Creating and Disseminating Reproducible Research

- *R Packages*
- *RMarkdown Documents*
- *Writing in RMarkdown*

James Balamuta

IMSI Data Science Workshop Series Summer 2022

Lecture Objectives

- **Create** dynamic reports using *Quarto* or *RMarkdown*
- **Explore** the R ecosystem for extension packages**.**

R Packages

Base *R*

… default set of features that form the *R* language…

Data

trees

Girth	Height	Volume

mtcars

mpg	…	gear	carb

…

Help



…

Functions

mean(x, …)
sd(x, …)
sum(x, …)
prod(x, …)
plot(x, y, …)
…

R Packages

… ~18,571 extensions to *R* available on CRAN …

Available CRAN Packages By Name

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

CRAN
Mirrors
What's new?
Task Views
Search

About R
R Homepage
The R Journal

Software
R Sources
R Binaries
Packages
Other

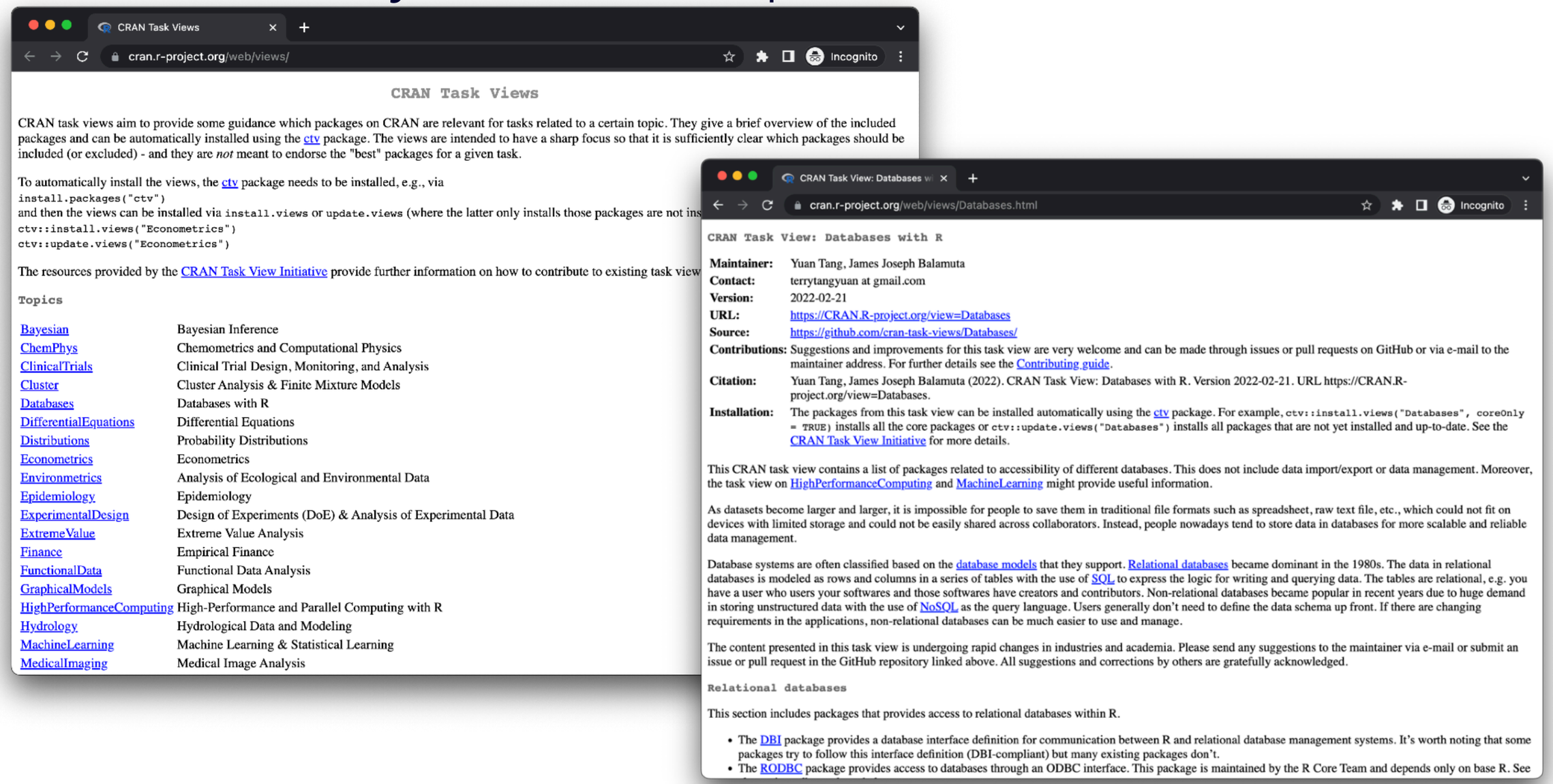
Documentation
Manuals
FAQs
Contributed

A3	Accurate, Adaptable, and Accessible Error Metrics for Predictive Models
abbyyR	Access to Abbyy Optical Character Recognition (OCR) API
abc	Tools for Approximate Bayesian Computation (ABC)
abc.data	Data Only: Tools for Approximate Bayesian Computation (ABC)
ABC.RAP	Array Based CpG Region Analysis Pipeline
ABCanalysis	Computed ABC Analysis
abcdeFBA	ABCDE_FBA: A-Biologist-Can-Do-Everything of Flux Balance Analysis with this package
ABCoptim	Implementation of Artificial Bee Colony (ABC) Optimization
ABCp2	Approximate Bayesian Computational Model for Estimating P2
abcrf	Approximate Bayesian Computation via Random Forests
abctools	Tools for ABC Analyses
abd	The Analysis of Biological Data
abe	Augmented Backward Elimination
abf2	Load Gap-Free Axon ABF2 Files
ABHgenotypeR	Easy Visualization of ABH Genotypes
abind	Combine Multidimensional Arrays
abjutils	Useful Tools for Jurimetrical Analysis Used by the Brazilian Jurimetrics Association
abn	Modelling Multivariate Data with Additive Bayesian Networks
abnormality	Measure a Subject's Abnormality with Respect to a Reference Population

https://cloud.r-project.org/web/packages/available_packages_by_name.html

Task Views

Subject Matter Expert Curated Task Views



https://cran.r-project.org/web/views/
https://cran.r-project.org/web/views/Databases.html

Using R Packages

… installing and loading an *R* package from CRAN …

1.
```
# Required once per R computer / Cloud project
install.packages(c("rmarkdown","quarto"))
```

```
trying URL 'https://cran.rstudio.com/bin/macosx/el-capitan/contrib/4.2/rmarkdown_1.8.tgz'
Content type 'application/x-gzip' length 2227986 bytes (2.1 MB)
==================================================
downloaded 2.1 MB

The downloaded binary packages are in
/var/folders/b0/vt_1hj2d6yd8myx9lwh81pww0000gn/T/Rtmp66L8EO/downloaded_packages
```

2.
```
# Required once per R session
library("rmarkdown"); library("quarto");
```

---

* Installing an *R* package is only required **once per *R* installation**. Useful to include a **# install.packages("pkgname")** prior to the library loads to emphasize

** Inside of all R Markdown documents or R Scripts, you will need to call library.

Loading an *R* Package

… what happens when **library()** is used to load a package …

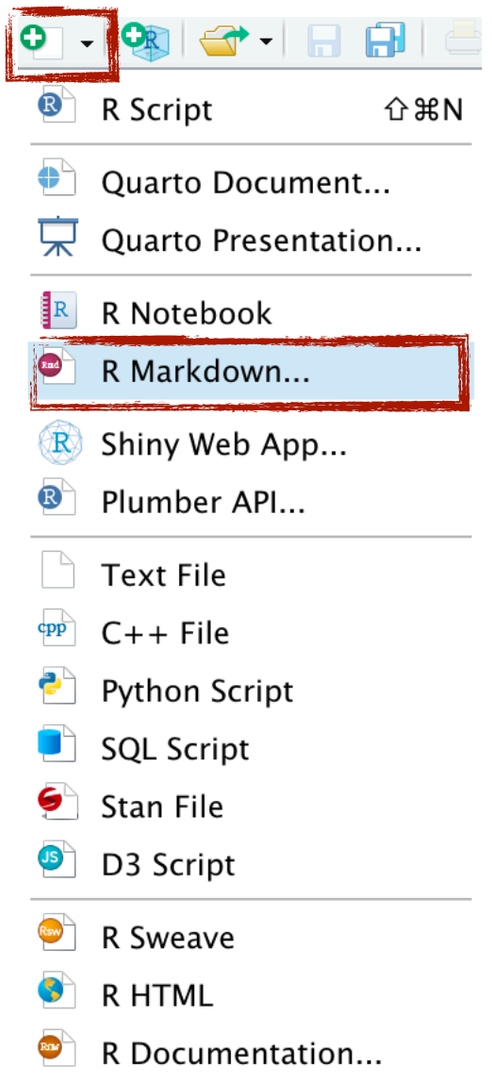
Ask Yourself: *"Is the package loaded?"*

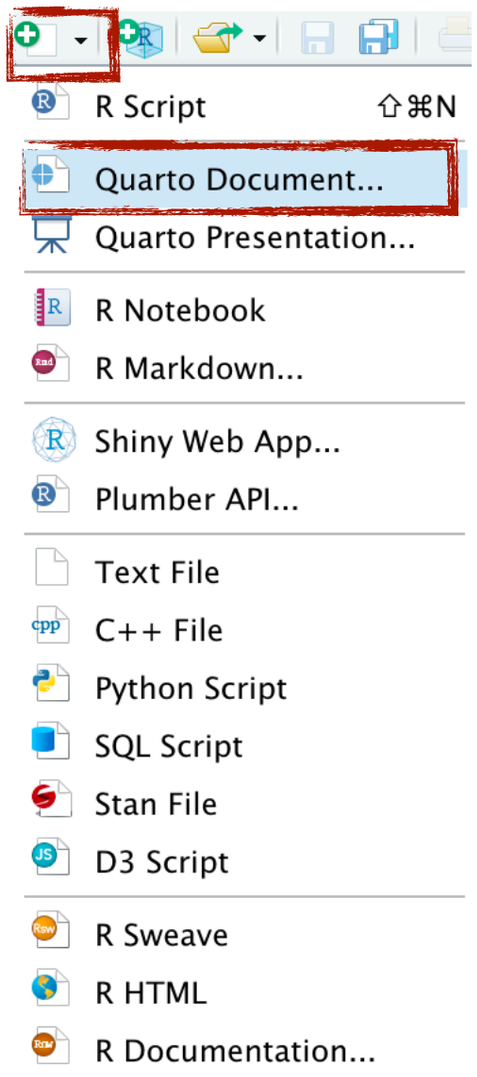
→

RMarkdown and Quarto Documents

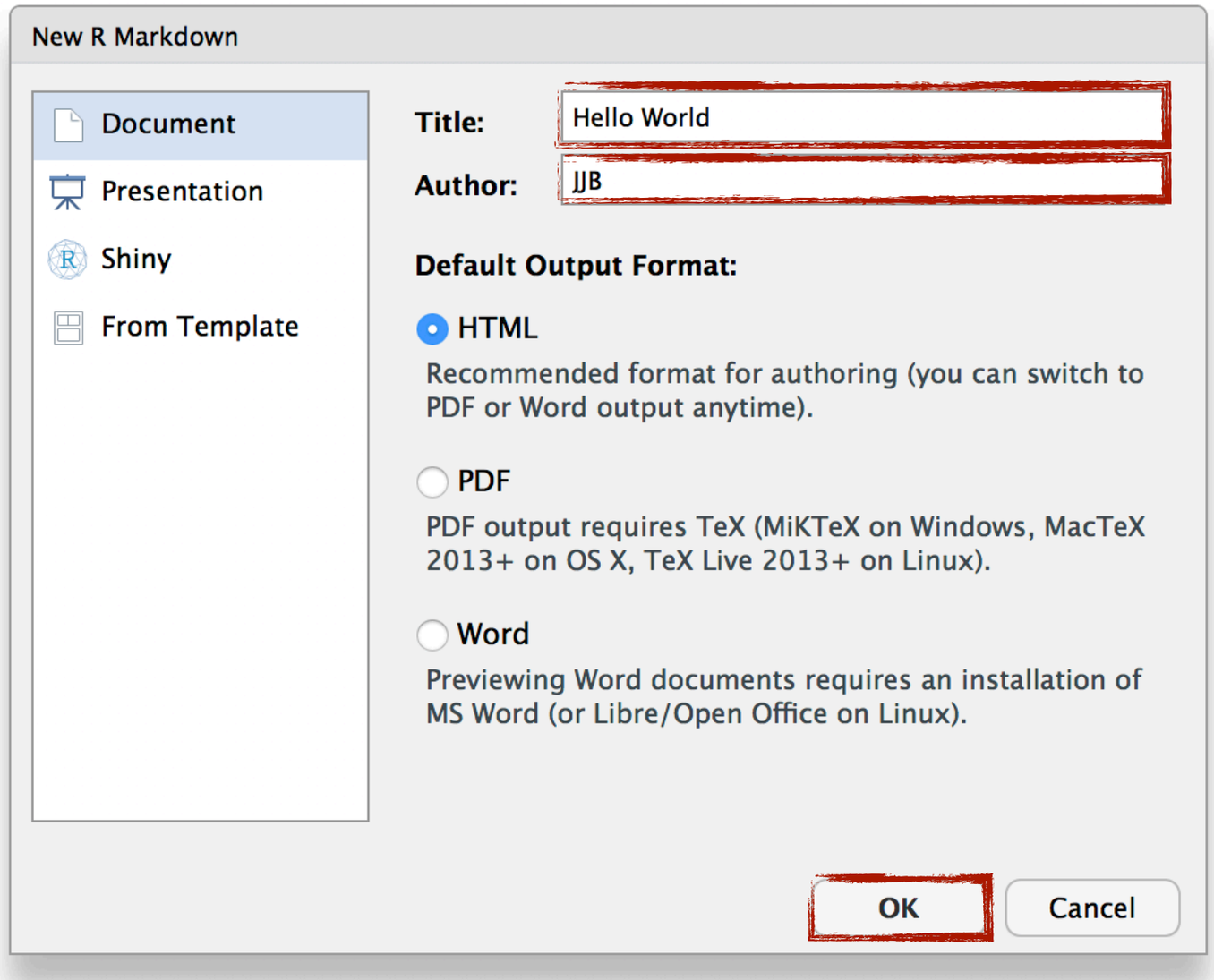
Create an RMarkdown Document

Click the White Plus
Select **R Markdown** or **Quarto** ...



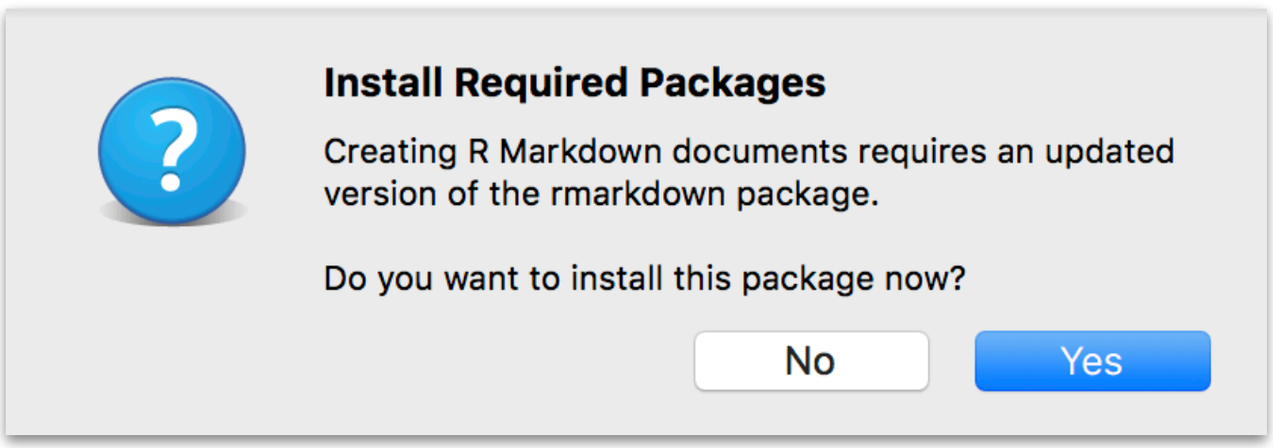


Enter Document Title and Author
Press **OK**



Oopsie

… something went wrong …




```
library("rmarkdown")
```

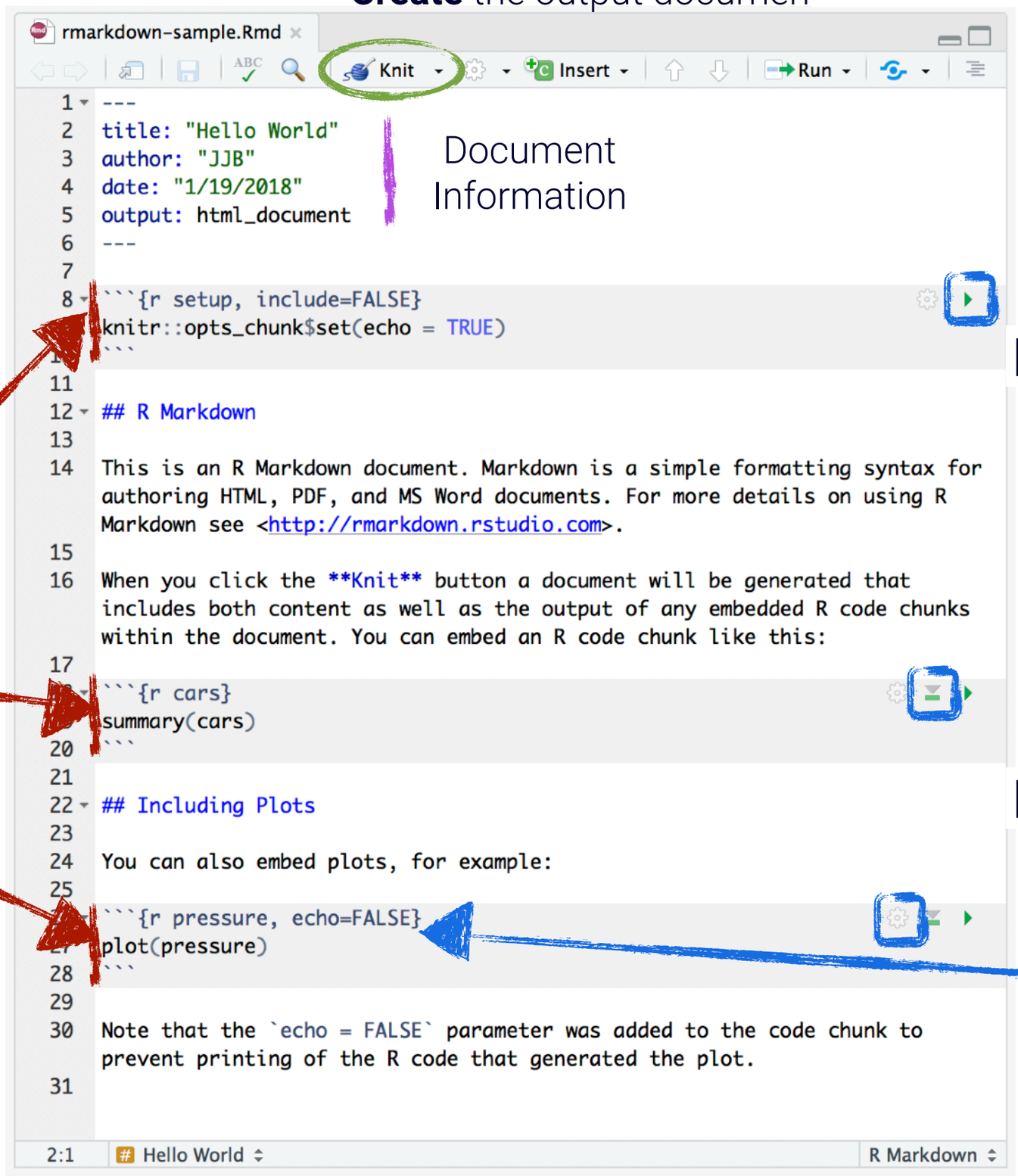
```
Error in library("rmarkdown") :
 there is no package called
 'rmarkdown'
```

… the package wasn't *installed*, so use …

```
install.packages("rmarkdown")
```

RMarkdown Document Sections

Create the output documen **[Cmd/Cntrl + Shift + K]**



Document Information

Run *this* Code Chunk **[Cmd/ Cntrl + Shift + Enter]**

Text with formatting that describes what is contained in the document

Code Chunks

Run Code Chunks up to *here* **[Cmd/Cntrl + Shift/Opt + P]**

Change Code Chunk *options*

RMarkdown Output

rmarkdown-sample.Rmd

Knit * | Insert | Run

```
1  ---
2  title: "Hello World"
3  author: "JJB"
4  date: "1/19/2018"
5  output: html_document
6  ---
7
8  ```{r setup, include=FALSE}
9  knitr::opts_chunk$set(echo = TRUE)
10 ```
11
12 ## R Markdown
13
14 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>.
15
16 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:
17
18 ```{r cars}
19 summary(cars)
20 ```
21
22 ## Including Plots
23
24 You can also embed plots, for example:
25
26 ```{r pressure, echo=FALSE}
27 plot(pressure)
28 ```
29
30 Note that the `echo = FALSE` parameter was added to the code chunk to
   prevent printing of the R code that generated the plot.
31
```

2:1 | Hello World | R Markdown

Knit

[Cmd/Cntrl + Shift + K]

Hello World

JJB

1/19/2018

R Markdown

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.

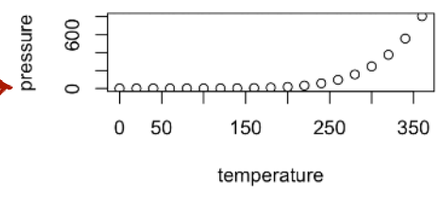
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
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
```

Including Plots

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 plot.

---

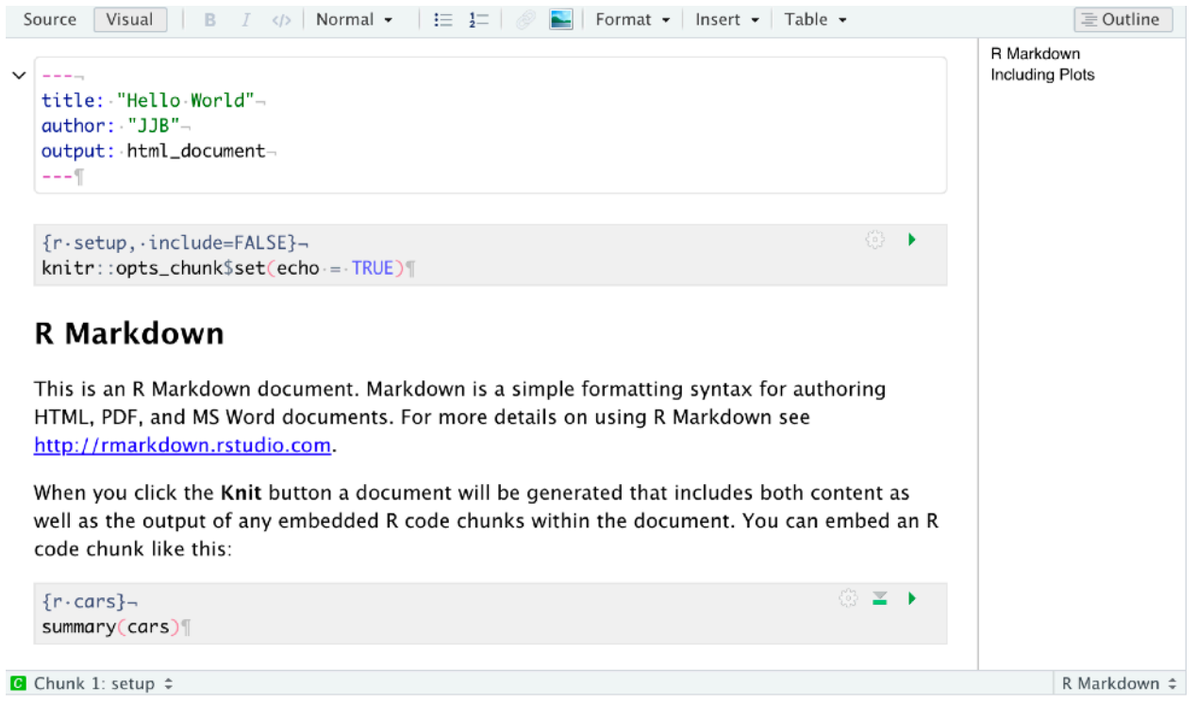
* Press "Knit" is the same as typing: **rmarkdown::render("rmarkdown-sample.Rmd")**

What You See Is What You Get

Write in a Word-like style, but retain the Markdown syntax.

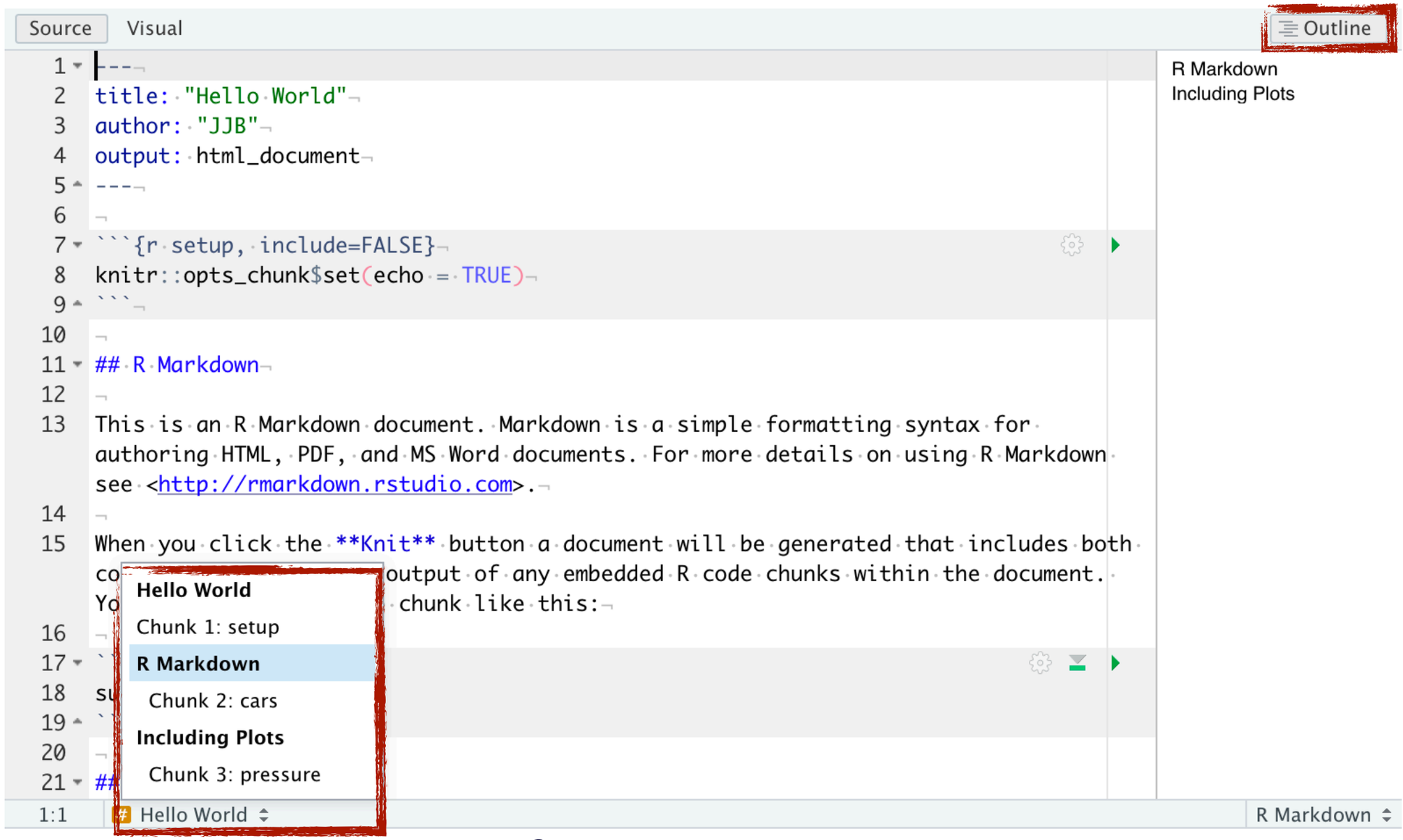
Click

Markdown Language Editor



WYSIWYG Editor

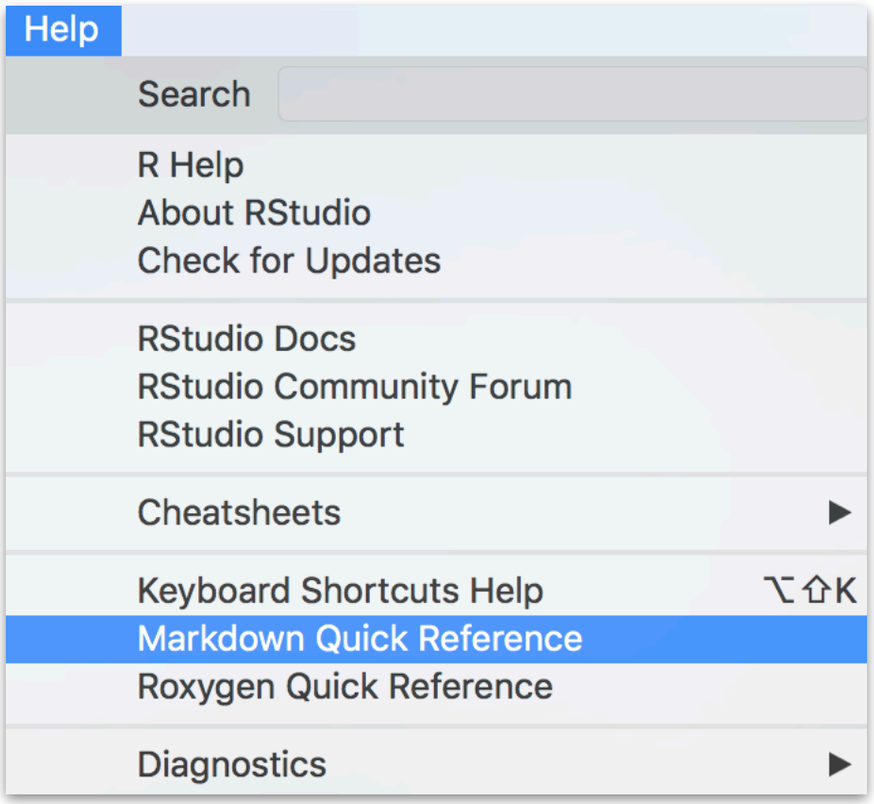
Navigate Document

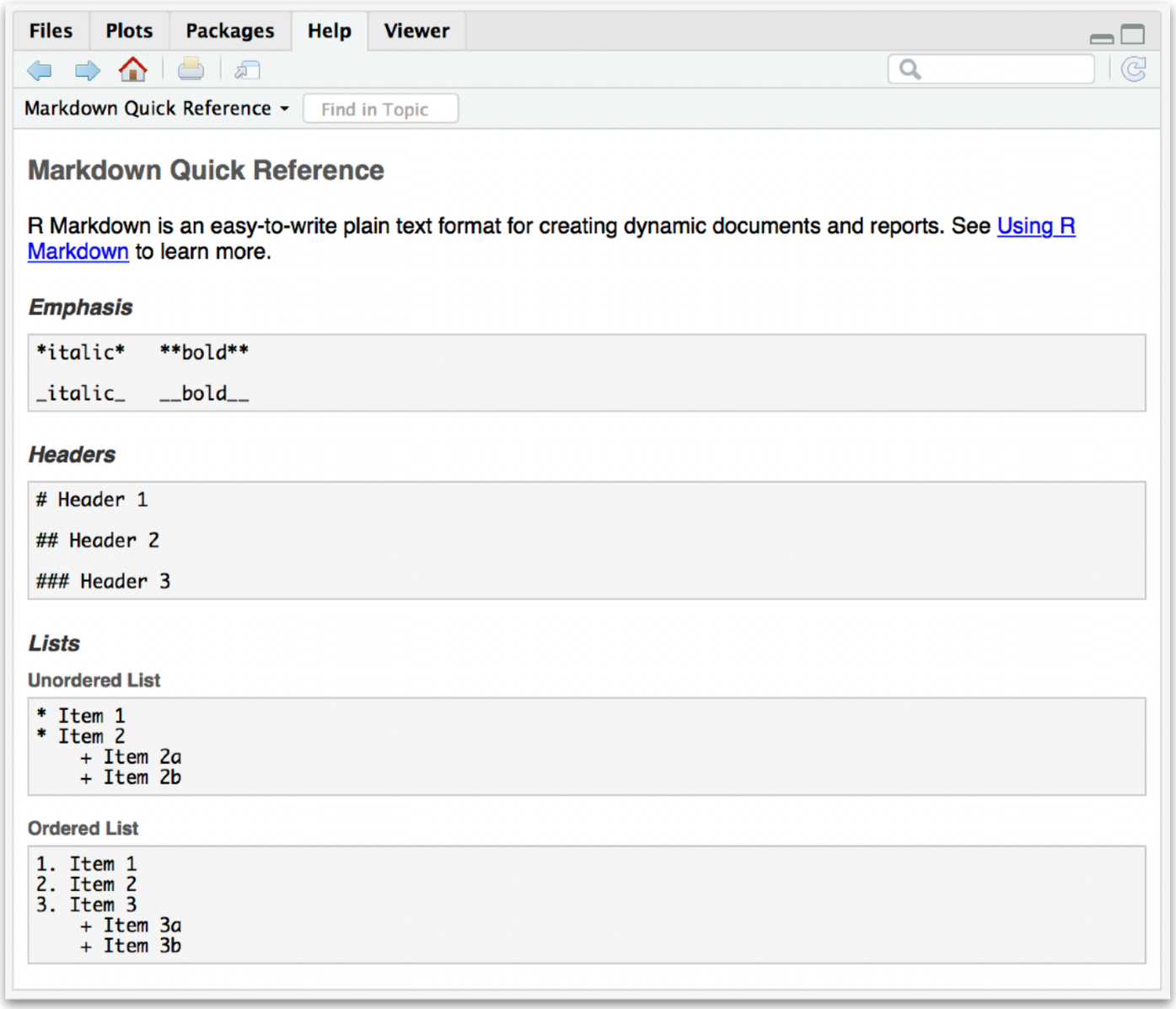


Writing in RMarkdown

Help Writing

… markdown syntax guide built-in to RStudio …





Code Demo

... writing in rmarkdown ...

Text Emphasis & Links

… making text stand out and pointing to external resources …

Code

```
Writing text with emphasis in
 *italics*, **bold** and `code style`.
```

→

Output

Writing text with emphasis in
italics, **bold** and code style.

```
Line breaks create a new paragraph.

Links can be hidden e.g.
[illinois](www.illinois.edu) or
not <http://illinois.edu> .
```

→

Line breaks create a new paragraph.

Links can be hidden e.g.
illinois or not http://illinois.edu.

Lists

… ordered and unordered …

Code

My **un**ordered list:

- Write Selection Simulation
- Conference Abstracts
    - UseR
    - Learning at Scale

→

Output

My **un**ordered list:

- Write Selection Simulation
- Conference Abstracts
    - UseR
    - Learning at Scale

My **ordered** list:

1. Apples
1. Bananas
1. Chobani
    1. Pineapple
    1. Everything else

→

My **ordered** list:

1. Apples
2. Bananas
3. Chobani
    1. Pineapple
    2. Everything else

---

* Make sure a new line (space) exists between text and the first list item.

** Indent four spaces to create a new level in the list.

Quotes & Math

… imparting wisdom and funky symbols …

Code

```
> "Never gonna give you up,
>  never gonna let you down..."
>
> --- Rick Astley
```

→

Output

"Never gonna give you up, never gonna let you down..."

– Rick Astley

```
Inline math $a^2 + b^2 = c^2$

Display math (centered math) $$1 - x = y$$
```

→

Inline math $a^2 + b^2 = c^2$

Display math (centered math)

$$a^2 + b^2 = c^2$$

LaTeX help: https://artofproblemsolving.com/wiki/index.php/LaTeX:Symbols

Tables [*] [**]

… organizing output in a rectangular form …

Code

```
| Left                      | Center                | Right  |
|---------------------------|:----------------------|---------:|
| Hey, check it out         | Colons provide  |    873  |
| its **Markdown**  | alignment thus  |   1000 |
| right in the table        | *centered* text |          |
```

Output

Left	Center	Right
Hey, check it out	Colons provide	873
its **Markdown**	alignment thus	1000
right in the table	*centered* text	

---

* Markdown table generator: https://www.tablesgenerator.com/markdown_tables

** Other supported markdown table styles:
http://rmarkdown.rstudio.com/authoring_pandoc_markdown.html#tables

Code Chunks

… how we embed code into the document …

Example

```
```{r chunk-label}
# Your R code here
```
```

Code

```
We're embedding _R_ code
**into** a report!!
```

```
```{r add_nums}
1 + 2
```
```

Output

We're embedding *R* code **into** a report!!

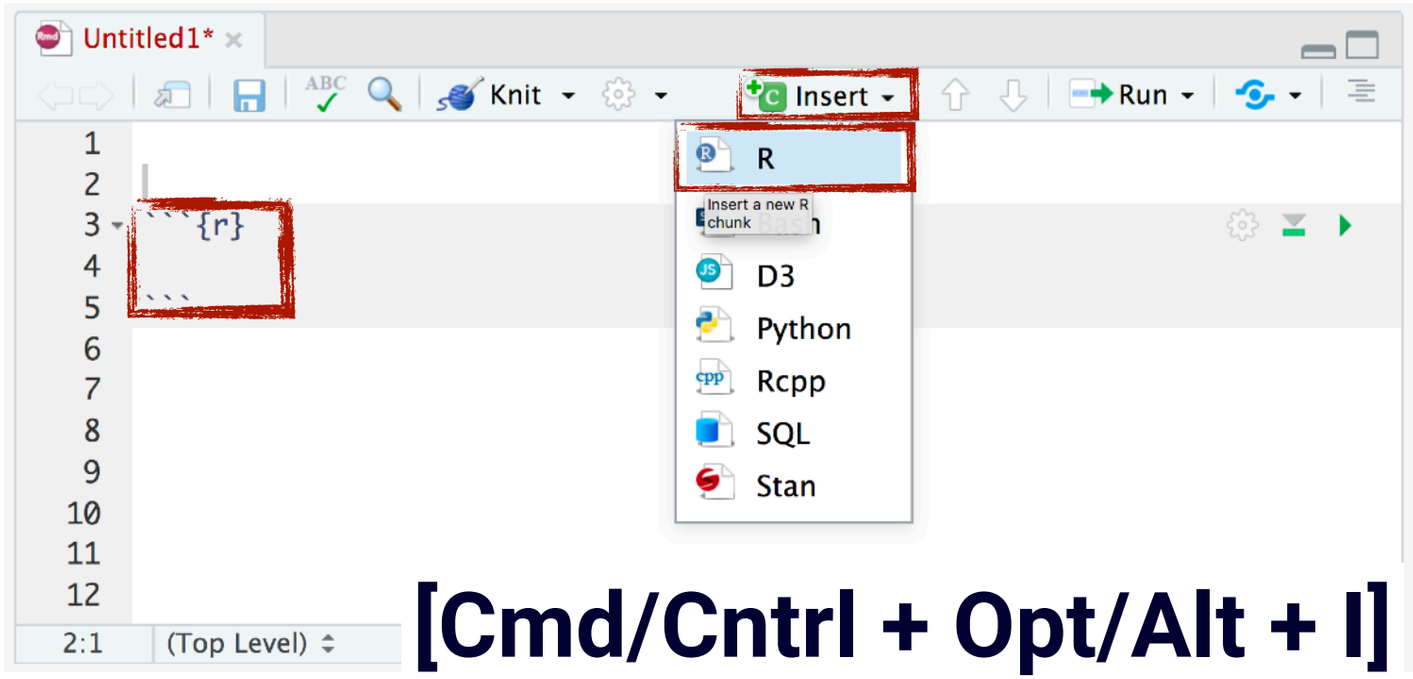
```
1 + 2
```

```
## [1] 3
```

* Syntax or code highlighting is retained.

Ways to Create a Code Chunk

… dropdown menus, short cuts, and keyboard movements …



[Cmd/Cntrl + Opt/Alt + I]

https://support.apple.com/en-us/HT207358

Pets or Livestock

… to label or not to label …

Console | Terminal | **R Markdown** | Jobs

```
processing file: rmarkdown-sample-names.Rmd
  |.....................                                 |  25%
  ordinary text without R code

  |.....................................                 |  50%
label: add_num
  |.............................................         |  75%
  ordinary text without R code

  |.....................................................| 100%
label: unnamed-chunk-1
```

```
```{r add_nums}
1 + 2
```
```

```
```{r}
1 + 2
```
```

Code Chunk Options

… customizing how the embedded code/results appear …

```
```{r chunk_label, chunkopt1 = value1, chunkopt2 = value2}
# Code here
```
```

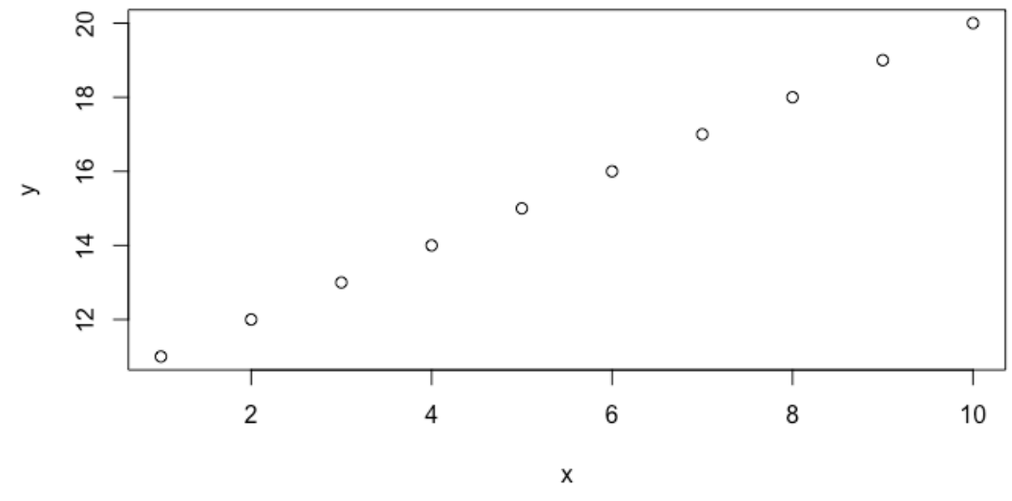
Code

Let's hide the _R_ code from showing up in the report!

```
```{r ex-hide, echo = FALSE}
x = 1:10
y = 11:20
plot(x, y)
```
```

Output

Let's hide the *R* code from showing up in the report!



Common Chunk Options *

… most frequently used options …

	Chunk Option	Description
**	cache = TRUE	Stores to file code chunk results
	eval = FALSE	Doesn't evaluate/run the code
	eval = 2:3	Evaluate only lines 2 and 3
	echo = FALSE	Doesn't display the code
	echo = -c(1,3,4)	Hide only lines 1, 3 and 4
	results = 'hide'	Doesn't display code output
	include = FALSE	Doesn't display code or output
	fig.path = 'img/'	File path to store images
	fig.align = 'center'	Align image in the center
	fig.width = #	Sets width of figure
	fig.height = #	Sets height of figure
	warning = FALSE	Hides R's warning messages
	message = FALSE	Hides R's note messages

---

* More options: http://yihui.name/knitr/options
** Caching is very useful and dangerous, more on this later

eval vs. **echo**

… difference between code options …

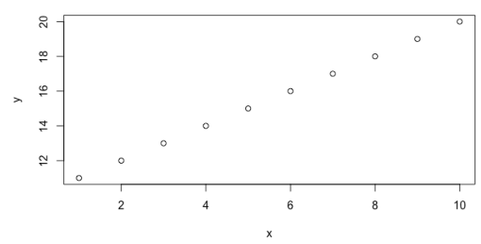
Code

````
```{r ex-orig}
x = 1:10; y = 11:20
plot(x, y)
```
````

→

Output

x = 1:10; y = 11:20
plot(x, y)



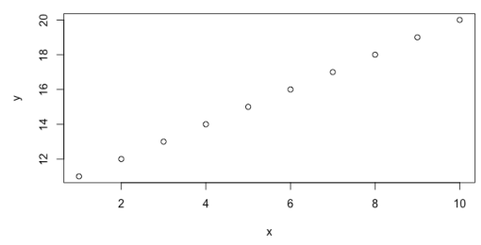

````
```{r ex-not-run, eval = FALSE}
x = 1:10; y = 11:20
plot(x, y)
```
````

→

x = 1:10; y = 11:20
plot(x, y)

````
```{r ex-hide-code, echo = FALSE}
x = 1:10; y = 11:20
plot(x, y)
```
````

→



Embedding In-line *R* Code

… evaluating an *R* expression without a code chunk …

Inline _R_ code is possible with **`r expression_here`**

Code

````
```{r calc-values, echo = FALSE}
x = 1:10
x_mu = mean(x)
x_sd = sd(x)
```
````

Pre-compute values in code chunks

Text

```
The _mean_ of **x** is `r x_mu` and
the _standard deviation_ is `r x_sd`.
```

Output

The *mean* of **x** is 5.5 and the *standard deviation* is 3.02765.

Access variable contents

Inserting Images

… adding non-*R* produced images …

Code

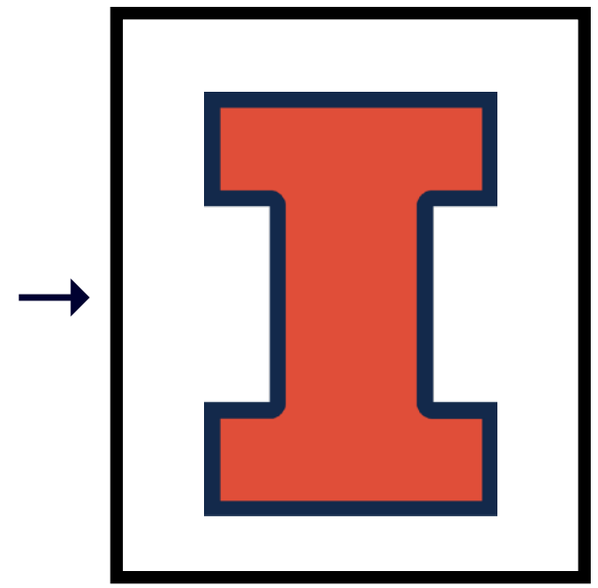
Output

Path to File

```
![Here is a Relative Path](img/block-i.png)
```

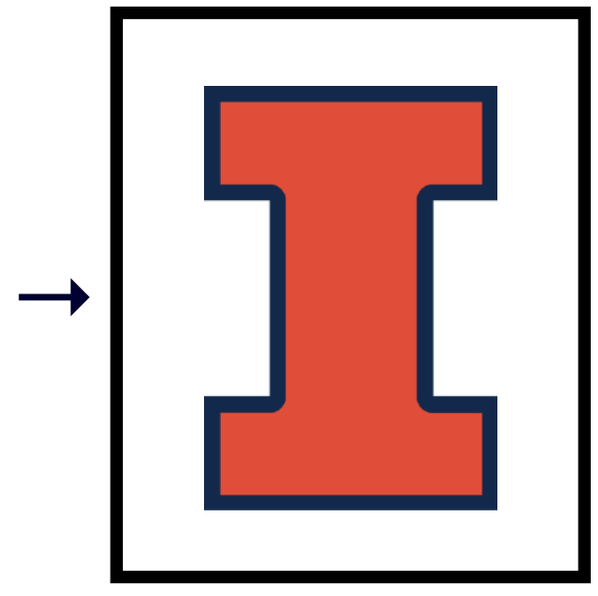
Relative Path

Caption



```
![Here is an Absolute path](C:/jjb/img/block-i.png)
```

Absolute Path



---

* Relative paths are the *best* to use if the goal is to share your work with others as they are operating system independent. More next!

** For instance, do you have a user called "jjb" on your computer with a folder "img"?

Recap

- **R Packages are extensions to *R***
    - install.packages('pkg') run **once per** computer / cloud project.
    - library('pkg') use for **every** analysis to enable the extension.
- **RMarkdown**
    - Combine *R* code with analysis in a reproducible manner.