Using R with SAP Analytics Cloud

Best practices for connecting and running your R environment

Run Simple

TABLE OF CONTENTS

ABOUT THIS DOCUMENT

This document is intended to assist system administrators and data professionals with the installation and configuration of R with the SAP Analytics Cloud. This document is for informational purposes only. SAP does not warrant or represent that the information provided herein is complete or error-free. This document does not create or modify any legal relationship between you and SAP. You are encouraged to notify SAP of any inaccurate information or errors found in this document.

Legal Notice:

Required Core Components

ABOUT GCC
The GNU Compiler Collection (GCC) is the standard open-source compiler for most Unix-like operating systems. GCC is licensed under the GNU General Public License v. 3 with a GCC Runtime Library Exception (see **https://gcc.gnu.org/onlinedocs/libstdc++/manual/license.html**).

ABOUT OPENSSL
OpenSSL is an open-source toolkit for the Transport Layer Security and Secure Sockets Layer protocols. OpenSSL is also a general-purpose cryptography library. OpenSSL is licensed under the OpenSSL\SSLeay license(s) (see **https://www.openssl.org/source/license.txt**).

ABOUT R AND RSERVE

R is an open-source programming language that includes packages for advanced visualizations, statistics, Machine Learning and much more. The R language is popular among statisticians and data miners for developing statistical software and is widely used for advanced data analysis. RServe is a TCP/IP server that allows other programs to use facilities of R without the need to initialize R or link with the R library. R and RServe are products of the R Foundation which consist of components licensed under a variety of licenses (https://www.r-project.org/Licenses/), including the GNU General Public License.

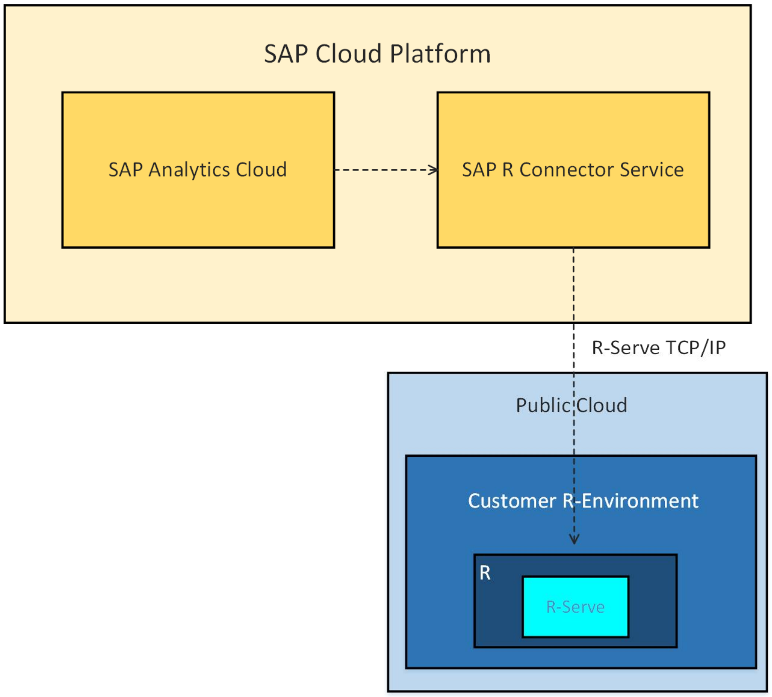
The goal of integration with R is to enable the embedding of visualizations based on R scripts into SAP Analytics Cloud Stories.

GCC, OpenSSL, R and RServe (the “Required Core Components”) are not part of SAP Analytics Cloud nor are they provided by SAP. To use the SAP Analytics Cloud with R, please download the Required Core Components and configure them as described in this document. Your use of the Required Core Components is governed by the applicable project licenses and not by any agreement between you and SAP. SAP does not provide support for the Required Core Components.

PREPARING THE R ENVIRONMENT

You need to deploy a Linux-based operating system on a virtual machine (VM) or a stand-alone machine dedicated to hosting the R environment. Note down the IP address for the host as you will need to provide this information later.

The system landscape diagram summarizes how the R environment works with the SAP Analytics Cloud.



The sections below summarize what resources are required to proceed with the setup and methods for setting up these resources.

Note: All the subsequent actions should be performed using the root or admin user account.

Required core components

Resource	Minimum supported version
GNU Compiler Collection (GCC)	4.8.4
R source	3.3.3 ("Another Canoe")
OpenSSL	1.0.1f

Access and installation via SUSE/ and Zypper

Using a Suse Linux Enterprise Server and the Zypper command line package manager, you can access and initially setup the required resources.

OpenSSL

1. Run:

```
zypper install openssl
zypper install openssl-devel
```

GCC

Note: If your system includes a C compiler you can ignore next two steps.

1. To determine the version go to: http://download.opensuse.org/repositories/devel:/gcc/. You will need to specify the correct version in the following step.
2. Run:

```
zypper ar
http://download.opensuse.org/repositories/devel:/gcc/<version>/
devel-gcc
zypper refresh
zypper install gcc
```

R source

1. For a listing for released R versions go to: http://download.opensuse.org/repositories/devel:/languages:/R:/released/. You will need to specify the version in the following step.

2. Run:

```
zypper addrepo -f http://download.opensuse.org/repositories/
devel:/languages:/R:/released/SLE_12/ R-base
zypper install R-base R-base-devel
```

Access and installation using Ubuntu

You can also use Ubuntu's APT (Advanced Packaging Tool) library to perform the installation of the R environment core components.

1. To install OpenSSL run:

```
apt-get install openssl libssl-dev
```

2. To determine what Ubuntu code name best suits your specific Linux environment refer to:

   https://wiki.ubuntu.com/Releases

3. To access the required R resources, pick a Cran mirror from:

   https://cran.r-project.org/mirrors.html

4. In the `/etc/apt/sources.list directory` specify the Ubuntu code name and mirror URL. For example:

```
deb https://<my.favorite.cran.mirror>/bin/linux/ubuntu trusty/
```

5. Run:

```
apt-get update
apt-get install build-essential
apt-get install r-base
apt-get install r-base-dev
```

For more information refer to: https://cran.rstudio.com/bin/linux/

SETTING UP THE R ENVIRONMENT

The following is an overview of the specific tasks required to set up your R environment to connect and work with SAP Analytics Cloud.

a. Create an Rserve certificate.
b. Create an Rserve configuration file.
c. Install Rserve.
d. Create a secondary user account for the SAP Analytics Cloud-R connection.
e. Setup a connection to Rserve in SAP Analytics Cloud.
f. Install your R packages.

What follows are more detailed instructions for each task.

A. Creating the certificate for Rserve

To enable a connection to your R environment (specifically Rserve) you need to create a generate a certificate key. You will need to provide the key when setting up the connection to your Rserve.

1. Navigate to the /temp directory.

2. To create directories for Rserve and the certificate file, run:

```
mkdir Rserv
cd Rserv
mkdir CA
cd CA
```

3. To generate the certificate run:

```
openssl genrsa -out Rserve.key 2048
openssl req -new -key Rserve.key -out Rserve.csr
openssl x509 -req -days <numdays> -in Rserve.csr -signkey
Rserve.key -out Rserve.crt
```

Note: You need to specify a valid number for <numdays>.

4. Save the Rserve.crt file to the computer from which you access SAP Analytics Cloud. You can also copy all the contents of the file to the machine's clipboard.

B. Creating the Rserve configuration file

The configuration file contains all the information Rserve needs to establish a secure connection.

1. To create the workspace directory for Rserve, run:

```
cd /tmp/Rserv
mkdir workspace
```

2. Create a new file /etc/Rserv.conf containing the following information:

```
        workdir /tmp/Rserv/workspace
        remote enable
        auth required
        plaintext disable
        port (port numberA)
        maxsendbuf 0
        tls.key /tmp/Rserv/CA/Rserve.key
        tls.cert /tmp/Rserv/CA/Rserve.crt
        tls.port (port numberB)
        qap disable (port numberA)
```

Note: The `tls.port` refers to encrypted port that will be used in the connection to SAP Analytics Cloud.. The port referenced in `port` and `qap disable` should be the same.

The table below explains the configuration details included in the file:

Setting	Explanation
`workdir`	Directory used to store any temporary R files.
`remote`	Specifies if remote access is required.
`auth`	Specifies is user access credentials will be checked. It is recommended that this be set to ‘`required`’.
`plaintext`	Specifies that plaintext can be used to pass credentials. It is recommended that this be set to ‘`disable`’.
`port`	Default port for non-encrypted connection.
`Maxsendbuffer`	Specifies the maximum size for any sent buffer. Default is ‘`0`’ (unlimited).
`tls.key`	Path to the Rserve key generated in the previous step.
`tls.cert`	Path to the Rserve certificate.
`Tls.port`	Port used for an encrypted connection.
`qap.disable`	Specifies which default unencrypted port should be disabled.

C. Installing Rserve

1. To open the R shell run the following command:

```
R
```

2. To install Rserve, in the R shell, run:

```
Install packages (“Rserve”)
```

Note: When prompted to select a mirror for downloading, for best results, select the closest geographical location.

3. Close the R shell.

D. Creating a secondary user account

You need to create a user account that will be used for connecting SAP Analytics Cloud and the Rserve you installed in the previous step. To limit any damage from a potential malicious user, the secondary account should have reduced privileges for running and accessing Rserve. To add R scripts in SAP Analytics Cloud, the user account needs to be able to write to the Rserve work directory.
Before you proceed, ensure the R shell is closed.

1. To create a new user account run:

```
useradd <username>
passwd  <username>
```

2. Use the displayed prompts to create a password.

3. To provide the user account with access to the Rserve workspace folder, run:

```
cd/temp/Rserv
chown -R <username> workspace
chgrp   -R
```

4. To restrict the new user from accessing the authentication keys, run:

```
chmod -R 700 CA
```

5. You can now test the encrypted connection for new user account.

   a. To change the Rserve configuration to execute commands as the newly created user account, run:

```
id <username>
```

   b. Save the gid and uid numbers.
   c. Add the following lines to the etc/Rserv.conf file:

```
gid <group-ID>
uid <user–ID>
```

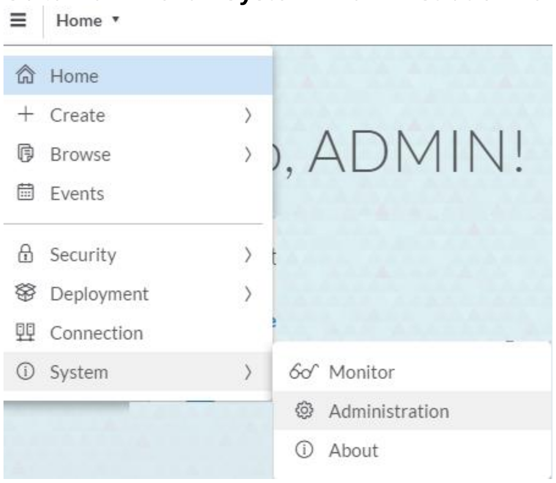
   d. Run:

```
R CMD Rserve
```

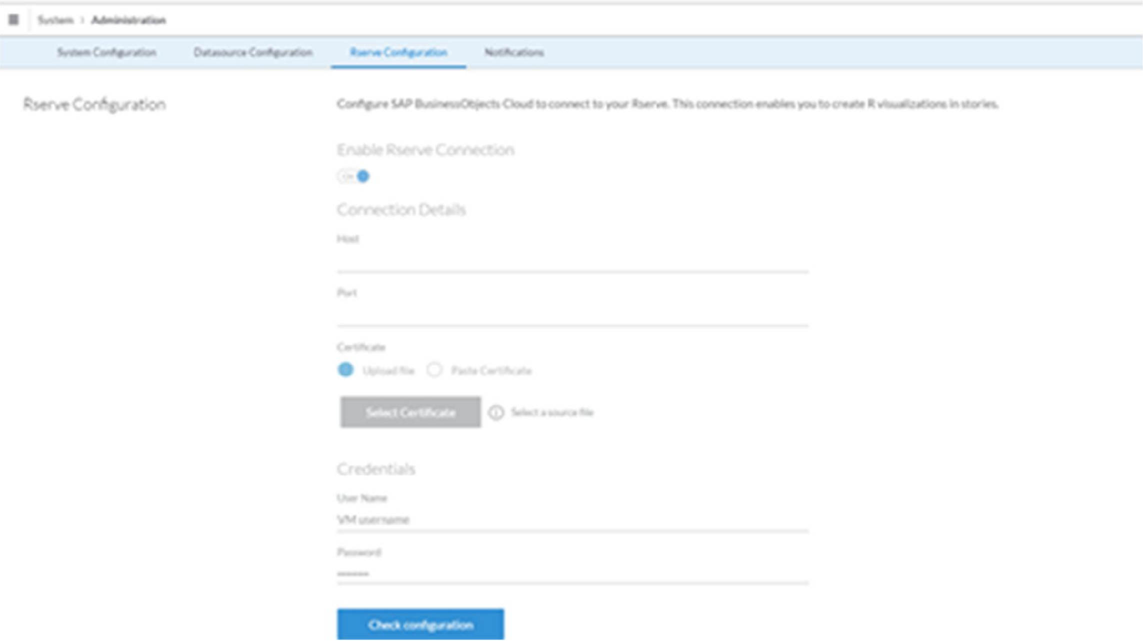
Connecting from Sap Analytics Cloud to Rserve

You can now need to provide the Rserve certificate, user account credentials, IP address for the machine hosting your Rserve, as well the connection port specified in the Rserve configuration file.

1. Launch your SAP Analytics Cloud instance.
2. Go to **Main Menu**> **System**> **Administration**> and select **Rserve Configuration.**



The following configuration screen is displayed:

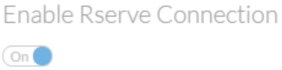


Configure SAP BusinessObjects Cloud to connect to your Rserve. This connection enables you to create R visualizations in stories.

User Name: VM username

Password: •••••••

3. Set **Enable Rserve Connection** to **On**.



Note: All configuration from this point on will impact all your users working on SAP Analytics Cloud.

4. Under **Connection Details** provide the following information:

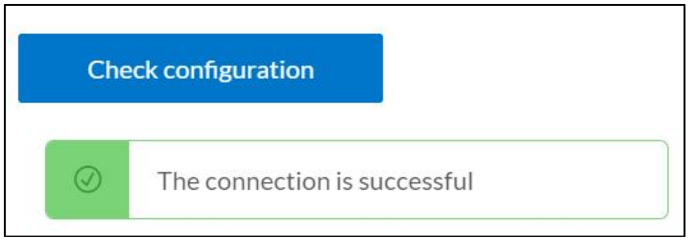
- **Host**: Enter the address for the machine hosting your R environment is running.
- **Port**: Enter the port you specified when you created the Rserve configuration file.

Note: You need to provide the certificate your generated in “Creating the certificate for Rserve” section using one of the following options:

- **Upload File**: select this option to upload the source file containing the certificate.
- **Paste Certificate**: select this option if you have copied the certificate to your clipboard.

5. Under **Credentials** provide the **User Name** and **Password** you created in the “Creating a secondary account” section.
6. Select **Check Configuration** to validate that your connection has been successful.

   The confirmation message shown below is displayed.



The connection is successful

Installing the R packages

You need to install the following packages in your R environment. These packages are required by SAP Analytics Cloud:

Package	Notes
`ggplot2`	A system for 'declaratively' creating graphics, based on "The Grammar of Graphics".
`jsonlite`	Json parser and generator.
`bit64`	Allows support for 64-bit integer numbers.
`data.table`	Used for data analysis and provides aggregation of large data (e.g. 100GB in RAM), fast ordered joins, fast add/modify/delete of columns by group using no copies at all, list columns, a fast-friendly file reader and parallel file writer.

Note: SAP Analytics Cloud currently only supports SVG and PNG outputs for visualizations. Other output types might not be displayed.

If you are using SUSE on Linux to run your R environment, you need to zipper in gcc-c++ to enable ggplot2.

1. To open the R shell run:

```
R
```

2. To install the packages,run:

```
install.packages(c("ggplot2", "jsonlite", "bit64", "data.table"))
```

WORKING WITH R IN SAP ANALYTICS CLOUD

High Level Summary of R Capabilities

The R Visualization feature allows users to integrate their own R environment into SAP Analytics Cloud.

Once integrated you can perform the following:

- Insert R visualizations into stories.
- Interact with your visualizations using controls such as filters.
- Edit your R scripts and preview visualizations.
- Share stories containing R visualizations with other users.

With the R visualization capability, users can perform statistical and analytical analyses and create charts reflecting these analyses. These visualizations remain interactive and consider the row-level security of users.

The following resources are provided to help you produce high end visualizations using R:

- A dashboard containing an R script editor, a snapshot of the environment after the R script is run, a console containing the output of the executed script, and a preview screen for the visualization.

- In the editor, you have the option to insert R-code from the menu. This code includes:
  - The code snippet for listing all the available packages installed on your R system.
  - Sample R scripts for visual statistical analysis.

Adding R Visualizations to Stories

You can insert an R visualization into a story canvas.

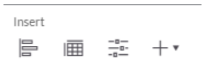
Prerequisite: You first need to acquire data in an SAP Analytics Cloud model before the data can be accessed in your R-script.

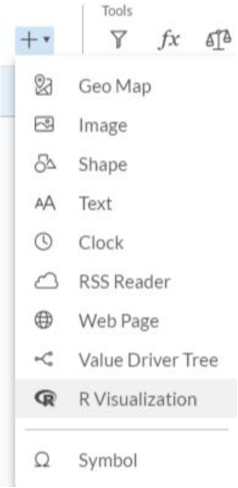
> Note: You cannot use remote or key figure models as input data. All hierarchies in models are flattened when working with R.

An R visualization is structured on the following:

- Input data based on a model, specified dimensions, and filters.
- Input parameters based on a calculation input control.
- An R script executed to generate a visualization.

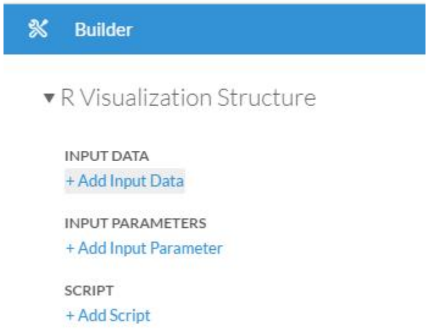
1. From the **Insert** menu of a story canvas page select **+ Add**> **R Visualization**.





An empty R visualization is added to the canvas page.

2. To configure the input data for the R visualization, select +**Add Input Data** under **INPUT DATA.**

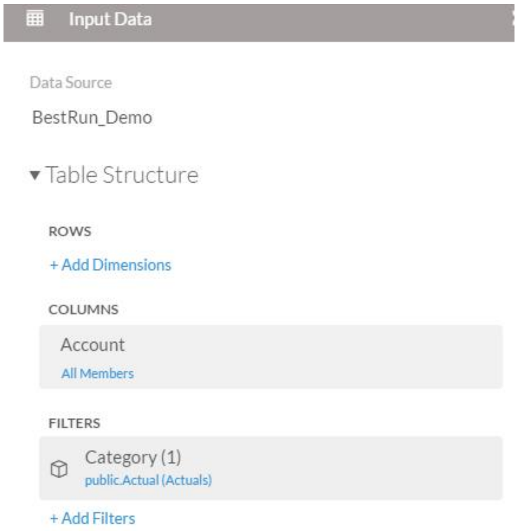


The **Select Datasource** dialog is displayed.

Note: If a data source has already been added to the story, it will be used by default.

3. Choose a model from the **Select an Existing Model** list and select **OK**.

4. Use the **Table Structure** settings to specify the table structure for your input data.



Input Data

Data Source

BestRun_Demo

▾ Table Structure

ROWS

+ Add Dimensions

COLUMNS

Account

All Members

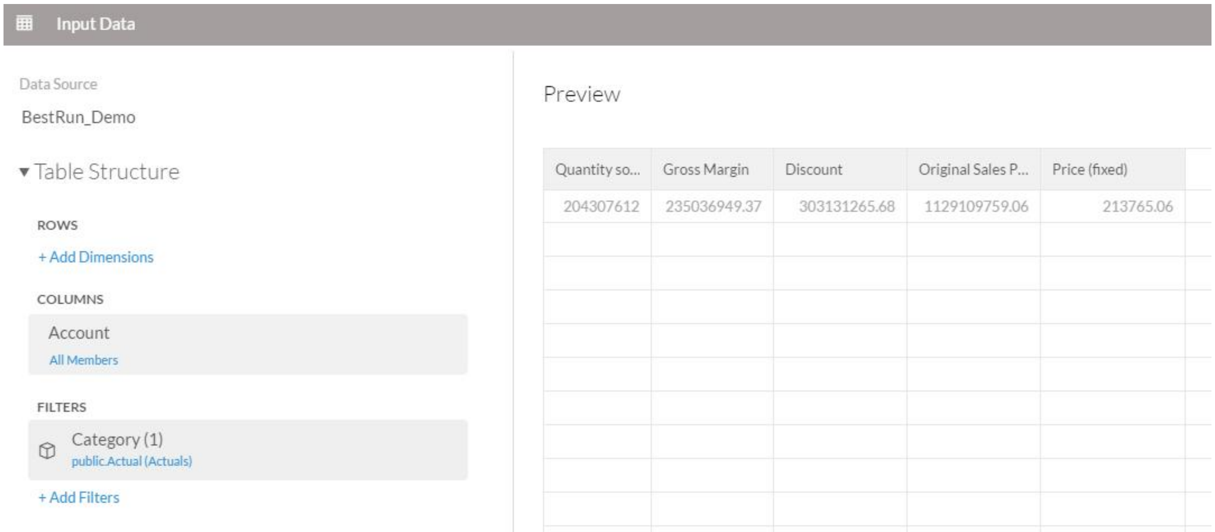
FILTERS

Category (1)

public.Actual (Actuals)

+ Add Filters

Note: Select the **Expand** icon under Designer to view a **Preview** for the input data table structure.



Input Data

Data Source

BestRun_Demo

▾ Table Structure

ROWS

+ Add Dimensions

COLUMNS

Account

All Members

FILTERS

Category (1)

public.Actual (Actuals)

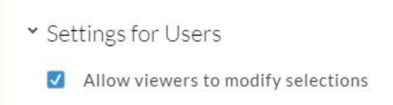
+ Add Filters

Preview

Quantity so...	Gross Margin	Discount	Original Sales P...	Price (fixed)
204307612	235036949.37	303131265.68	1129109759.06	213765.06

a. **ROWS**: select **+Add Dimensions** to add dimensions from the list of available options for your input data. Hover over a dimension to display the ··· **More** icon if you want to specify display options or attributes. Use the **Manage Filters** icon to specify filters for the dimension.

Note: To create a dynamic input control for a given dimension, select **Manage Filters** and enable the **Allow viewers to modify selections** option.



˅ Settings for Users

☑ Allow viewers to modify selections

b. **COLUMNS**: is used to manage the available measures. Select **All Members** to display the currently available measures. Hover over a column entry to display the ··· **More** icon if you want to specify display options or attributes. Use the **Manage Filters** icon to specify filters for the measures.

c. **FILTERS**: is used to manage the filters for the input data. Select **+Add Filters** to specify filters.

d. When you have finished setting up the table structure select **OK**.

5. To configure an input parameter for your visualization:

a. Under **INPUT PARAMETERS** select an input control from the following options:

- If the story already contains calculation input controls, select +**Add Input Parameter** and choose any of the listed parameters.

- To create a new calculation input control: select **+Add Input Parameter** > **+Create a New Calculation Input Control**:

  The **Calculation Input Control** dialog is displayed.

Note: Input controls can currently only be based on a **Static List** and contain the **Number** Data Type.

b. Select **Click to Add Values** and choose either **Select by Range** or **Select by Member**.

i. If you have selected **Select by Range**, enter the **Min** and **Max** values for your range in the **Set Values for Custom Range** dialog. You can optionally set an **Increment** value.

ii. Select **OK**.

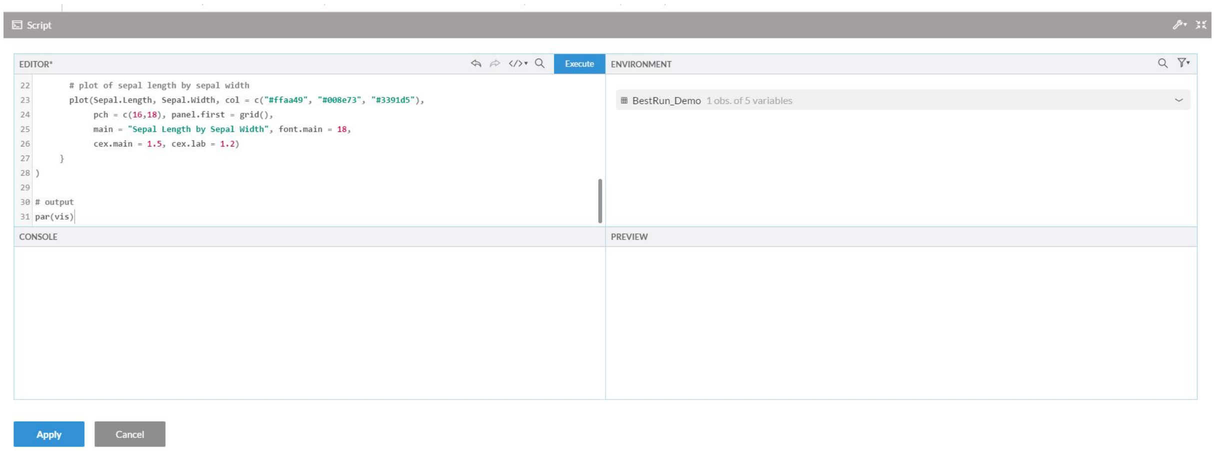
iii. To create a member based input control, add numeric values to the **Custom Members** area in the **Select Values from Custom LOV dialog**, and then select **Update Selected Members**.

iv. Expand the **Settings for Users** section, and then choose whether users can do the following in the input control: **Single Selection**, or **Multiple Selection.**

v. Select **OK.**

6. Under Script select **+ Add Script** to create a new script or **Edit Script** if a script has previously been specified.

7. Select the **Expand** icon under **Designer** to view the **Editor**, **Environment**, **Console**, and **Preview** panes.

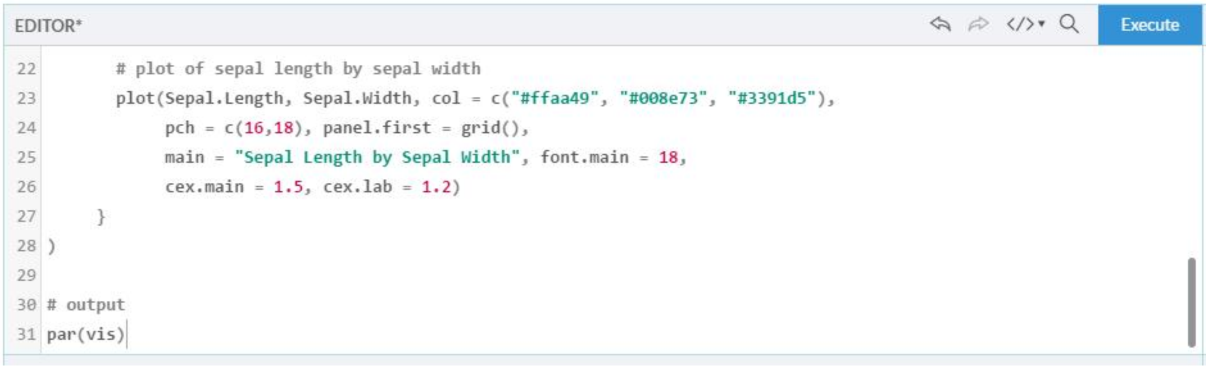


```
        # plot of sepal length by sepal width
        plot(Sepal.Length, Sepal.Width, col = c("#ffaa49", "#008e73", "#3391d5"),
             pch = c(16,18), panel.first = grid(),
             main = "Sepal Length by Sepal Width", font.main = 18,
             cex.main = 1.5, cex.lab = 1.2)
    }
)

# output
par(vis)
```

BestRun_Demo 1 obs. of 5 variables

8. Enter your script in the **Editor** and select **Execute**.

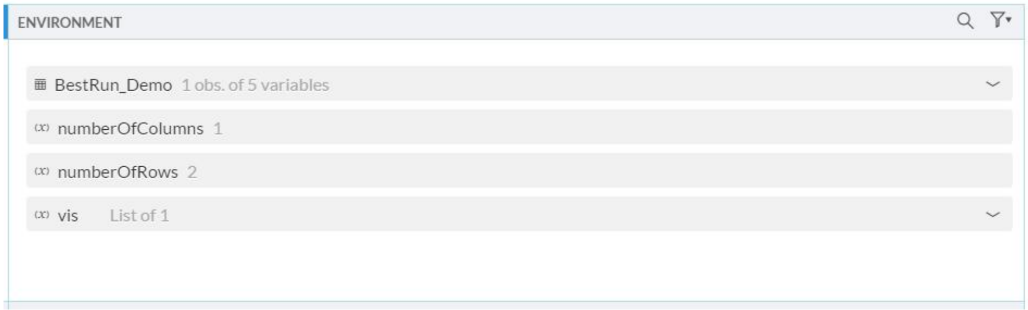


```
        # plot of sepal length by sepal width
        plot(Sepal.Length, Sepal.Width, col = c("#ffaa49", "#008e73", "#3391d5"),
             pch = c(16,18), panel.first = grid(),
             main = "Sepal Length by Sepal Width", font.main = 18,
             cex.main = 1.5, cex.lab = 1.2)
    }
)

# output
par(vis)
```

If the script is processed successfully, the following is displayed:

The variables from the script output are listed under **Environment.**



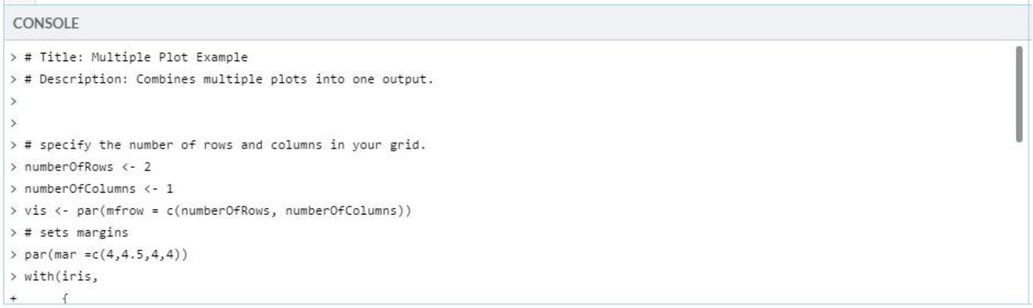
BestRun_Demo 1 obs. of 5 variables

numberOfColumns 1

numberOfRows 2

vis List of 1

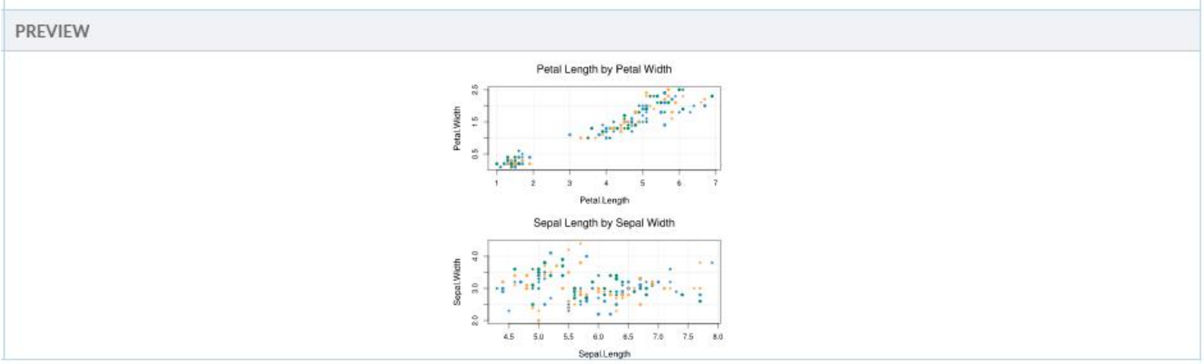
The outputted code from the executed script is displayed under **Console**.



CONSOLE

```
> # Title: Multiple Plot Example
> # Description: Combines multiple plots into one output.
>
>
> # specify the number of rows and columns in your grid.
> numberOfRows <- 2
> numberOfColumns <- 1
> vis <- par(mfrow = c(numberOfRows, numberOfColumns))
> # sets margins
> par(mar =c(4,4.5,4,4))
> with(iris,
+       {
```

Any visualization rendered by the executed script is displayed under **Preview**.



9. Select **Apply** to insert the R visualization into the current canvas page.

Note: Sample scripts to render R based visualizations are provided for your convenience. In the **Script Editor**, go to </>▾ **Snippets**> **Examples** to access the samples. The code and visualizations rendered by these samples are for instructional purposes. If you have the requisite R programming background, you can use these samples to generate visualizations based on your data.

www.sap.com/contactsap

The information contained herein may be changed without prior notice. Some software products marketed by SAP SE and its distributors contain proprietary software components of other software vendors. National product specifications may vary.

These materials are provided by SAP SE or an SAP affiliate company for informational purposes only, without representation or warranty of any kind, and SAP or its affiliated companies shall not be liable for errors or omissions with respect to the materials. The only warranties for SAP or SAP affiliate company products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.

In particular, SAP SE or its affiliated companies have no obligation to pursue any course of business outlined in this document or any related presentation, or to develop or release any functionality mentioned therein. This document, or any related presentation, and SAP SE's or its affiliated companies' strategy and possible future developments, products, and/or platform directions and functionality are all subject to change and may be changed by SAP SE or its affiliated companies at any time for any reason without notice. The information in this document is not a commitment, promise, or legal obligation to deliver any material, code, or functionality. All forward-looking statements are subject to various risks and uncertainties that could cause actual results to differ materially from expectations. Readers are cautioned not to place undue reliance on these forward-looking statements, and they should not be relied upon in making purchasing decisions.

Run Simple