review

From Data to Insight

Dr. Çetinkaya-Rundel

July 11, 2016

Terminology

- R: statistical programming language
- RStudio: front-end software for R that allows you to organize your files and plot, keeps a history of your command, and provides an environment for creating reports with R Markdown
    - It's much more than that, but for our purposes, this should be a sufficient definition
- R Markdown: Authoring format for dynamic documents including your R code and your write-up

R Markdown

- R code goes in chunks, marked by three backticks and the letter `r` in curly braces to begin and three backticks to end
    - Within a chunk **#** is used to mark a comment, any text following this sign on the same line will not get processed as code.
- Interpretations, i.e. your write-up “in English”, goes outside of R chunks

review.Rmd

Knit HTML | Run

```
---
title: "review"
author: "mine"
date: "7/11/2016"
output: html_document
---

Write-up goes outside the chunk

```{r}
# R code goes inside the chunk
x <- 2
x * 3
```
```

← input

Files | Plots | Packages | Help | Viewer

Publish

review

mine

7/11/2016

Write-up goes outside the chunk

```
# R code goes inside the chunk
x <- 2
x * 3
```

```
## [1] 6
```

← output

Independent environments

- Your Console uses one working environment
- Your R Markdown document uses a different (independent) working environment
- If you define an object in your Console, but do not define it in your R Markdown document, you will get an error when you try to knit your document saying that the object is not found

Console | R Markdown

```
Loading required package: grDevices
Loading required package: stats
Loading required package: utils
Please visit openintro.org for free statistics materials
> x <- 2
> x
[1] 2
>
```

review.Rmd

Knit HTML | Run

```
1  ---
2  title: "review"
3  author: "mine"
4  date: "7/11/2016"
5  output: html_document
6  ---
7
8  ```{r}
9  x * 3
10 ```
```

Console | R Markdown

~/review.Rmd

```
processing file: review.Rmd
Quitting from lines 9-10 (review.Rmd)
Error in eval(expr, envir, enclos) : object 'x' not found
Calls: <Anonymous> ... handle -> withCallingHandlers -> withVisible -> eval ->
 eval
Execution halted
```

Deciphering errors

- This is a skill you'll develop over time, so do not get discouraged if initially the errors seem too cryptic

- Approach deciphering what the error is saying methodically — you don't need to understand everything printed in the error to figure out what the issue is

    - First see which line of code is causing the error, noting that the error will point you to the first line of the R chunk

    - Go to that chunk to see if you can figure out what the issue is (maybe spelling error?)

    - Read the error further to see if there are other clues like "object not found" or "could not find function" etc.

Common erros in code

- Spelling!
  - Spelling of objects you create as well as spelling of functions
- Non-matching parantheses and quotation marks

`ggplot2` (+)
`dplyr` (%>%)

- **`ggplot2`**: Package we are using for plotting
  - Plots are comprised of layers
  - Layers are separated by **+**
  - Stylistic requirement: End lines of **`ggplot2`** code with **+**, move to the next line for the next layer

- **`dplyr`**: Package we are using for data wrangling
  - Pipes is comprised of chains
  - Lines of chains are separated by **`%>%`**
  - You read a pipe as take the output of the preceding line and use it as the first argument of the next line
  - Stylistic requirement: End lines of **`dplyr`** code with **`%>%`**, move to the next line for the next step in the chain