

Master Tutorial Introduction to Reproducible Research using R, RStudio, and R Markdown

Rob Stilson

April 16, 2016

Introduction

- What is Reproducible Research and why is it important?
- **Learning Objective #1:** Using R, RStudio, and R Markdown to setup a document for Reproducible Research
- **Learning Objective #2:** Publishing your Reproducible Research Document in HTML, Word, or PDF
- **Learning Objective #3:** Updating the data used in your document
- **Learning Objective #4:** Publishing your findings on the web via RPubS

What is Reproducible Research? FirstPerson.

- Fully documenting the steps that were taken to conduct a research study
- This enables:
 - Replication of the study's findings
 - Understanding exactly what analyses were conducted
 - Updating an analysis efficiently.
- Full transparency of the data and the process

Areas RR is applicable

FurstPerson.

- For Academics
 - Studies
 - Theses and Dissertations
- For Practitioners
 - Tech Reports
 - White Papers
 - Tech Manuals
 - Assessment Reports

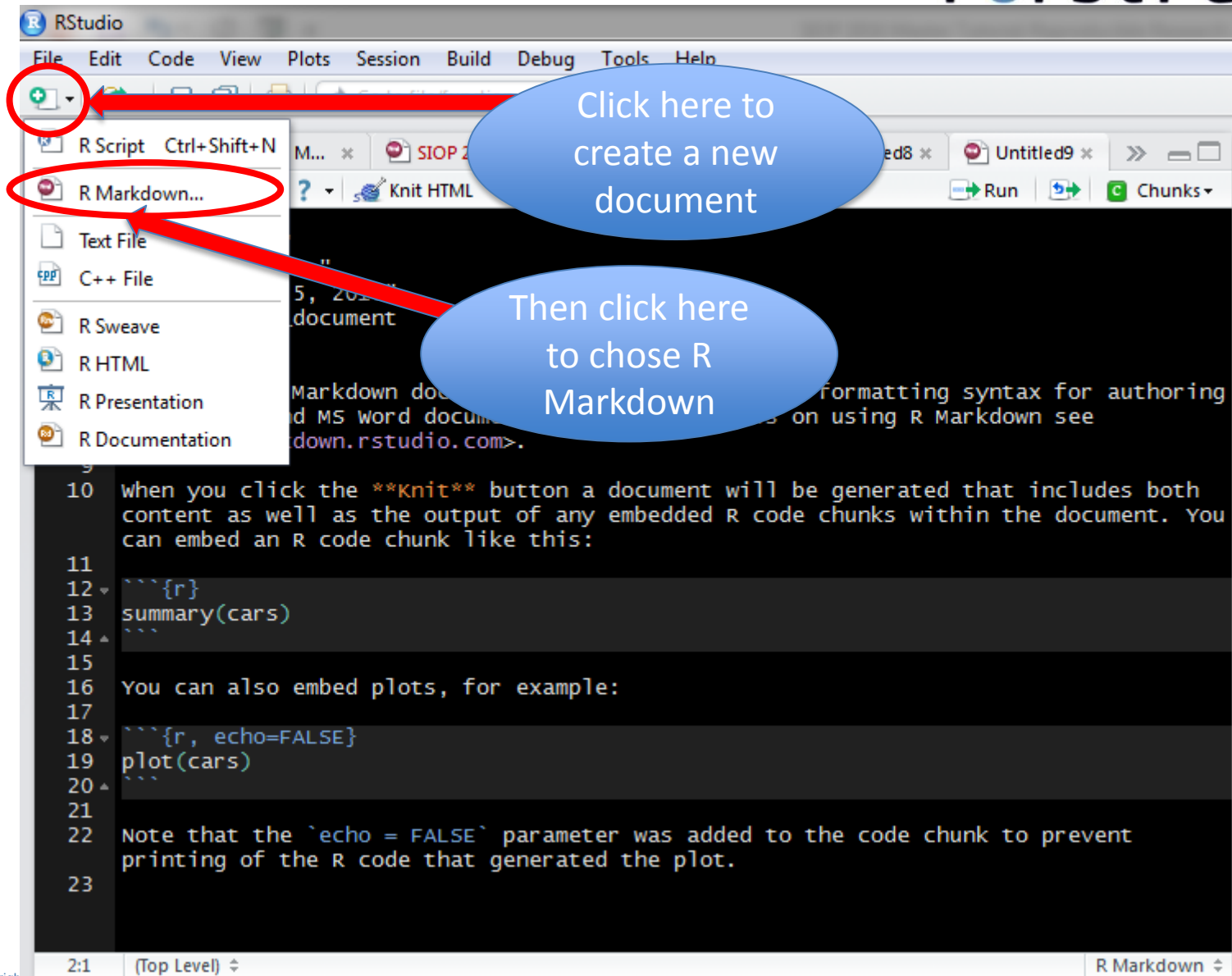
Things to Remember

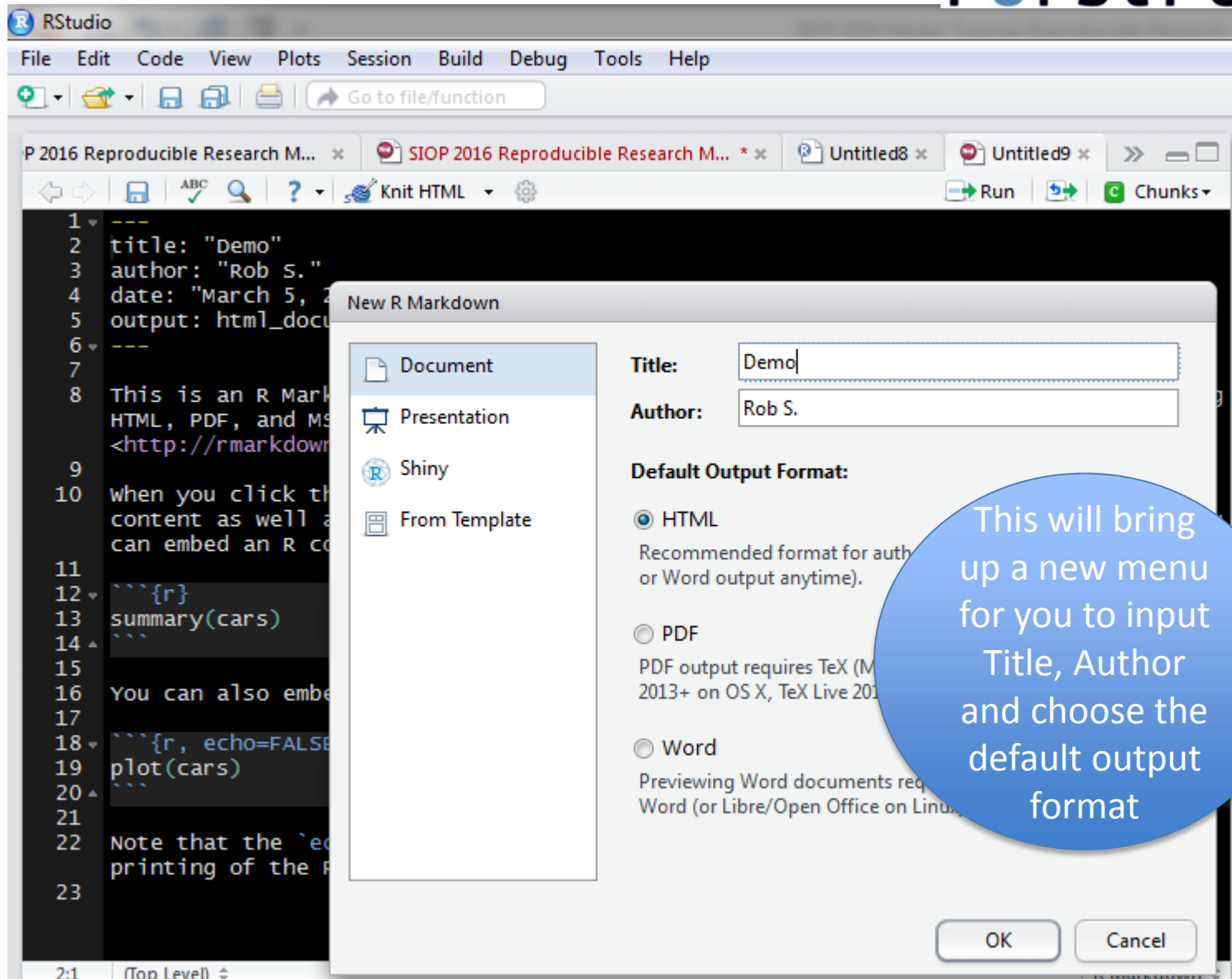
FurstPerson.

- This Master Tutorial is NOT about learning R
 - If you are new to R, this tutorial will give you an incentive to learn it
 - A lot of code (don't worry, all materials available at the end)
- This Master Tutorial IS about Reproducible Research Documents
 - Will be demonstrated in R
 - Markdown can also be used with
 - Python
 - JavaScript
 - Many others

Learning Objective #1: Using R, RStudio, and R Markdown to setup a document for Reproducible Research

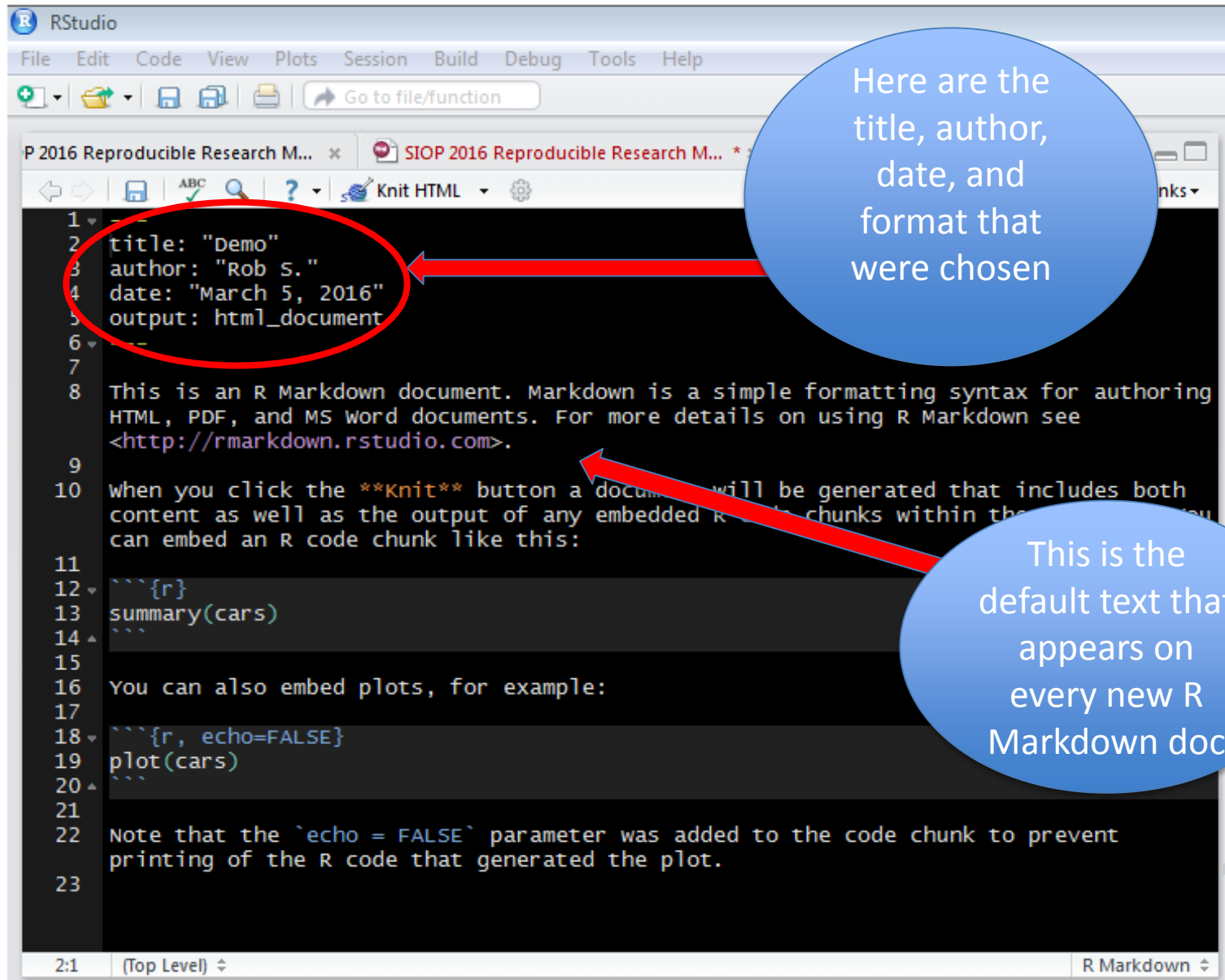
- Download from <https://www.rstudio.com/products/RStudio/#Desktop>
- R Studio is an Integrated Development Environment (IDE) for R
- Includes integration with R Markdown
- Additionally, it facilitates
 - Importing data
 - Installing packages
 - Updating packages
 - Syntax highlighting, code completion, and smart indentation
 - Quickly jumping to function definitions
 - Many other areas R





R Studio-R Markdown

FurstPerson.



The screenshot shows the RStudio interface with an R Markdown document open. The document content is as follows:

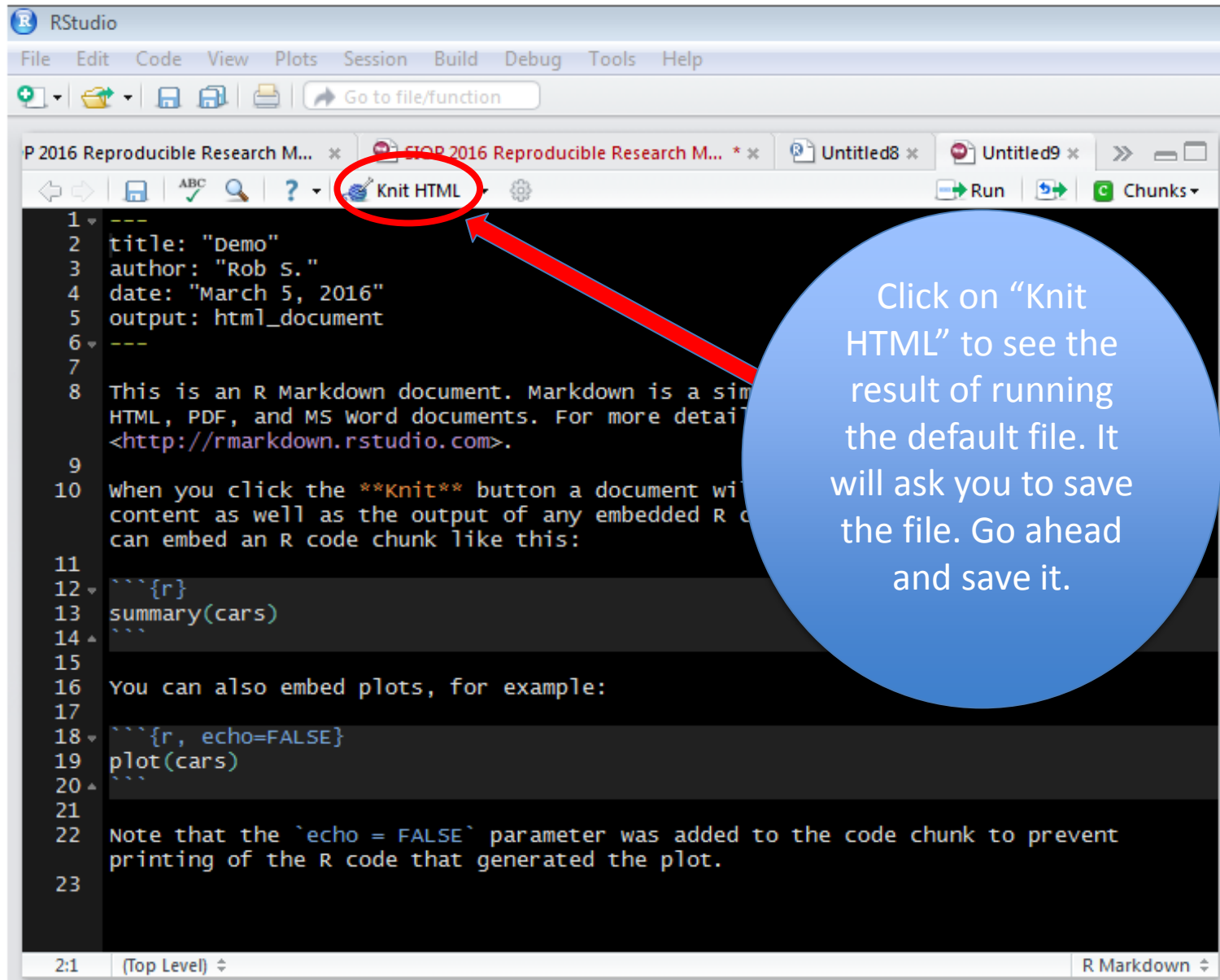
```
1 title: "Demo"
2 author: "Rob S."
3 date: "March 5, 2016"
4 output: html_document
5
6 ---
7
8 This is an R Markdown document. Markdown is a simple formatting syntax for authoring
9 HTML, PDF, and MS Word documents. For more details on using R Markdown see
10 <http://rmarkdown.rstudio.com>.
11
12 when you click the **knit** button a document will be generated that includes both
13 content as well as the output of any embedded R code chunks within the document.
14 You can embed an R code chunk like this:
15
16 ```{r}
17 summary(cars)
18 ```
19
20 You can also embed plots, for example:
21
22 ```{r, echo=FALSE}
23 plot(cars)
24 ```
25
26 Note that the `echo = FALSE` parameter was added to the code chunk to prevent
27 printing of the R code that generated the plot.
```

Two blue callout bubbles with red arrows provide additional context:

- A bubble pointing to lines 2-4: "Here are the title, author, date, and format that were chosen"
- A bubble pointing to line 10: "This is the default text that appears on every new R Markdown doc"

R Studio-R Markdown

FurstPerson.



```
processing file: Demo.Rmd
|.....| 20%
ordinary text without R code

|.....| 40%
label: unnamed-chunk-1
|.....| 60%
ordinary text without R code

|.....| 80%
label: unnamed-chunk-2 (with options)
List of 1
 $ echo: logi FALSE

|.....| 100%
ordinary text without R code

output file: Demo.knit.md

"C:/Users/Computer/AppData/Local/Pandoc/
-from markdown+autolink_bare_uris+asciif
gures --output Demo.html --smart --ema
-section-divs --template "C:/Users/Com
efault.html" --variable "theme:boo
Local/Temp/RtmpKIXVLA/rmarkdown-str30
ttps://cdn.mathjax.org/mathjax/latest
ight --variable "highlightjs=C:/Users\
h\highlightjs"

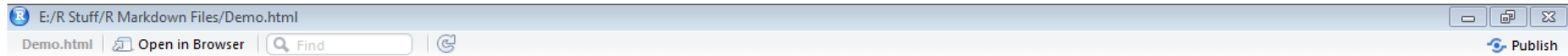
mo.utf8.md --to html -
backslash-implicit_fi
tained --standalone -
.2\rmarkdown\rmd\h\d
rs\Computer\AppData\
variable "mathjax-url:h
TMLorMML" --no-highl
y\3.2\rmarkdown\rmd\

Output created: Demo.html
```

You should see the status at 100% and that the output file has been created

R Studio-HTML Output

FurstPerson.



Demo

Rob S.

March 5, 2016

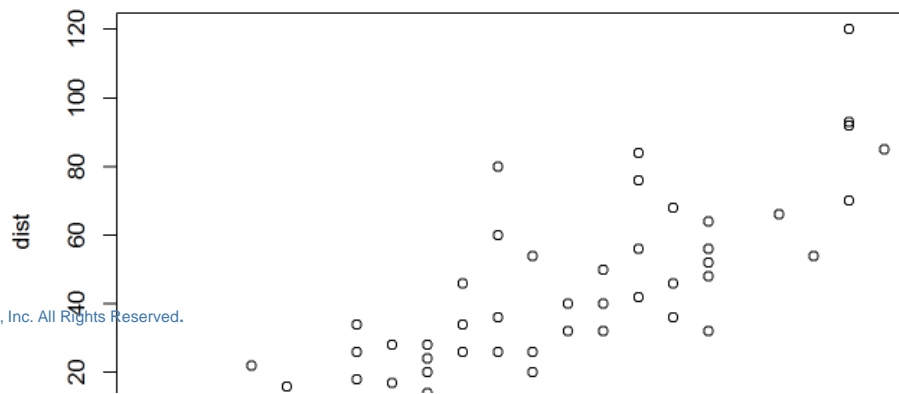
This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   : 2.00
##  1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##  Mean   :15.4    Mean   : 42.98
##  3rd Qu.:19.0    3rd Qu.: 56.00
##  Max.   :25.0    Max.   :120.00
```

You can also embed plots, for example:



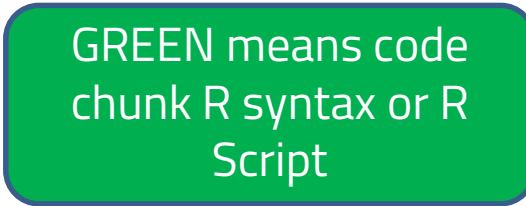
This is the
HTML file that
was created

Syntax Color Scheme

FurstPerson.

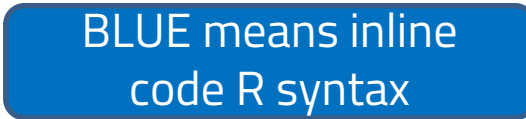
- Any syntax in the presentation will be color coded

- `~r{}`
- `summary(cars)`
- `~`



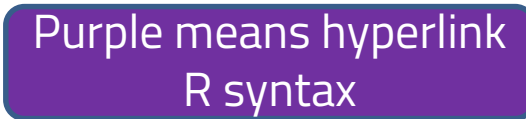
GREEN means code chunk R syntax or R Script

- ``r nrow (mydata)``



BLUE means inline code R syntax

- [Click here](#)



Purple means hyperlink R syntax

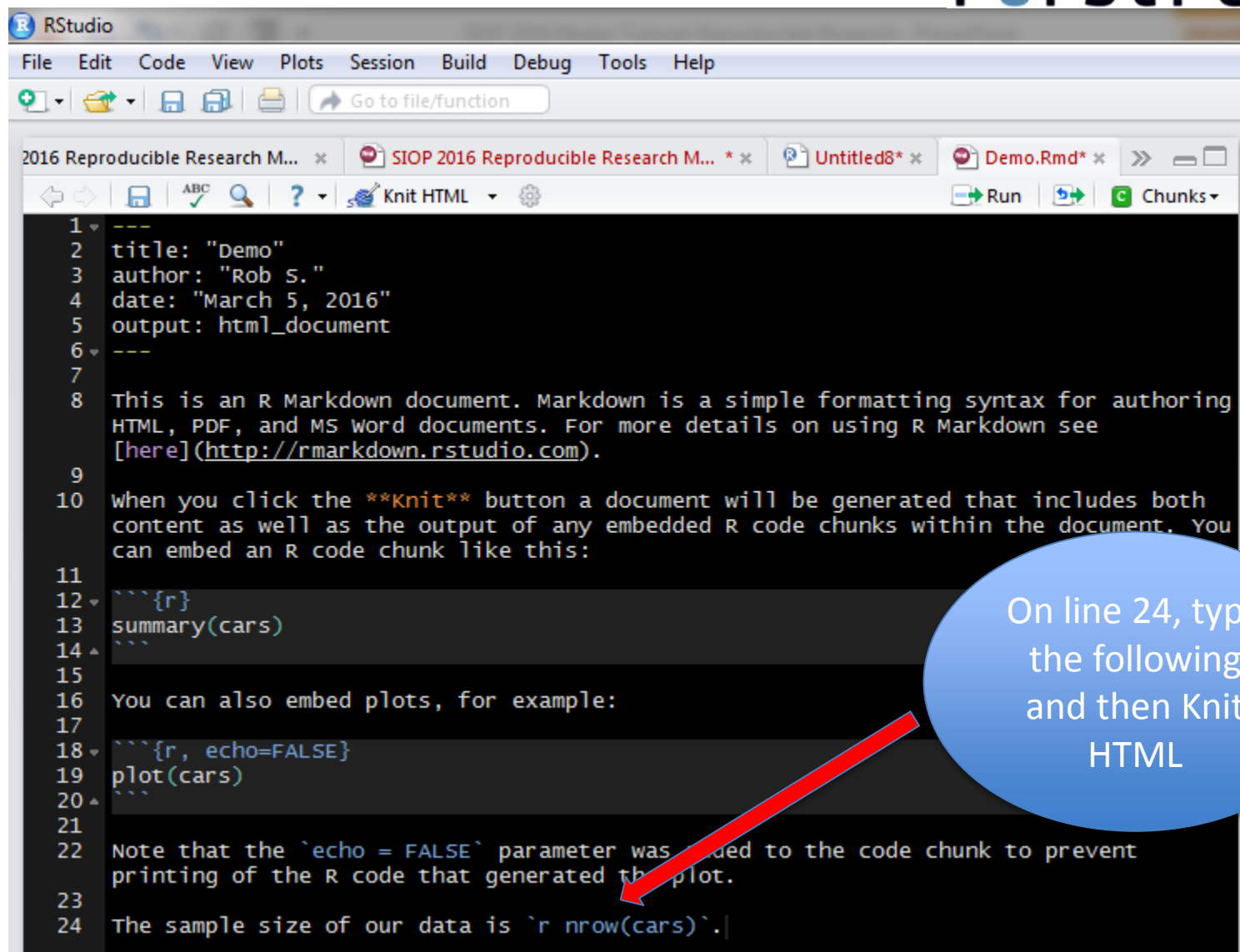
- In R Markdown, there are two type of code
 - Inline Code
 - Code Chunks
- Inline code is where you embed R code directly into your text
 - Uses include:
 - Keeping sample size fluid (e.g., ``r nrow(mydata)``)
 - Letting reader know text specifically references something in R (e.g., the ``psych`` package)
- Code chunks separate out R code from your text and stand alone
 - Uses include:
 - The bulk of the analysis of a project
 - Keeping comments hidden

- `{r}` designates a code chunk.
- This is where your R code goes
- The short cut to put in a code chunk is CTRL+ALT+ i (remember this)
- Line 13, `summary(cars)` is R code for run summary statistics on the data set "cars"

- Go back to your R Markdown Demo File in R Studio
- On line 8, notice the `<>` around the URL
- These `<>` symbols tell R Markdown to make it a hyperlink
- If you wanted it to say “here” with the hyperlink embedded in the word, it would look like the following with square brackets around the word in which the URL is embedded:
 - For more details on using R Markdown see [\[here\]](http://rmarkdown.rstudio.com)(http://rmarkdown.rstudio.com).
- On line 10 of your script, notice the `**` around `**Knit**`. This tells R Markdown to make the word “Knit” **bold** like CTRL+B in MS Word
- Let's do some examples

R Markdown-Inline Code

FurstPerson.



```
1 ---
2 title: "Demo"
3 author: "Rob S."
4 date: "March 5, 2016"
5 output: html_document
6 ---
7
8 This is an R Markdown document. Markdown is a simple formatting syntax for authoring
9 HTML, PDF, and MS word documents. For more details on using R Markdown see
10 [here](http://rmarkdown.rstudio.com).
11
12 When you click the Knit button a document will be generated that includes both
13 content as well as the output of any embedded R code chunks within the document. You
14 can embed an R code chunk like this:
15
16 ```{r}
17 summary(cars)
18 ```
19
20 You can also embed plots, for example:
21
22 ```{r, echo=FALSE}
23 plot(cars)
24 ```
25
26 Note that the `echo = FALSE` parameter was added to the code chunk to prevent
27 printing of the R code that generated the plot.
28
29 The sample size of our data is `r nrow(cars)`.
```

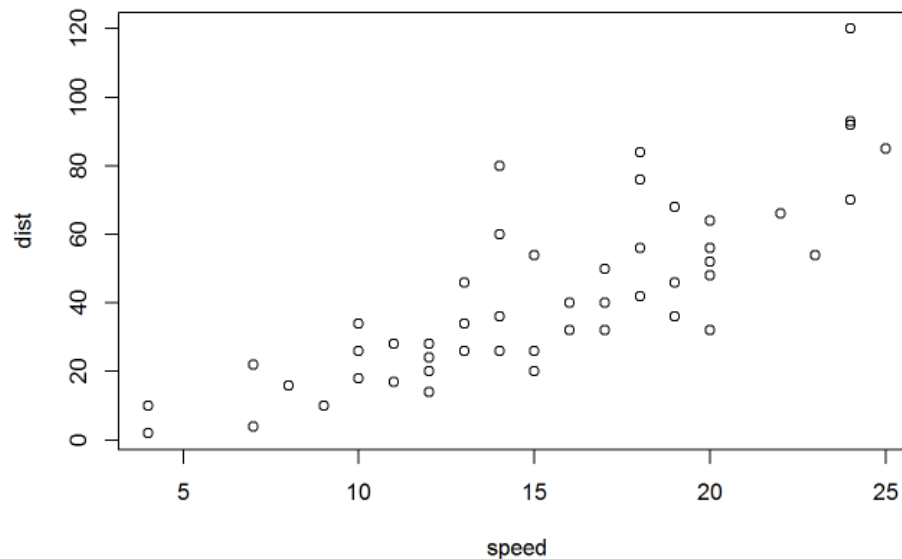
On line 24, type
the following
and then Knit
HTML

R Studio-HTML Output

FurstPerson.

```
##      speed      dist
## Min.   : 4.0    Min.   : 2.00
## 1st Qu.:12.0    1st Qu.: 26.00
## Median :15.0    Median : 36.00
## Mean   :15.4    Mean   : 42.98
## 3rd Qu.:19.0    3rd Qu.: 56.00
## Max.   :25.0    Max.   :120.00
```

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the

The sample size of our data is 50.

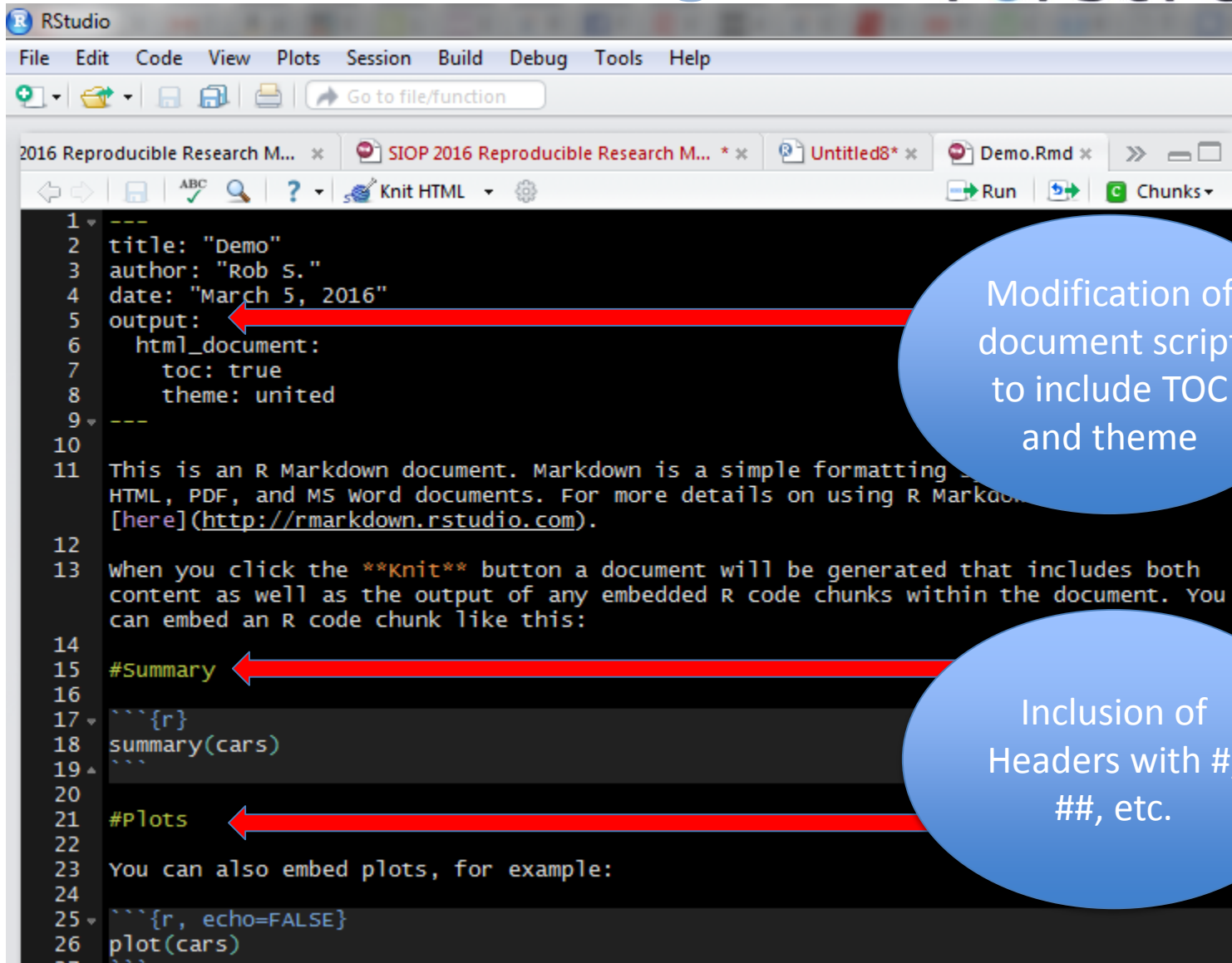
Instead of
printing ``r
nrow(cars)`` we
get the actual
number of
rows, 50.

- Now, you try.
- On line 26, type in the following:
- The mean of speed is ``r mean(cars$speed)`` and the mean of distance is ``r mean(cars$dist)``.
- Now click Knit HTML
- Notice how the document now includes the means of your variables
- If you received more “cars” data, you could simply amend the current cars file with the data, rerun the document, and the sample size and means would instantly update. (more on this later)

- Within R Markdown documents you can add:
 - Table of Contents
 - Headers
 - Italics and Bold
 - Superscripts and Subscripts
 - Strikethrough
 - Links
 - Inline equations
 - Images
 - Ordered and unordered list
 - Tables
- Here is a link to the [cheat sheet](#)

R Markdown-Formatting

FurstPerson.



The screenshot shows the RStudio interface with a file named 'Demo.Rmd' open. The editor displays R Markdown code. Two blue callout bubbles with red arrows point to specific parts of the code:

- The first bubble points to the `output:` section (lines 5-8), which is configured to generate an HTML document with a table of contents and the 'united' theme.
- The second bubble points to the `#Summary` header (line 15) and the `#Plots` header (line 21), illustrating how to include headers in the document.

```
1 ---
2 title: "Demo"
3 author: "Rob S."
4 date: "March 5, 2016"
5 output:
6   html_document:
7     toc: true
8     theme: united
9 ---
10
11 This is an R Markdown document. Markdown is a simple formatting
12 HTML, PDF, and MS Word documents. For more details on using R Markdown
13 [here](http://rmarkdown.rstudio.com).
14
15 when you click the **knit** button a document will be generated that includes both
16 content as well as the output of any embedded R code chunks within the document. You
17 can embed an R code chunk like this:
18
19 #Summary
20
21 {r}
22 summary(cars)
23
24 #Plots
25
26 You can also embed plots, for example:
27
28 {r, echo=FALSE}
29 plot(cars)
```

E:/R Stuff/R Markdown Files/Demo.html

Demo.html

Open in Browser

Find



Demo

Rob S.

March 5, 2016

- Summary
- Plots
 - Inline Code Example 1
 - Inline Code Example 2

TOC now in document

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word using R Markdown see [here](#).

When you click the **Knit** button a document will be generated that includes both content as well as the output of within the document. You can embed an R code chunk like this:

Summary

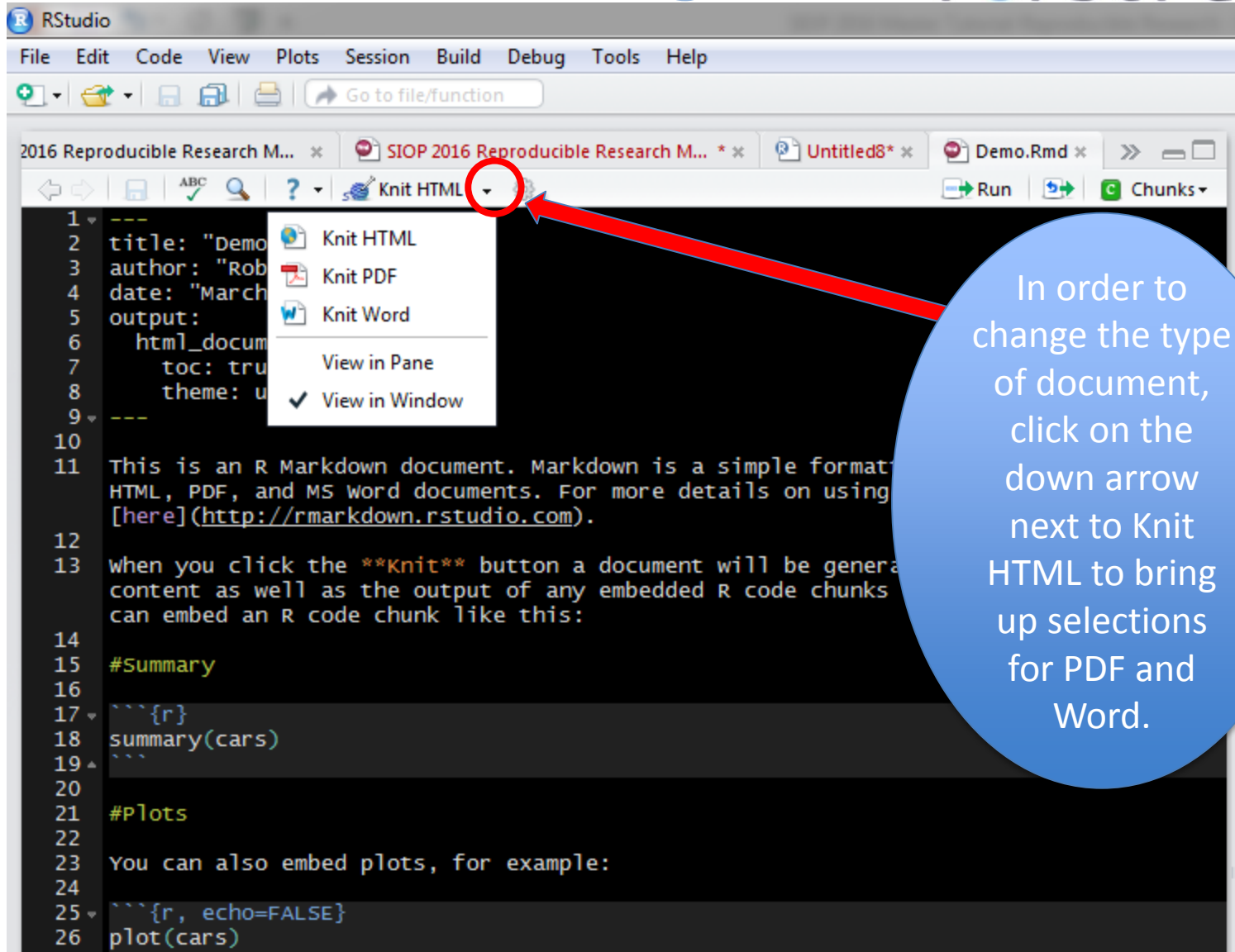
Headings get smaller as # are added

```
summary(mtcars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   : 2.00
## 1st Qu.:12.0    1st Qu.:26.00
## Median :15.0    Median :36.00
## Mean   :15.4    Mean   :42.98
## 3rd Qu.:19.0    3rd Qu.:56.00
## Max.   :25.0    Max.   :120.00
```


R Markdown-Formatting

FurstPerson.



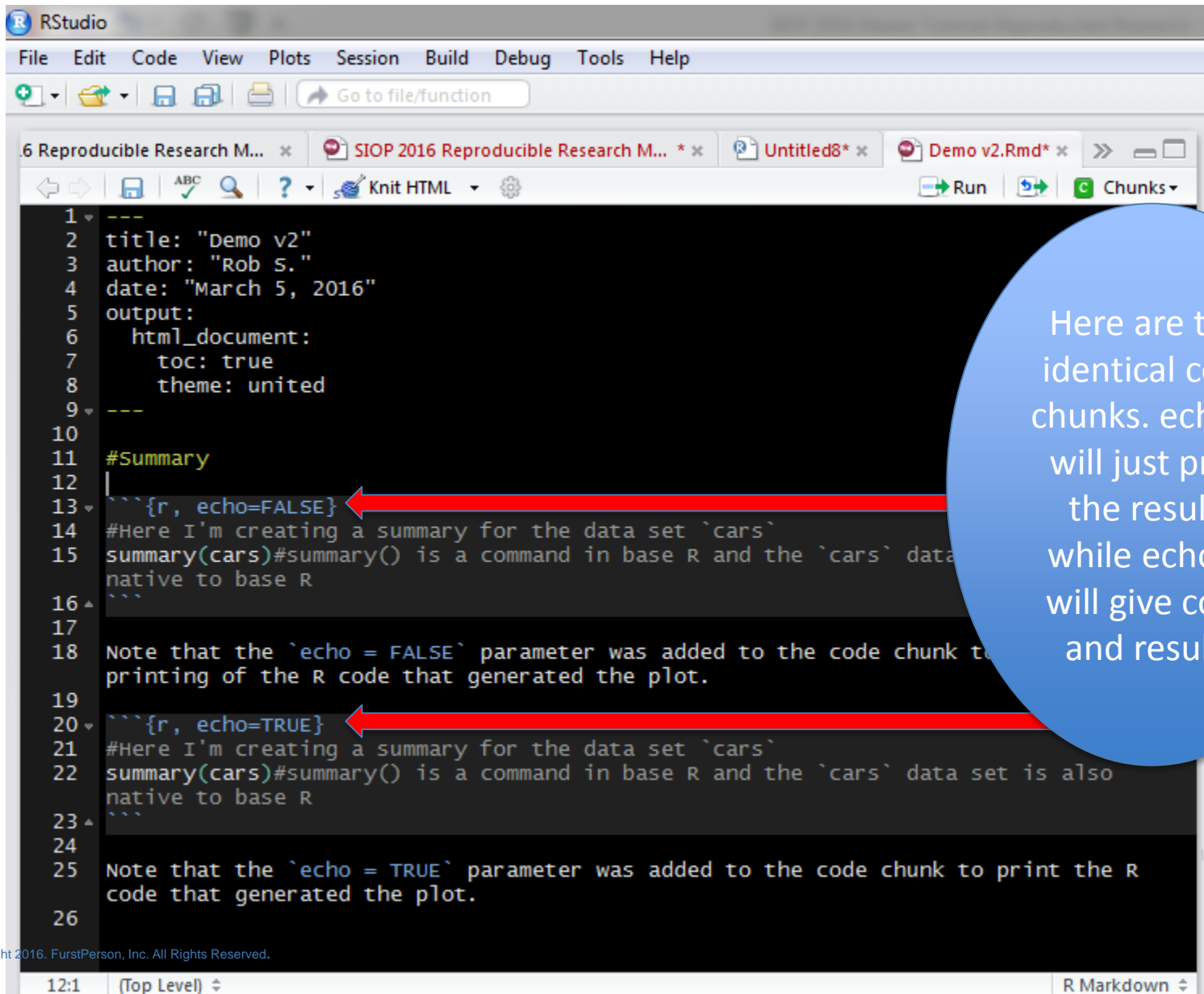
In order to change the type of document, click on the down arrow next to Knit HTML to bring up selections for PDF and Word.

- Much of the formatting is interchangeable between document types
- Tables sometimes behave differently between HTML and Word and PDF

- Code Chunks can be formatted to display many different layers
- Within the ````r{}````, set the following to either TRUE or FALSE
 - `eval`=run the code or not
 - `echo`=display the code or not
 - `error`=display error messages or not
 - `message`=display messages or not
 - `warning`=display warnings or not
- `results`-set to 'hide' if you want to hide the results
- For example, you may set `echo=TRUE` to display the code for a tutorial but `echo=FALSE` for a report to a client
- You may set `eval=FALSE` within a tutorial if you are giving many examples but don't want to crowd the document with output

R Markdown-Formatting

FurstPerson.

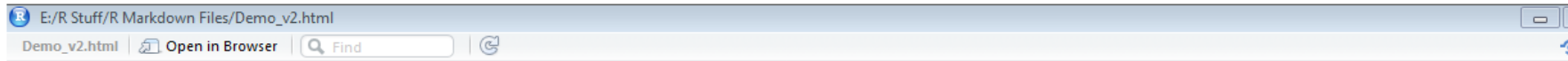


```
1 ---
2 title: "Demo v2"
3 author: "Rob S."
4 date: "March 5, 2016"
5 output:
6   html_document:
7     toc: true
8     theme: united
9 ---
10
11 #Summary
12 |
13 {r, echo=FALSE}
14 #Here I'm creating a summary for the data set `cars`
15 summary(cars)#summary() is a command in base R and the `cars` data
16   native to base R
17
18 Note that the `echo = FALSE` parameter was added to the code chunk to
19   printing of the R code that generated the plot.
20 {r, echo=TRUE}
21 #Here I'm creating a summary for the data set `cars`
22 summary(cars)#summary() is a command in base R and the `cars` data set is also
23   native to base R
24
25 Note that the `echo = TRUE` parameter was added to the code chunk to print the R
26   code that generated the plot.
```

Here are two identical code chunks. echo=F will just print the results while echo=T will give code and results

R Markdown-Formatting

FurstPerson.



Demo v2

Rob S.

March 5, 2016

- Summary

Summary

```
##      speed      dist
## Min.   : 4.0    Min.   : 2.00
## 1st Qu.:12.0    1st Qu.: 26.00
## Median :15.0    Median : 36.00
## Mean   :15.4    Mean   : 42.98
## 3rd Qu.:19.0    3rd Qu.: 56.00
## Max.   :25.0    Max.   :120.00
```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

```
#Here I'm creating a summary for the data set `cars`
summary(cars)#summary() is a command in base R and the `cars` data set is also native to base R
```

```
##      speed      dist
## Min.   : 4.0    Min.   : 2.00
## 1st Qu.:12.0    1st Qu.: 26.00
## Median :15.0    Median : 36.00
## Mean   :15.4    Mean   : 42.98
## 3rd Qu.:19.0    3rd Qu.: 56.00
## Max.   :25.0    Max.   :120.00
```

Note that the `echo = TRUE` parameter was added to the code chunk to print the R code that generated the plot

No code visible
below
“Summary”
with echo=F

With echo=T
you can see the
entire contents
of my Code
Chunk

Questions?

Learning Objective #2: Publishing your Reproducible Research Document in HTML, Word, or PDF

Locally Publishing R Markdown **F**urst**P**erson.

- So far we've seen how to generate a HTML, Word, or PDF file from R Markdown
- This section will focus how to go from initial analysis via R Script to R Markdown document to distribute to others
- We will answer the following questions
 - How do I add in bibliography information?
 - How do I go from the very beginning of a project to having something I can revisit later and remember what I did?
 - How do I make tables that look nice and are easy to read?
 - When sharing my document, how can I have R automatically load all of the necessary packages for the analysis?
 - What are the R packages I probably don't know about but are essential to nice looking documents?

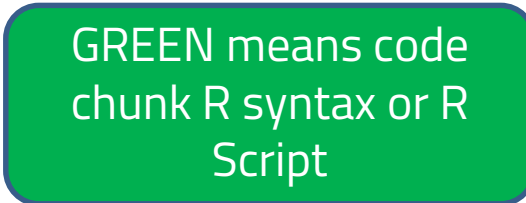
- This study will use a subset of the data from:
- Johnson, J. A. (2014). Measuring thirty facets of the five factor model with a 120-item public domain inventory: Development of the IPIP-NEO-120. J. of Research in Personality, 51, 78-89.
- Original study data can be found [here](#)
 - <https://osf.io/wxvth/>
- Data for the Mock Study can be found [here](#)
 - https://github.com/RobStilson/SIOP_2016_Master_Tutorial

Syntax Color Scheme

FurstPerson.

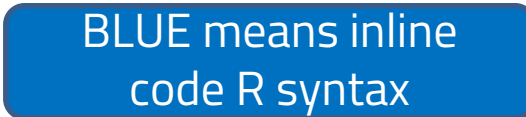
- Any syntax in the presentation will be color coded

- `~r{}`
- `summary(cars)`
- `~`



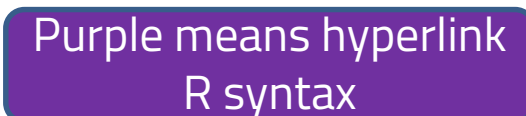
GREEN means code chunk R syntax or R Script

- ``r nrow (mydata)``



BLUE means inline code R syntax

- [Click here](#)



Purple means hyperlink R syntax

- Load into R with the following code:
- `#tryCatch code to load all necessary packages (Thanks Mike G.)`
- `tryCatch(require(RCurl),finally=utils:::install.packages(pkgs='RCurl',repos='http://cran.r-project.org'));`
- `require(RCurl)`
- `library(RCurl)`
- `x <-
getURL("https://raw.githubusercontent.com/RobStilson/SIOP_2016_Master_Tutorial/master/mydata_10k.csv")`
- `mydata_10k <- read.csv(text = x)`
- Now run (not Knit) in your R script

- We will take a random sample of 1000
- Use the `set.seed()` command to make sure it takes the same random sample each time
- `set.seed(1)`
- `mydata <- mydata_10k[sample(which(mydata_10k$country=='USA'), 1000, replace=FALSE),]`
- #By setting the seed to a specified number it makes the sampling reproducible
- #Below we will take another sample of 1000 and save to a different data set. The set seed will be the same so it will take the exact same
- `set.seed(1)`
- `mydata2 <- data[sample(which(data$country=='USA'), 1000, replace=FALSE),]`
- `rm(mydata2)` #This removes the data set since we no longer need it

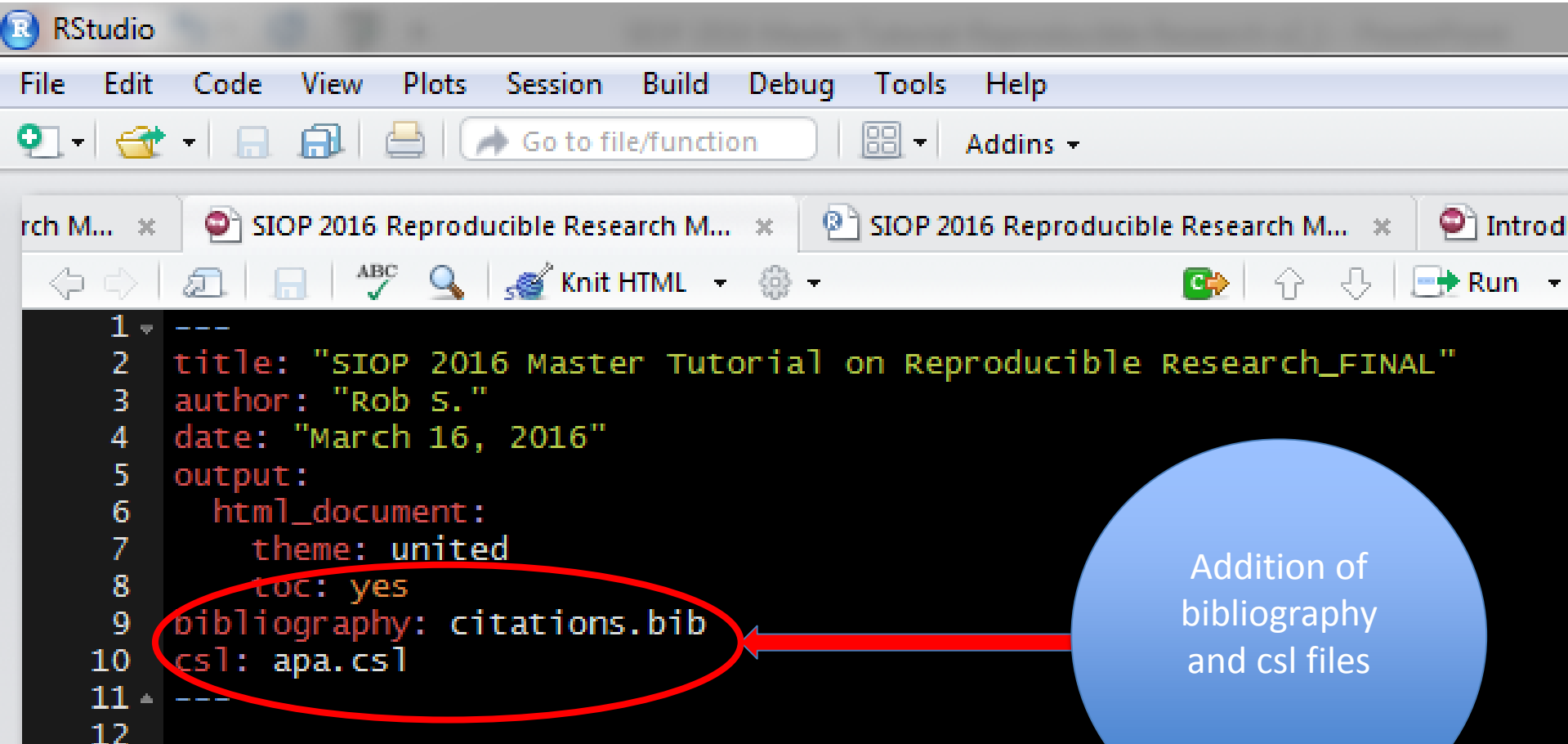
- The rest of the R code will be available on Github
- It will not be reproduced here, but will be demonstrated
- Overview of mock study
 - Take subset of only items (no ID info)
 - Rename items
 - Create keys
 - Score items and get alpha reliability
 - Create overall scale scores
 - Present results

- Now we need to cite where we got data
- Need additional files for bibliography
- Add them into Title portion of R Markdown document
- Additional files are
 - .bib-bibliography file
 - .csl-citation style file
- Overview of process on R Studio site [here](#)
 - http://rmarkdown.rstudio.com/authoring_bibliographies_and_citations.html
- Tutorial using APA format is [here](#)
- Make sure to save .bib and .csl files into same folder as R Markdown document

Mock Study-Bibliography

FurstPerson.

- In the current file, it will now look like this:

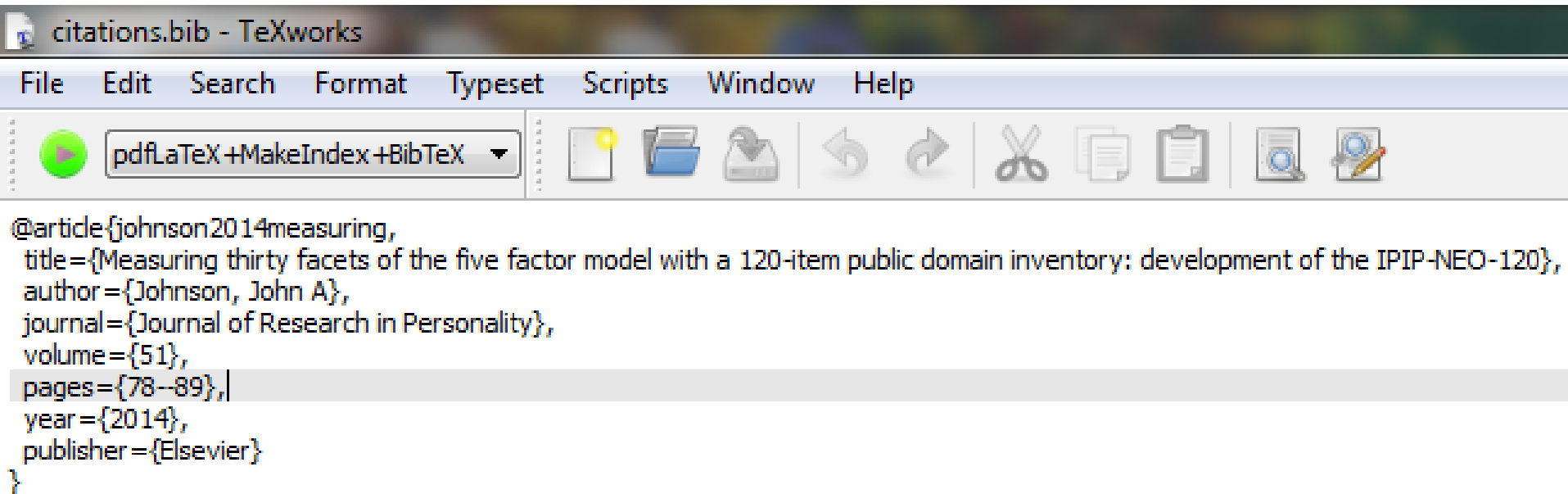


```
1 ---
2 title: "SIOP 2016 Master Tutorial on Reproducible Research_FINAL"
3 author: "Rob S."
4 date: "March 16, 2016"
5 output:
6   html_document:
7     theme: united
8     toc: yes
9   bibliography: citations.bib
10  csl: apa.csl
11 ---
12
```

Mock Study-.bib file

FurstPerson.

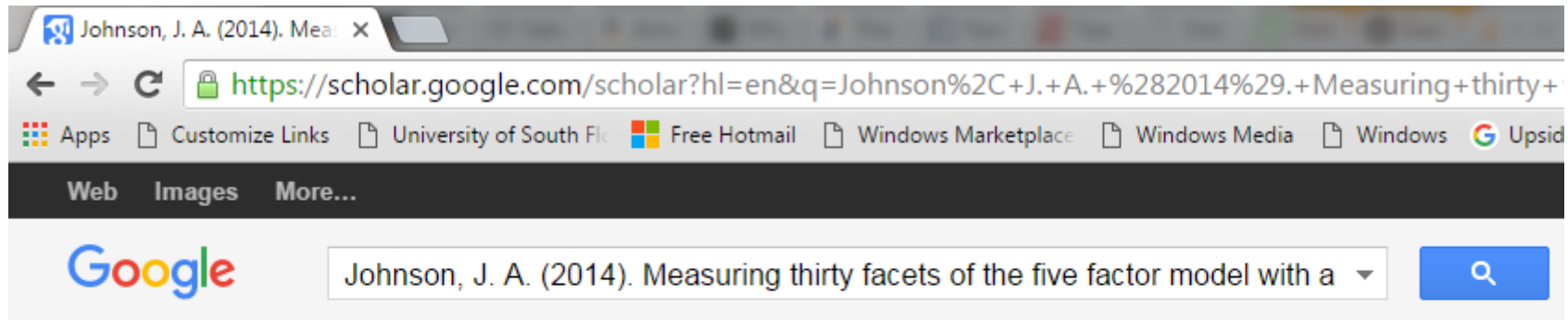
- Specific format needed for .bib file
- Scholar.google.com makes this easy



Mock Study-Citation for Mock Study

FurstPerson.

- Look up study on Google Scholar



Scholar

Articles

Case law

My library

Any time

Measuring thirty facets of the five factor model with a 120-item public domain inventory: development of the IPIP-NEO-120

JA Johnson - Journal of Research in Personality, 2014 - Elsevier

Abstract The IPIP-NEO (Goldberg, 1999) is a 300-item inventory that measures constructs similar to those in the NEO Personality Inventory (NEO PI-R; Costa & McCrae, 1992). Despite evidence for its reliability and validity, the IPIP-NEO is even longer than the ...

Cited by 12 Related articles All 3 versions [Cite](#) [Save](#)

Click on "Cite"

Mock Study-Citation for Mock Study

FurstPerson.

- Brings up "Cite" window

Cite

Copy and paste a formatted citation or use one of the links to import into a bibliography manager.

MLA Johnson, John A. "Measuring thirty facets of the five factor model with a 120-item public domain inventory: development of the IPIP-NEO-120." *Journal of Research in Personality* 51 (2014): 78-89.

APA Johnson, J. A. (2014). Measuring thirty facets of the five factor model with a 120-item public domain inventory: development of the IPIP-NEO-120. *Journal of Research in Personality*, 51, 78-89.

Johnson, John A. "Measuring thirty facets of the five factor model with a 120-item public domain inventory: development of the IPIP-NEO-120." *Journal of Research in Personality* 51 (2014): 78-89.

Johnson, J.A., 2014. Measuring thirty facets of the five factor model with a 120-item public domain inventory: development of the IPIP-NEO-120. *Journal of Research in Personality*, 51, pp.78-89.

Johnson JA. Measuring thirty facets of the five factor model with a 120-item public domain inventory: development of the IPIP-NEO-120. *Journal of Research in Personality*. 2014 Aug 31;51:78-89.

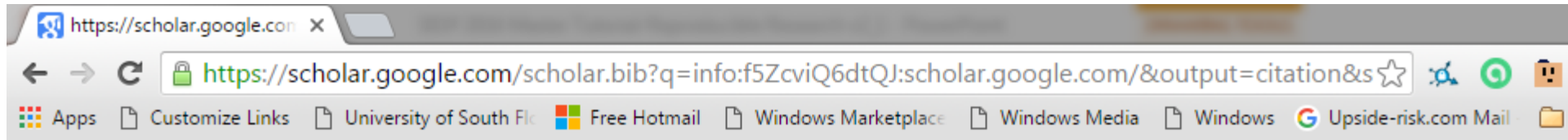
[BibTeX](#) [EndNote](#) [RefMan](#) [RefWorks](#)

Now, click on
"BibTeX"

Mock Study-Citation for Mock Study

FurstPerson.

- This should look familiar



```
@article{johnson2014measuring,  
  title={Measuring thirty facets of the five factor model with a 120-item public domain inventory: development of the IPIP-NEO-120},  
  author={Johnson, John A},  
  journal={Journal of Research in Personality},  
  volume={51},  
  pages={78--89},  
  year={2014},  
  publisher={Elsevier}  
}
```

Copy and paste all of
this into your .bib file.
Cite in R Markdown by
putting “@” in front of
cited name

- R Markdown automatically puts .bib output at end of document
- Simply end document with ##References and .bib output will be put below


References

Costa, P. T., & McCrae, R. R. (1992). Neo pi-r professional manual.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.

Johnson, J. A. (2014). Measuring thirty facets of the five factor model with a 120-item public domain inventory: Development of the iPIP-nEO-120. *Journal of Research in Personality*, 51, 78–89.

Nunnally Jr, J. C. (1970). Introduction to psychological measurement.



Here is the
citation among
others

- Base R can present tables and graphs in R Markdown
- However, not very visually appealing
- How to fix this?
 - sjPlot
 - pander
 - Stargazer
 - Xtable
 - Kable

- Will focus on sjPlot for HTML and pander for Word and PDF for the tutorial
- Base R can present tables in R Markdown
- Much more presentable with sjPlot (see below)
- Learn more about sjPlot [here](#)

```
## Call: scoreItems(keys = keys, items = items)
##
## (Unstandardized) Alpha:
##      AGR  CON  EXT  NEU  OPN
## alpha 0.59 0.46 0.63 0.69 0.12
##
```

Here is a more succinct table in a better looking format.

Table 3: Alpha reliabilities

Alpha Reliability of Scales

<i>Variable</i>	<i>AGR</i>	<i>CON</i>	<i>EXT</i>	<i>NEU</i>	<i>OPN</i>
alpha	0.587	0.463	0.635	0.691	0.122

- Here is the code to create the sjPlot table
- Below, the data is rounded to 3 digits and transformed to a df
- `scores_alpha <- round(as.data.frame(scores$alpha), 3)`
- Below is the specific way to wrap the sjPlot (sjt.df) code to get it to present correctly in R Markdown
- `scores_alpha_table <- sjt.df(scores_alpha, title = "Alpha Reliability of Scales", alternateRowColors=TRUE, useViewer=F, describe=FALSE, encoding = "UTF-8", no.output=TRUE, hideProgressBar = TRUE)$knitr`
- The above code is in the ````r{}```` code chunk.
- We then use the following inline code to actually present the table
- ``r scores_alpha_table``
- Adding “####Table 3: Alpha reliabilities” as a third level header

- This method can be used for any of the many sjPlot commands like:
 - `sjp.frq` – plot frequencies of (count) variables
 - `sjp.glm` – plot odds ratios (forest plots) of generalized linear models
 - `sjp.glmer` – plot odds ratios (forest plots) of generalized linear mixed effects models
 - `sjp.grpfrq` – plot grouped or stacked frequencies
 - `sjp.int` – plot interaction effects of regression models
 - `sjp.likert` – plot Likert scales
 - `sjp.lm` – plot estimates of linear models
 - `sjp.lmer` – plot estimates (forest plots) of linear mixed effects models
 - `sjp.pca` – plot principal component analysis
 - `sjp.scatter` – plot (grouped) scatter plots
 - `sjp.stackfrq` – plot stacked proportional bars
 - `sjp.xtab` – plot contingency tables

- This method can be used for any of the many sjPlot commands like:
 - `sjt.corr` – correlations as HTML table
 - `sjt.df` – (description of) data frames as HTML table
 - `sjt.frq` – frequencies of (count) variables as HTML table
 - `sjt.lm` – linear models as HTML table
 - `sjt.pca` – principal component analysis as HTML table
 - `sjt.stackfrq` – stacked frequencies as HTML table
 - `sjt.xtab` – contingency tables as HTML table

- There may be a way to render sjPlot in PDF and Word and still have it look nice
- I have not found it
- Enter pander!
- Pander makes tables and figures pretty in:
 - HTML
 - PDF
 - Word
- Not as customizable as sjPlot (that I know of)

Questions?

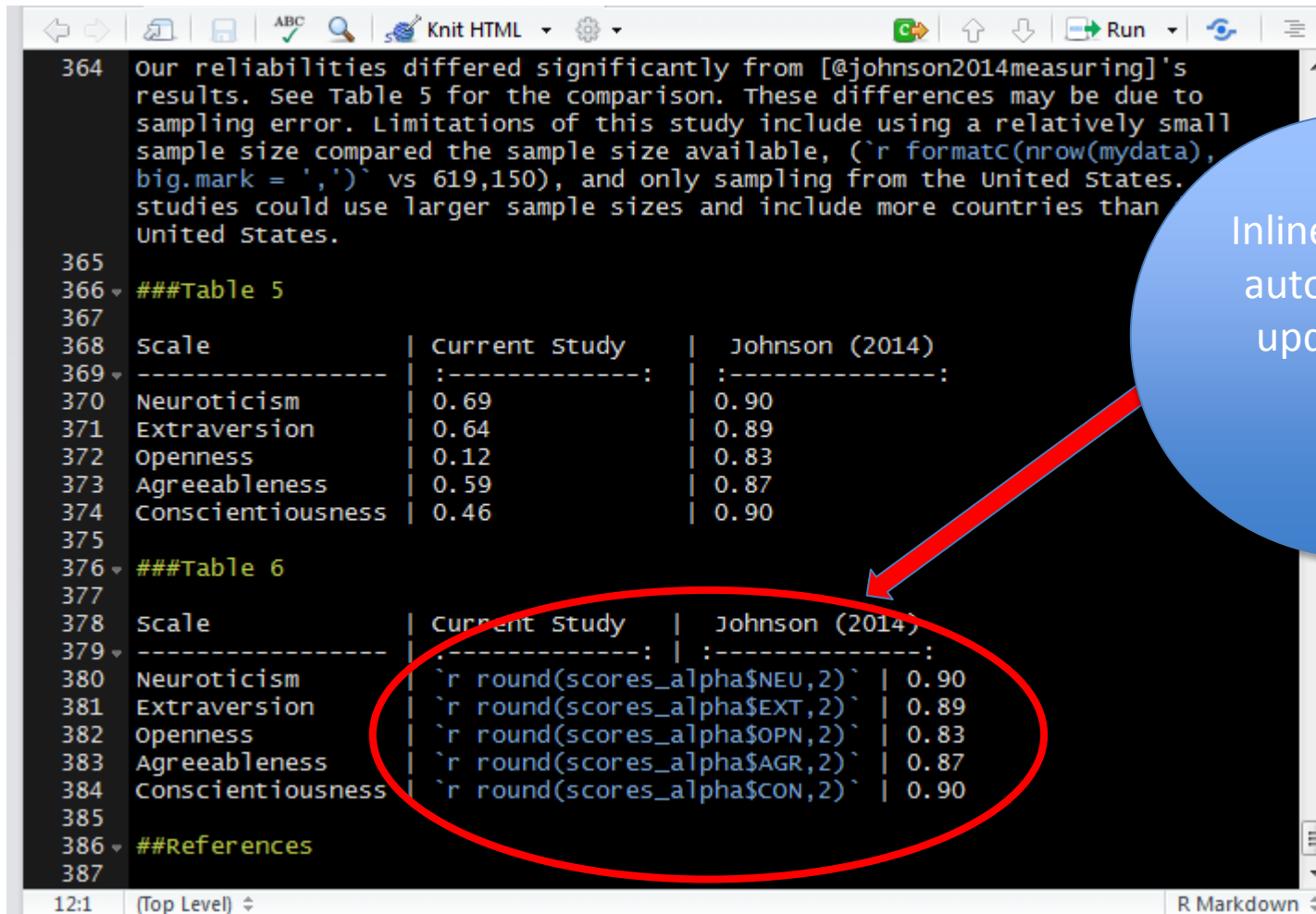
Learning Objective #3: Updating the data used in your document

- Now that we know how to create aesthetically pleasing tables what happens when new data comes in?
 - Simply add the data to your original file
 - Rename new data to original data file name
- Knit again and everything will be updated
 - Render any number in text with code if possible

Updating data in R Markdown

FurstPerson.

- You can use inline code within tables



```
364 our reliabilities differed significantly from [Johnson2014measuring]'s
365 results. See Table 5 for the comparison. These differences may be due to
366 sampling error. Limitations of this study include using a relatively small
367 sample size compared the sample size available, (r formatC(nrow(mydata),
368 big.mark = ',') vs 619,150), and only sampling from the United States.
369 studies could use larger sample sizes and include more countries than
370 United States.
371
372 Table 5
373
374 

| Scale             | Current Study | Johnson (2014) |
|-------------------|---------------|----------------|
| Neuroticism       | 0.69          | 0.90           |
| Extraversion      | 0.64          | 0.89           |
| Openness          | 0.12          | 0.83           |
| Agreeableness     | 0.59          | 0.87           |
| Conscientiousness | 0.46          | 0.90           |


410
411 Table 6
412
413 

| Scale             | Current Study                             | Johnson (2014) |
|-------------------|-------------------------------------------|----------------|
| Neuroticism       | <code>r round(scores_alpha\$NEU,2)</code> | 0.90           |
| Extraversion      | <code>r round(scores_alpha\$EXT,2)</code> | 0.89           |
| Openness          | <code>r round(scores_alpha\$OPN,2)</code> | 0.83           |
| Agreeableness     | <code>r round(scores_alpha\$AGR,2)</code> | 0.87           |
| Conscientiousness | <code>r round(scores_alpha\$CON,2)</code> | 0.90           |


449
450 References
```

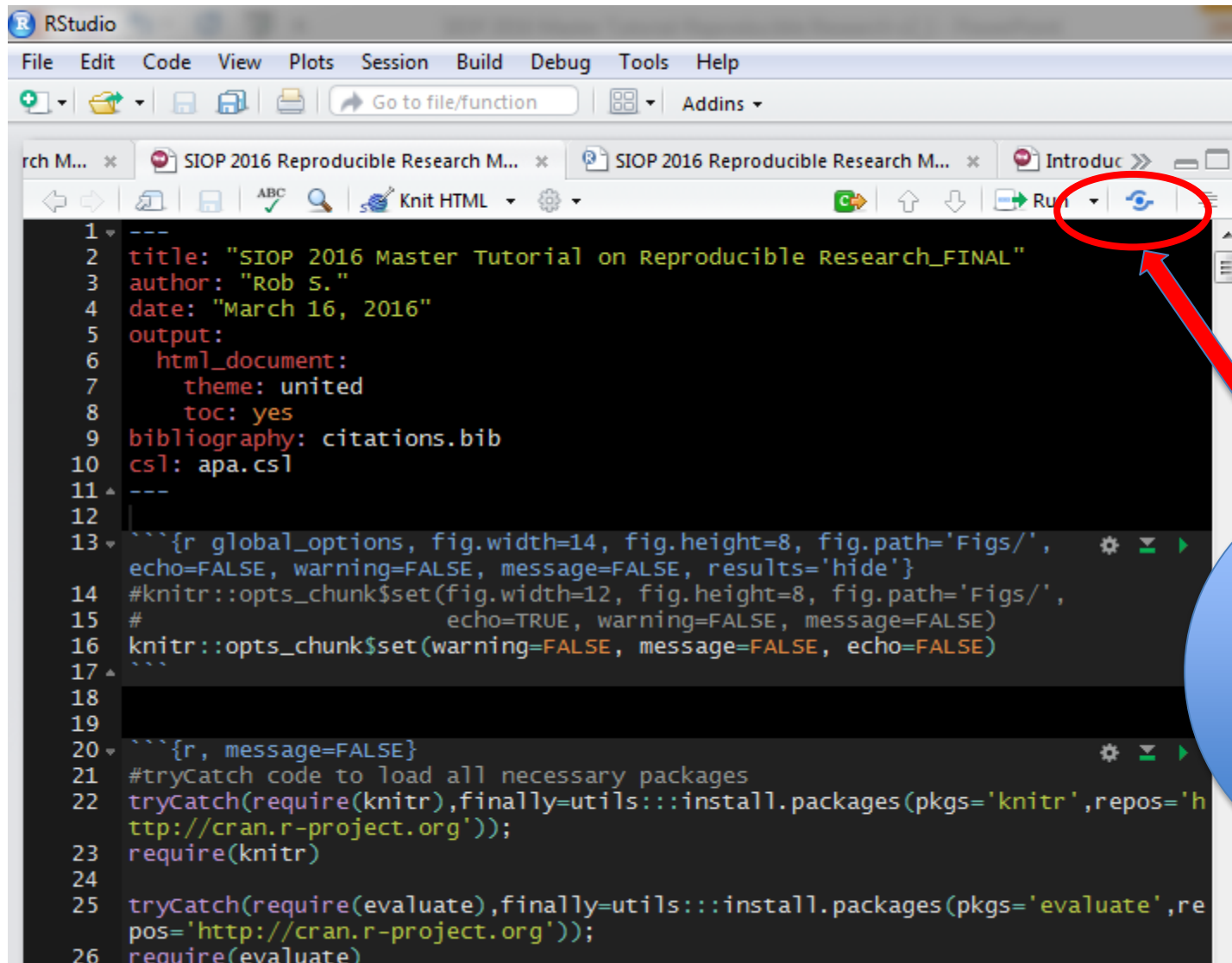
Inline code will automatically update your table

Questions?

Learning Objective #4: Publishing your findings on the web via RPubS

- Share your documents on RPubS.com
- You will need to create a free account
- Remember that **ANYBODY** can see this work
- Good place to go to see potential of R Markdown documents

- Can publish directly from R Studio



Click on this
button to
publish to
Rpubs.com

Questions?

Summary and Wrap Up

What did we learn

FurstPerson.

- How to create R Markdown Documents in R Studio
 - Inline code
 - Formatting text
 - Code chunks
 - Formatting Code chunks
 - sjPlot and pander within R Markdown
 - Publishing in HTML, Word, and PDF
 - How to create a bibliography in R Markdown
 - How to update existing docs with fresh data
 - How to publish to Rpubs.com

How you can get started

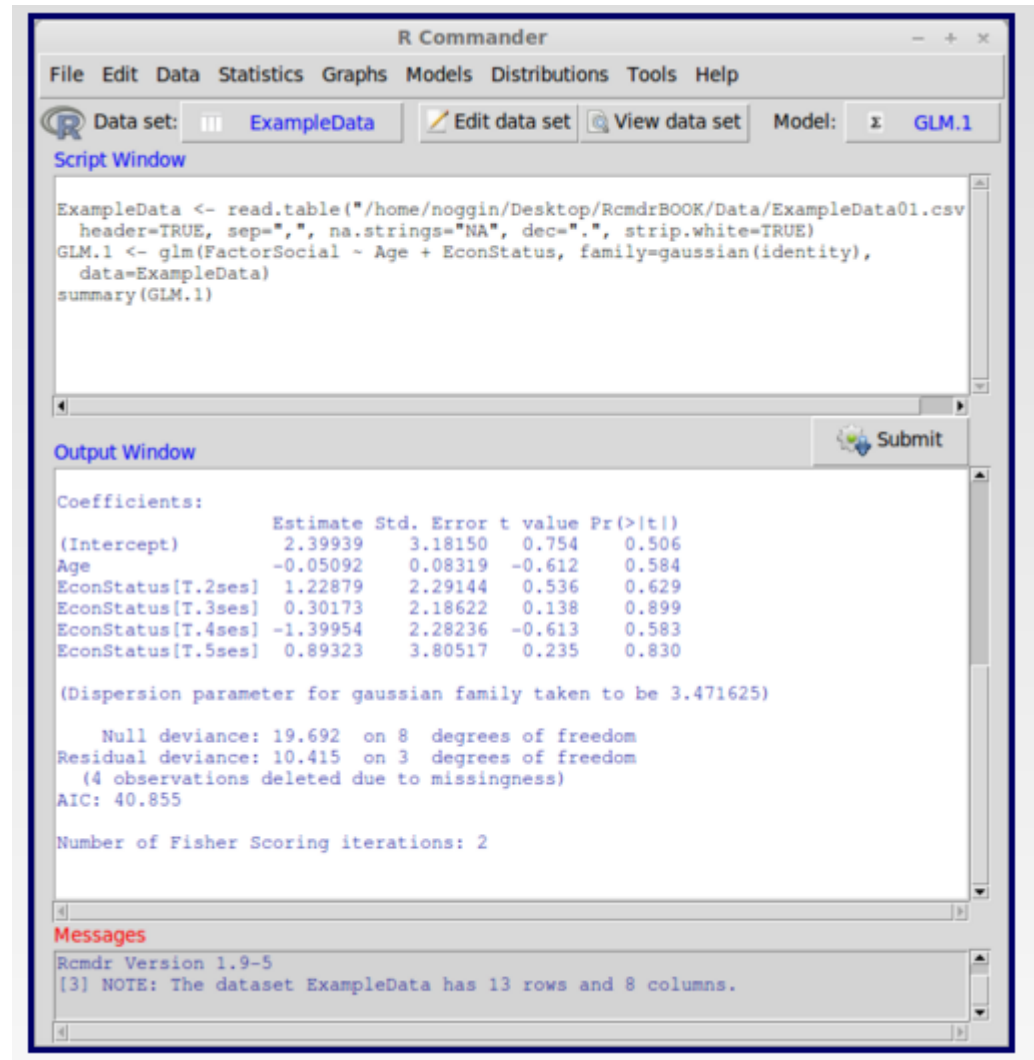
FurstPerson.

- If you know nothing about R
- [Data Camp](#)
- Introduction to R Class
 - Walks you through R from the beginning
 - Provides helpful hints and exercise solutions
- Create R Markdown documents of the class
 - Add you own comments to the docs
 - Try to exactly recreate the format

How you can get started

FurstPerson.

- If you know nothing about R and you don't like syntax
- [R Commander](#)
- Menu driven like SPSS
- Creates R Markdown script as you do analysis
- Integrates with R Studio



The screenshot shows the R Commander application window. At the top is a menu bar with options: File, Edit, Data, Statistics, Graphs, Models, Distributions, Tools, and Help. Below the menu bar is a toolbar with buttons for 'Data set: ExampleData', 'Edit data set', 'View data set', and 'Model: GLM.1'. The main window is divided into two panes. The top pane is the 'Script Window' containing R code:

```
ExampleData <- read.table("/home/noggin/Desktop/RcmdrBOOK/Data/ExampleData01.csv", header=TRUE, sep=".", na.strings="NA", dec=".", strip.white=TRUE)
GLM.1 <- glm(FactorSocial ~ Age + EconStatus, family=gaussian(identity), data=ExampleData)
summary(GLM.1)
```

 The bottom pane is the 'Output Window' displaying the results of the GLM model. It includes a table of coefficients, a dispersion parameter, deviance statistics, and the number of Fisher Scoring iterations. A 'Submit' button is located in the top right corner of the Output Window. At the very bottom is a 'Messages' pane showing version information and a note about the dataset dimensions.

```
ExampleData <- read.table("/home/noggin/Desktop/RcmdrBOOK/Data/ExampleData01.csv",
  header=TRUE, sep=".", na.strings="NA", dec=".", strip.white=TRUE)
GLM.1 <- glm(FactorSocial ~ Age + EconStatus, family=gaussian(identity),
  data=ExampleData)
summary(GLM.1)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.39939	3.18150	0.754	0.506
Age	-0.05092	0.08319	-0.612	0.584
EconStatus[T.2ses]	1.22879	2.29144	0.536	0.629
EconStatus[T.3ses]	0.30173	2.18622	0.138	0.899
EconStatus[T.4ses]	-1.39954	2.28236	-0.613	0.583
EconStatus[T.5ses]	0.89323	3.80517	0.235	0.830

(Dispersion parameter for gaussian family taken to be 3.471625)

Null deviance: 19.692 on 8 degrees of freedom
Residual deviance: 10.415 on 3 degrees of freedom
(4 observations deleted due to missingness)
AIC: 40.855

Number of Fisher Scoring iterations: 2

Messages

```
Rcmdr Version 1.9-5
[3] NOTE: The dataset ExampleData has 13 rows and 8 columns.
```

How you can get started

FurstPerson.

- Start your next project in R
- This is a great way to learn
- Get as far as you can before switching to a format you know better (e.g., loaded and cleaned data, but don't know how to do LM)
- Make notes of where you got stuck in R so you can figure out for next project
- Lather, rinse, repeat

How you can get started

FurstPerson®

- If you know some R or
- If no project springs to mind
- Go to www.twotutorials.com
 - Put together by Anthony Damico
 - 90+ ~2 minute video tutorials on R
 - Examples include
 - 013 how to read spss, stata, and sas files into r
 - 014 how to read an excel file (dot xls and dotxlsx) into a data frame with r
 - 016 how to add comments, save a script file, and make your work reproducible in r
 - 022 how to generate basic descriptive statistics - means, medians, sums, quantiles - on data tables in r

How you can get started

FurstPerson.

- Follow along with the videos
- Make a transcript of what is being done
- Copy the code being created
- Now you have a mini-tutorial
- Make these your own
 - Comments
 - Multiple iterations of code (e.g., with and without missing data, etc.)
 - Multiple iterations of graphs (e.g., different color schemes, legend or not, etc.)

Questions?

[https://github.com/
RobStilson](https://github.com/RobStilson)

Thank You!

Rob Stilson
Sr. Manager, R&D
Rob.Stilson@furstperson.com

Appendix

- [Embedding RData files in R Markdown files for more reproducible analyses](#)
- [Packrat](#)-packrat is a package dependency management program for R
- [Making Reproducible Research Enjoyable](#)