



Introduction to R and Predictive Analysis with Power BI

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
- SharePoint MVP since 2004



Critical Path Training

- Advanced technical training on Power BI, SharePoint & others
- Offering **Power BI Bootcamp** as 3-day deep dive into Power BI

My Upcoming Power BI Training Events

- PBIUG Focus Event in Chicago on June 7-8
- Power BI Bootcamp in Tampa on Aug 29



Agenda

- What is R and why do you need it?
- Getting Started with R Studio
- World's Fastest R Language Primer
- Creating Visuals with R
- Working with R Visuals in 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
 - Libraries 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 optimize financial portfolios
 - Packages analyze genomic sequences
 - Packages predict component failure times



CRAN

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

The Comprehensive R Archive Network

Download and Install R

Precompiled binary distributions of the base system and contributed packages. **Windows and Mac** users most likely want one of these versions of R:

- [Download R for Linux](#)
- [Download R for \(Mac\) OS X](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your **Linux package management system** in addition to the link above.

Source Code for all Platforms

Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

Navigation links: [Home](#), [About R](#), [R Homepage](#), [The R Journal](#), [Software](#), [R Sources](#), [R Binaries](#), [Packages](#), [Other](#)

Mirror

Thailand	http://mirrors.psu.ac.th/pub/cran/	Prince of Songkla University, Hatyai
Turkey	http://cran.pau.edu.tr/	Pamukkale University, Denizli
	http://cran.uct.ac.za/pub/cran/	Middle East Technical University Northern Cyprus Campus, Mersin
UK	https://www.data.bris.ac.uk/R/	University of Bristol
	http://www.stats.bris.ac.uk/R/	University of Bristol
	https://mirrors.dti.ac.uk/CRAN/	EMBL-EBI (European Bioinformatics Institute)
	http://mirrors.ebi.ac.uk/CRAN/	EMBL-EBI (European Bioinformatics Institute)
	http://cran.msi.imperial.ac.uk/	Imperial College London
	http://mirror.mfu.ac.uk/R/	Medlinex University London
	http://mirrors.uct.ac.za/pub/cran/	St Andrews University
USA	https://cran.cnr.berkeley.edu/	University of California, Berkeley, CA
	http://cran.cnr.berkeley.edu/	University of California, Berkeley, CA
	http://cran.umd.ac.edu/	University of California, Los Angeles, CA
	http://mirror.lasalle.edu/CRAN/	Iowa State University, Ames, IA



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
- Current version of Microsoft Open R
 - Snapshot of CRAN as of April 1, 2016
 - Based on R Language version of R-3.2.4-Revised



Stages of R Awareness

- Standing Up
 - Installing the environment and playing with data
- Walking
 - Creating impressive graph and charts
- Jogging
 - Crunching numbers to generate advanced statistics
- Running
 - Creating a domain-specific predictive model
- Sprinting
 - Distributing your predictive model as a CRAN package



Agenda

- ✓ What is R and why do you need it?
- Getting Started with R Studio
 - World's Fastest R Language Primer
 - Creating Visuals with R
 - Working with R Visuals in Power BI Desktop



Install Microsoft R Open

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

MRAN About R Microsoft R Open Community Download

Download Microsoft R Open 3.2.5, the enhanced R distribution

Download Microsoft R Open 3.2.5 & MKL today. While the installation of MKL, which is used for multithreaded performance, is **optional**, we recommend you install both Microsoft R Open & MKL for optimal performance on Windows and Linux. Mac OS X users have Math Libraries by default. For a list of all system requirements, [click here](#).

[System Requirements & Install Docs](#) | [Forum](#) | [News File](#) | [Past Releases](#)

Microsoft R Open & MKL Downloads

This release of Microsoft R Open is supported on **64-bit platforms only**.

Note: CRAN R 3.2.5 was not released for OS X, so neither was Microsoft R Open 3.2.5. [Get v3.2.4 for the Mac](#).

Choose your download(s) of Microsoft R Open & MKL: [Verify Your Downloads](#)

64-Bit Platform	Software	Math Library
Windows 7.0 (SP1), 8.1, 10, Windows Server® 2008 R2 (SP1) and 2012	Microsoft R Open	MKL
Ubuntu 15.04	Microsoft R Open	MKL
Ubuntu 14.04	Microsoft R Open	MKL

Microsoft R Server Users:

- Get [R Open for Microsoft R Server 2016](#)
- Get [RRO 8.0.3](#) for RRE 7.4.1



Installing R Studio

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

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](#)

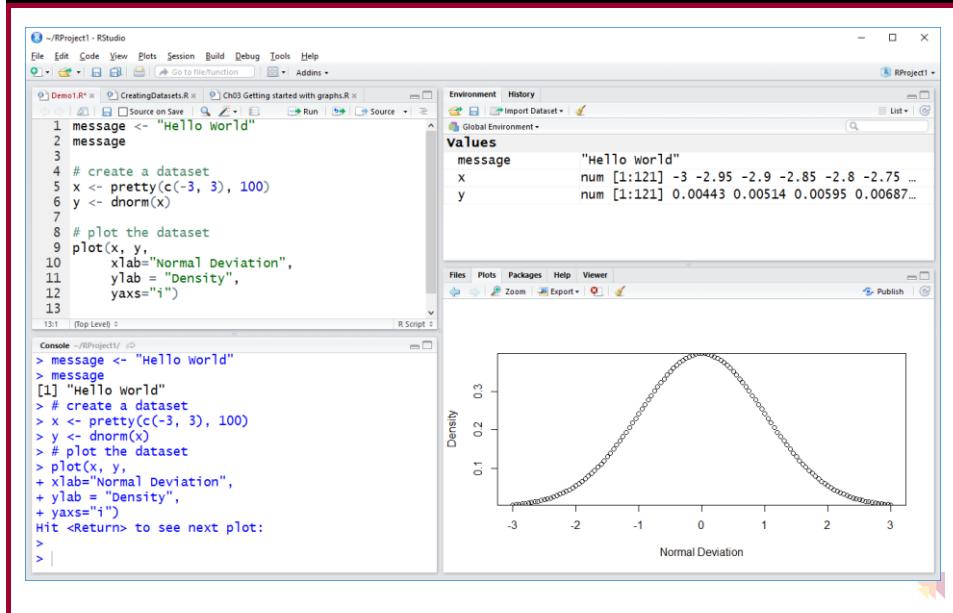
RStudio Desktop 0.99.902 — [Release Notes](#)

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	f741e9a1345985c16e692967adbad210
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	6f14d4717b01e7763d18f1c0ad8e6474
RStudio 0.99.902 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)	81.9 MB	2016-05-14	a79e80d992a497a92aafee8629f90e90

The RStudio IDE



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"
>
```





DEMO

Getting Started with R Studio

Agenda

- ✓ What is R and why do you need it?
- ✓ Getting Started with R Studio
- World's Fastest R Language Primer
- Creating Visuals with R
- Working with R Visuals in Power BI Desktop



Data Structures Supported in R

- Vector
 - One-dimensional, single-mode array
- Matrix
 - Two-dimensional, single-mode array
- Array
 - N-dimensional, single-mode array
- Factor
 - Integer-backed list of categorical values
- List
 - Ordered collection of multi-mode objects
- Data frame
 - Two-dimensional, multi-mode array
- Factor
 - Integer-backed list of categorical values (*e.g. enumerations in C#*)



DEMO

Working with Data Structures in R

Agenda

- ✓ What is R and why do you need it?
- ✓ Getting Started with R Studio
- ✓ World's Fastest R Language Primer
- Creating Visuals with R
 - Working with R Visuals in Power BI Desktop



R Graphics Packages

- There are three core graphics packages
 - graphics
 - lattice
 - ggplot2



Graphics Functions

- barplot – bar and column charts
- dotchart – Cleveland dot plots
- hist - Histogram
- density – Kernel density plot
- stripchart – Strip charts
- qqnorm – Quantile-quantile plots
- xplot – Scatter plots
- smoothScatter – Smooth scatter plots



DEMO

Creating Graphs using RStudio

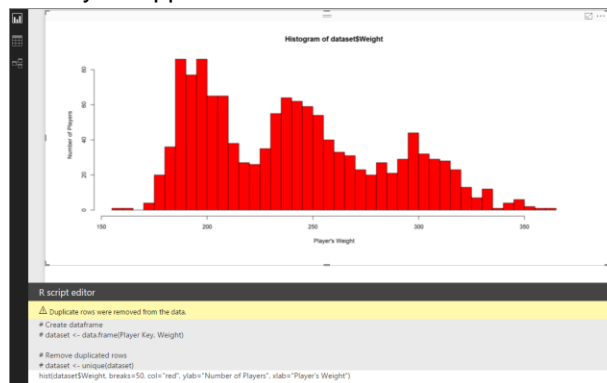
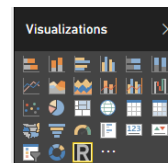
Agenda

- ✓ What is R and why do you need it?
- ✓ Getting Started with R Studio
- ✓ World's Fastest R Language Primer
- ✓ Creating Visuals with R
- Working with R Visuals in Power BI Desktop



R Visuals in Power BI

- Power BI Desktop supports R Visuals
 - Feature currently in preview
 - Allows you to create graphics from Power BI dataset
 - Feature not yet supported in Power BI service



Summary

- ✓ What is R and why do you need it?
- ✓ Getting Started with R Studio
- ✓ World's Fastest R Language Primer
- ✓ Creating Visuals with R
- ✓ Working with R Visuals in Power BI Desktop

