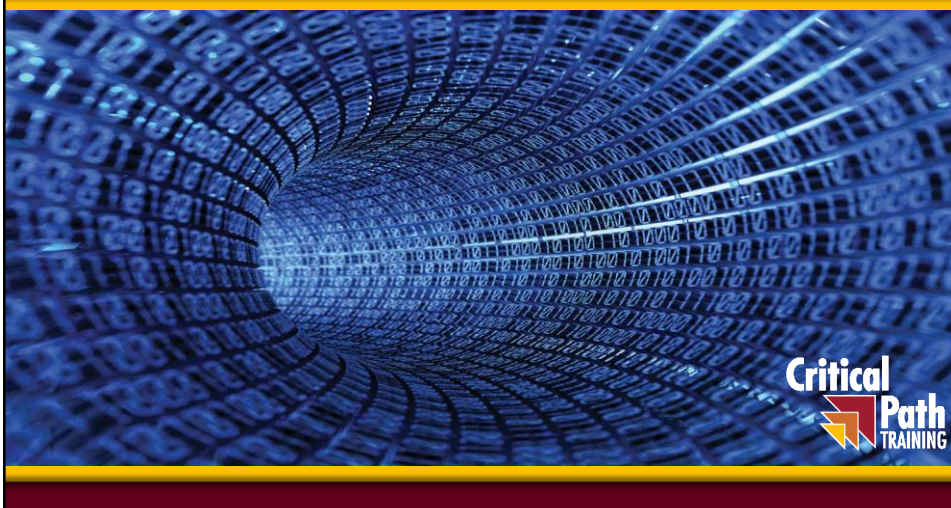
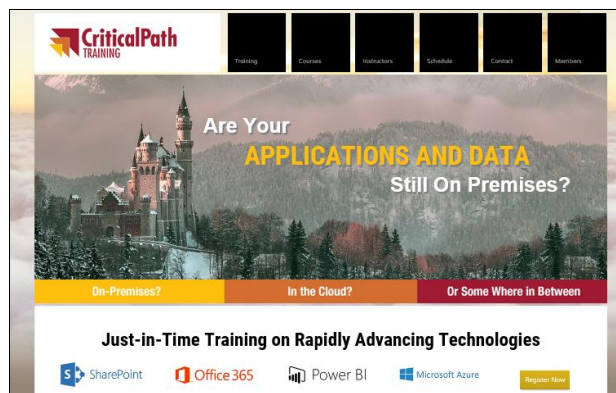


## Introduction to R with Power BI



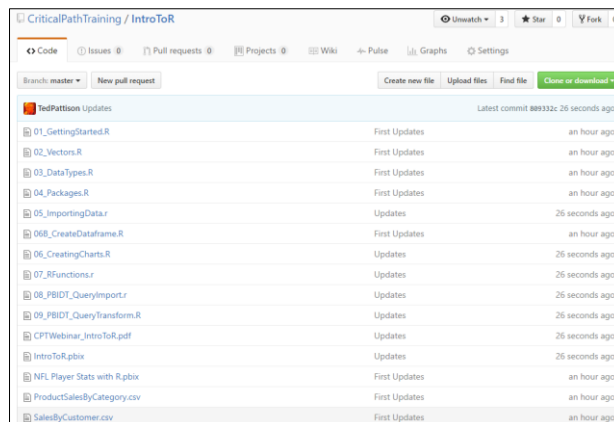
## Critical Path Training

- Upcoming Power BI Bootcamp Classes (4 days)
  - Nov 30<sup>th</sup> – attend in Tampa, FL or remote
  - Jan 23<sup>rd</sup> – attend in Tampa, FL or remote
- Visit us at <http://CriticalPathTraining.com> for more info



## Slides and Sample Code from this Webinar

- Webinar materials are available in GitHub
- <https://github.com/CriticalPathTraining/IntroToR>



## Agenda

- Understanding R as an Analytics Platform
- Installing the 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

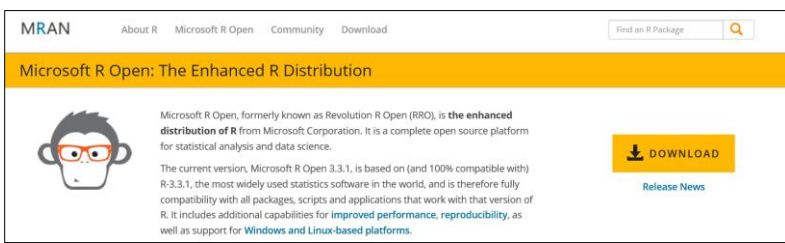
- The Comprehensive R Archive Network
  - Public archive with over 8,000 downloadable packages
  - <http://cran.us.r-project.org/>



The screenshot shows the CRAN website. On the left, there is a sidebar with links: 'CRAN Mirrors', 'Task Views', 'Search', 'About R', 'R Homepage', 'The R Journal', 'Software', 'R Sources', 'R Binaries', 'Packages', and 'Other'. A red arrow points from the 'CRAN Mirrors' link to a list of mirrors. The list includes mirrors for Thailand, Turkey, UK, and USA, with links to their respective CRAN mirrors. On the right, the 'Download and Install R' section provides instructions for downloading and installing R, including links for Linux, Mac OS X, and Windows.

## 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 website. The header includes 'MRAN' and navigation links: 'About R', 'Microsoft R Open', 'Community', and 'Download'. A search bar is also present. The main section is titled 'Microsoft R Open: The Enhanced R Distribution'. It features a cartoon monkey wearing glasses and a description of Microsoft R Open, formerly known as Revolution R Open (RRO), as the enhanced distribution of R from Microsoft Corporation. It mentions that the current version, Microsoft R Open 3.3.1, is based on (and 100% compatible with) R-3.3.1, the most widely used statistics software in the world, and is therefore fully compatible with all packages, scripts and applications that work with that version of R. It also includes additional capabilities for improved performance, reproducibility, as well as support for Windows and Linux-based platforms. A 'DOWNLOAD' button and a 'Release News' link are visible.

## 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

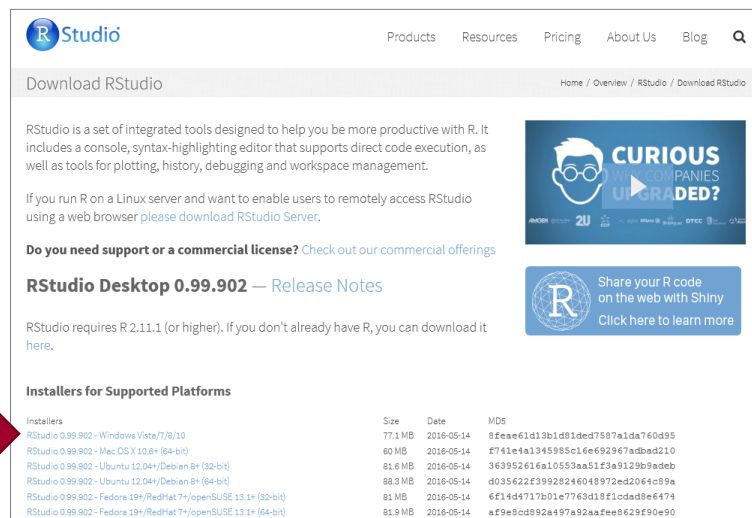
- ✓ Understanding R as an Analytics Platform
- Installing the Microsoft R Open and RStudio
  - Writing R Code in RStudio
  - Integrating R with Power BI Desktop



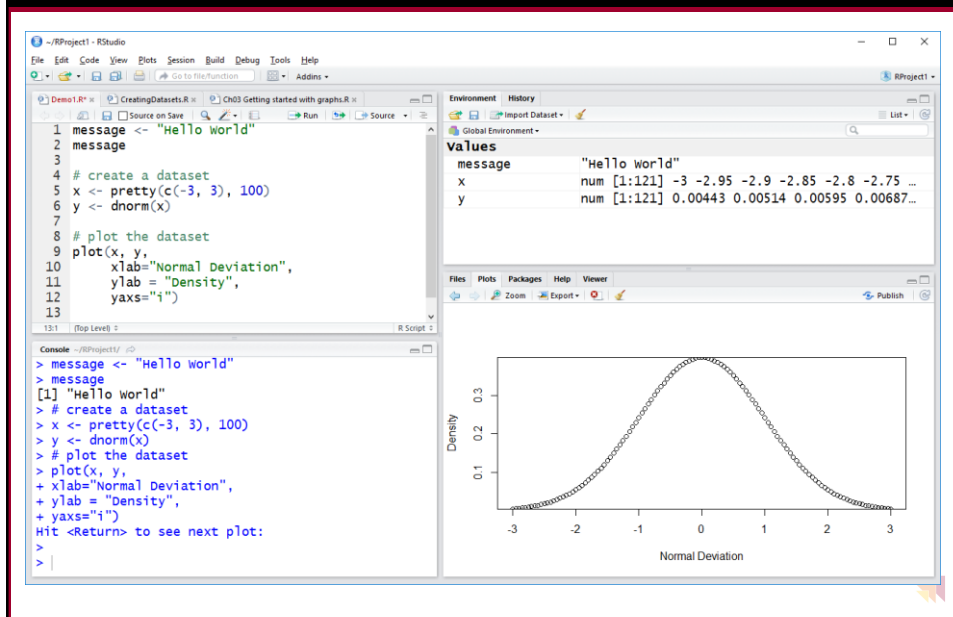
- <https://mran.microsoft.com/download/>



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



## The RStudio IDE



## Agenda

- ✓ Understanding R as an Analytics Platform
- ✓ Installing the 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

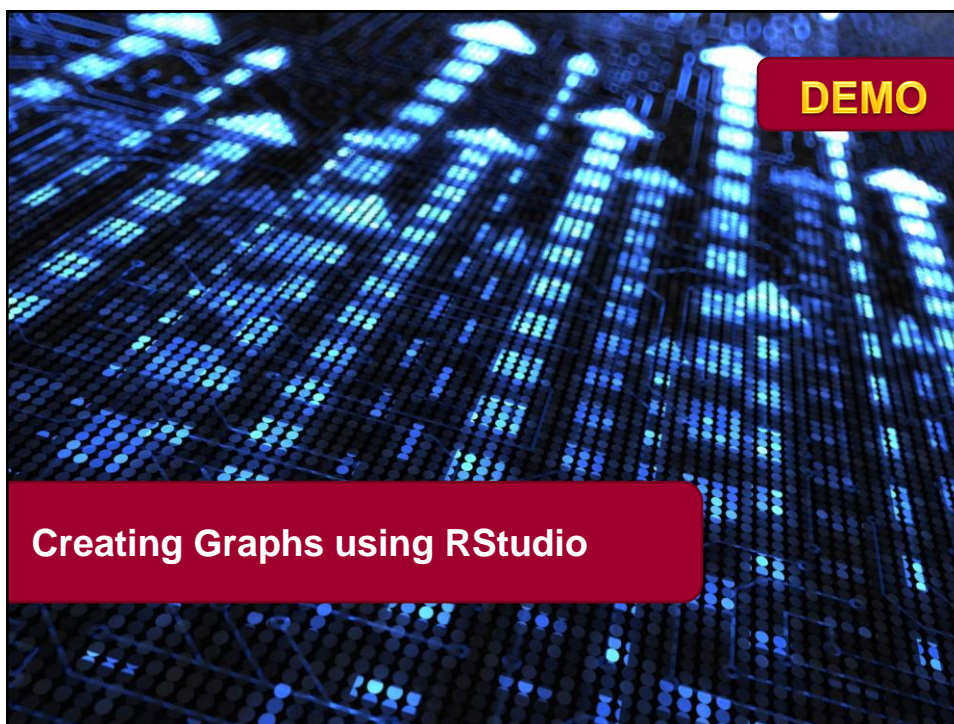
- In R, variables represent named objects
- Object names can contain
  - Letters
  - Numbers
  - Underscores (`_`)
  - Dots (`.`)



## 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 the Microsoft R Open and RStudio
- ✓ Writing R Code in RStudio
- Integrating R with Power BI Desktop



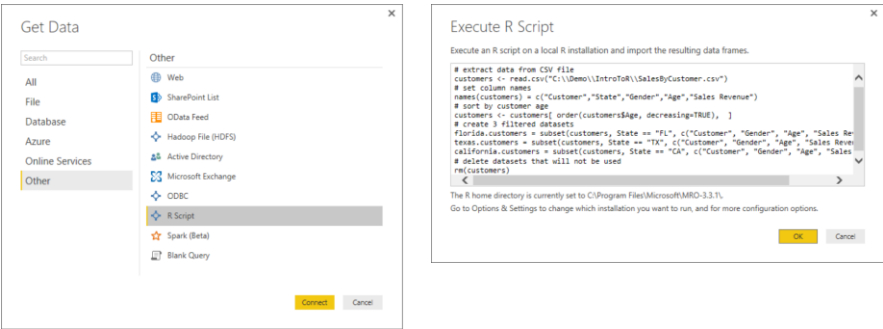
## Where Can You Use R Code in PBIDT?

- 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



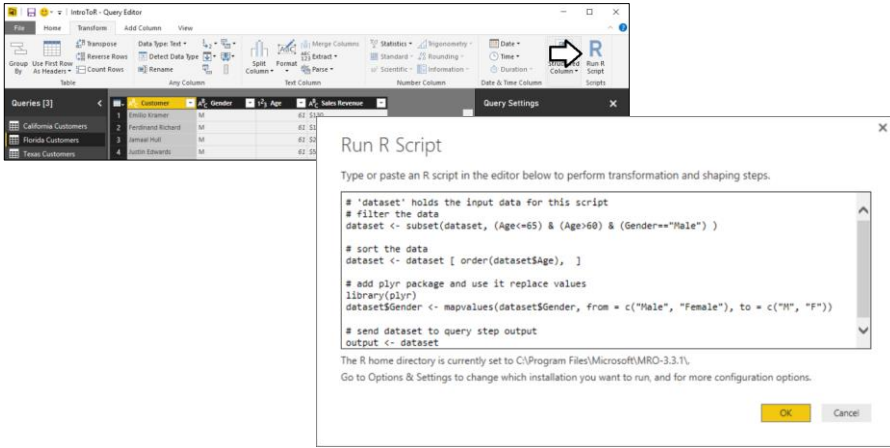
## Using R Code as a Query Data Source

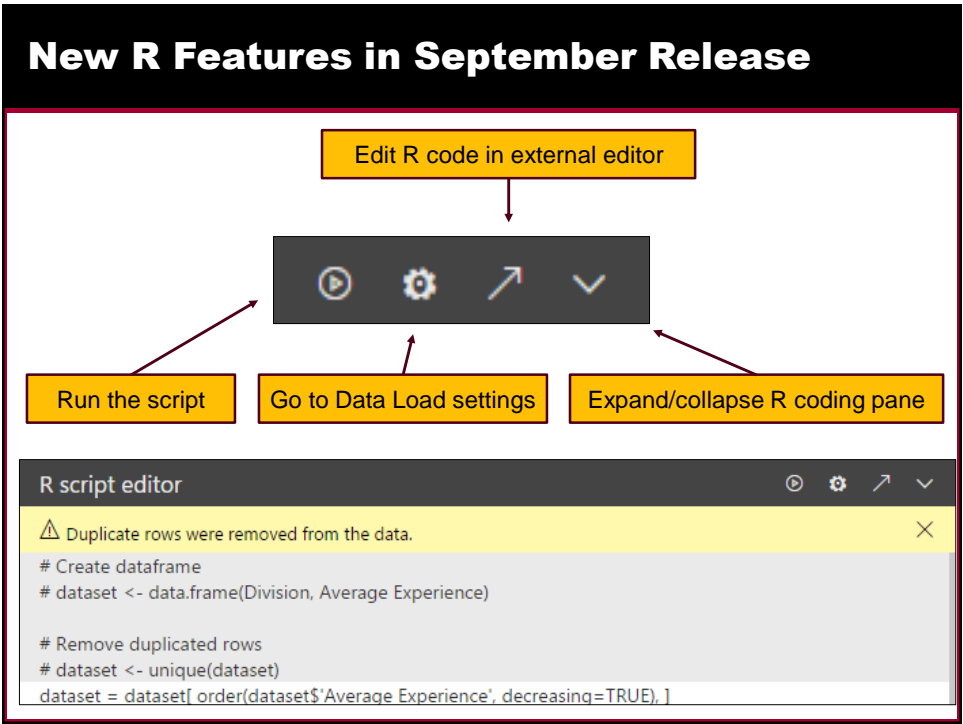
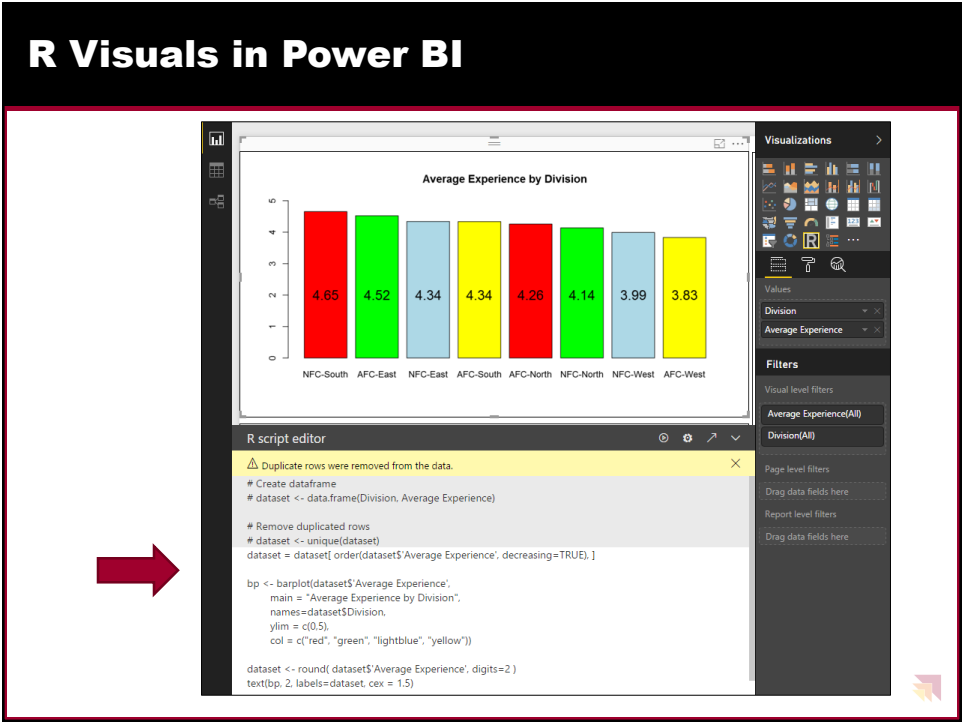
- Create new query based on R script
  - Copy and paste code from RStudio into PBIDT



## 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 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



## Power BI Team Blog

- Power BI Team Blog is an Essential Resource

<https://powerbi.microsoft.com/en-us/blog/>

- Be on the lookout for monthly updates



## Power BI Bootcamp

- Upcoming Power BI Bootcamp Classes (4 days)
  - Nov 30<sup>th</sup> – attend in Tampa, FL or remote
  - Jan 23<sup>rd</sup> – attend in Tampa, FL or remote
  - Visit us at <http://CriticalPathTraining.com> for more info
- Drilldown into Essential Topics & Skills for Data Professionals

+	Module 01: Introducing the Power BI Platform
+	Module 02: Getting Started with Power BI Desktop
+	Module 03: Using the Query Features of Power BI Desktop
+	Module 04: Using the Data Modeling Features of Power BI Desktop
+	Module 05: Modeling Data with Dimensional Hierarchies and Time Intelligence
+	Module 06: Designing Interactive Reports in Power BI Desktop
+	Module 07: Designing and Deploying Dashboards
+	Module 08: Designing Dashboards for Mobile Devices
+	Module 09: Managing Power BI Gateways and On-premises Data
+	Module 10: Integrating Power BI with Excel 2016 and Excel Online
+	Module 11: Integrating Power BI with SQL Server 2016
+	Module 12: Introduction to R and Predictive Analysis with Power BI



## Summary

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

Get the slides and sample R code for this webinar  
<https://github.com/CriticalPathTraining/IntroToR>

