

# An Introduction to knitr and RMarkdown

<https://github.com/sahirbhatnagar/knitr-tutorial>

---

Sahir Bhatnagar

August 12, 2015

McGill University

# Acknowledgements

- Toby, Matthieu, Vaughn, Ary
- [Maxime Turgeon](#) (Windows)
- [Kevin McGregor](#) (Mac)
- Greg Voisin
- [Don Knuth](#) (T<sub>E</sub>X)
- [Friedrich Leisch](#) (Sweave)
- [Yihui Xie](#) (knitr)
- [John Gruber](#) (Markdown)
- [John MacFarlane](#) (Pandoc)
- You



# Disclaimer #1



R Markdown v2



L<sup>A</sup>T<sub>E</sub>X

*I don't work for, nor am I an author of any of these packages. I'm just a messenger.*



# Objectives for today

- Create a reproducible document (pdf, html)

# Objectives for today

- Create a reproducible document (pdf, html)



C'est parti

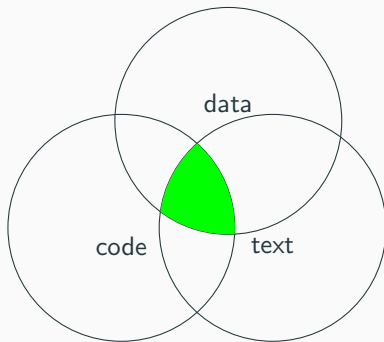


What?

---

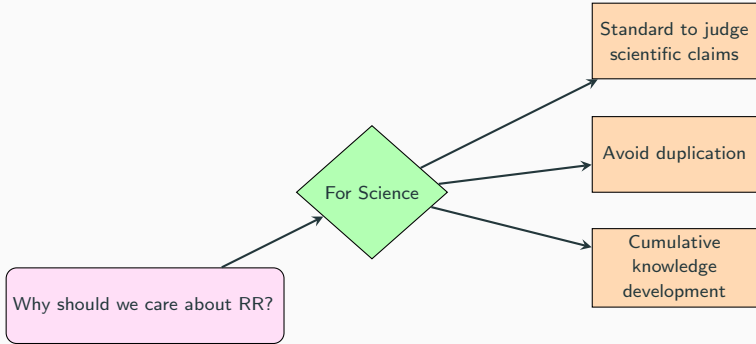


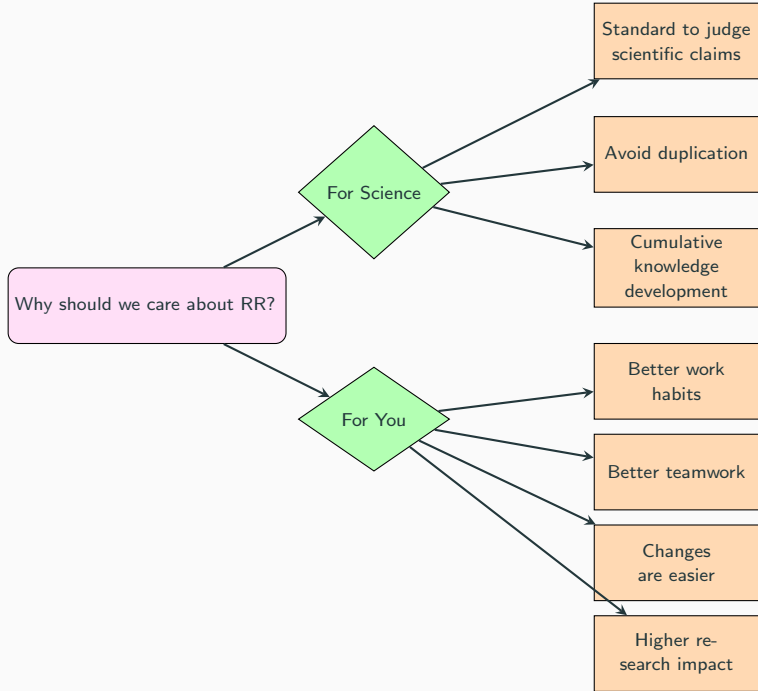
# What is needed for Reproducible research?



Why?

---





001-motivating-example

---

# A Motivating Example

*Demonstrate:* [001-motivating-example](#)

How?

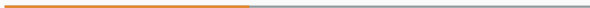
---

## Free and Open Source Software

- RStudio: Creating, managing, compiling documents
- $\text{\LaTeX}$ : Markup language for typesetting a pdf
- Markdown: Markup language for typesetting an html
- R: Statistical analysis language
- knitr: Integrate  $\text{\LaTeX}$  and R code. Based on Prof. Friedrich Leisch's [Sweave](#)

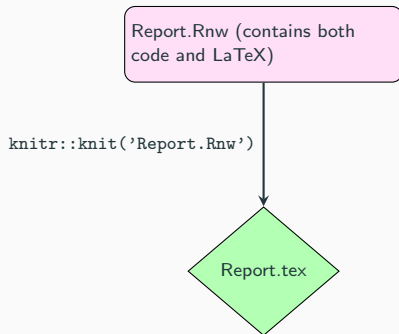


knitr



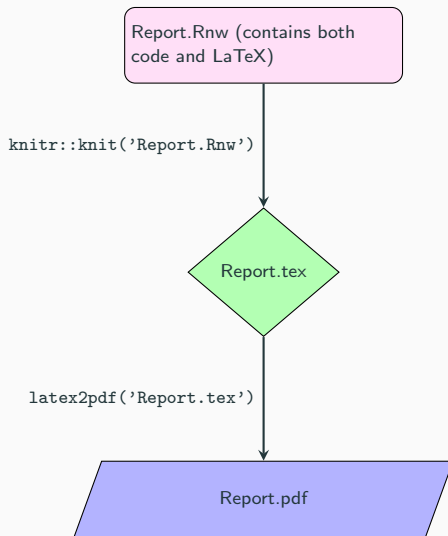
# What knitr does

**LaTeX** example:



# What knitr does

**LaTeX** example:

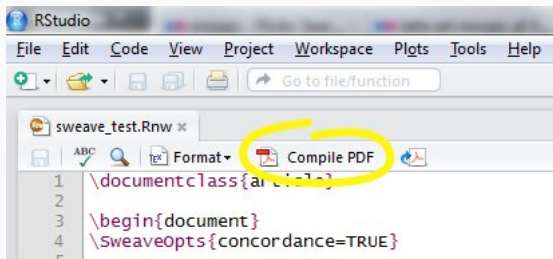


# Compiling a .Rnw document

The two steps on previous slide can be executed in one command:

```
knitr::knit2pdf()
```

or in RStudio:



# Incorporating R code

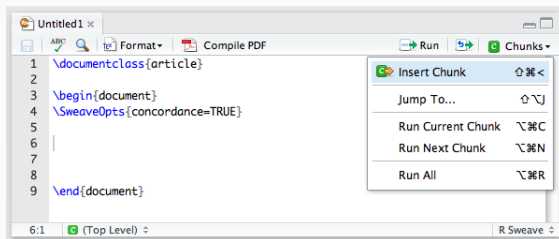
- Insert R code in a **Code Chunk** starting with

`<< >>=`

and ending with

`@`

In RStudio:



## Example 1: Show code and results

```
<<example-code-chunk-name, echo=TRUE>>=  
x <- rnorm(50)  
mean(x)  
@
```

produces

```
x <- rnorm(50)  
mean(x)  
  
## [1] 0.031
```

## Example 2: Tidy code

```
<<example-code-chunk-name2, echo=TRUE, tidy=TRUE>>=  
for(i in 1:5){ print(i+3)}  
@
```

produces

```
for (i in 1:5) {  
  print(i + 3)  
}
```

```
## [1] 4  
## [1] 5  
## [1] 6  
## [1] 7  
## [1] 8
```

## Example 2.2: don't show code

```
<<example-code-chunk-name3, echo=FALSE>>=  
for(i in 1:5){ print(i+3)}  
@
```

produces

```
## [1] 4  
## [1] 5  
## [1] 6  
## [1] 7  
## [1] 8
```



## Example 2.3: don't evaluate and don't show the code

```
<<example-code-chunk-name4, echo=FALSE, eval=FALSE>>=  
for(i in 1:5){ print(i+3)}  
@
```

produces

# R output within the text

- Include R output within the text
- We can do that with “S-expressions” using the command `\Sexpr{...}`

## Example:

The iris dataset has `\Sexpr{nrow(iris)}` rows and  
`\Sexpr{ncol(iris)}` columns

produces

The iris dataset has 150 rows and 5 columns

# Include a Figure

```
<<fig.ex, fig.cap='Linear Regression',fig.height=3,fig.width=3>>=  
plot(mtcars[ , c('disp','mpg')])  
fit <- lm(mpg ~ disp , data = mtcars)  
abline(fit,lwd=2)  
@
```

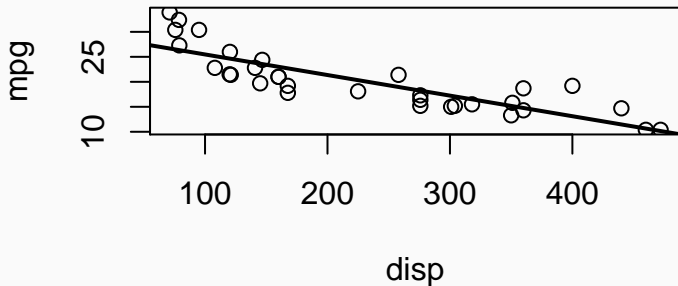


Figure 1: Linear regression

# Include a Table

```
<<table.ex, results='asis'>>=
library(xtable)
tab <- xtable(iris[1:5,1:5],caption='Sample of Iris data')
print(tab, include.rownames=FALSE)
@
```

```
library(xtable)
tab <- xtable(iris[1:5,1:5], caption = 'Sample of Iris data')
print(tab, include.rownames = F)
```

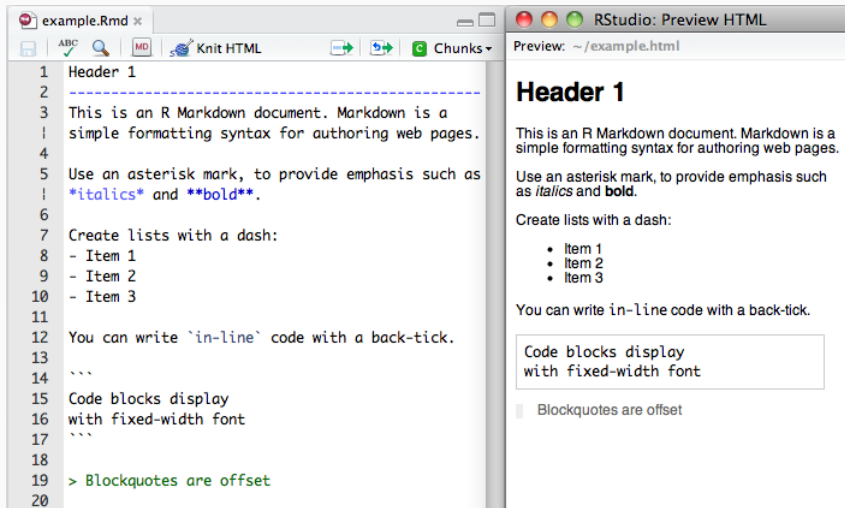
Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.10	3.50	1.40	0.20	setosa
4.90	3.00	1.40	0.20	setosa
4.70	3.20	1.30	0.20	setosa
4.60	3.10	1.50	0.20	setosa
5.00	3.60	1.40	0.20	setosa

Table 1: Sample of Iris data

## RMarkdown

---

# Markdown: HTML without knowing HTML



The image shows a side-by-side comparison of an R Markdown document and its rendered HTML output in RStudio. The left pane, titled 'example.Rmd', shows the source text with line numbers 1 through 20. The right pane, titled 'RStudio: Preview HTML', shows the rendered result with the title 'Preview: ~/example.html'.

**Source Document (example.Rmd):**

```
1 Header 1
2 -----
3 This is an R Markdown document. Markdown is a
4 | simple formatting syntax for authoring web pages.
5 Use an asterisk mark, to provide emphasis such as
6 | italics and bold.
7 Create lists with a dash:
8 - Item 1
9 - Item 2
10 - Item 3
11
12 You can write `in-line` code with a back-tick.
13
14 ```
15 Code blocks display
16 with fixed-width font
17 ```
18
19 > Blockquotes are offset
20
```

**Rendered HTML Preview:**

## Header 1

This is an R Markdown document. Markdown is a simple formatting syntax for authoring web pages.

Use an asterisk mark, to provide emphasis such as *italics* and **bold**.

Create lists with a dash:

- Item 1
- Item 2
- Item 3

You can write in-line code with a back-tick.

Code blocks display with fixed-width font

Blockquotes are offset

# R + Markdown = RMarkdown

chunks.Rmd x

Knit HTML Chunks

```
1 R Code Chunks
2 =====
3
4 With R Markdown, you can insert R code chunks including plots:
5
6 ```{r qplot, fig.width=4, fig.height=3,
7   message=FALSE}
8 # quick summary and plot
9 library(ggplot2)
10 summary(cars)
11 qplot(speed, dist, data=cars) +
12   geom_smooth()
13
```

RStudio: Preview HTML

Preview: ~/chunks.html Save As Publish

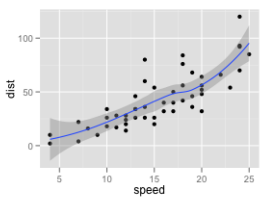
## R Code Chunks

With R Markdown, you can insert R code chunks including plots:

```
# quick summary and plot
library(ggplot2)
summary(cars)
```

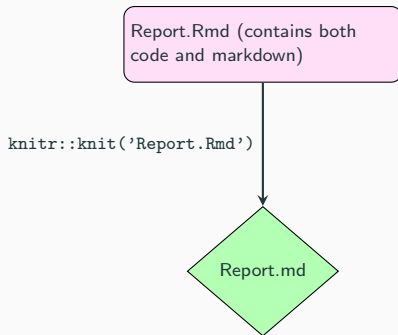
##	speed	dist
##	Min. : 4.0	Min. : 2
##	1st Qu.:12.0	1st Qu.: 26
##	Median :15.0	Median : 36
##	Mean :15.4	Mean : 43
##	3rd Qu.:19.0	3rd Qu.: 56
##	Max. :25.0	Max. :120

```
qplot(speed, dist, data = cars) + geom_smooth()
```



# What rmarkdown does

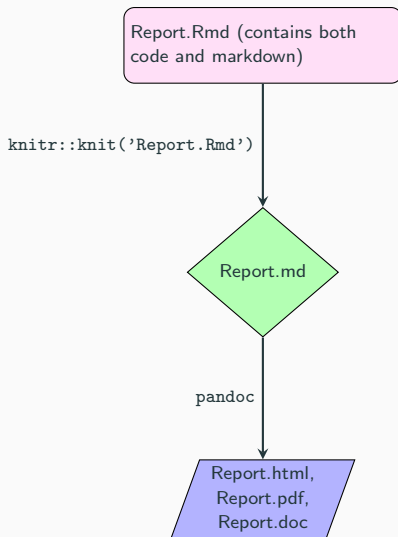
RMarkdown example:





# What rmarkdown does

RMarkdown example:

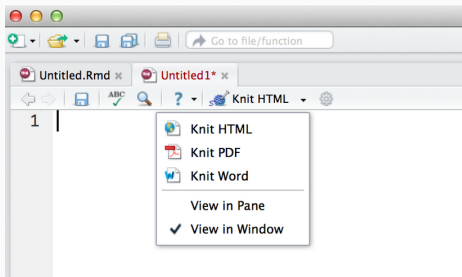


# Compiling a .Rmd document

The two steps on previous slide can be executed in one command:

```
rmarkdown::render()
```

or in RStudio:



## Final Remarks

---

# How to choose between $\text{\LaTeX}$ and Markdown ?

$\text{\LaTeX}$  {  
math/stat symbols  
beamer presentations  
customized documents  
publish to journals, arXiv

Markdown {  
quick and easy reports  
use javascript libraries  
interactive plots  
publish to websites

# Opinion: Reproducible research can still be wrong: Adopting a prevention approach

**Jeffrey T. Leek<sup>a,1</sup> and Roger D. Peng<sup>b</sup>**

<sup>a</sup>Associate Professor of Biostatistics and Oncology and <sup>b</sup>Associate Professor of Biostatistics,  
Johns Hopkins University, Baltimore, MD

---

computational tools such as knitr, iPython notebook, LONI, and Galaxy (8) have simplified the process of distributing reproducible data analyses.

# Always Remember ...

$$\text{Reproducibility} \propto \frac{1}{\text{copy paste}}$$

# Is the juice worth the squeeze?



# Session Info

- R version 3.2.1 (2015-06-18), x86\_64-pc-linux-gnu
- Base packages: base, datasets, graphics, grDevices, methods, stats, utils
- Other packages: data.table 1.9.4, dplyr 0.4.1, ggplot2 1.0.1, knitr 1.10.5, xtable 1.7-4
- Loaded via a namespace (and not attached): assertthat 0.1, chron 2.3-45, colorspace 1.2-6, DBI 0.3.1, digest 0.6.8, evaluate 0.7, formatR 1.2, grid 3.2.1, gtable 0.1.2, highr 0.5, magrittr 1.5, MASS 7.3-43, munsell 0.4.2, parallel 3.2.1, plyr 1.8.3, proto 0.3-10, Rcpp 0.12.0, reshape2 1.4.1, scales 0.2.5, stringi 0.5-5, stringr 1.0.0, tools 3.2.1

Slides made with Beamer [mtheme](#)