROC FEDSQL runs in CAS (if associated with a CAS Session)



# Resources



## **Documentation**



Viya Programming Documentation



## **Video Tutorials**









Viya

> Accessing and Manipulating Data in SAS Viya

Sign In United States Contact Us SOBICH

You may have heard of SAS Viya, a modernization of the SAS platform, but you might not be familiar with how to access and manage your data in SAS Viva

> Leveraging SAS Viya to Improve Base SAS Processes

Explore technical insights and implementation details to get your code CAS-enabled, achieving faster results

> SAS Grid Manager and SAS Viya: A Perfect Pair

Join us for this webinar and learn how SAS Grid Manager works with SAS Viya to process large volumes of data for rapid results.

> SAS Optimization on SAS Viya: Getting Started

In this presentation, you'll get coding examples for linear programming, mixed integer linear programming, nonlinear programming and network optimization.

> SAS Visual Data Mining and Machine Learning: Getting Started

The presentation and accompanying demonstration will offer a view into SAS Visual Data Mining and Machine Learning.

> SAS Visual Statistics on Viya: Getting Started

Quickly create predictive models through the visual point-and-click interface or through programming using SAS Studio or other programming interfaces.

> SAS Viya Architecture and Data Movement: A SAS Programmer's Overview

Experienced SAS programmers can learn more about SAS Viya architecture and SAS Viya programming.

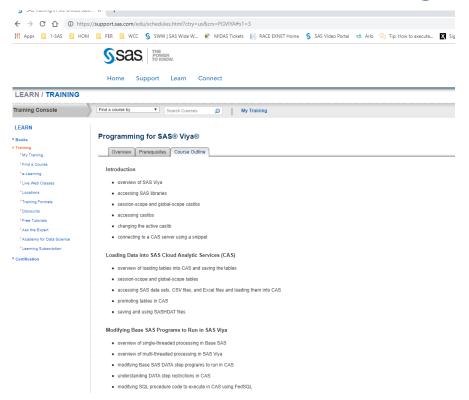
> SAS Viya: Working With APIs

Experienced and novice SAS users can join us to learn how and when to use APIs to run machine learning models in SAS Viya.

Free Webinars

Ask The Expert Webinars

## **Training**



To access free e-learning courses:

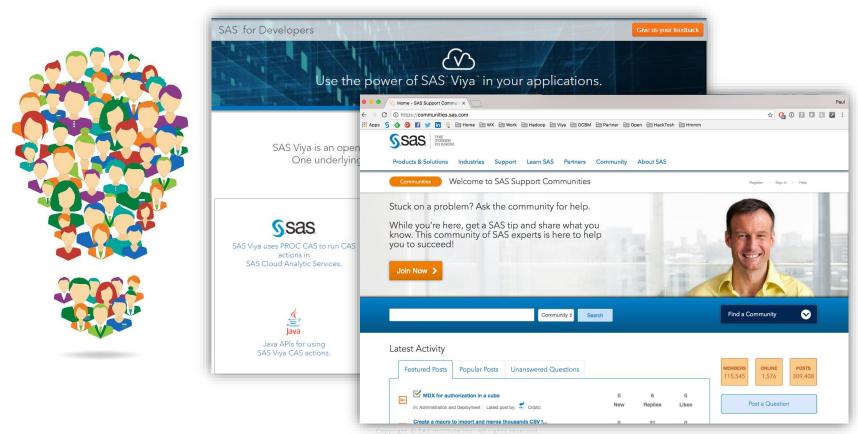
Free e-Learning for Employees

**Programming for SAS Viya** 



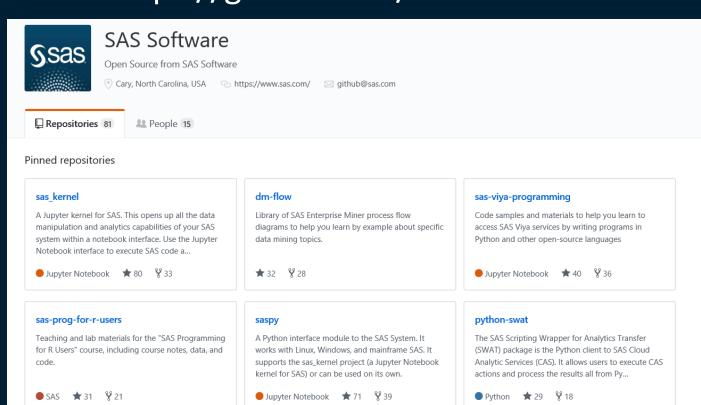
## **Useful Websites**

Developer.sas.com, Communities.sas.com





# https://github.com/sassoftware





# Resources Programming

- SAS Studio
- CAS actions documentation
- SAS Github page for SWAT-Python
- SAS Github page for SWAT-R
- More example scripts for using SWAT-R & SWAT-Python



## SAS® Viya Resources

### SAS Visual Statistics User's Guide

http://support.sas.com/software/products/visual-statistics/index.html#s1=2

## SAS Visual Data Mining and Machine Learning User's Guide

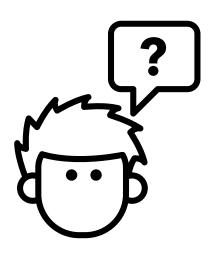
http://support.sas.com/software/products/visual-data-mining-machine-learning/index.html#s1=1

## SAS Visual Text Analytics User's Guide

http://support.sas.com/software/products/visual-text-analytics/index.html

## Overview, Training, Samples and Tips

- SAS Viya Overview
- SAS Viya Training
- A Beginner's Guide to Programming in the SAS® Cloud Analytics Services Environment





Q&A



## **Explore Helpful Resources**

### Ask the Expert

View other user webinars that provide insights into using SAS products to make your job easier.

### **FREE Training**

Learn from home - free for 30 days. Get software labs to practice and online support if needed.

### **SAS Support Communities**

Ask questions, get answers and share insights with SAS users.

### **SAS Analytics Explorers**

An exclusive platform to collaborate, learn and share your expertise. Gain access to a diverse network to advance your career. Special rewards and recognition exclusively for SAS users.

#### SAS Users YouTube Channel

A plethora of videos on hundreds of topics, just for SAS users.

#### **Newsletters**

Get the latest SAS news plus tips, tricks and more.

### **Users Groups**

Meet local SAS users, network and exchange ideas – virtually.

#### **SAS Profile**

If you haven't already done so, create your SAS Profile to access free training, SAS Support Communities, technical support, software downloads, newsletters and more.

