

#sqlpass

# Using R Services in SQL Server

Hylke Peek en Yorick Bouma | June 16, 2016 | Veenendaal

# Topics

Advanced Analytics with SQL Server 2016

Data Scientist – R Services using R Gui

SQL Server Professional – R Services using SQL Server Management Studio

#sqlpass

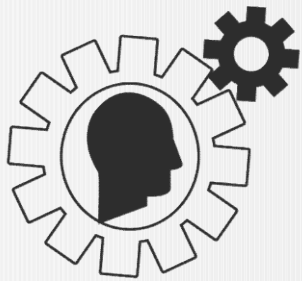


## Advanced Analytics with SQL Server 2016

SQL Server 2016 | June 16, 2016 | #sqlpass

**infoSupport**  
Solid Innovator

# Intelligent Applications



Make decisions without coding



Automatically update themselves

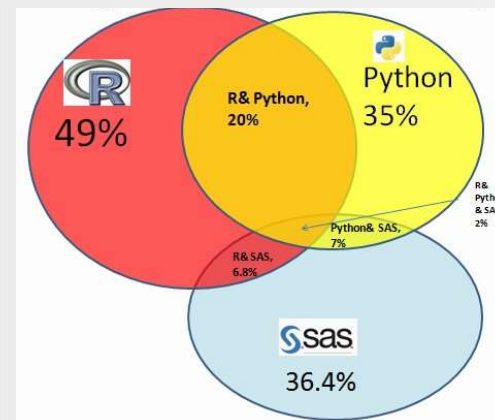


Customized for each client

## R - What and Why?



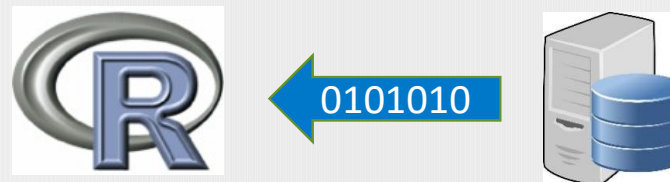
- Tool for Statistical Analysis
- Open Source
- Free
- Community
- Used by Data Analysts (a.k.a. Data Scientist)



- Language with many statistical capabilities
- Many data analysts use R
- R's popularity is growing

# Challenges

## Data Movement



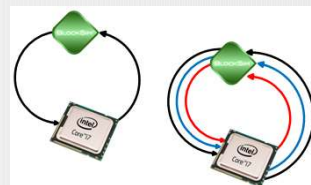
In-Database Analytics

## Consuming models



Deploy model to SQL Server

## Scale / Performance

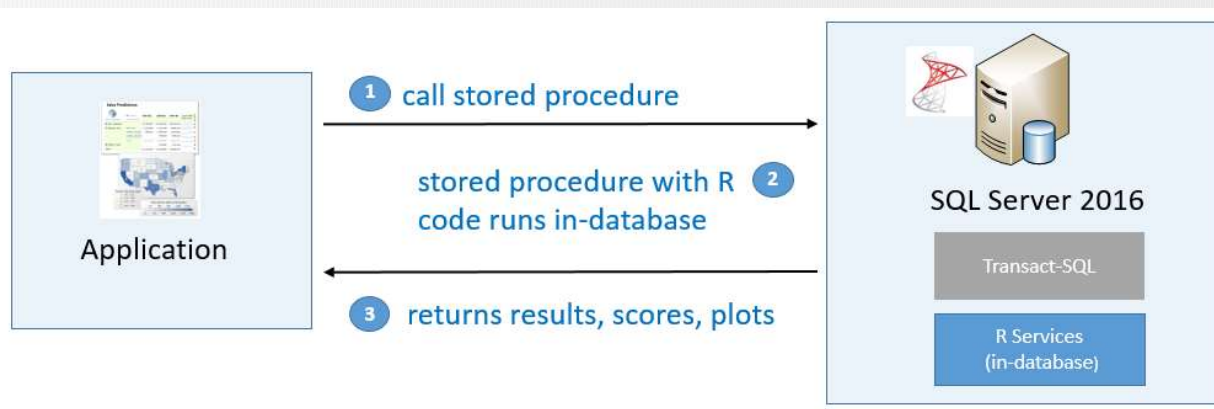
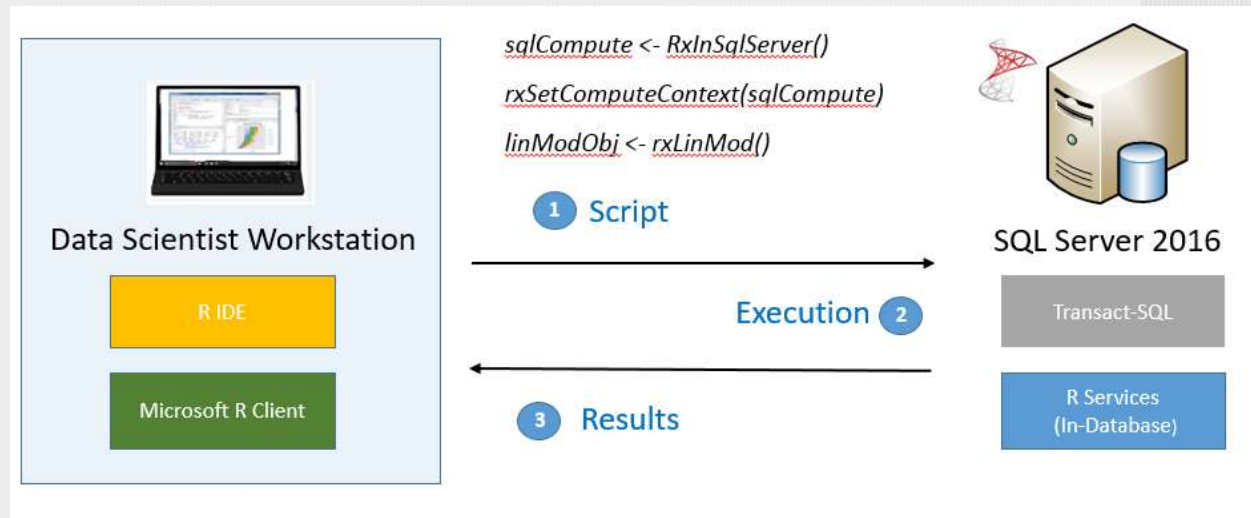


Distributed processing functions

# Key Concepts

## Data Scientist

Data Exploration and Predictive Modelling



## SQL Developer and Administrator

Operationalizing models and managing runtime execution

#sqlpass



## Data Scientist – R Services using R GUI



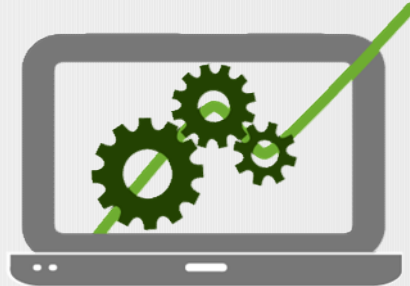
# Data Science Process



Collect



Explore



Prepare  
variables



Model

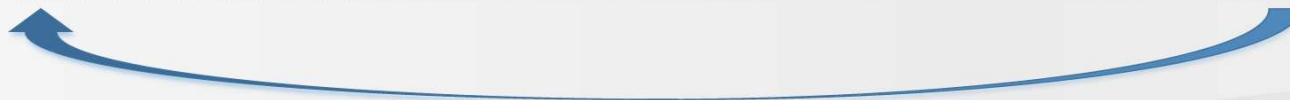
Variables

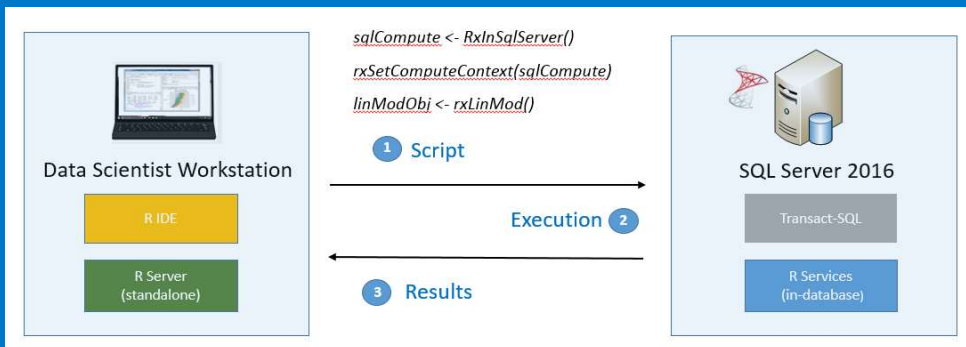


Prediction



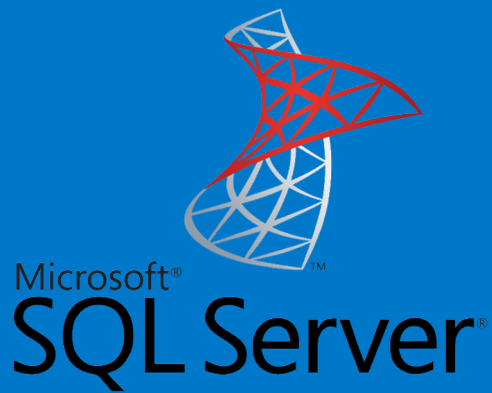
App





## Demo

#sqlpass



## SQL Server Professional – R Services using SSMS

# Prerequisites



**R script from your Data Scientist**



**Install R Services (In-Database) on SQL Server 2016**



**Enable R Services feature**

# Prerequisites – R script from your Data Scientist

```
testDataQuery <- "select wheelBase, engineSize, horsepower,
  peakRPM, highwayMPG from dbo.carprice_test"

testDataSource <- RxSqlServerData(
  sqlQuery = testDataQuery
  , connectionString = connStr
  , rowsPerRead=500
  , colClasses = colClasses
  , colInfo = colInfo
)

# predict and write the prediction results back to SQL Server table
scoredOutput <- RxSqlServerData(
  connectionString = connStr,
  table = "carprice_scores"
)

predict <- rxPredict(
  modelObject = linmodel, data = testDataSource,
  outData = scoredOutput, predVarNames = "Score",
  type = "link", writeModelVars = TRUE, overwrite = TRUE
)
```

# Prerequisites - Install R Services (In-Database) on SQL Server 2016

The image displays two side-by-side screenshots of the SQL Server 2016 Setup wizard.

**Left Screenshot: Feature Selection**

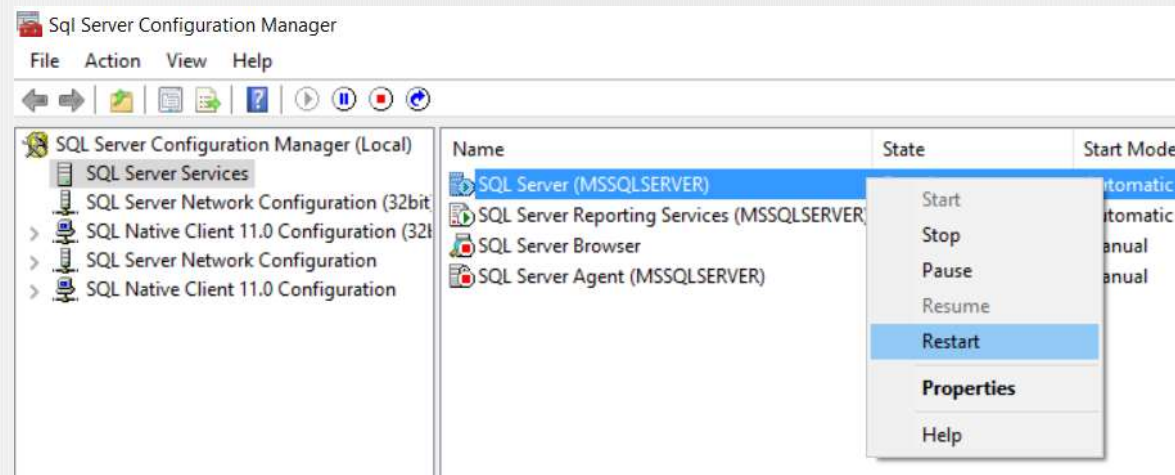
- Title Bar:** SQL Server 2016 Setup
- Section Header:** Feature Selection
- Text:** Select the Evaluation features to install.
- Left Pane:** A list of installation steps including Product Key, License Terms, Global Rules, Install Setup Files, Install Rules, Installation Type, **Feature Selection** (highlighted), Feature Rules, Instance Configuration, Server Configuration, Database Engine Configuration, Consent to install Microsoft R ..., Feature Configuration Rules, Ready to Install, Installation Progress, and Complete.
- Main Area:**
  - Features:** A list of features with checkboxes. Under the 'Instance Features' section, 'Database Engine Services' and 'R Services (In-Database)' are checked. Other features like 'SQL Server Replication', 'Full-Text and Semantic Extractions for Search', 'Data Quality Services', 'PolyBase Query Service for External Data', 'Analysis Services', and 'Reporting Services - Native' are unchecked. Under 'Shared Features', 'R Server (Standalone)', 'Reporting Services - SharePoint', and 'Reporting Services Add-in for SharePoint Products' are also unchecked.
  - Feature description:** The description for the selected 'R Services (In-Database)' feature: 'The configuration and operation of each instance feature of a SQL Server instance is isolated from other SQL Server instances. SQL Server instances can operate side-by-side on the same server.' Below this, it lists prerequisites: 'Already installed: Windows PowerShell 3.0 or higher, Microsoft Visual Studio 2010 Redistributable'. It also shows disk space requirements: 'Drive C: 2320 MB required, 1080 MB available'.
  - Buttons:** 'Select All' and 'Unselect All' buttons are present.
  - Directories:** Three text boxes for directory paths are shown, all pointing to 'C:\Program Files\Microsoft SQL Server\':
    - Instance root directory:
    - Shared feature directory:
    - Shared feature directory (x86):
- Bottom:** '< Back', 'Next >', and 'Cancel' buttons.

**Right Screenshot: Consent to install Microsoft R Open**

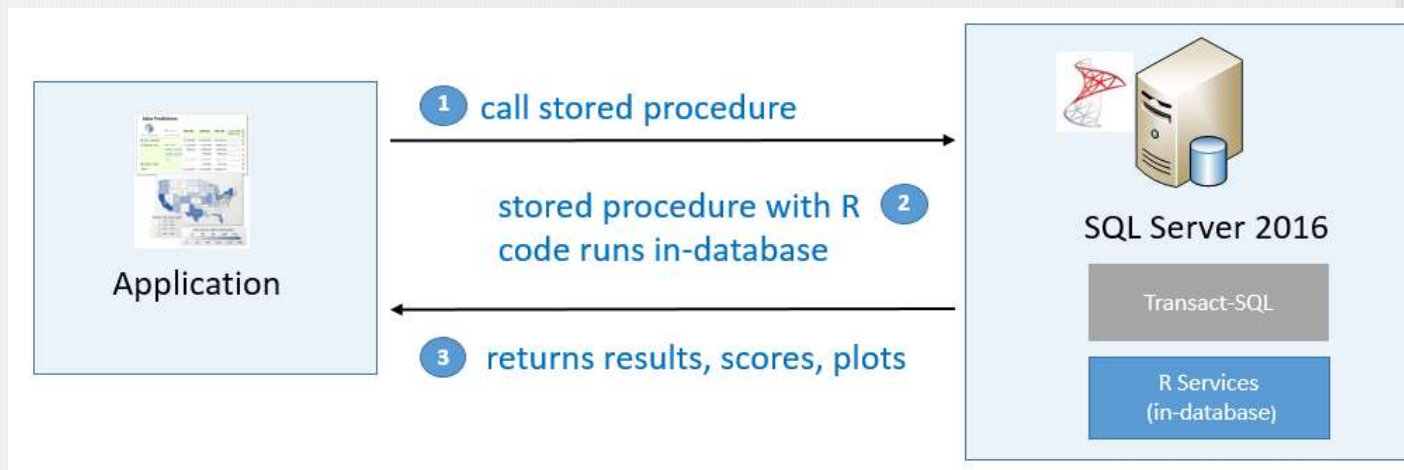
- Title Bar:** SQL Server 2016 Setup
- Section Header:** Consent to install Microsoft R Open
- Text:** Download and install necessary pre-requisite.
- Left Pane:** The same list of installation steps as the left screenshot, with 'Consent to install Microsoft R ...' highlighted.
- Main Area:**
  - Text:** 'Microsoft R Open is an enhanced distribution of R made available by Microsoft under the GNU General Public License v2.' and 'R is © the R Foundation for Statistical Computing. For more information on R-related products and services, visit <http://r-project.org>.'
  - Text:** 'By clicking "Accept" you are choosing to download Microsoft R Open and install it on your machine, and agreeing to accept patches and updates to this software according to your SQL Server update preferences.'
  - Button:** An 'Accept' button is highlighted with a red rectangle.
- Bottom:** '< Back', 'Next >', and 'Cancel' buttons.

# Prerequisites - Enable R Services feature

```
Exec sp_configure 'external scripts enabled', 1  
Reconfigure with override
```



# Operationalizing





# Operationalizing – sp\_execute\_external\_script

- Executes provided script at an external location
- Only supports R (at this moment)
- Takes a transact-SQL query as input
- A R data frame can be outputted as a result
- Support for parameters

# Operationalizing – sp\_execute\_external\_script example

```
testDataQuery <- "select wheelBase, engineSize, horsepower,
peakRPM, highwayMPG from dbo.carprice_test"

testDataSource <- RxSqlServerData(
  sqlQuery = testDataQuery
  , connectionString = connStr
  , rowsPerRead=500
  , colClasses = colClasses
  , colInfo = colInfo
)

# predict and write the prediction results back to SQL Server table
scoredOutput <- RxSqlServerData(
  connectionString = connStr,
  table = "carprice_scores"
)

predict <- rxPredict(
  modelObject = linmodel, data = testDataSource,
  outData = scoredOutput, predVarNames = "Score",
  type = "link", writeModelVars = TRUE, overwrite = TRUE
)
```

```
DECLARE @linmodel varbinary(max) = (SELECT TOP 1 model FROM carprice_models)

EXEC sp_execute_external_script
  @language = N'R',
  @input_data_1 = N'
    SELECT wheelBase, engineSize, horsepower,
    peakRPM, highwayMPG FROM dbo.carprice_test',
  @script = N'
    linmodel <- unserialize(as.raw(model));

    OutputDataSet <- rxPredict(
      modelObject = linmodel, data = InputDataSet,
      outData = NULL, predVarNames = "Score",
      type = "link", writeModelVars = TRUE, overwrite = TRUE
    )',
  @params = N'@model varbinary(max)',
  @model = @linmodel;
```

# Operationalizing – call from application

Dataset Properties

Query  
Fields  
Options  
Filters  
Parameters

Choose a data source and create a query.

Name:  
PriceHistogram

☐ Use a shared dataset.  
☒ Use a dataset embedded in my report.

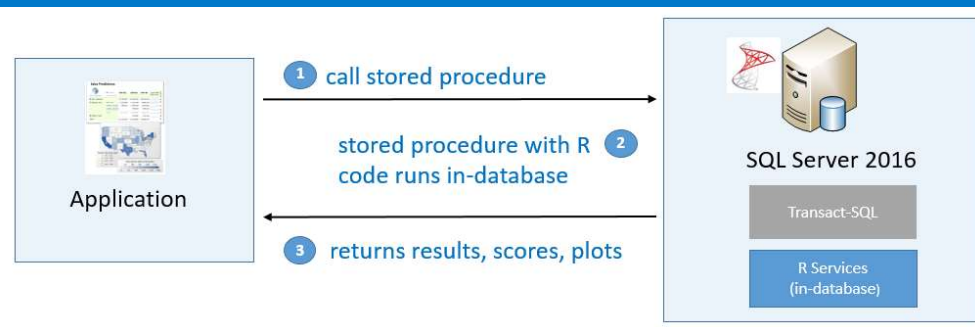
Data source:  
Localhost New...

Query type:  
☐ Text ☐ Table ☒ Stored Procedure

Select or enter stored procedure name:  
PlotHistogram

Query Designer... Refresh Fields





## Demo

# Data Solutions Trainings

- |   |   |                        |
|---|---|------------------------|
| ▪ <b>Updating Your Skills to SQL Server 2016</b>        | ➔ | <b>28 - Sep - 2016</b> |
| ▪ <b>SQL Server Performance Tuning and Optimization</b> | ➔ | <b>24 - Aug - 2016</b> |
| ▪ <b>The Essentials of Programming in R</b>             | ➔ | <b>22 - Aug - 2016</b> |
| ▪ <b>Windows Powershell Desired State Configuration</b> | ➔ | <b>29 - Aug - 2016</b> |
| ▪ <b>Microsoft Azure Machine Learning</b>               | ➔ | <b>08 - Sep - 2016</b> |
| ▪ <b>Microsoft Azure Data Solutions</b>                 | ➔ | <b>06 - Okt - 2016</b> |
| ▪ <b>Mastering and Optimizing DAX</b>                   | ➔ | <b>17 - Okt - 2016</b> |

**[training.infosupport.com](http://training.infosupport.com)**

#sqlpass

Questions?

# Evaluation

