

Where Does R Fit Into Your SQL Server Stack?

Stacia Varga

Data Inspirations

stacia@datainspirations.com

blog.datainspirations.com

[@_Stacia_V](#)

Stacia (Misner) Varga

- Over 30 years of IT experience, 17 years of BI experience
- Frequently speaks, writes, and teaches about Microsoft BI technologies
- Principal Consultant and Founder, Data Inspirations
- Data Platform MVP and SSAS Maestro
- Las Vegas User Group Chapter Leader and SQL Saturday Organizer
- Wants you to make sure you can *understand what your data is telling you*





Agenda

- Introducing R
- Using R in SQL Server
- Preparing Data with R
- Reporting and Visualizing Data with R



INTRODUCING R

What is R?

R is a **programming language** and software environment for **statistical computing and graphics...** [and] is widely used among statisticians and data miners for.. data analysis.

1992 Conceived as implementation of S

1994 Released

2000 Stable beta version

Who don't necessarily have coding skills

[https://en.wikipedia.org/wiki/R_\(programming_language\)](https://en.wikipedia.org/wiki/R_(programming_language))



Why Use R?

Comprehensive statistical routines

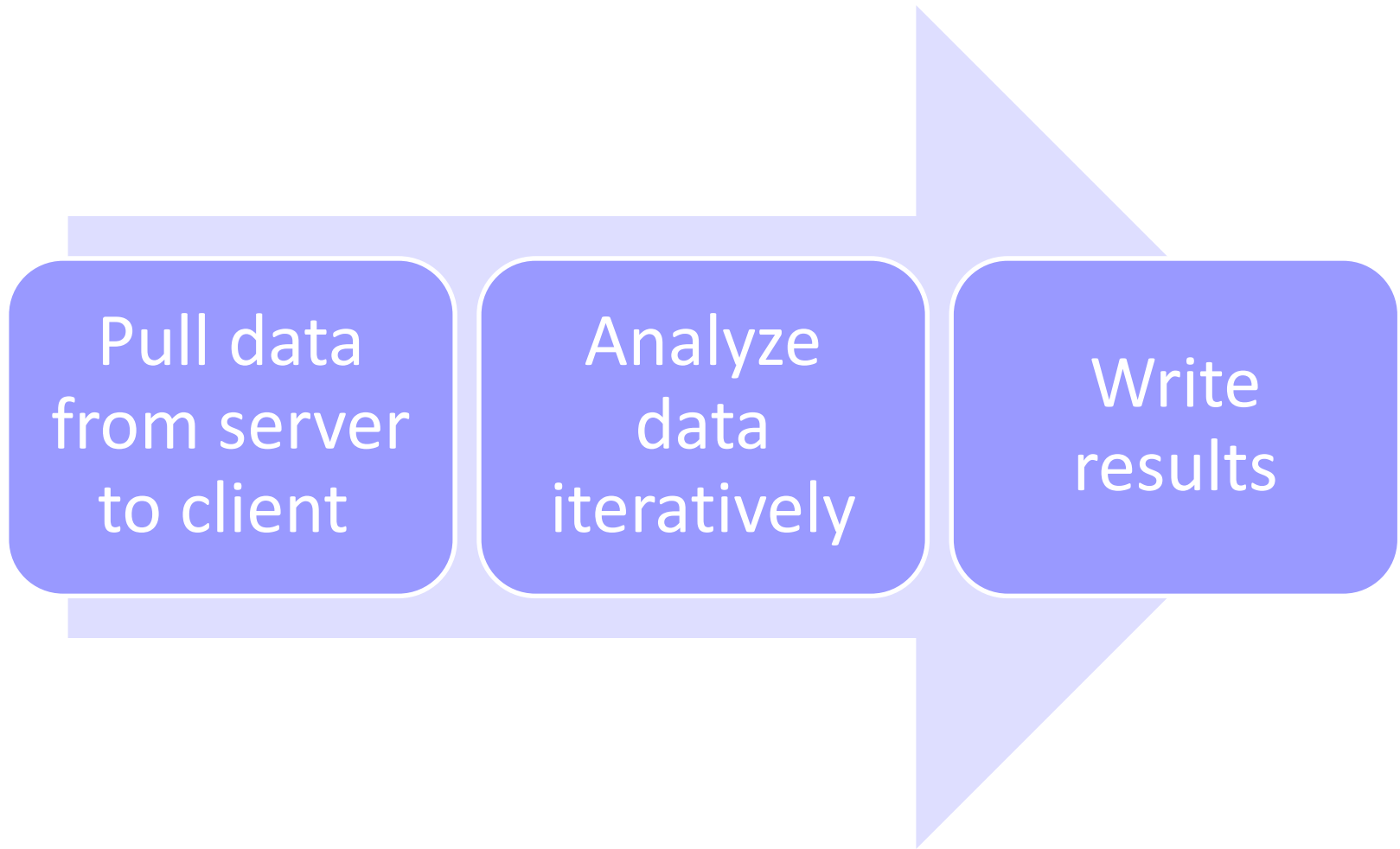
Extensibility via packages

Integration with other applications

Powerful data visualization capabilities

Free!

Typical R Workflow



R Distributions

CRAN



<https://cran.r-project.org>

Comprehensive R Archive Network
Code + Documentation

Microsoft R Open



<https://mran.revolutionanalytics.com/rro>

Multithreading capabilities
Parallelism

R Development Tools

RStudio



<https://www.rstudio.com>

R Tools for Visual Studio



<https://microsoft.github.io/RTVS-docs/>

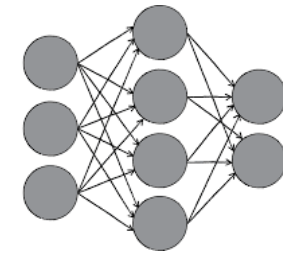
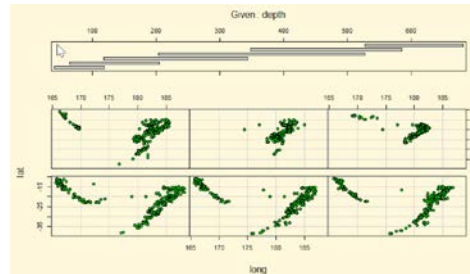
Common R Use Cases

```
> rxSummary(., data = InDataSource2)
Call:
rxSummary(formula = ~., data = InDataSource2)
Summary statistics Results for: ~.
Data: InDataSource2 (Kxq)ServerData Data Source
Number of valid observations: 18484

  Name      Mean      StdDev   Min   Max  ValidObs MissingObs
CustomerKey 2.024150e+04 5.336016e+03 11000 29483 18484 0
YearlyIncome 5.730378e+04 3.228584e+04 10000 170000 18484 0
TotalChildren 1.844352e+00 1.612408e+00 0 5 18484 0
NumberChildrenatHome 1.004058e+00 1.522660e+00 0 5 18484 0
NumberCarsOwned 1.502705e+00 1.128294e+00 0 4 18484 0
Age 4.621732e+01 1.153755e+01 29 99 18484 0
BikeBuyer 4.940489e-01 4.999781e-01 0 1 18484 0

Category Counts for MaritalStatus
Number of categories: 2
Number of valid observations: 18484
Number of missing observations: 0
MaritalStatus Counts
M 10011
S 8473

Category Counts for Gender
Number of categories: 2
Number of valid observations: 18484
Number of missing observations: 0
Gender Counts
M 9351
F 9133
```



Exploratory
data analysis

Advanced data
visualization

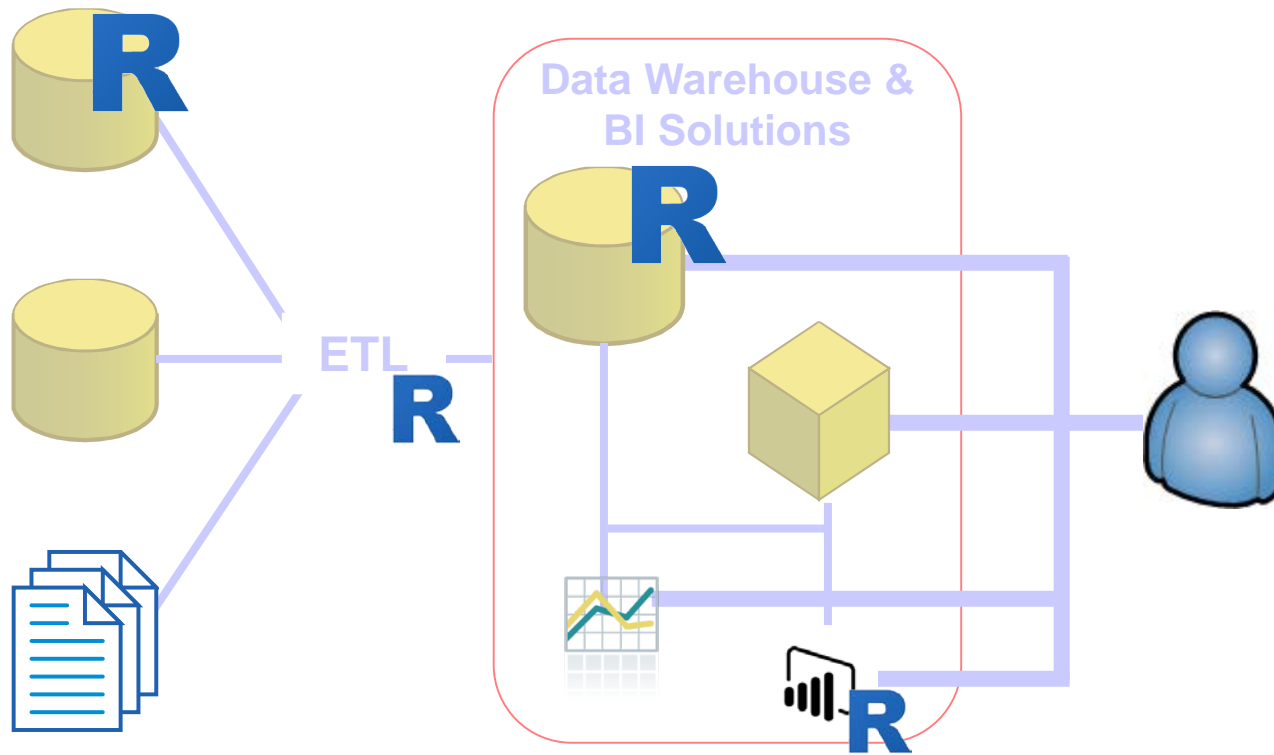
Predictive
modeling

□ Introducing the Iris Data Set



	id	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	1	5.1	3.5	1.4	0.2	setosa
2	2	4.9	3	1.4	0.2	setosa
3	3	4.7	3.2	1.3	0.2	setosa
4	4	4.6	3.1	1.5	0.2	setosa
5	5	5	3.6	1.4	0.2	setosa
6	6	5.4	3.9	1.7	0.4	setosa
7	7	4.6	3.4	1.4	0.3	setosa
8	8	5	3.4	1.5	0.2	setosa

R in the Microsoft Stack





USING R IN SQL SERVER

Why SQL Server 2016 R Services?

Pull data
from server
to client

Moving data
removes security
controls

Analyze
data
iteratively

Local desktop has
memory & CPU
constraints

Write
results



Installing R Services

Server

- R Services (In-Database)
- R Services (Standalone)

Client Workstation (Dev)

- R Distribution
- R Services (Standalone)
- R IDE

Configuring R Services

```
exec sp_configure 'external scripts enabled', 1  
reconfigure with override  
  
-- restart SQL Server service
```

Server

Configuring R Services (Optional)

- Add login for SQLRUserGroup to instance
- Grant EXECUTE ANY EXTERNAL SCRIPT to each user of R in a database
- Reduce maximum SQL Server memory to 80 % of physical memory



Server

❑ Installing Packages on SQL Server

Internet Access

```
lib.SQL <- "C:\\Program  
Files\\Microsoft SQL  
Server\\MSSQL13.MSSQLSERVER\\R  
_SERVICES\\library"
```

```
install.packages("e1071",  
lib = lib.SQL)
```

```
<InstallPath>\\R_Services\\bin\\x64
```

No Internet Access

1. On server with access, install minicran and igraph packages, then create repository
2. On SQL Server, install packages from repository

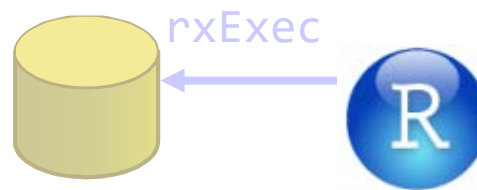
<https://docs.microsoft.com/en-us/sql/advanced-analytics/r/create-a-local-package-repository-using-minicran>

Do NOT install packages in R script!

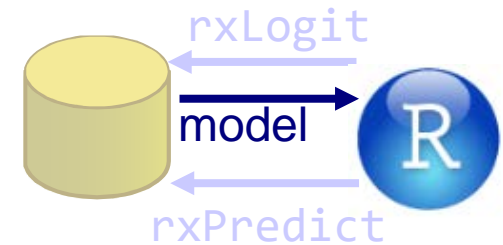
Using rx Functions



Explore and
visualize



Use server
resources only



Train and
store model
output



□ Operationalizing R Models

- Get R code from developer
- Use `sp_execute_external_script` T-SQL wrapper for R code
- Store – Serialize and convert model to data frame
- Use – Deserialize model and use in `rxPredict`



PREPARING DATA WITH R

❑ Executing R in SSIS

Run R in Batch Mode



```
R CMD BATCH  
[options] infile [outfile]
```

Call Stored Procedure



```
EXEC <stored_proc>
```

Getting Data for Power BI

Get Data

- All
- File
- Database
- Azure
- Online Services
- Other

Other

- Web
- SharePoint List
- OData Feed
- Hadoop File (HDFS)
- Active Directory
- Microsoft Exchange
- ODBC
- R Script
- Spark (Beta)
- Blank Query

Execute R Script

Execute an R script on a local R installation and import the resulting data frames.

R is installed at C:\Program Files\Microsoft\MRO\R-3.3.0.

Go to Options & Settings to change which installation you want to run, and for more configuration options.

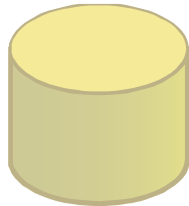
OK

Cancel



REPORTING AND VISUALIZING DATA WITH R

Using Reporting Services



Stored procedure

- Create temporary file for graphic
- Use R function to plot data as graphic
- Output result as R data.frame to SQL Server varbinary(max)



Dataset

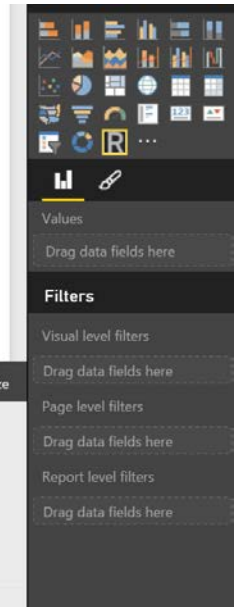
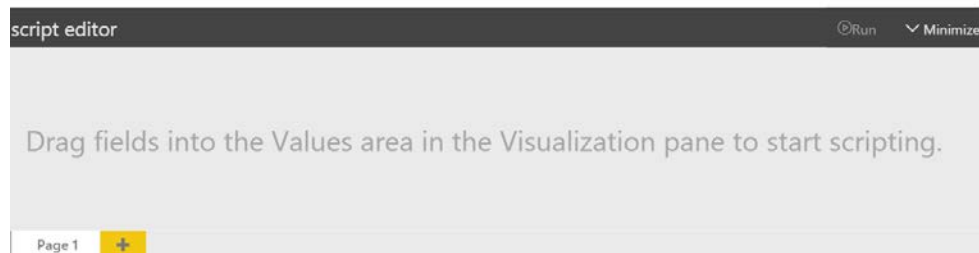
- Call stored procedure

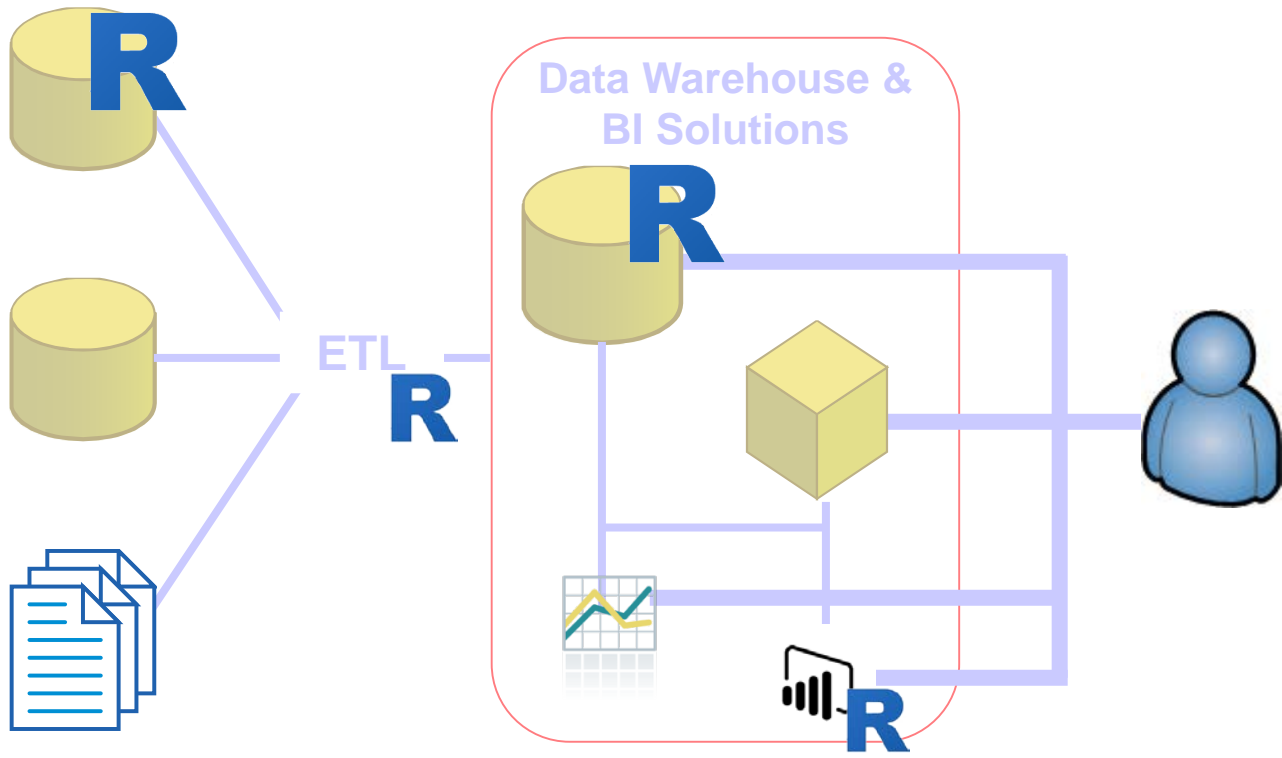
Report Body

- Add image to report with database image as source

Visualizing in Power BI

- Identify R distribution location in Options
- Enable script visuals when prompted
- Select fields
- Reference generic dataset in R script







Resources

- Summary of rx functions
 - <https://msdn.microsoft.com/en-us/library/mt652103.aspx>
- Download demo scripts
 - <http://tinyurl.com/sqlr-sqlsatcolumbus>