

# Scientific reports

Making research results easily accessible and  
reproducible

Bioinformatics Workshop

November 23, 2016

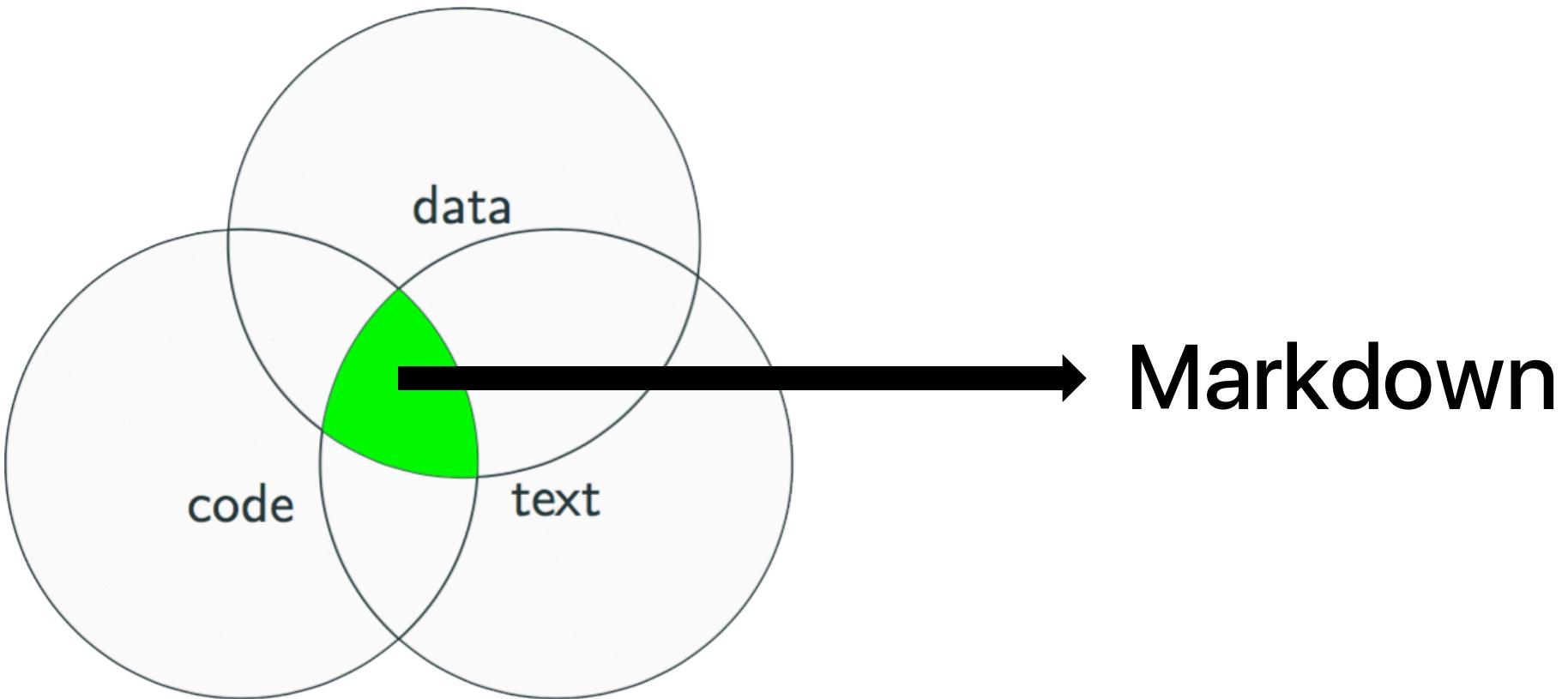
Omar Wagih

[wagih@ebi.ac.uk](mailto:wagih@ebi.ac.uk)

# Slides and exercises

<https://github.com/omarwagih/eipp/>

# What do you mean by scientific reports?

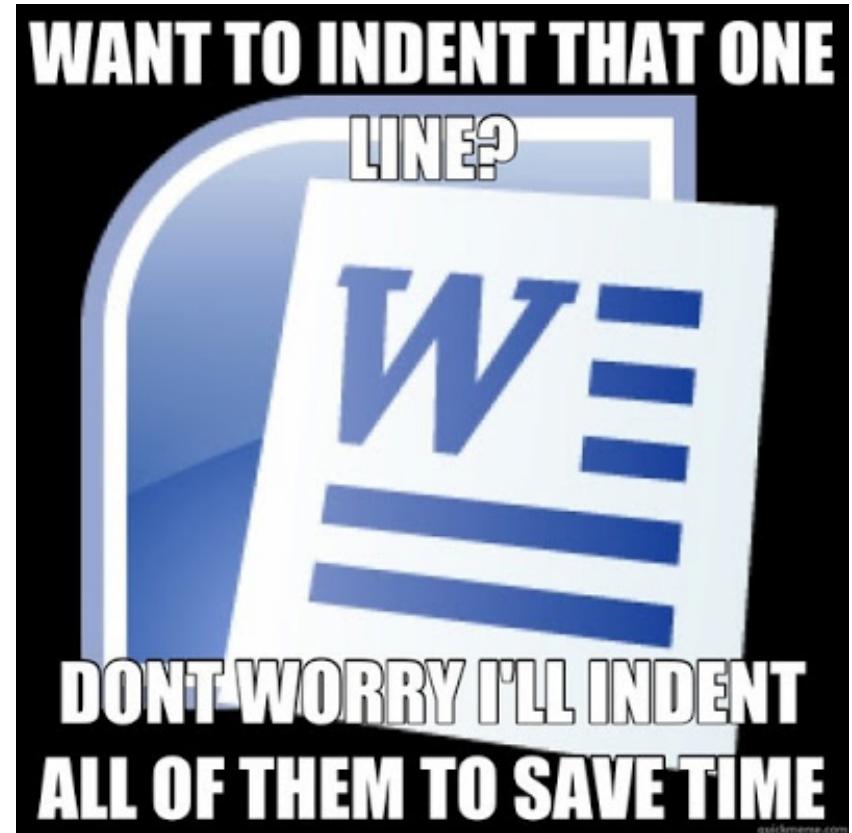


# What is markdown?

- Markdown is a particular type of markup language
- Markup languages are designed for producing documents from plain text
- Some of you may be familiar with *LaTeX*. This is another (less human friendly) markup language for creating documents
- *LaTeX* more flexible, but it is restricted to pdf and has a **much** greater learning curve

# Why markdown?

- Easy to learn and use
- Focus on content, rather than coding and debugging errors
- Designed to be converted to HTML, but can be converted to other formats like PDF... and if you really want, Microsoft Word



# Markdown to HTML example

Markdown syntax

```
Header 1  
=====
```

```
Header 2  
-----
```

```
### Header 3
```

```
This is regular  
text.
```

```
> This is a  
blockquote.  
>  
> This is the  
second paragraph  
in the blockquote.  
>  
> ## This is an H2  
in a blockquote
```

Resulting HTML

```
<h1>Header 1</h1>
```

```
<h2>Header 2</h2>
```

```
<h3>Header 3</h3>
```

```
<p>This is regular  
text.</p>
```

```
<blockquote>
```

```
<p>This is a  
blockquote.</p>
```

```
<p>This is the second  
paragraph in the  
blockquote.</p>
```

```
<h2>This is an H2 in a  
blockquote</h2>
```

```
</blockquote>
```

# Markdown demo

<http://dillinger.io/>

# Rmarkdown

- Rmarkdown is an adaptation of markdown
- Allows you to dynamically integrate R code, plots and tables with traditional markdown
- The result: pretty pretty reports

The screenshot shows the RStudio interface with two panes. The left pane displays an RMarkdown file named 'chunks.Rmd' with the following content:

```
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,
| message=FALSE}
7 # quick summary and plot
8 library(ggplot2)
9 summary(cars)
10 qplot(speed, dist, data=cars) +
11     geom_smooth()
12 ...
13
```

The right pane shows the rendered output of the R code:

## 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()
```

A scatter plot is displayed, showing the relationship between 'speed' (x-axis, ranging from 5 to 25) and 'dist' (y-axis, ranging from 0 to 100). The plot includes black dots representing individual data points, a solid blue line representing a smooth fit, and a light gray shaded area representing the confidence interval of the fit.

# Why should you care?

- Better work habits
- Reproducible
- Easier to make changes
- Sharable
- Impress your boss!



Higher  
research  
impact

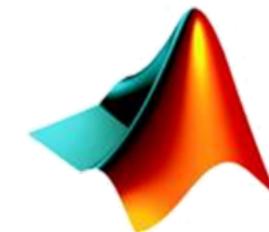
# How to generate these reports?

- iPython (now Jupyter)
- knitr
- MATLAB Guide
- Sage Notebook
- ...

IP[y]: Notebook



sage



# An introduction to Rmarkdown

- Basics

## Syntax

```
Plain text  
  
End a line with two spaces  
to start a new paragraph.  
  
*italics* and _italics_  
  
**bold** and __bold__  
  
superscript^2^  
  
~~strikethrough~~  
  
[link] (www.rstudio.com)  
  
# Header 1  
  
## Header 2  
  
### Header 3  
  
#### Header 4  
  
##### Header 5  
  
##### Header 6
```

## Becomes

Plain text  
  
End a line with two spaces to start a new paragraph.

*italics* and *italics*  
**bold** and **bold**  
superscript<sup>2</sup>  
strikethrough  
[link](#)

## Header 1

## Header 2

## Header 3

Header 4  
Header 5  
Header 6

# An introduction to Rmarkdown

- Basics (continued)

Syntax	Becomes
endash: --	endash: –
emdash: ---	emdash: —
ellipsis: ...	ellipsis: ...
inline equation: \$A = \pi * r^2\$	inline equation: $A = \pi * r^2$
image:	image: 
horizontal rule (or slide break):	horizontal rule (or slide break):
***	
> block quote	block quote

# An introduction to Rmarkdown

- Lists and tables

Syntax	Becomes												
<pre>* unordered list * item 2   + sub-item 1   + sub-item 2</pre>	<ul style="list-style-type: none"><li>• unordered list</li><li>• item 2<ul style="list-style-type: none"><li>◦ sub-item 1</li><li>◦ sub-item 2</li></ul></li></ul>												
<pre>1. ordered list 2. item 2   + sub-item 1   + sub-item 2</pre>	<ol style="list-style-type: none"><li>1. ordered list</li><li>2. item 2<ul style="list-style-type: none"><li>◦ sub-item 1</li><li>◦ sub-item 2</li></ul></li></ol>												
<table><thead><tr><th>Table Header</th><th>Second Header</th></tr></thead><tbody><tr><td>Table Cell</td><td>Cell 2</td></tr><tr><td>Cell 3</td><td>Cell 4</td></tr></tbody></table>	Table Header	Second Header	Table Cell	Cell 2	Cell 3	Cell 4	<table><thead><tr><th>Table Header</th><th>Second Header</th></tr></thead><tbody><tr><td>Table Cell</td><td>Cell 2</td></tr><tr><td>Cell 3</td><td>Cell 4</td></tr></tbody></table>	Table Header	Second Header	Table Cell	Cell 2	Cell 3	Cell 4
Table Header	Second Header												
Table Cell	Cell 2												
Cell 3	Cell 4												
Table Header	Second Header												
Table Cell	Cell 2												
Cell 3	Cell 4												

# An introduction to Rmarkdown

## Code chunks

- Separate “chunks” of code into groups
- Each group has specific output (e.g. a single plot)

Chunk 1

```
```{r}
print("hello world")
````
```

Chunk 2

```
```{r}
plot(1:10)
````
```

# An introduction to Rmarkdown

- Basic R code “chunks”
- Specify R code using
- Example:

```
```{r}  
<code>  
```
```

## Syntax

Make a code chunk with three back ticks followed by an r in braces. End the chunk with three back ticks:

```
```{r}  
paste("Hello", "World!")  
```
```

## Becomes

Make a code chunk with three back ticks followed by an r in braces. End the chunk with three back ticks:

```
paste("Hello", "World!")
```

```
## [1] "Hello World!"
```

# An introduction to Rmarkdown

- Basic R code “chunks”
- Specify options:
  - **eval** should the code be run? TRUE or FALSE
  - **echo** should the code be printed? TRUE or FALSE

```
```{r eval=TRUE, echo=FALSE}  
<code>  
```
```

## Syntax

Add chunk options within braces. For example, `echo=FALSE` will prevent source code from being displayed:

```
```{r eval=TRUE, echo=FALSE}  
paste("Hello", "World!")  
```
```

## Becomes

Add chunk options within braces. For example, `echo=FALSE` will prevent source code from being displayed:

```
## [1] "Hello World!"
```

# An introduction to Rmarkdown

- Basic R code “chunks”
- Specify inline R code using ``r <code>``
- Example:

## Syntax

Place code inline with a single back ticks. The first back tick must be followed by an R, like this ``r paste("Hello", "World!")``.

## Becomes

Place code inline with a single back ticks. The first back tick must be followed by an R, like this `Hello World!`

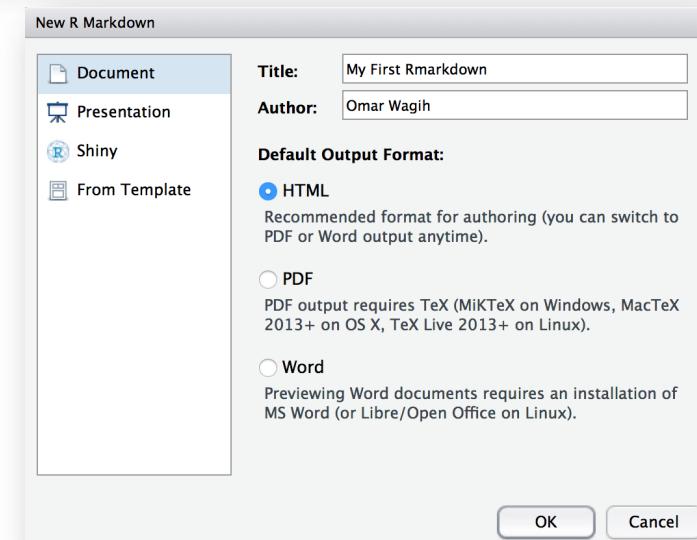
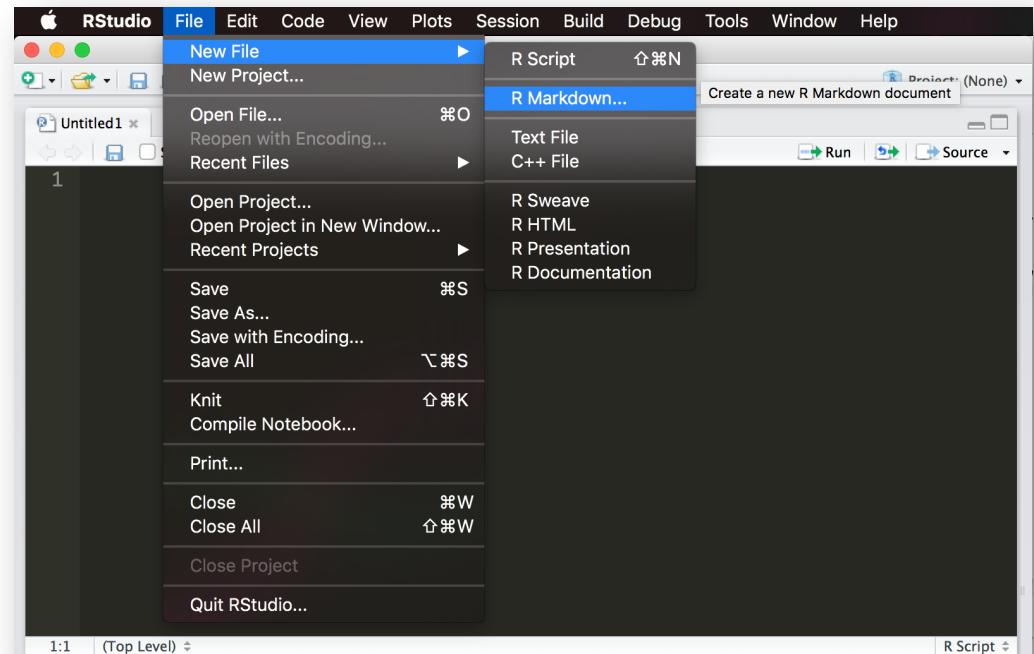
# Getting started

- Make sure you have installed the following if you haven't already done so
  - R (<https://cran.rstudio.com/>)
  - Rstudio (<https://www.rstudio.com/products/rstudio/download/>)
- **Get into pairs**



# Getting started

- Create a new Rmarkdown (Rmd) document
- **File → Newfile → R markdown...**
- Follow the prompt to install knitr (if it appears)
- Set a title for your document
- Enter your name in the author field and hit “OK”

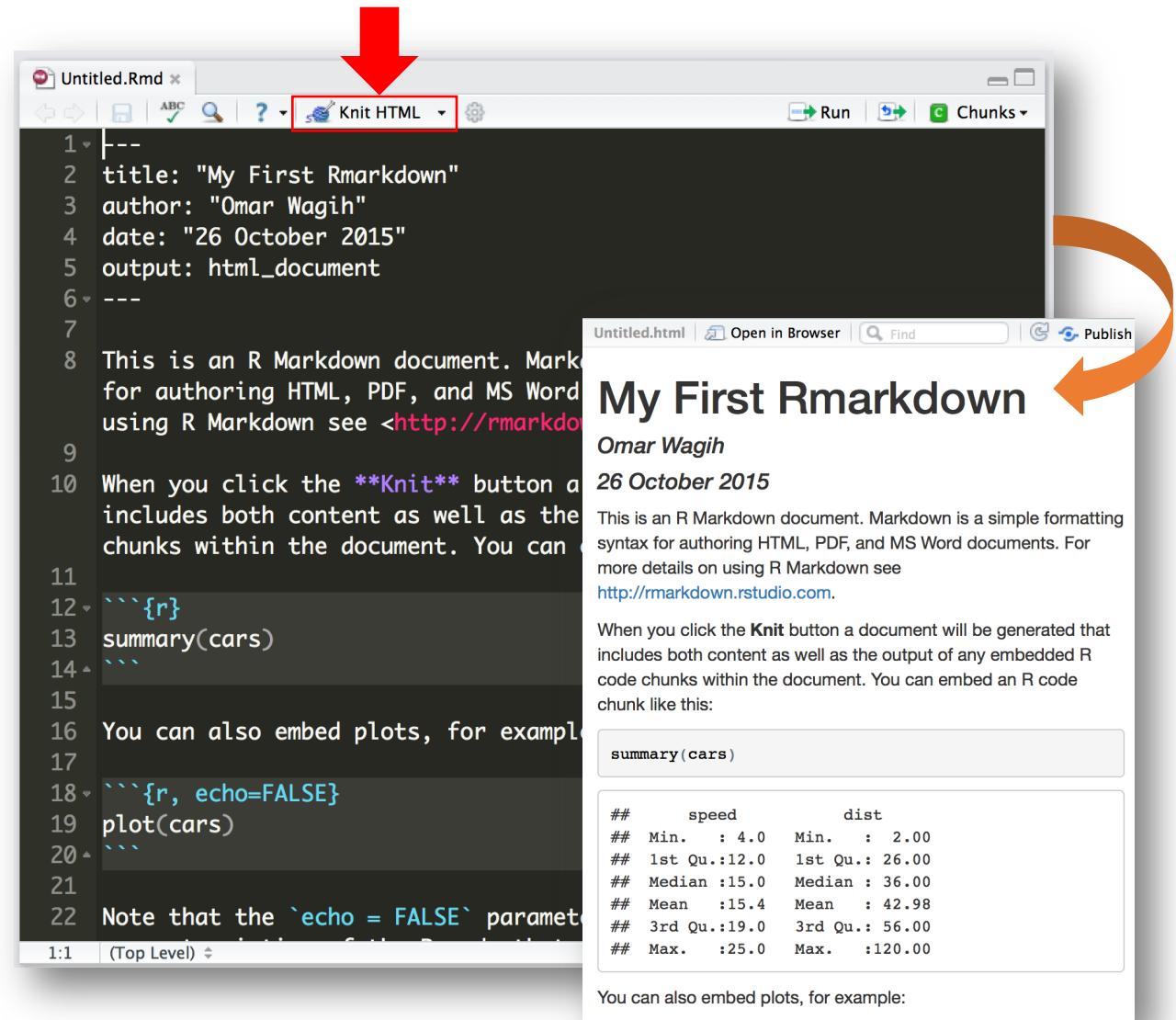
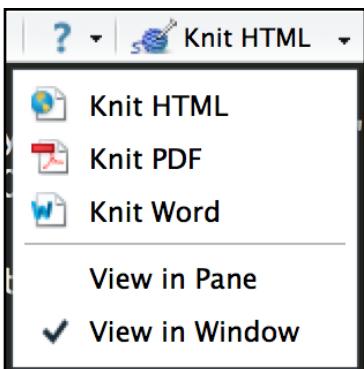


# Getting started

- Sample Rmarkdown
- Click Knit HTML to generate the HTML version



- Try generating different formats (e.g. pdf)



The screenshot shows the RStudio interface with an R Markdown file named "Untitled.Rmd". The "Knit HTML" button in the toolbar is highlighted with a red box and has a red arrow pointing down to it. The code editor shows R Markdown syntax, including code chunks and a summary table. The generated HTML output is shown in a separate window, also with a "Knit HTML" button highlighted with a red box and an orange arrow pointing back to it. The HTML output includes a title, author, date, and a summary table of car data.

```
1--  
2title: "My First Rmarkdown"  
3author: "Omar Wagih"  
4date: "26 October 2015"  
5output: html_document  
6---  
7  
8This 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>  
9  
10When 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:  
11

```
summary(cars)
```

  
12`

```
summary(cars)
```

`  
13  
14  
15  
16You can also embed plots, for example:  
17  
18`

```
plot(cars)
```

`  
19  
20  
21  
22Note that the `echo = FALSE` parameter can be used to suppress the  
23output of R code chunks.  
24  
25
```

Untitled.html | Open in Browser | Find | Publish

## My First Rmarkdown

Omar Wagih

26 October 2015

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:

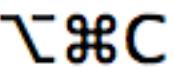
```
summary(cars)
```

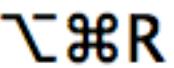
|            | speed | dist           |
|------------|-------|----------------|
| ## Min.    | 4.0   | Min. : 2.00    |
| ## 1st Qu. | 12.0  | 1st Qu.: 26.00 |
| ## Median  | 15.0  | Median : 36.00 |
| ## Mean    | 15.4  | Mean : 42.98   |
| ## 3rd Qu. | 19.0  | 3rd Qu.: 56.00 |
| ## Max.    | 25.0 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

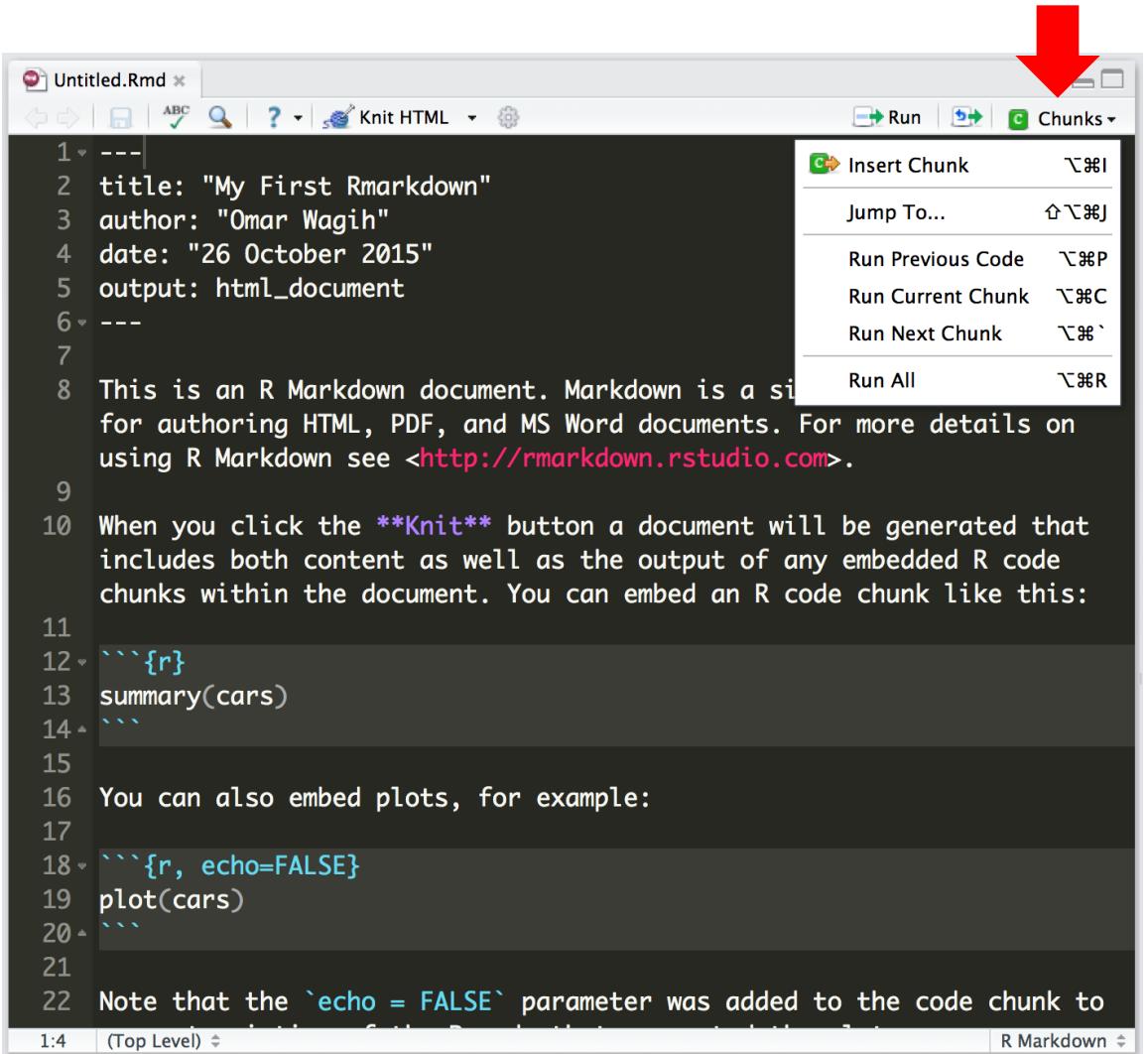
You can also embed plots, for example:

# Getting started

- You can run code from each chunk without generating the report

Run Current Chunk 

Run All 



The screenshot shows the RStudio interface with an R Markdown document titled "Untitled.Rmd". The Chunks menu is open, displaying options: Insert Chunk (⌘I), Jump To... (⇧⌘J), Run Previous Code (⌘P), Run Current Chunk (⌘C), Run Next Chunk (⌘`), and Run All (⌘R). A red arrow points to the Chunks icon in the top right corner of the window.

```
1 ---  
2 title: "My First Rmarkdown"  
3 author: "Omar Wagih"  
4 date: "26 October 2015"  
5 output: html_document  
6 ---  
7  
8 This is an R Markdown document. Markdown is a simple syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.  
9  
10 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:  
11  
12 ```{r}  
13 summary(cars)  
14 ```  
15  
16 You can also embed plots, for example:  
17  
18 ```{r, echo=FALSE}  
19 plot(cars)  
20 ```  
21  
22 Note that the `echo = FALSE` parameter was added to the code chunk to
```

# Hands-on tutorial

## **Exercise 1:** Simple markdown syntax (10 mins)

Aims:

- Familiarise yourself with basic markdown syntax
- Get comfortable with generating reports, and running code chunks

- Play around with basic rmarkdown
- Download **exercise 1 HTML**  
<https://github.com/omarwagih/btm-2015/>
- Try to generate a similar report
- There is also a **cheat sheet** there to help you
- You can play around with <http://dillinger.io/> for further help too

## Excercise 1

Your name

27 October 2015

### Rmarkdown

#### Excercise 1

##### Basic markdown syntax

Markdown is a text format that can be converted to formats such as HTML. The syntax is really easy. Here are some lists. Remember to use four spaces to indent.

- Markdown
  - is
  - awesome
- and easy
- to use

Numbered lists are also similar

1. We are
2. the EMBL
  - predocs

Here is a quote block

This text is written in markdown. Here is some *italic* text, you can also use *underscores*. Here is some **bold text**, you can also use **double underscores**.

I can superscript like this 27<sup>th</sup> and ~~strikeout~~ like this.

Creating a link is very easy. Here is one to [EMBL's awesome website yay](#).

Here is a simple table

| Table Header | Second Header |
|--------------|---------------|
| Table Cell   | Cell 2        |
| Cell 3       | Cell 4        |

It's also very easy to include an image from the internet or your computer



European Molecular  
Biology Laboratory

I can also write some code like this `print("hello world")`

End of excercise 1

# Hands-on tutorial

## **Exercise 2:** Basic code chunks and in-line code (10 mins)

Aims:

- Introduce code chunks with print statements, tables and plots
- Introduce basic code chunk options

# Hands-on tutorial

**Exercise 3:** Learning extra options for plotting (10 mins)

Aims:

- Learning more advanced chunk options for plotting

# Caching

- *caching*, to store away in hiding or for future use
- What do you do if running a chunk takes too long? Cache the results of the chunk so you don't need to run it each time
- Easy to enable
- If you make any changes to the chunk, R will re-cache

```
```{r cache=TRUE}  
Sys.sleep(10)  
paste("hello world")  
```
```

```
```{r cache=FALSE}  
Sys.sleep(10)  
paste("hello world")  
```
```

# Hands-on tutorial

## **Exercise 4:** Caching chunks and other tips

Aims:

- Learn how to cache chunks

# Hands-on tutorial

## **Exercise 5:** Putting it all together

Aims:

- Try to combine everything you've learned into one

# Bonus exercise: generating slides

- You can even generate nifty html slides with Rmarkdown
- Just add **“output: ioslides\_presentation”** to the header

```
---
```

```
title: "My awesome Rmarkdown presentation"
author: "EMBL predocs"
date: October 27, 2015
output: ioslides_presentation
---
```

- **#** = Header slide with no data
- **##** = Header for slide with content

# Hands-on tutorial

## **Exercise 6:** Generating slides

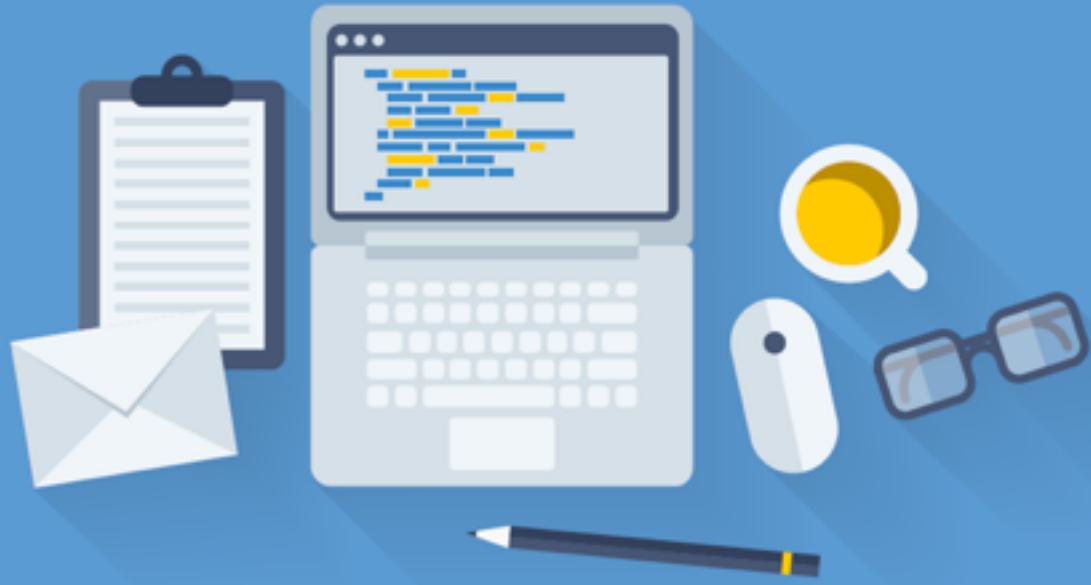
Aims:

- Learn how to make basic html slides using Rmarkdown
- Integrate R code and images with slides

# Bonus exercise: generating slides

- Tips and tricks:
- You can use the following shortcuts on your slides

**f** - enable fullscreen mode  
**w** - toggle widescreen mode  
**o** - enable overview mode  
**h** - enable code highlight mode  
**p** - show presenter notes



The end