2017-01-23-intro-rstudio-rmarkdown

Rick O. Gilmore 2017-01-06

Today's topics

- Introduction to RStudio
- Introduction to R Markdown

RStudio

• An integrated development environment (IDE) for R

Components of RStudio

- Code editor
- R console
- Integrated help
- Image viewer
- Integration with git and other version-control packages
- Project management

RStudio in the cloud

- RStudio can be run in a browser from a server running RStudio Server
- Demo
 - Running this under Amazon Web Services (AWS) free tier
- Instructions for doing this yourself can be found here

R Markdown

- Markdown
 - Mark-up language to make it easy to write HTML
- R Markdown special type of Markdown
 - Allows for "literate" programming, mixing text, analysis, figures
 - Adds to Markdown syntax

Markdown syntax

- Text formatting
 - italics by surrounding text with single asterisks or underscores: *italics* or _italics_
 - boldface by surrounding text with double asterisks or underscores: **boldface** or __boldface__
 - strikethrough by surrounding text with double tildes: ~~strikethrough~~
 - Clickable URLs by surrounding link text with square brackets and URL with parentheses: [Clickable URLs](http://www.psu.edu)

Markdown syntax

- Paragraph formatting
 - Headings with level specified by the number of hash (#) marks
 - Lists (bullet and enumerated)
 - Block quotes
 - Code blocks

```
# This is a Heading 1
## This is a Heading 2
### This is a Heading 3
```

I put these in quotes here because these hash marks also separate different slides (more on this later).

- An item
 - A nested item
 - * A doubly-nested item
- Another item

Code:

- An item
 - A nested item
 - A doubly-nested item
- Another item

- 1. An enumerated item
 - A nested item
- 2. A second enumerated item

Code:

- 1. An enumerated item
 - A nested item
- 1. A second enumerated item

Notice how the numbers are incremented automatically!

Four score and seven years ago, some famous President spoke infamous words that would live on throughout history. These words are famous enough that I want to highlight them with a block quote.

- > Four score and seven years ago, some famous President
- > spoke infamous words that would live on throughout history.
- > These words are famous enough that I want to highlight them with a block quote.

More on Markdown syntax

- Images can be inserted using this syntax ![Alt text](/path/to/img.jpg)
- Clickable links using http://www.psu.edu
- Comments won't print in rendered output <!- This is a comment ->

R Markdown additions

- .Rmd extension
- Combine text, code, images, figures, video
- "Computable" reports, documents, slide shows, notebooks
- Output in multiple formats from the same file (next week)

Make some data

```
x = rnorm(n = 100, mean = 0, sd = 1) # N(0,1)

y = rnorm(n = 100, mean = 2, sd = 0.5) # N(2, 0.5)
```

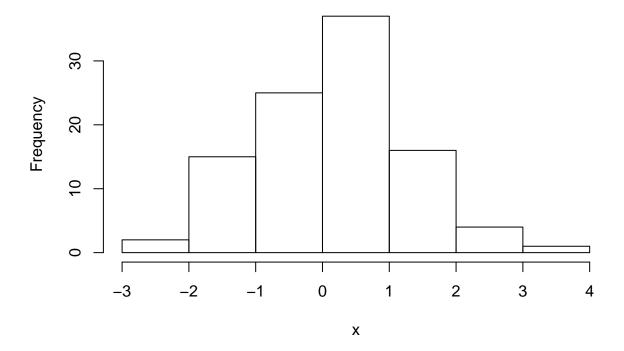
Summary of x, y

```
summary(x)
     Min. 1st Qu. Median
                             Mean 3rd Qu.
                                             Max.
## -2.3860 -0.6766 0.2201 0.1499 0.9344 3.5480
summary(y)
##
     Min. 1st Qu. Median
                             Mean 3rd Qu.
                                             Max.
##
     1.105
            1.769
                    1.985
                            2.030
                                  2.224
                                            4.048
```

Histogram of x

```
hist(x)
```

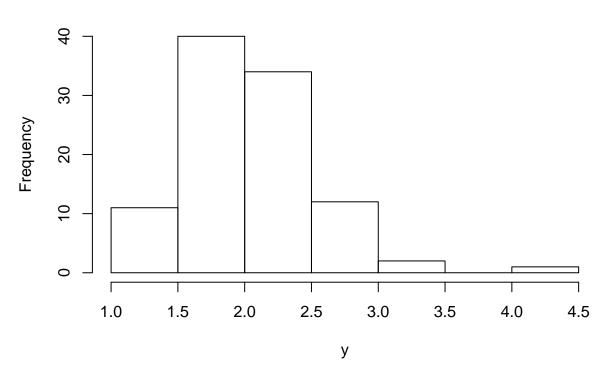
Histogram of x



Histogram of y

hist(y)

Histogram of y



Embed figure saved locally

```
<img src="../img/my-img.jpg" height=500px>
height parameter is optional, but useful.
```

Embed figure from the web

```
<img src="http://cdn.abclocal.go.com/content/wpvi/images/cms/280434_1280x720.jpg" width=900px>
```

Embed YouTube video

• YouTube gives you code to cut and paste.

Printing computed variables

```
summ.x = summary(x)
summ.y = summary(y)
names(summ.x) # Figure out variable names for indexing
## [1] "Min." "1st Qu." "Median" "Mean" "3rd Qu." "Max."
```

Index by variable name: X lies within the range of [-2.386, 3.548].

Index by numeric index: The (y-x) difference in means is 1.8801.

Calculate and report: The correlation between x and y is 0.0535384.