Getting the Most Out of RStudio

Josh Paulson

Product Manager
josh@rstudio.org

June 26, 2012

Setup For Day 2

- Confirming everyone is setup
- Install the knitr package (v0.5+)

  https://github.com/yihui/knitr
- Install TeX

  http://www.latex-project.org/ftp.html
- Obtain tutorial files:
  - GitHub - Download repository as Zip file

    https://github.com/jwpaulson/rstudio-training
  - USB Drive - Ask Josh

Overview - Day 2

- Customized Reports and Presentations (Sweave/knitr)
- Interactive R Documents (R Markdown/Notebook)
- Publishing Code, Analysis, and Results (RPubs)
- Debugging R Code
- Troubleshooting and Getting Help with R
- Best Practices for R Programming
- Introduction to RStudio Server

Reproducible Research

- Current workflow and what's the problem?
  - Selection and Response Interface
  - Copy and Paste to MS Word
  - Publish Results
- How does this lead to problems?
- John Chamber's Prime Directive:

  "Program in such a way that the computations can be understood and trusted."

Reproducible Research

- LaTeX is a Reproducible Document Format
- Sweave combines LaTeX and R Code:
  (See "sweave/demo.pdf" example)
- Couple Code with Analysis and Results
- Create Reusable and Repeatable Research
- Improved Transparency of Results and Publications
- Trustworthy Analysis That is Fully Inspectable
- LaTeX Help:
  http://www.stdout.org/~winston/latex/

Sweave/knitr and RStudio

- Step Through Your R Code
- Auto-completion for Chunk Options
- Chunk Navigation and Syntax Highlighting
- One-Click Compile PDF
- Error Navigation
- SyncTeX and Concordance
- Spell Check Your Document
- RStudio runs Sweave in a separate process for reproducibility

Hands On

- Creating an Sweave File
- http://rpubs.com/josh/activity-7
- Insert R Code Chunks
- Customize Chunk Options
- Compile to PDF
- Use SyncTeX and Other Features
- Ask Questions

Beamer Presentations

- You can also make presentations with Sweave
- Beamer is a specific class of LaTeX.
- Open file: "beamer/beamer.pdf"
  Look famiilar? These slides were created with beamer
- Now let's look at the source code. Open file:
  "beamer/beamer.tex"
- Beamer Help:
  http://www.math.umbc.edu/~rouben/beamer/
  https://bitbucket.org/rivanvx/beamer/wiki/Home

Using knitr

- See "sweave/knit-minimal.pdf" Example
- Differences between Sweave and knitr
- New Chunk Options
- http://yihui.name/knitr/options
- Customizing Figures

Hands On

- Switching to knitr
- http://rpubs.com/josh/activity-8
- New Chunk Options
- Compile to PDF
- Ask Questions

Create an HTML Notebook

- knitr allows us to do a lot more in RStudio
- Wrap R Script into a Sharable HTML File
- Quickly print and share your code and its output
- Publish online or share on an internal network

Hands On

- Create an HTML Notebook
- http://rpubs.com/josh/activity-9
- Turn R Scripts to an HTML Notebook
- Note: knitr Package (version 0.5 or later) Required
- Ask Questions

R Markdown

- Reproducible Research Without the TeX-pertise
- Accessible Syntax (Markdown) for Easy Report Generation
- Interweave R Code and Contextual Analysis
- Standalone HTML File - Easy to share
- Support for Dynamic and Interactive Content

R Markdown with RStudio

- Learning Markdown:
  - Default R Markdown Template
  - Markdown Quick Reference (Click **MD** Button)
- Insert R Code Chunk
- Customize Chunk Options
- Run R Code and Evaluate Chunks
- Chunk Navigation
- Knit HTML

Hands On

- Create an R Markdown File
- http://rpubs.com/josh/activity-10
- Insert an R Code Chunk, Commands, and Plots
- Customize Chunk Options (figures, etc.)
- Knit HTML
- http://yihui.name/knitr/options
- Ask Questions

Publishing Over the Web with RPubs

- Easy Web Publishing From R
- One Click to Share Results and Analysis:
  - Ad-Hoc Sharing of Work
  - Classroom Environment (Lectures and Homework)
  - More Examples at RPubs website
- Free for Public Documents
- Check it out: https://rpubs.com/

Hands On

- Publish Your First Document to RPubs
- http://rpubs.com/josh/activity-11
- Use R Markdown to Publish
- Make Changes and Republish
- Publish a HTML Notebook
- Note: knitr Package (version 0.5 or later) Required
- Ask Questions

Debugging R Code

- `traceback()` - identifies the errror (prints call stack)
- `browser()` - pauses R to allow "interactive" debugging
- `debug()` - Stats debugger with browser() statements for an expression
- `options(error = recover)` - automatically starts debugger on error
- `options(warn = 2)` - converts all warnings into errors (easier identification)

Debugging R Code

- Advantage of RStudio
- Utalize workspace with `browser()`
- Code Navigation
- View the Source with **Go to Function Definition**
- Future Interactive Debugger?

Hands On

- Debugging your R code
- http://rpubs.com/josh/activity-12
- Use features of RStudio for easier debugging
- Ask Questions

Profile your code

- Allows you to see what is called and how long it took
- Use `Rprof()`?

  > Rprof("profile-qplot.txt")

  > qplot(carat, price, color=clarity, data=diamonds)

  > Rprof(NULL)
- Better from profr package

  > library(profr)

  > p ¡- profr(print(qplot(carat, price, color=clarity, data=diamonds)))

Troubleshooting and Getting Help with R

- Use Code Completion to Mitigate Errors
- Learning a New Package
- Press **F1** for Function Help
- Press **F2** to View Function Definition
- Advantage of Side-By-Side Help
- Inspect Source Code and View Function Definitions

Everything in R is an Object

- Variables, Data, Functions, R expressions, etc.
- show object "x" as a variable and function
- Note that objects are copied in assignment statements

```
> x <- 1:10
> y <- x
> x <- 10 %% 3
> x
> y
```

Problems with Graphics

- Example:

  > *source("plots/chart.R")*
  > *plot(cars)*

- Likely leftover graphics settings/options
- Use `dev.off()` to reset graphics device

Problems with Graphics

- Difference between what see and when save
- Example:

```
> png()
> plot(cars)
> dev.off()
```

- Use the export plot feature to finalize

De-Mystifying Errors

- R returns a + ... now what?
- Missing a parentheses or quote?
- Press **Esc**
- Error: No such file or directory
- Problem with path or working directory?
- see current working directory in title bar
  > getwd()

De-Mystifying Errors

- Likely a bug in the code - look at what you executed
- Check your working directory
- Check your version of R and packages
- 32-bit vs 64-bit problem?
- Permissions problem?
- Network problem - proxies, etc.?
- Try on another machine

De-Mystifying Errors - Resources

- There are lots of online resources to help you:
- RSeek meta search engine:
  http://www.rseek.org/
- Stack Overflow - R Tag:
  http://stackoverflow.com/questions/tagged/r
- R-help mailing list:
  https://stat.ethz.ch/mailman/listinfo/r-help
- R-help list archives:
  http://tolstoy.newcastle.edu.au/R/

Read Error Messages

ERROR: dependency parser is not available for package formatR
** removing /home/josh/R/library/formatR*
Warning in install.packages :
installation of package formatR had non-zero exit status
ERROR: dependencies formatR, highlight are not available for package knitr
** removing /home/josh/R/library/knitr*
Warning in install.packages :
installation of package knitr had non-zero exit status

RStudio Problems

- Is this an R Problem or RStudio Problem?
- RStudio is only an interface for R, it's not R
- R session failed to start?
- Version, Startup files, Can't find R?
  Firewall and proxy settings?

  http://support.rstudio.org/help/kb/troubleshooting/rstudio-will-not-start
- log files: Help Menu | Diagnostics | Show Log Files
  http://support.rstudio.org/help/kb/troubleshooting/rstudio-application-logs

RStudio Problems

- Reset RStudio's state?

  http://support.rstudio.org/help/kb/troubleshooting/resetting-rstudios-state

- Report it to us:

  http://support.rstudio.org/

- R version and other information:

  `> sessionInfo()`

- Version of RStudio

  `> RStudio.version()`

RStudio Daily Version

- We're constantly updating RStudio
- New Features
- We fix bugs and they become in the daily
- http://www.rstudio.org/download/daily/
- Upgrading and reverting back to an older version

Best Practices

- Don't use `attach()` for data
- Think reproducibility
- Example:

  > cars$speed
  > speed
  > attach(cars)
  > speed
  > detach(cars)

- Ambiguous, not sharable, and hard to look back at code

Best Practices

- With knitr, cache results for longer computations
- Example:

```
> fibo ¡- function(n) {
+ if ( n ¡ 2 ) n
+ else fibo(n-1) + fibo(n-2)
+ }
> for(i in 0:28) fibo(i)
```

Best Practices

- Startup files - see `?Startup`
- .Rprofile() - runs R script at beginning of session
- Rprofile.site - global profile for all users
- .Renviron() - defines environemnt variables

Best Practices

- Speed can be important
- Breakup scripts into functions
- Use vectorized operations not loops
- Are you making R slow (too many copies)?
  Example: ggplot2 package
  > library(ggplot2)
  > qplot(carat, price, color=clarity, data=diamonds)

Working on the Web

- RStudio also runs over the web (via RStudio Server)
- Why run over the web?
  - Multiple users can share single high-capacity Linux server
  - Seamlessly switch computers and preserve full session context
  - Collaboration (easy sharing of code, data, and configuration)
- RStudio's user interface is built using JavaScript so the desktop and web experiences are identical
- Sync between desktop and web with version control

RStudio Server

- Requires Linux
- Download and Install
  - `$ wget http://download2.rstudio.org/rstudio-server-0.96.304-amd64.deb`
  - `$ sudo dpkg -i rstudio-server-0.96.304-amd64.deb`
- Accessing the Server
  - `http://<server-ip>:8787`
- Configuration and Management
  - http://www.rstudio.org/docs/server/getting_started

Recap and Review

- Reproducible Research with Sweave and knitr
- R Markdown for easier/flexible reports
- Publish and share results on RPubs
- Debugging your R code
- Troubleshooting
- Best Practices
- RStudio Server

Thank You

- Feedback and Questions
- See more on our website:
  - http://www.rstudio.org
- Contact me:
  - josh@rstudio.org