

Tools and Techniques that can facilitate Data-Driven Decisions and Data Automatisaton

Daniel Christen

Advisory Life Sciences Industry Consultant

SAS Institute Inc.

PHUSE Single Day Events



Heidelberg, Germany

Thursday 27 June 2024

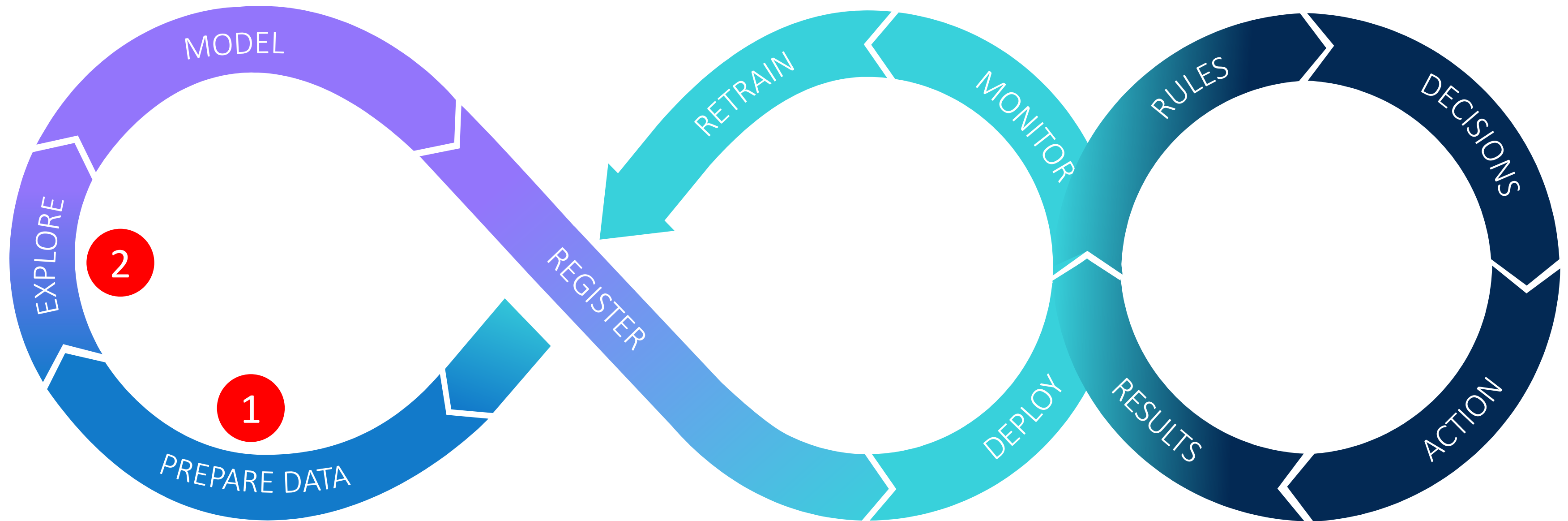
Abstract

Data-driven decisions and data automatization are two key aspects of modern clinical trials that can improve efficiency, quality and reliability. In this presentation, we will start with improving your data quality assurance process, and how with high quality data your team can gather robust insights of safety and efficacy indicators of your ongoing clinical trials. We will also showcase some examples of tools and techniques that can facilitate data-driven decisions and data automatization, such as building easy to use applications. The presentation will highlight the benefits and challenges of these approaches.

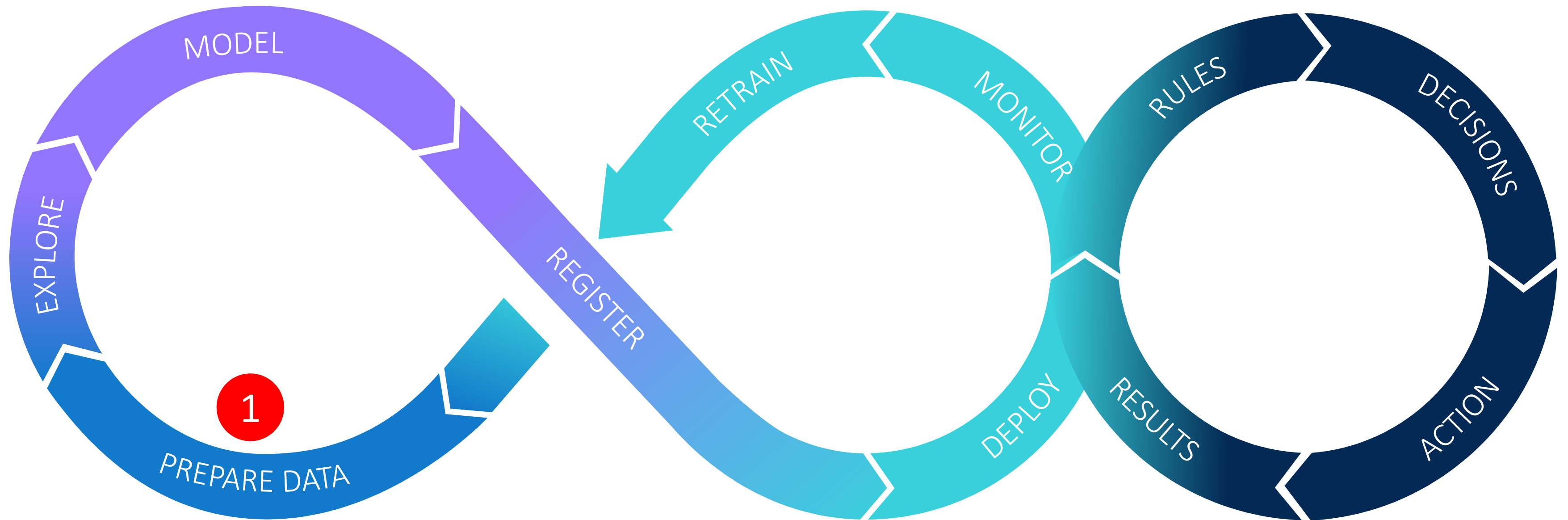
[Data Source for demo :](#)

[GitHub - jfjelstul/worldcup: A Comprehensive Database on the FIFA World Cup \(Men's and Women's\)](#)

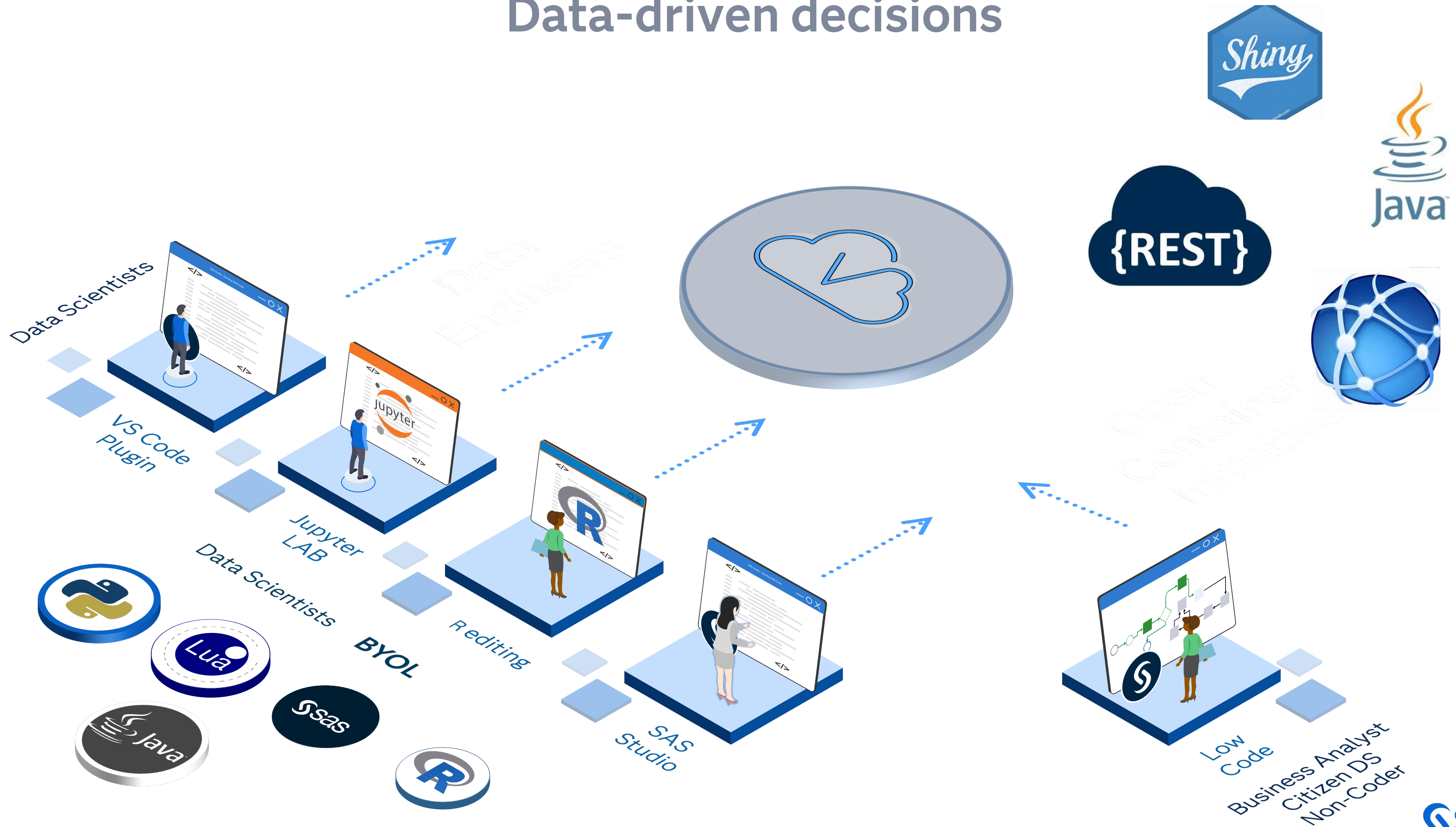
Improving your data quality assurance process



Prepare Data



Data-driven decisions



Data Preparation



Data Preparation

The screenshot displays the RStudio IDE interface. The main editor window shows an R script with the following code:

```
1 # Load the SWAT package
2 library(swat)
3 # Load the readxl package
4 library(readxl)
5 # install.packages('openxlsx')
6 library(openxlsx)
7
8 # Create a connection object
9 conn <- swat::CAS('https://viya4.globalhls.sashq-d.openstack.sas.com/cas-shared-default-http/', 443, 'schdac', 'c
10
11 # Check the connection
12 print(conn)
13
14 # Specify the path to the Excel file
15 file_path <- "C:/Users/schdac/OneDrive - SAS/PhUSE/2024/SDE Heidelberg/worldcup-master.zip/worldcup-master/data-
16
17 # Use the read.xlsx function from the openxlsx package to read the Excel file
18 # Install the openxlsx package if not already installed
19
20
21 excel_data <- read.xlsx(file_path)
22
23
```

The Environment pane on the right shows the connection object:

```
library(swat);
conn <- swat::CAS('https://viya4.globalhls.sashq-d.openstack.sas.com:443')
```

Below the connection object, it indicates "(Not connected)".

The Packages pane at the bottom right shows the User Library:

Name	Description	V...
admi...	ADaM in R Asset Library	1.0.1
admi...	Utility Functions and Development Tools for the Admiral Package Family	1.0.0
askn	Password Entry	1.2.0

A blue circular icon with a white arrow pointing upwards is overlaid on the console area. In the bottom left corner, there is a small illustration of a person standing next to a laptop displaying the R logo.

Loading Data To CAS

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R code for connecting to CAS, reading an Excel file, and loading it into a CAS table named 'goals'.
- Environment:** Lists datasets available in the CAS connection.
- Console:** Shows the output of the R script, including connection details and a note about generating function signatures.
- Files:** A sidebar showing the user library with installed packages.

```
# Load the readxl package
library(readxl)
# install.packages('openxlsx')
library(openxlsx)

# Create a connection object
conn <- swat::CAS('https://viya4.globalhls.sashq-d.openstack.sas.com/cas-shared-default-http/', 443, 'schdac', 'cas')

# Check the connection
print(conn)

# Specify the path to the Excel file
file_path <- "C:/Users/schdac/OneDrive - SAS/PhUSE/2024/SDE Heidelberg/worldcup-master.zip/worldcup-master/data-goals.xlsx"

# Use the read.xlsx function from the openxlsx package to read the Excel file
# Install the openxlsx package if not already installed
excel_data <- read.xlsx(file_path)

# Loading 'goals' is the dataframe loaded from the .RData file
ctbl <- as.casTable(conn, excel_data, casOut = list(name = "goals", caslib = "public", promote = TRUE))
```

Environment: viya4.globalhls.sashq-d.openstack.sas.com CAS

- ENROLLMENT DATES
- EX_DM
- FREQ_HEURE
- HACKATHON_CKD_TRAIN
- HACKATHON_CKD_TRAIN_DC
- HEART
- HEART DATA
- HEART FAILURE
- HMEQ
- JOURS
- LB
- LB_DM
- LB_FLAT
- MAR_SYNTHETIC
- MASTER_PF_LIST
- MATCHINGVARS
- MH_DM
- MODELINGVARS
- NATIONAL PARK SERVICE
- NEW HEART
- NEXTGEN_CPV_DATASET_A
- NHANES_PAD

Console:

```
R 4.3.2 ~/>
n123')
NOTE: Connecting to CAS and generating CAS action functions for loaded
action sets..
NOTE: To generate the functions with signatures (for tab completion), set
options(cas.gen.function.sig=TRUE).
```

Files:

Name	Description	V...
admi...	ADaM in R Asset Library	1.0.1
admi...	Utility Functions and Development Tools for the Admiral Package Family	1.0.0
askn	Password Entry	1.2.0

Accessing Data In CAS

The screenshot displays the SAS Drive web interface. The browser address bar shows the URL `https://viya4.globalhls.sashq-d.openstack.sas.com/SASDrive/`. The page header includes the SAS Drive logo and navigation options. A sidebar on the left contains a 'Filter' section with expandable categories: 'Item Name', 'Date Modified' (with start and end date pickers), and 'Modified By' (listing 'Me' with 39 items and 'Samiul Haque' with 1 item). The main content area shows a grid of SAS programs, including '_CAR_Set_Up_Environ...', '_Dummy_Program.sas', 'Abbreviated_Demo.sas', 'ADOS Histogram Aut', and 'ADOS Histogram sas'. A right sidebar contains a vertical list of application icons and a search bar.

Data Correlations And Outliers

SAS® Information Catalog

https://viya4.globalhls.sashq-d.openstack.sas.com/SASInformationCatalog/

SAS® Information Catalog - Discover Information Assets

5 ?

Discover Information Assets

Ignite your analytics journey

What assets are you looking for?

Import data

CATALOG AT A GLANCE

Total assets

9.1 K

Files

3.9 K

Data sets

2.8 K

1.5 K

Reports

437

WELCOME

Take a tour or visit our SAS Information Catalog User's Guide.

Also check out our latest updates to SAS Information Catalog and see what's new!

COLLECTIONS

Recent

Favorites

My Data

Search indexes: (no filter)

Actions

	Name	★	Status	Date Created	Created By
☒	GOALS	☆	--	Jun 19, 2024 12:31 PM	Daniel Christe
☐	CAR Create Studv Worksp...	☆	--	Apr 26, 2024 2:23 PM	Daniel Christe

36 items

Data Privacy

SAS® Information Catalog

https://viya4.globalhls.sashq-d.openstack.sas.com/SASInformationCatalog/details/~fs~catalog~fs~inst...

SAS® Information Catalog - Discover Information Assets

Catalog Home > GOALS

GOALS

Public

Completeness: 99%

Columns: 27

Rows: 3.6 K

Size: 1.8 MB

Status: None

Actions

Overview

Column Analysis

Sample Data

Date analyzed: Jun 19, 2024 12:33 PM

Private

Information Privacy

Jul 14, 1990 - Dec 19, 2082

Time Period Covered

DEU, Brazil, BRA, Ger...

Top Areas Covered

(none found)

Top Languages

Summary

This dataset describes information about the following entities: phone, family name, given name and 1 other

varchar

double

Filter

Name/Label	Le...	Semantic Type	Infor...	Terms
key_id	8	Generic...	None	(none)
goal_id	6	Generic...	None	(none)
tournam...	7	Generic...	None	(none)

Contacts (0)

No contacts are assigned.

Tags (0)

No tags are assigned.

Properties

Asset type: In-men

Date modified: Jun 19

Modified by: --

Date created: Jun 19

Data Completeness

SAS® Information Catalog

https://viya4.globalhls.sashq-d.openstack.sas.com/SASInformationCatalog/details/~fs~catalog~fs~inst...

SAS® Information Catalog - Discover Information Assets

5 ?

Catalog Home > Search Results > GOALS

New search

GOALS

Public

Completeness: 99%

Columns 27

Rows 3.6 K

Size 1.8 MB

Status None

Actions

Date analyzed: Jun 19, 2024 12:33 PM

Overview

Column Analysis

Sample Data

Private

Information Privacy

Jul 14, 1990 - Dec 19, 2082

Time Period Covered

DEU, Brazil, BRA, Ger...

Top Areas Covered

(none found)

Top Languages

Summary

This dataset describes information about the following entities: phone, family name, given name and 1 other

varchar

double

Filter

Name/Label	Le...	Semantic Type	Info...	Terms
key_id	8	Generic...	None	(none
goal_id	6	Generic...	None	(none
tourna...	7	Generic...	None	(none

Contacts (0)

No contacts are assigned.

Tags (0)

No tags are assigned.

Properties

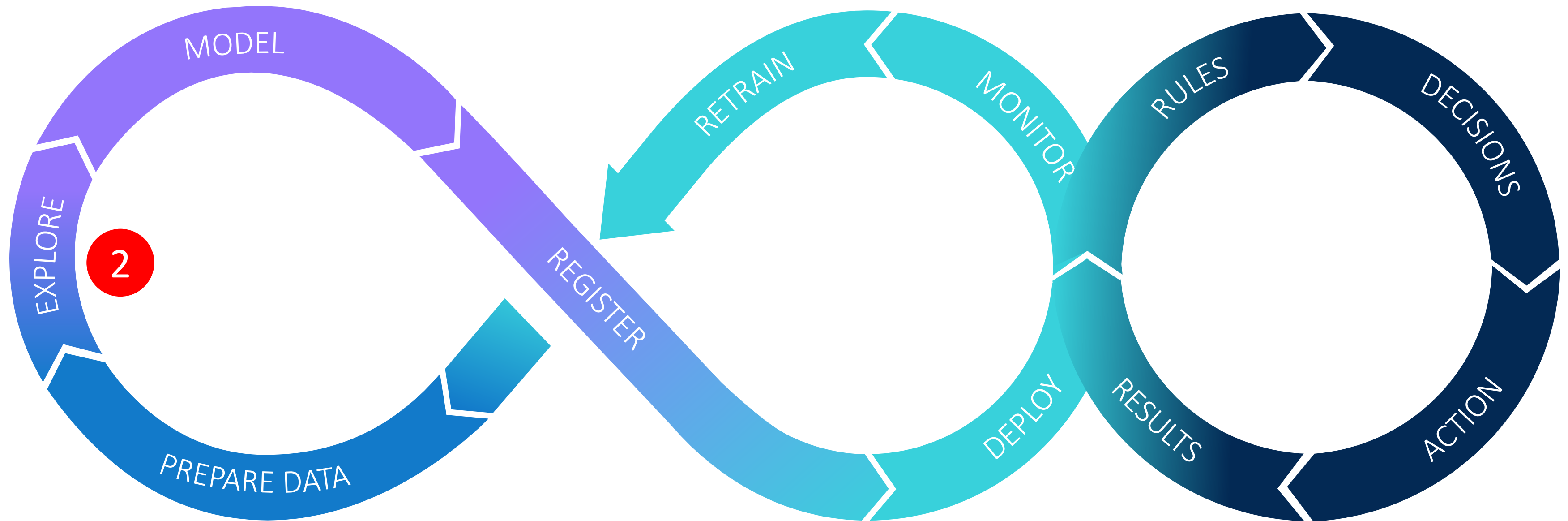
Asset type: In-memory data

Date modified: Jun 19, 2024 12:31 PM

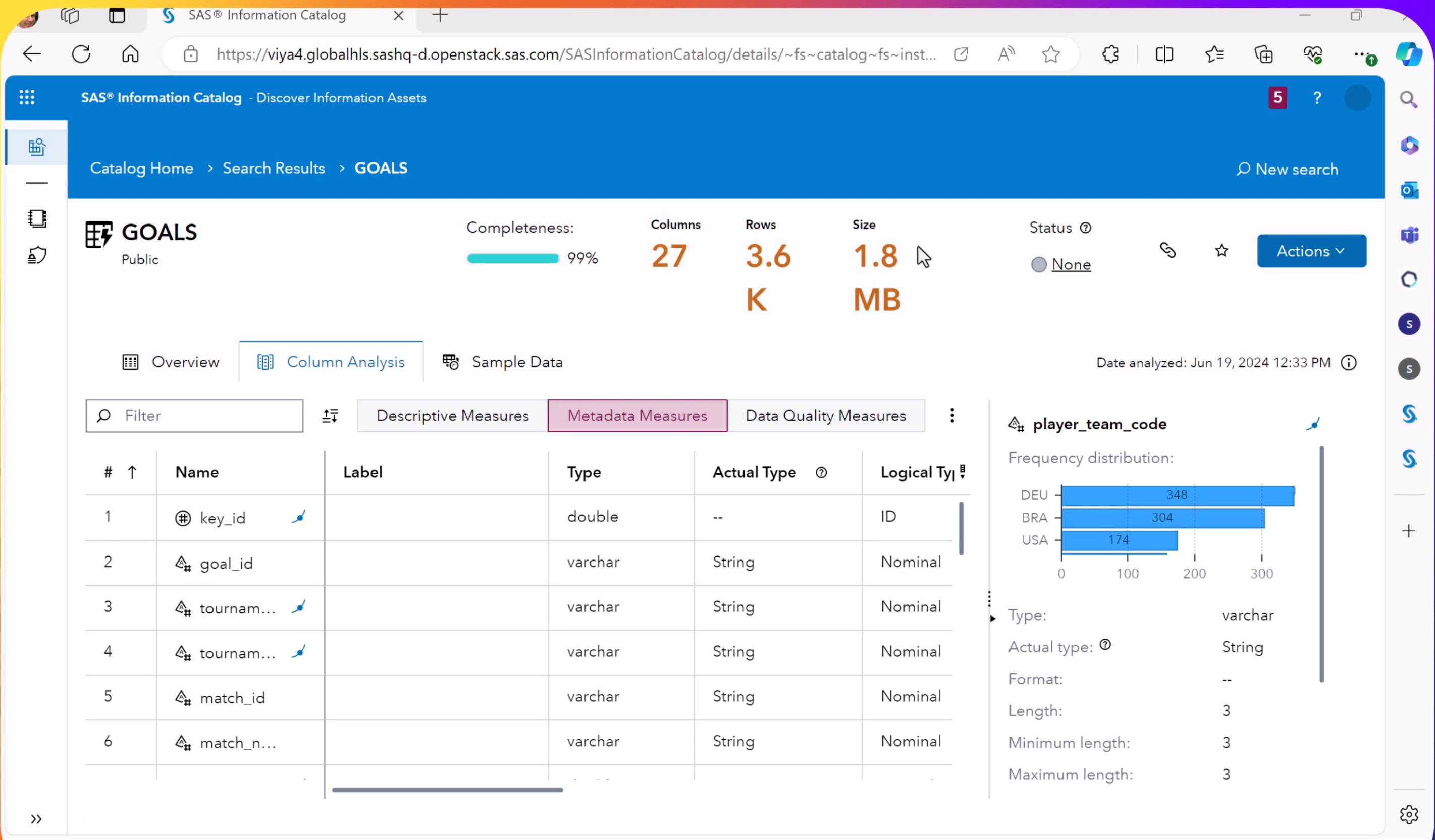
Modified by: --

Date created: Jun 19, 2024 12:31 PM

Improving your data quality assurance process



Explore And Visualize



Enriching The Data

SAS® Information Catalog

https://viya4.globalhls.sashq-d.openstack.sas.com/SASInformationCatalog/

SAS® Information Catalog - Discover Information Assets

5 ?

Discover Information Assets

Ignite your analytics journey

What assets are you looking for?

Import data

CATALOG AT A GLANCE

WELCOME

Take a tour or visit our SAS Information Catalog User's Guide.

Also check out our latest updates to SAS Information Catalog and see what's new!

9.1 K

3.9 K

2.8 K

1.5 K

437

COLLECTIONS

Recent

Favorites

My Data

Search indexes: (no filter)

Actions

	Name	★	Status	?	Date Created	Created By
☐	CAR_Upload_To_CAS	☆	--		Apr 19, 2024 11:33 AM	Daniel Christe
☐	GOALS	☆	--		Jun 19, 2024 12:31 PM	Daniel Christe

36 items



Managing Access

SAS® Information Catalog - Discover Information Assets

Catalog Home > Search Results

Goals

Search indexes: (no filter)

Top 5 Results

	Name	★	Status	Date Analyzed	As: ↓
☐	GOALS	☆	●	(not analyzed)	Tabl
☐	GOALS	☆	--	Jun 19, 2024 12:33 PM	In-m
☐	GOALS_ENHANCED	☆	--	(not analyzed)	In-m
☐	STENT FAILURE	☆	✓	May 16, 2024 4:29 PM	In-m
☐	STEPS	☆	●	Jul 27, 2023 6:16 PM	CAS

(5 of 5)

GOALS

PUBLIC

Details Columns Tags

Properties

Asset type: Table

Date modified: Jun 19, 2024 2:17 PM

Modified by: --

Date created: Jun 19, 2024 2:17 PM

Created by: --

Physical region: --

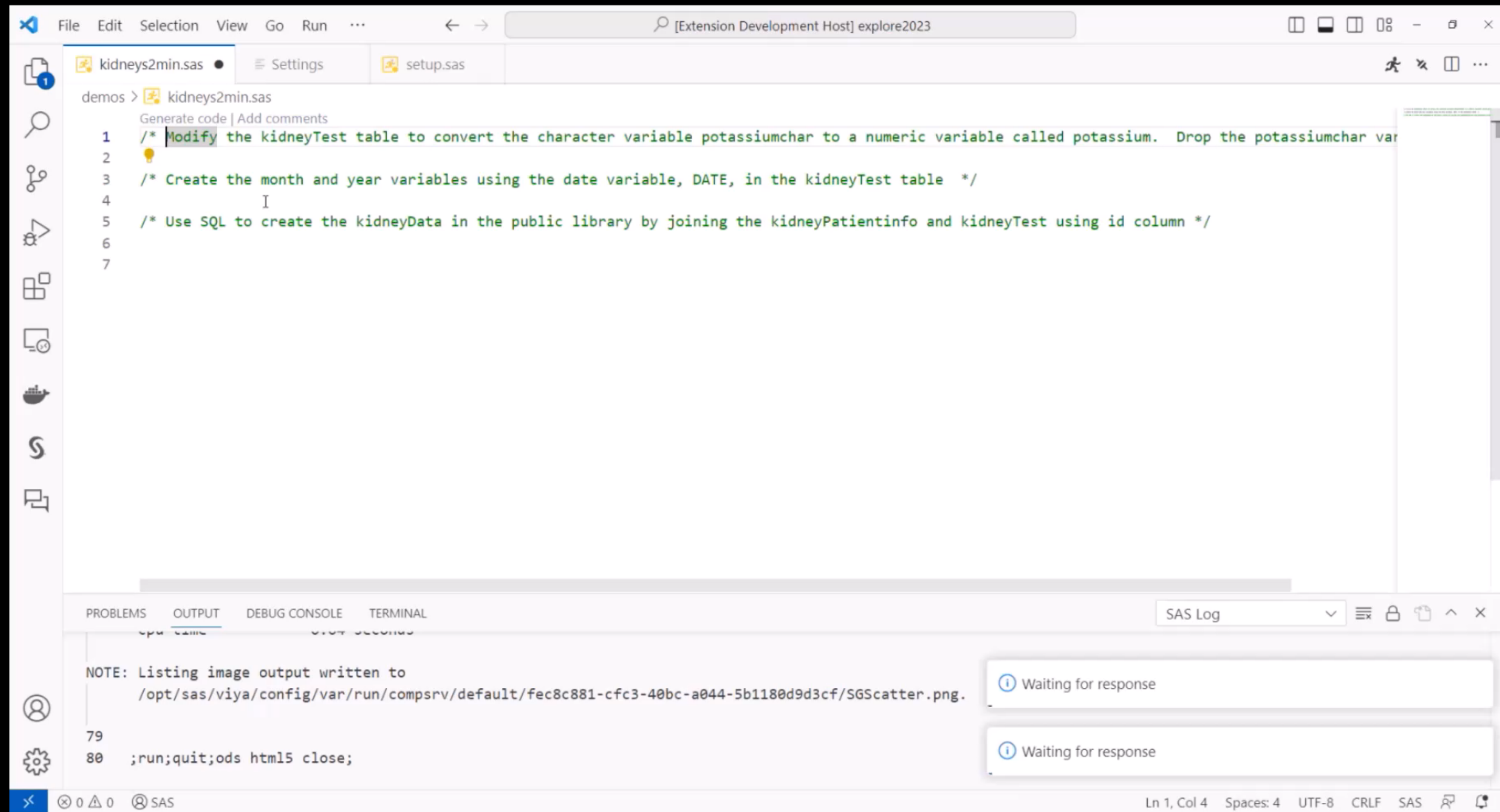
Rows: --

Size: --

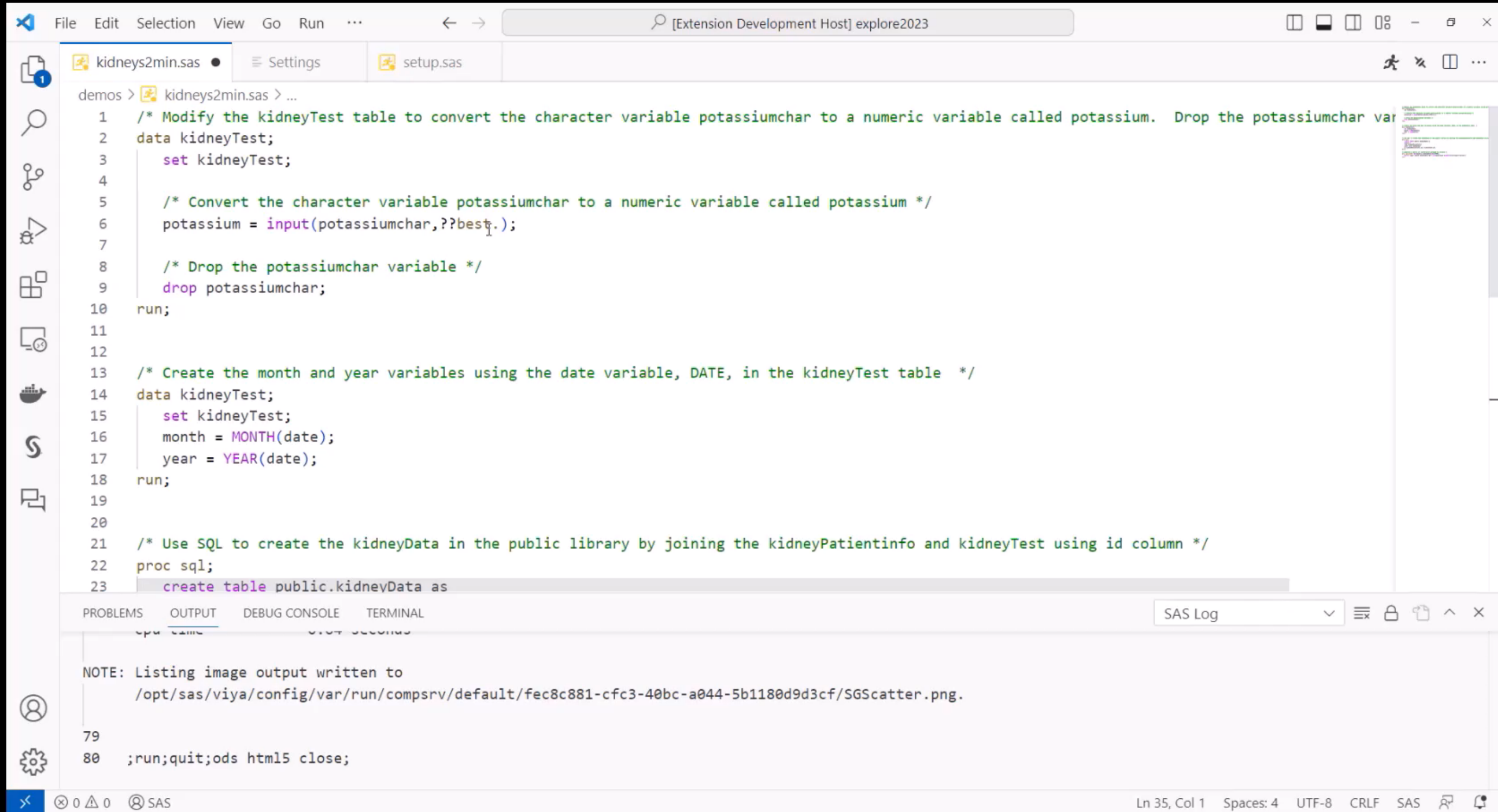
SAS Viya Copilot – Features

	Generate Code	Explain Code	Add Comments
Input	Plain English Prompt written in plain English inside <code>/* */</code> + context code where available	SAS Code SAS Code (may include inline comments)	SAS Code SAS Code (may include inline comments)
Output	SAS Code SAS Code in Editor Window	Plain English English Explanation in Conversation Window	SAS Comments English inline comments in Editor Window

SAS Viya Co-Pilot



SAS Viya Co-Pilot



The screenshot displays the SAS Viya Co-Pilot interface. The top menu bar includes File, Edit, Selection, View, Go, Run, and a search bar. The left sidebar contains icons for Explorer, Search, Source Control, Run and Debug, Extensions, and Settings. The main editor area shows a SAS script with the following code:

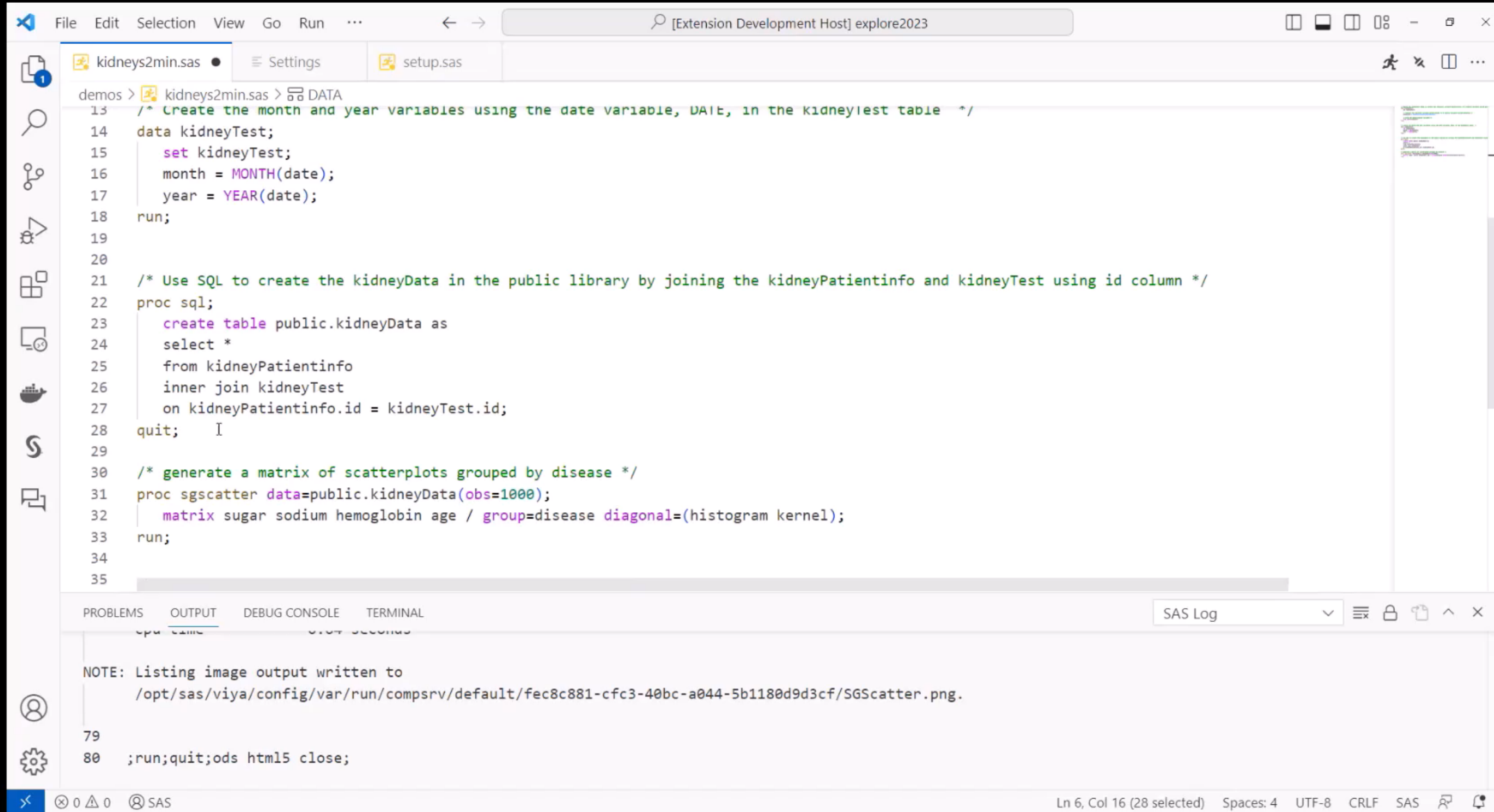
```
1 /* Modify the kidneyTest table to convert the character variable potassiumchar to a numeric variable called potassium. Drop the potassiumchar var
2 data kidneyTest;
3     set kidneyTest;
4
5     /* Convert the character variable potassiumchar to a numeric variable called potassium */
6     potassium = input(potassiumchar, ??best.);
7
8     /* Drop the potassiumchar variable */
9     drop potassiumchar;
10 run;
11
12
13 /* Create the month and year variables using the date variable, DATE, in the kidneyTest table */
14 data kidneyTest;
15     set kidneyTest;
16     month = MONTH(date);
17     year = YEAR(date);
18 run;
19
20
21 /* Use SQL to create the kidneyData in the public library by joining the kidneyPatientinfo and kidneyTest using id column */
22 proc sql;
23     create table public.kidneyData as
```

The bottom panel shows the SAS Log window with the following output:

```
NOTE: Listing image output written to
      /opt/sas/viya/config/var/run/compsrv/default/fec8c881-cfc3-40bc-a044-5b1180d9d3cf/SGScatter.png.
79
80 ;run;quit;ods html5 close;
```

The status bar at the bottom indicates the current position is Ln 35, Col 1, with 4 spaces, UTF-8 encoding, CRLF line endings, and the SAS environment.

SAS Viya Co-Pilot



The screenshot displays the SAS Viya Co-Pilot interface. The top menu bar includes File, Edit, Selection, View, Go, Run, and a search bar. The left sidebar contains icons for Explorer, Search, Source Control, Run and Debug, Data Explorer, and a Co-Pilot icon. The main editor shows a SAS script with the following code:

```
13 /* Create the month and year variables using the date variable, DATE, in the kidneytest table */
14 data kidneyTest;
15     set kidneyTest;
16     month = MONTH(date);
17     year = YEAR(date);
18 run;
19
20
21 /* Use SQL to create the kidneyData in the public library by joining the kidneyPatientinfo and kidneyTest using id column */
22 proc sql;
23     create table public.kidneyData as
24     select *
25     from kidneyPatientinfo
26     inner join kidneyTest
27     on kidneyPatientinfo.id = kidneyTest.id;
28 quit;
29
30 /* generate a matrix of scatterplots grouped by disease */
31 proc sgscatter data=public.kidneyData(obs=1000);
32     matrix sugar sodium hemoglobin age / group=disease diagonal=(histogram kernel);
33 run;
34
35
```

The bottom panel shows the SAS Log output, which includes a note about the image output being written to a file:

```
NOTE: Listing image output written to
/opt/sas/viya/config/var/run/compsrv/default/fec8c881-cfc3-40bc-a044-5b1180d9d3cf/SGScatter.png.
79
80 ;run;quit;ods html5 close;
```

The status bar at the bottom indicates the current position is Ln 6, Col 16 (28 selected), with 0 errors and 0 warnings. The interface also shows a sidebar with icons for Explorer, Search, Source Control, Run and Debug, Data Explorer, and a Co-Pilot icon.

SAS Data Maker

SAS® Data Maker

Your gateway to seamless synthetic data generation

Select preprocessed tabular source data:

Select Data

Search area: ⓘ

3

Number of records to generate: ⓘ

10000

Generate Data

How To



Overview

See an overview of SAS Data Maker.

[Watch video](#)



Get Started

Learn about Data Maker capabilities and how to get started with SAS.

[View documentation](#)

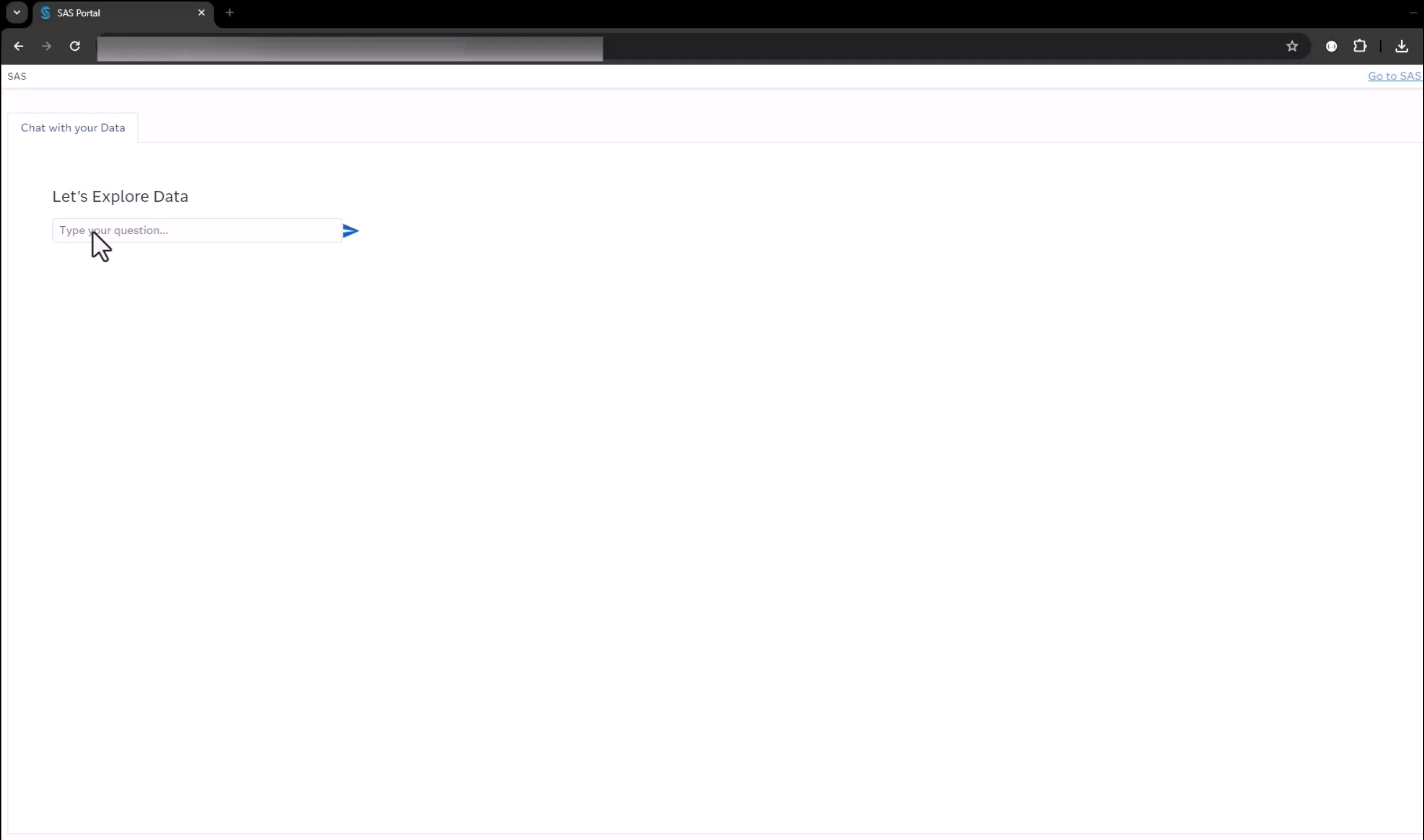


Provide Feedback

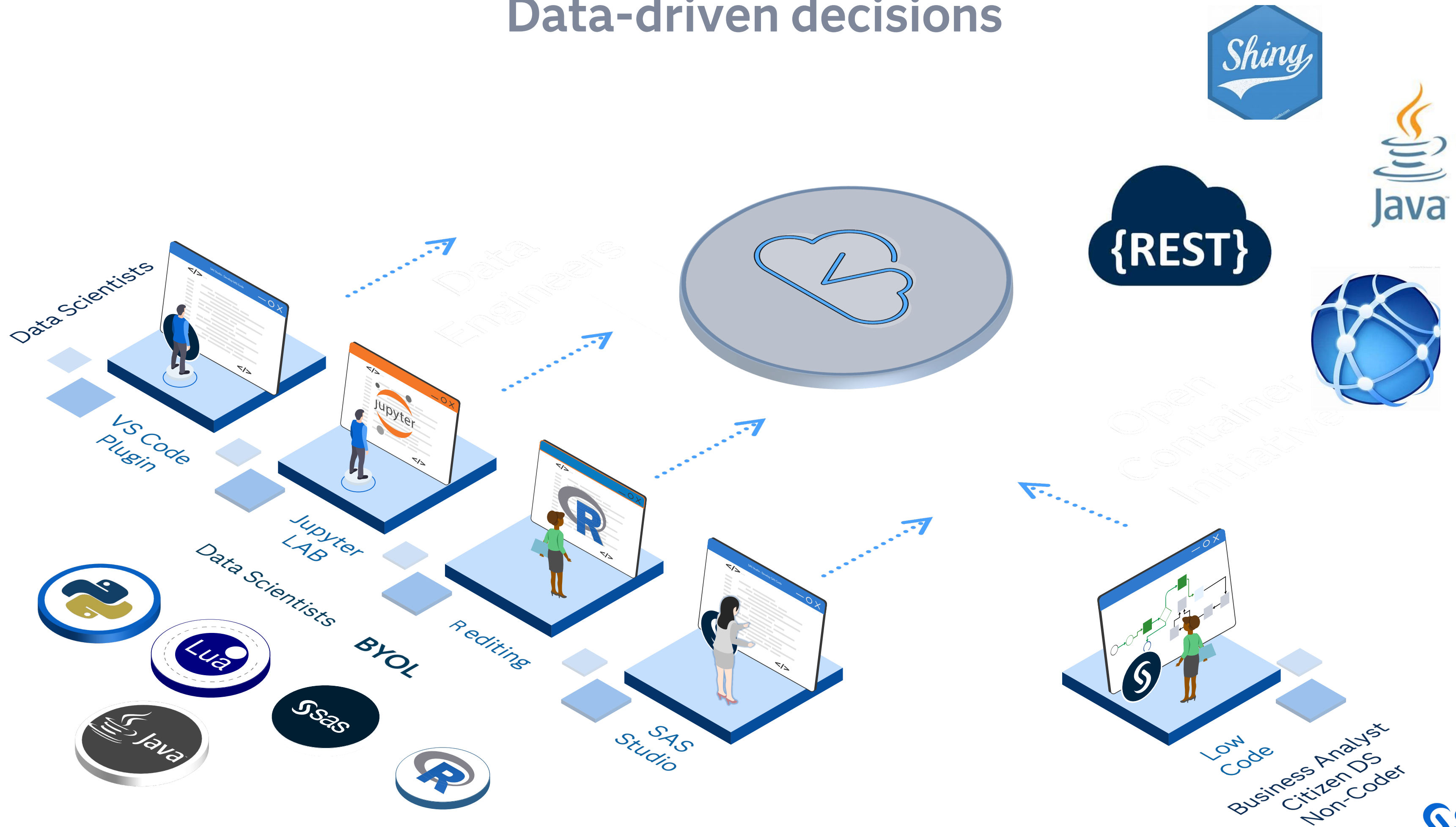
Tell us about your experience.

[Submit feedback](#)

Chat With Your Data



Data-driven decisions



Thank You !

PHUSE Single Day Events



Heidelberg, Germany

Thursday 27 June 2024

