Install R packages with SQLMLUTILS

Last updated by | Vitor Tomaz | Feb 24, 2023 at 3:31 AM PST

Contents

Issue

Issue when trying to run the sp_execute_external_script command to add R packages to Azure SQL Managed Instance, for specific scenario the readr package. But issue can be to install any package.

Command used

```
execute sys.sp_execute_external_script @language =N'r',@script=N'install.packages("readr")'
```

Error Message

Msg 39004, Level 16, State 20, Line 0 A 'R' script error occurred during execution of 'sp_execute_external_script' with HRESULT 0x80004004. External script request id is 18C93558-86F6-4741-A639-81B286B793BA. Msg 39019, Level 16, State 2, Line 0 An external script error occurred: Warning in install.packages("readr") : 'lib = "C:/WFRoot/Ext/R.9.4.7.1185/library"' is not writable Error in install.packages("readr") : unable to install packages Calls: source -> withVisible -> eval -> eval -> install.packages

Error in execution. Check the output for more information. Error in eval(ei, envir) : Error in execution. Check the output for more information. Calls: runScriptFile -> source -> withVisible -> eval -> eval -> .Call Execution halted

Impacted Regions

ALL

Investigation Analysis

Based on public documentation https://learn.microsoft.com/en-us/sql/machine-learning/install/sql-machine-learning-services-windows-install?view=sql-server-ver15&viewFallbackFrom=sql-server-ver16 and https://learn.microsoft.com/en-us/sql/machine-learning/package-management/install-additional-r-packages-

on-sql-server?view=sql-server-ver16 ☐ **sqlmlutils** should be used to install R packages on Azure SQL Managed
Instance.

Mitigation

To install **sqlmlutils** required to install packages, the machine learning feature should be enabled, below are the
details steps to Install the readr package on SQL MI:

1. Machine learning should be enabled as described here https://learn.microsoft.com/en-us/azure/azure-sql/managed-instance/machine-learning-services-overview?view=azuresql ☐

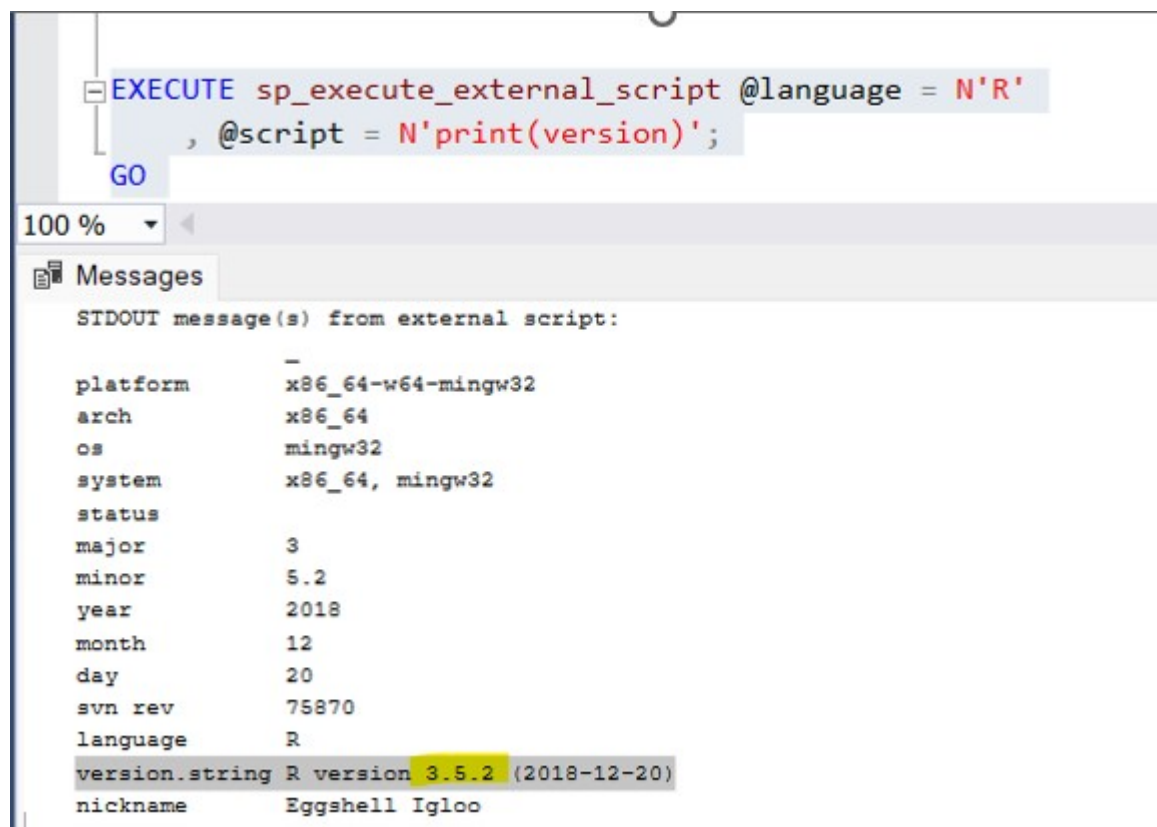   To enable Machine Learning run the T-SQL Command:

   ```
   exec sp_configure 'external scripts enabled' , 1
   reconfigure with override
   ```

   This will cause a restart for the SQL MI so disconnections are expected.

2. When Machine learning is enabled R is installed with default packages, to review the R version for the
   Azure SQL Managed Instance the following T-SQL command should be run preferable on the desired
   database to install the package.

   ```
   EXECUTE sp_execute_external_script @language = N'R'
       , @script = N'print(version)';
   GO
   ```

   For this example the R version for SQL MI founded was: "version.string R version 3.5.2 (2018-12-20)"

3. You can validate the R packages installed by running the follow command:

```
EXECUTE sp_execute_external_script
@language=N'R',
@script = N'str(OutputDataSet);
packagematrix <- installed.packages();
Name <- packagematrix[,1];
Version <- packagematrix[,3];
OutputDataSet <- data.frame(Name, Version);',
@input_data_1 = N'
'
WITH RESULT SETS ((PackageName nvarchar(250), PackageVersion nvarchar(max) ))
GO
```

Here is a result example for the above query, by now the readr package is not installed.

```
EXECUTE sp_execute_external_script
    @language=N'R',
@script = N'str(OutputDataSet);
packagematrix <- installed.packages();
Name <- packagematrix[,1];
Version <- packagematrix[,3];
OutputDataSet <- data.frame(Name, Version);',
@input_data_1 = N'
    '
WITH RESULT SETS ((PackageName nvarchar(250), PackageVersion nvarchar(max) ))
GO
```

100 %

Results   Messages

	PackageName	PackageVersion
25	MicrosoftML	9.4.7
26	MicrosoftR	3.5.2
27	nlme	3.1-137
28	nnet	7.3-12
29	parallel	3.5.2
30	png	0.1-7
31	R6	2.3.0
32	RevoMods	11.0.1
33	RevoPemaR	10.0.0
34	RevoScaleR	9.4.7
35	RevoUtils	11.0.2
36	RevoUtilsMath	11.0.0
37	RODBC	1.3-15
38	rpart	4.1-13
39	RUnit	0.4.26
40	spatial	7.3-11
41	splines	3.5.2
42	stats	3.5.2
43	stats4	3.5.2
44	survival	2.43-3
45	tcltk	3.5.2
46	tools	3.5.2

4. Download and install R based on the SQL MI R version, I downloaded the 3.5.2 R version R: https://www.r-project.org/ ⊡

5. Download and Install Rstudio https://posit.co/download/rstudio-desktop/ ⊡

6. Download and install Azure Data Studio https://learn.microsoft.com/en-us/sql/azure-data-studio/what-is-azure-data-studio?view=sql-server-ver16 ⊡

7. Install SQLMLUTILS, for this exercise I installed online, but also an offline option is available:

To install sqlmlutils first ODBC packages should be installed as described here https://learn.microsoft.com/en-us/sql/machine-learning/package-management/install-additional-r-packages-on-sql-server?view=sql-server-ver16#add-the-package-online ⊡ To install ODBC packages cmd should be used, based on review with PG team the right command should be:

```
R.exe -e "install.packages('odbc', type='binary', repos='https://cran.us.r-project.org')
```

The above command is not well documented on the public link. To run the above command you need to make sure to be on R path by default should be "C:\Program Files\R\R-3.5.2\bin" here is the example

```
C:\Program Files\R\R-3.5.2\bin>R.exe -e "install.packages('odbc', type='binary', repos='https://cran.us.r-project.org')"
```

This is the output generated.

```
> install.packages('odbc', type='binary', repos='https://cran.us.r-project.org')
Installing package into 'C:/Users/glomor/Documents/R/win-library/3.5'
(as 'lib' is unspecified)

  There is a binary version available (and will be installed) but the
  source version is later:
     binary source
odbc  1.2.2  1.3.3

also installing the dependency 'BH'

probando la URL 'https://cran.us.r-project.org/bin/windows/contrib/3.5/BH_1.72.0-3.zip'
Content type 'application/zip' length 18270616 bytes (17.4 MB)
==================================================
downloaded 17.4 MB

probando la URL 'https://cran.us.r-project.org/bin/windows/contrib/3.5/odbc_1.2.2.zip'
Content type 'application/zip' length 1079260 bytes (1.0 MB)
==================================================
downloaded 1.0 MB

package 'BH' successfully unpacked and MD5 sums checked
package 'odbc' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\glomor\AppData\Local\Temp\RtmpO8hoo1\downloaded_packages
>
>
```

After ODBC packages are installed, the next step is to install sqlmlutils by download the sqlmlutils_1.2.1_R3.5.2.tar.gz file found here https://github.com/microsoft/sqlmlutils/releases ☐ The command also is not well documented and you need to make sure to be on R path and specify the path for the file downloaded.

```
R.exe CMD INSTALL "C:\Users\glomor\Downloads\sqlmlutils_1.2.1_R3.5.2.tar.gz
```
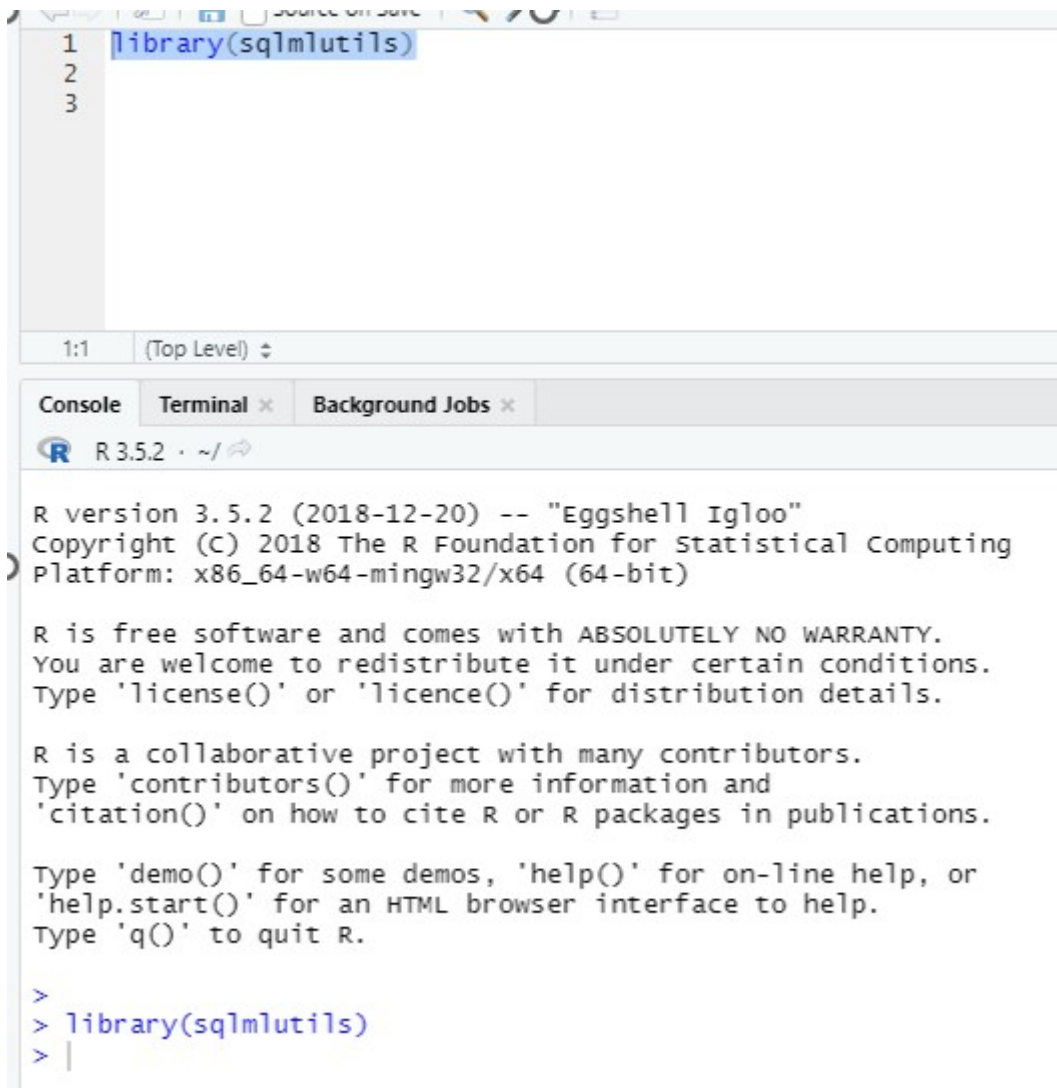
Here is the example:

```
C:\Program Files\R\R-3.5.2\bin>R.exe CMD INSTALL "C:\Users\glomor\Downloads\sqlmlutils_1.2.1_R3.5.2.tar.gz"
In R CMD INSTALL
* installing to library 'C:/Users/glomor/Documents/R/win-library/3.5'
* installing *source* package 'sqlmlutils' ...
** R
** byte-compile and prepare package for lazy loading
** help
*** installing help indices
  converting help for package 'sqlmlutils'
    finding HTML links ... done
    checkSproc                          html
    connectionInfo                      html
    createSprocFromFunction             html
    dropSproc                           html
    executeFunctionInSQL                html
    executeSQLQuery                     html
    executeScriptInSQL                  html
    executeSproc                        html
    sql_install.packages                html
    sql_installed.packages              html
    sql_remove.packages                 html
    sqlmlutils-package                  html
** building package indices
** testing if installed package can be loaded
*** arch - i386
*** arch - x64
* DONE (sqlmlutils)

C:\Program Files\R\R-3.5.2\bin>
```

8. If the sqlmlutils were successfully installed the R packages can be installed by following the described here https://learn.microsoft.com/en-us/sql/machine-learning/package-management/install-additional-r-packages-on-sql-server?view=sql-server-ver16#add-the-package-online ☐

A way to validate if sqlmlutils were successfully installed is trying to upload to R Studio by running library(sqlmlutils) command.

If the sqlmlutils were not succesfully istalled an error like follow will be shown:

To install the required R package on this example the readr package, the Rstudio tool should be used, this is the example of the command used, I connected to SQL MI using public endpoint there is also a useful youtube video https://www.youtube.com/watch?v=JOeVh0HCxgQ&t=57s ☑

```
library(sqlmlutils)


connection <- connectionInfo(
server   = "sqlmiserver.public.a92ad62f61b7.database.windows.net,3342", -- replace with your SQLMI
database = "dbname",   --Replace with your database name
uid      = "user",     --Replace with your SQL user
pwd      = "Pasword") --Replace with your SQL password

sql_install.packages(connectionString = connection, pkgs = "readr", verbose = TRUE, scope = "PUBLIC")
```

This is the output for the command executed

9. You can go to the SSMS studio and validate if the readr package is installed by running the command on the step 3 on SSMS

```sql
EXECUTE sp_execute_external_script
    @language=N'R',
    @script = N'str(OutputDataSet);
    packagematrix <- installed.packages();
    Name <- packagematrix[,1];
    Version <- packagematrix[,3];
    OutputDataSet <- data.frame(Name, Version);',
    @input_data_1 = N'
    '
    WITH RESULT SETS ((PackageName nvarchar(250), PackageVersion nvarchar(max) ))
    GO
```

	PackageName	PackageVersion
1	assertthat	0.2.1
2	cli	2.0.2
3	clipr	0.7.0
4	crayon	1.3.4
5	digest	0.6.25
6	ellipsis	0.3.0
7	fansi	0.4.1
8	glue	1.4.0
9	hms	0.5.3
10	lifecycle	0.2.0
11	magrittr	1.5
12	pillar	1.4.3
13	pkgconfig	2.0.3
14	Rcpp	1.0.4.6
15	readr	1.3.1
16	rlang	0.4.5
17	tibble	3.0.1
18	utf8	1.1.4
19	vctrs	0.2.4
20	base	3.5.2
21	boot	1.3-20
22	checkpoint	0.4.4
23	class	7.3-14

More Information

There is also a useful youtube video https://www.youtube.com/watch?v=JOeVh0HCxgQ&t=57s ⧉

Public Doc Reference

https://learn.microsoft.com/en-us/azure/azure-sql/managed-instance/machine-learning-services-overview?view=azuresql ⧉

https://www.r-project.org/ ⧉

https://posit.co/download/rstudio-desktop/ ⧉

https://learn.microsoft.com/en-us/sql/azure-data-studio/what-is-azure-data-studio?view=sql-server-ver16 ⧉

https://learn.microsoft.com/en-us/sql/machine-learning/package-management/install-additional-r-packages-on-sql-server?view=sql-server-ver16#add-the-package-online ⧉

https://github.com/microsoft/sqlmlutils/releases ⧉

https://learn.microsoft.com/en-us/sql/machine-learning/package-management/install-additional-r-packages-on-sql-server?view=sql-server-ver16#add-the-package-online ⧉

Internal Reference

There is an ICM with detailed about this
https://portal.microsofticm.com/imp/v3/incidents/details/343225001/home ⧉

How good have you found this content?

🙂 🙁