



Getting Started with R Scripting in Power BI Desktop

Presented by Ted Pattison

WELCOME TO THE POWER BI USER GROUP (PBIUG)

THE OFFICIAL USER GROUP FOR POWER BI USERS. OUR MISSION IS TO ENGAGE EVERY INDIVIDUAL USING POWER BI IN VALUE-ADDED NETWORKING, COLLABORATION AND KNOWLEDGE SHARING.

About Ted Pattison and Critical Path Training



Ted Pattison

- 25 years as an author, technical trainer & conference speaker
- Specializing in Power BI, Office 365, SharePoint & Azure
- Awarded as SharePoint MVP by Microsoft 12 years in a row



Critical Path Training

- Advanced technical training on Power BI, SharePoint & others
- Offering **Power BI Bootcamp** as 3-day deep dive into Power BI
- <https://www.criticalpathtraining.com>

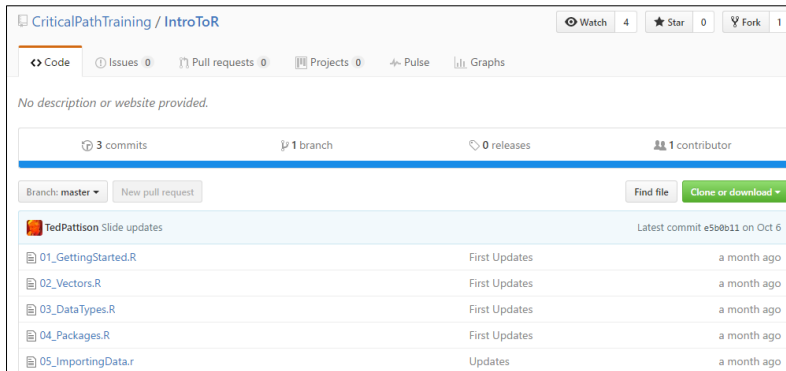
My Upcoming Power BI Training Events

- Power BI Bootcamp in Tampa on Nov 29 and Jan 23
- Dynamic Communities Distance Learning Course
- Dynamic Communities Events (e.g. Focus, Summit)



Sample Code from the Webinar

- All code maintained in GitHub repository
- <https://github.com/CriticalPathTraining/IntroToR>



Agenda

- Understanding R as an Analytics Platform
- Installing Microsoft R Open and RStudio
- Writing R Code in RStudio
- Integrating R with Power BI Desktop

What is R?

- What is R?
 - Platform for statistics, data analysis and visualization
 - Free, cross-platform, open source software
 - Programming language + Runtime layer + Libraries
 - R code distributed and versioned using packages
 - Flourishing ecosystem of R package authors
- Why do you need it?
 - Analyzing data and generating statistics
 - Creating rich graphs and charts
 - Fitting statistical models for predictive analysis



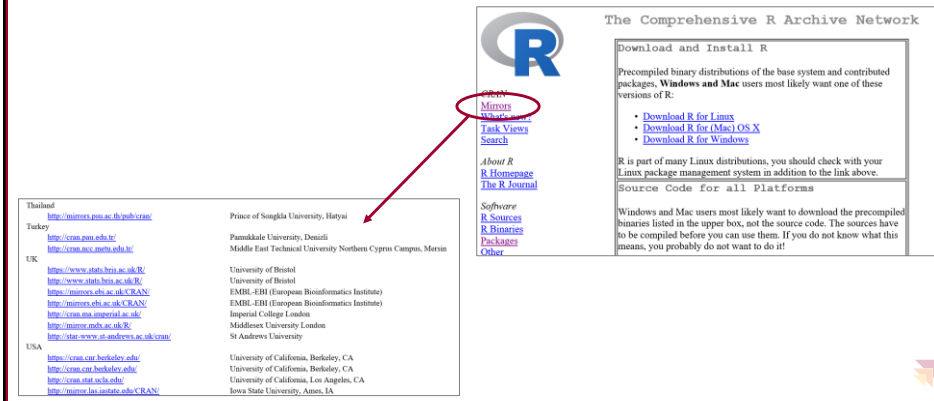
R Packages

- Package is versioned redistributable unit of code
 - Package contains functions, data and compiled code
 - R is installed with a default set of packages
 - Other packages can be downloaded and installed
- Examples of available domain-specific packages
 - Packages to download and unpack data in zip archive
 - Packages to create fancy charts and graphs
 - Packages to optimize financial portfolios
 - Packages predict component failure times
 - Packages to analyze genomic sequences



CRAN

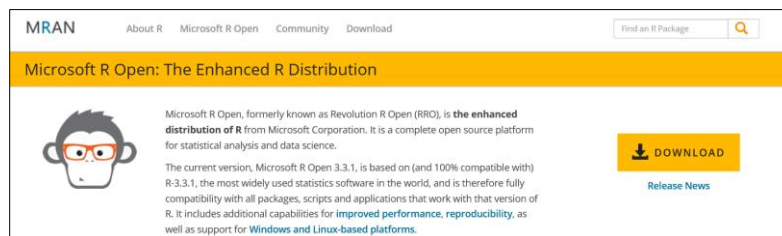
- The Comprehensive R Archive Network
 - Public archive with over 8,000 downloadable packages
 - <http://cran.us.r-project.org/>



The screenshot shows the CRAN website. The main heading is 'The Comprehensive R Archive Network'. Below it, there's a section 'Download and Install R' with links for 'Download R for Linux', 'Download R for (Mac) OS X', and 'Download R for Windows'. A red arrow points to the 'CRAN Mirrors' link in the navigation menu. Below the navigation menu, there's a list of mirrors for various countries: Thailand, Turkey, UK, and USA, each with a list of mirror URLs.

Microsoft R Open

- What is Microsoft R Open?
 - An enhanced distribution of R from Microsoft
 - Improved performance and multithreading
 - Reproducibility through package versioning stability
 - Free, cross-platform, open source software
 - Available at <https://mran.microsoft.com/open/>



The screenshot shows the Microsoft R Open (MRO) website. The main heading is 'Microsoft R Open: The Enhanced R Distribution'. Below it, there's a description of MRO: 'Microsoft R Open, formerly known as Revolution R Open (RRO), is the enhanced distribution of R from Microsoft Corporation. It is a complete open source platform for statistical analysis and data science.' There's a 'DOWNLOAD' button and a 'Release News' link.

Stages of R Awareness

- Stage 1: Standing Up
 - Installing the environment and playing with data
- Stage 2: Walking
 - Writing & testing R code and creating graphs and charts
- Stage 3: Jogging
 - Crunching numbers to generate advanced statistics
- Stage 4: Running
 - Creating a domain-specific predictive model
- Stage 5: Sprinting
 - Distributing your predictive model as a CRAN package



Agenda

- ✓ Understanding R as an Analytics Platform
- Installing Microsoft R Open and RStudio
 - Writing R Code in RStudio
 - Integrating R with Power BI Desktop



Install Microsoft R Open

- <https://mran.microsoft.com/download/>

Download Microsoft R Open 3.3.1, the enhanced R distribution

Microsoft R Open, the enhanced distribution of R from Microsoft, is a complete and free open source platform for statistical analysis and data science. Microsoft R Open 3.3.1 is based on (and 100% compatible with) the statistical language, R 3.3.1. It includes additional capabilities for performance, reproducibility and platform support. [Learn more...](#)

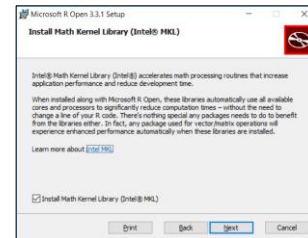
[Prerequisites & Install Docs](#) | [Forum](#) | [News](#) | [Past Releases](#)

Microsoft R Open & MKL Downloads

While the install of MKL, used for multithreaded performance, is optional, we recommend both Microsoft R Open & MKL for optimal performance on Windows and Linux. The OS X operating system has Math Libraries by default.

Platforms (64-bit only)	Downloads
Windows - Windows 7.0 (SP1), 8.1, 10 and Windows Server® 2008 R2 (SP1), 2012 SHA-256: D4F8D276A1705C24F8A0A0F7E7A7A33274A4F0F0B767A0C230B9797	R Open / MKL
Ubuntu - 14.04, 15.04 SHA-256: C258A8A42F29A7F7175A4D0A605778A27A4A0F0B767A0C230B9797	R Open / MKL
Red Hat Enterprise Linux - 6.5, 7.1 SHA-256: C258A8A42F29A7F7175A4D0A605778A27A4A0F0B767A0C230B9797	R Open / MKL

Microsoft R Server Users:
Get R Open for R Server 2016
Get R.R.O.B.3 for R.R.E.7.6.1



Installing R Studio

- <https://www.rstudio.com/products/rstudio/download/>

R Studio Products Resources Pricing About Us Blog

Download RStudio Home / Overview / RStudio / Download RStudio

RStudio is a set of integrated tools designed to help you be more productive with R. It includes a console, syntax-highlighting editor that supports direct code execution, as well as tools for plotting, history, debugging and workspace management.

If you run R on a Linux server and want to enable users to remotely access RStudio using a web browser [please download RStudio Server](#).

Do you need support or a commercial license? [Check out our commercial offerings](#)

RStudio Desktop 0.99.902 — Release Notes

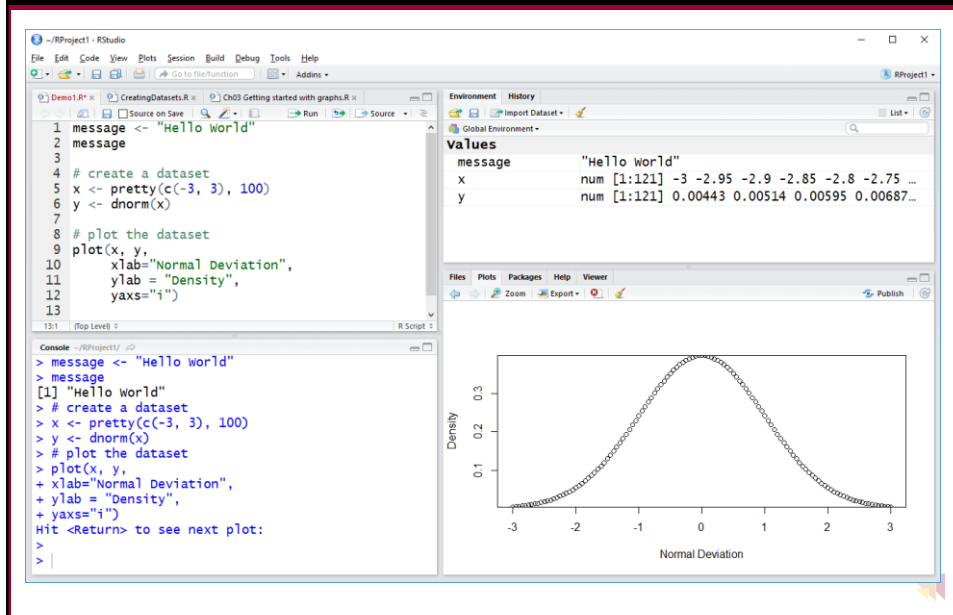
RStudio requires R 2.11.1 (or higher). If you don't already have R, you can download it [here](#).

Installers for Supported Platforms

Installers	Size	Date	MD5
RStudio 0.99.902 - Windows Vista/7/8/10	77.1 MB	2016-05-14	8f9ee61d13b1d81ded7587a1da760d95
RStudio 0.99.902 - Mac OS X 10.6+ (64-bit)	60 MB	2016-05-14	f741e4a1345985c16e692967adba210
RStudio 0.99.902 - Ubuntu 12.04+/Debian 8+ (32-bit)	81.6 MB	2016-05-14	363952616a10553aa51f3a9129b9adeb
RStudio 0.99.902 - Ubuntu 12.04+/Debian 8+ (64-bit)	88.3 MB	2016-05-14	d035622f39928246048972ed2064c89a
RStudio 0.99.902 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (32-bit)	81 MB	2016-05-14	6f14d4717b01e7763d18f1cdad8e6474
RStudio 0.99.902 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)	81.9 MB	2016-05-14	a79e80d952a497a92aafe629f90e90



The RStudio IDE



Agenda

- ✓ Understanding R as an Analytics Platform
- ✓ Installing Microsoft R Open and RStudio
- Writing R Code in RStudio
- Integrating R with Power BI Desktop



R Projects and Workspaces

- R projects based on folder structure
 - Data and scripts added to current working directory
- Each R project defines a workspace
 - Workspace tracks set of user-defined objects
 - Workspace defines set of loaded packages
 - Workspace data saved/loaded using .RData files

```
Console ~/RProject1/ ↵
> getwd()
[1] "C:/Users/Student/Documents/RProject1"
> .libPaths()
[1] "C:/Users/Student/Documents/R/win-library/3.2"
[2] "C:/Program Files/Microsoft/MRO/R-3.2.4/library"
> |
```



Writing and Testing R Code in Scripts

```
01_GettingStarted.R x
Source on Save
1 # use <- for variable assignment
2 message <- "Hello World"
3
4 print(message)
5
6 # create vector using the c function
7 vector1 <- c(2, 4, 6, 8)
8
9 # create vectors using sequence
10 vector2 <- 1:10
11 vector3 = letters[1:5]
12 vector4 = LETTERS[24:26]
13 vector6 = 2^(1:8)
14
15 # create vector with election years
16 election.years <- seq(from = 1996, to = 2016, by = 4)
17
18 # enumerate through election years using for loop
19 for (year in election.years){
20   print(paste(year, "is an election year"))
21 }
22
23 # remove all objects from workspace
24 rm(list=objects())
```



R Objects

- In R, variables represent named objects
- Object names can contain
 - Letters
 - Numbers
 - Underscores (_)
 - Dots (.)



Essential Data Structures in R

- Vector
 - One-dimensional, single-mode array
- Matrix
 - Two-dimensional, single-mode array
- Array
 - N-dimensional, single-mode array
- List
 - Ordered collection of multi-mode objects
- Data frame
 - Two-dimensional, multi-mode array
- Factor
 - Integer-backed list of categorical values



Agenda

- ✓ Understanding R as an Analytics Platform
- ✓ Installing Microsoft R Open and RStudio
- ✓ Writing R Code in RStudio
- Integrating R with Power BI Desktop



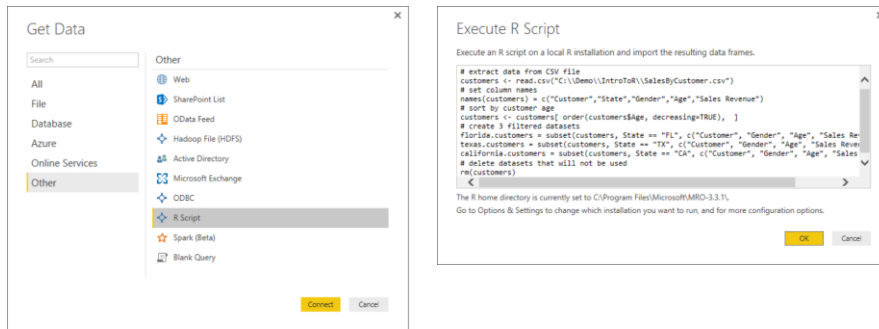
Where Can You Use R Code in PBIDT?

- As a data source to a query
 - You can use R code to import and reshape data
- Within a Query Applied Step
 - You can use R code to add transforms to a query
- Inside an R Visual in a Power BI Report
 - You can use R code to creates charts from your data



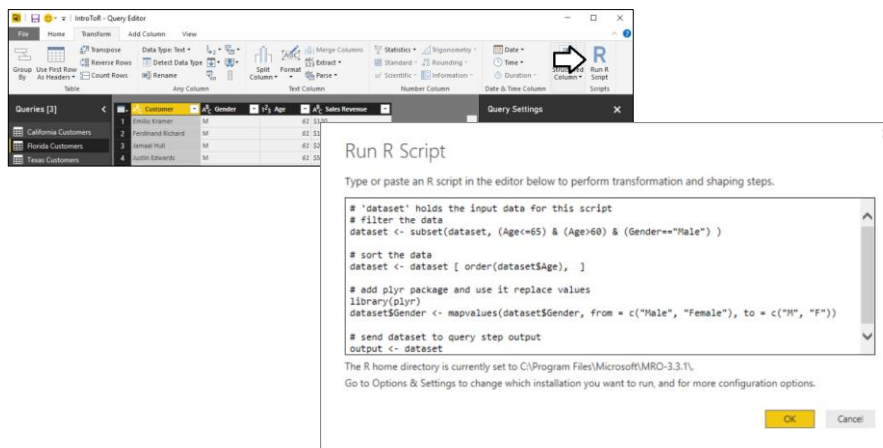
Using R Code as a Query Data Source

- Create new query based on R script
 - Copy and paste code from RStudio into PBIDT

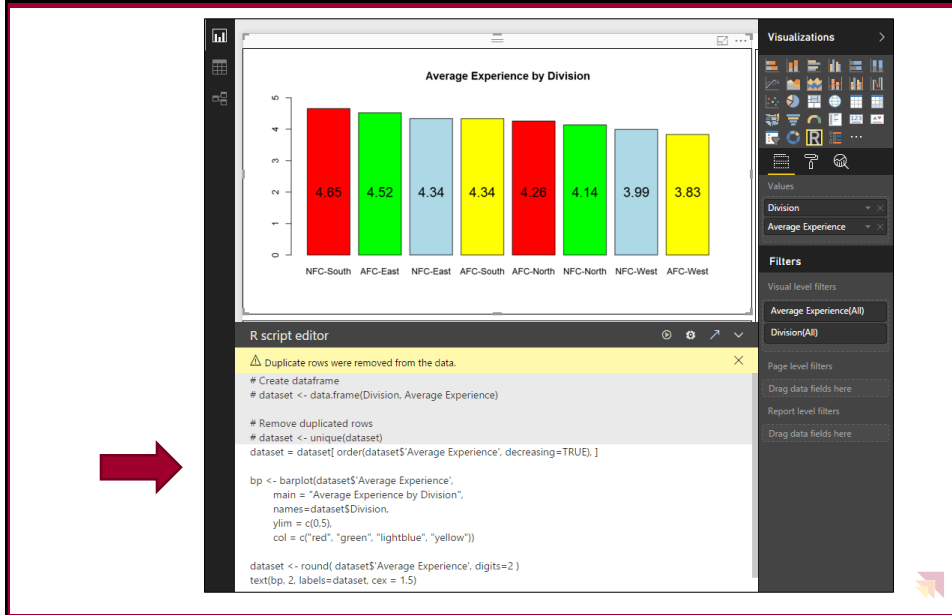


Using R Code as an Applied Query Step

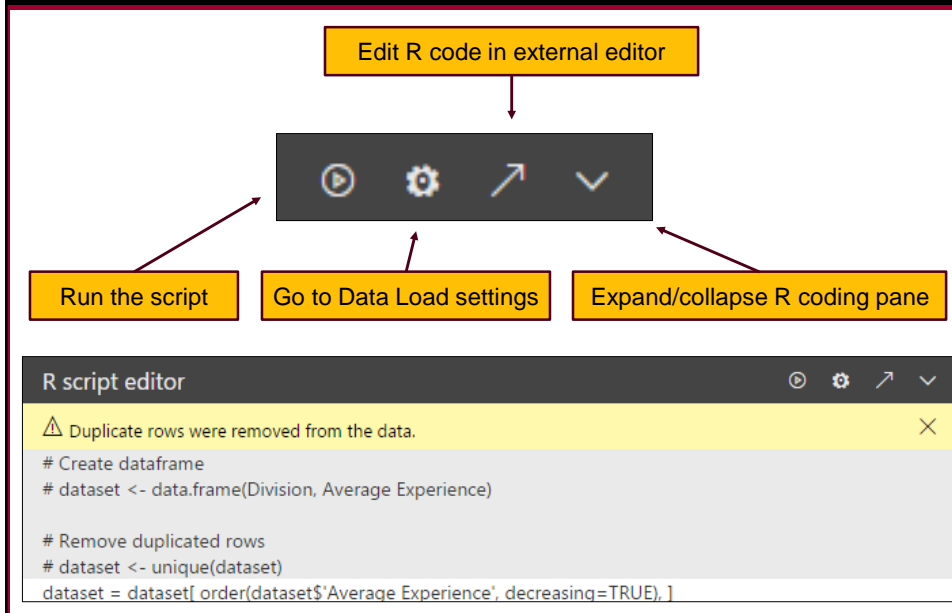
- Add new Run R Script step to query
 - Use R code and R packages to transform data



R Visuals in Power BI



New R Features in September Release



R Integration Limitations with Power BI

- Power BI Desktop R Limitations
 - Only data frames are imported
 - Complex columns and Vector columns are not imported
 - Values that are N/A are translated to NULL values
 - Prompting for user input halts script
 - R visual data for plotting is limited to 150,000 rows
 - R visual calculation times out with error after 5 minutes
 - R visual is not interactive – no highlighting support
 - Plots can only be displayed to R default display device



Summary

- ✓ Understanding R as an Analytics Platform
- ✓ Installing Microsoft R Open and RStudio
- ✓ Writing R Code in RStudio
- ✓ Integrating R with Power BI Desktop

- ❑ GitHub repository with code for this webinar

➤ <https://github.com/CriticalPathTraining/IntroToR>

- ❑ Check out critical Path Training

➤ <https://www.criticalpathtraining.com>

