Reproducible Reports

## Reproducible reporting?

Imagine you get an E-Mail

Dear Claas

I was working on an experiment over the last couple of months. Now  
I am currently analysing the data in R and write down everything  
for a paper. Could you be so nice and read through my statistics manuscript and give me feedback?  
-- *All the best Karl*

Attachment.Rmd

* Only one file? (.Rmd)
* Expected 2 files! (.R, .docx)
* RStudio can open it (please open it now)!
* You dont understand the syntax. (A mix of text and code?)

## Reproducible reporting?

* You ask him

Dear Claas

The ending (.Rmd) stands for R-Markdown. Markdown is a  
lightweight markup language for text. It can be combined  
with chunks of R-Code and then we speak of R-Markdown.  
It can be used for dynamic reproducible report creation.

The document can be turned into an nicely formatted Document (e.g. HTML)  
where the code chunks are executed and replaced by their output.  
Just install the knitr R package first and then  
hit the knit to HTML button in your RStudio!  
-- *All the best Karl*

* Rmd stands for R-Markdown
* Intermixes markdown markup + R-Code
* knitr can turn this innto HTML
* We follow his instructions and install knitr

--- bg:#EEE

## Reproducible reporting?

* R-Studio way: Packages > Install Packages
* On console:
* install.packages("knitr")

Get R-Studio

--- bg:#EEE

## Reproducible reporting?

* Knit his file:
* hit knit HTML button
* you get HTML output

## Reproducible reporting?

* You start to get the idea: That is Cool ;-)
* only 1 document!
* prevents copy + paste
* less error prone
* always up to date documents
* saves work time and your nerves
* BUT HTML?

## Reproducible reporting?

* Knitr R package (Yuhui Xie, 2012)
* dynamic report generator
* R-Markdown -> HTML
* Steps: .Rmd -> .md -> .html
* Generate HTML reports (e.g research blogs, etc)
* LaTeX -> PDF
* Steps: .Rnw -> .tex -> .PDF
* Generate PDF reports (for Ba./Ma. Thesis, etc.)
* R-Studio improves integrationn (Version 0.98.1028)
* From markdown -> .html, .PDF, word

Knitr Repro. reporting (full course with latex)

## R-Markdown and R-Studio

* Create a .Rmd file
* file > new > R-Markdown
* In wizward (~ since Version 0.98.1028)
* Pick: Document type/ Format (HTML)
* Insert: Name and Title

## R-Markdown and R-Studio

* The new document

## R-Markdown Syntax (Metadata)

* Optional and R-Studio specific (embraced by ---)
* Example:

---  
title: "This is my test title"  
author: "Claas-Thido Pfaff"  
date: "06.09.2014"  
output: html\_document  
---

--- bg:#EEE

## Your turn (Rmd file)

* Please Create a .Rmd file
* file > new > R-Markdown
* In wizward
* Pick: Document type/ Format (HTML)
* Insert: Name and Title ("Reproducible Reporting")
* Remove the example content (not the metadata)
* Keep and use this document through the exercises!

Failed? Your rescue!

## R-Markdown and R-Studio

* Content sensitive menu bar
* Knit to HTML
* Chunk Menu (Insert, Execute)
* Help (pops up in help browser)

## R-Markdown Syntax (headers)

* Headers for chapters and sections (Hash)

# Header ones ## Header two ### Header three ... ###### Header six

* Example

## R-Markdown Syntax (lists)

* Numbered list (digits + .)
* 1.,2.,3.,...,n

1. One 2. Two 3. Three 4. Four

* Example

## R-Markdown Syntax (lists)

* Bullet list (symb: \*, +, -; nesting 4 spaces)

\* One  
\* Two  
 + Two one  
 + Two two  
 - Two two one  
\* Three  
 1. Three one  
 2. Three two

## R-Markdown Syntax (emphasis)

This is \*italic\* (sometimes also \_italic\_)

* This is *italic*

This is \*\*bold\*\* (sometimies also \_\_bold\_\_)

* This is **bold**

This is \*\*\*bold and italic\*\*\* (sometimes also \_\_\_bat\_\_\_)

* This is ***bold and italic***
* Example

--- bg:#EEE

## Your turn (Lists and emphasis)

* Create 2nd level header (e.g Exercises)
* Add a 3rd level header below (e.g. Lists)
* Create a numbered list (hint: 1., 2. 3.)
* Create a bullet list (hint: \*, +, -)
* Try out nesting (4 spaces indent)

Failed? Your rescue!

## R-Markdown Syntax (quotes)

* Quoting with (greater than ">")

> 42 is the answer to the meaning of life,   
> the universe, and everything.  
> \*Douglas Adams\*

* Output:

## R-Markdown Syntax (tables)

* Source (align colon)
* Output

Table Generator

## R-Markdown Syntax (links)

* URL or path to files

[link text](url or relative path "Some Text")

* <https://github.com/>
* path/to/test.html
* Example

[GitHub](https://github.com/ "Build software better, together.")

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

* You can also simply paste in URLs https://github.com/ (see above)

## R-Markdown Syntax (links)

* Links as References
* Reference sections/figures/tables (Microsoft Word, Open Office, LaTeX)
* No native support in markdown
* But you can use HTML anchors
* Set anchor

<!-- set an html anchor --> <a name="abcd"></a>

* Set reference in text

<!-- refer in markdown syntax --> [link text](#abcd)

## R-Markdown Syntax (images)

![deep thought](path/to/image/deep\_thought.png "Deep Thought")

* No direct control over size (HTML)

<img src="assets/img/deep\_thought.png" alt="deep thought" title = "Deep Thought" style="width: 400px;"/>

Image Source

## R-Markdown Syntax (formulas)

* Not included in basic markdown
* Need external frameworks (mathjax)
* Inline formulas $ ... $
* Display formulas $$ ... $$ (<div> formula </div>)

$$V\_t(S\_t) = \max\_{x\_t \in \chi\_t} \left(C(S\_t, x\_t) +   
 \gamma \sum\_{s^{\prime} \in \mathcal{S}}   
 \mathbb{P}(s^{\prime} | S\_t^n, x\_t)   
 V\_{t+1}^{n-1} s^{\prime} \right)$$

Mathjax Home Formula Syntax Detexify

## R-Markdown Syntax (code)

* code blocks (3 Backticks)

```  
 Text or code goes here  
 ```

myfunction <- function(x,y){ z <- x+y; return(z) }

* In line code

In text `code` hightlight

## R-Markdown Syntax (code chunks)

A code blocks plus (curly braces, r, options)

```{r preparation, cache = TRUE, include = TRUE, eval = TRUE, echo = TRUE}  
x = c(1,2,3,4,5)  
y = summary(x)  
```

* Modify how the chunk is treated (chunk options)
* include = (T)/F (include chunk in output)
* echo = (T)/F (include source code)
* eval = (T)/F (evaluate chunk)
* cache = T/(F) (cache chunk)
* and others ...

Knitr chunk options

## R-Markdown Syntax (inline code)

Prepare Code:

Option: include = FALSE

```{r preparation, include = F}  
x = c(1,2,3,4,5)  
mean\_of\_x = mean(x)  
```

Use in your text with inline code chunks:

Lorem ipsum dolor sit ` r x + 10` amet, consetetur sadipscing   
elitr, the mean of x is ` r mean\_of\_x` ...

Lorem ipsum dolor sit 11, 12, 13, 14, 15 amet, consetetur sadipscing  
elitr, the mean of x is 3 ...

## R-Markdown Syntax (plots)

Example plot:

```{r how\_to\_include\_plots, fig.width = 8, fig.height = 5, fig.align = 'center'}  
require(ggplot2)  
qplot(Sepal.Length, Petal.Length, data = iris, color = Species,  
 xlab = "Sepal Length", ylab = "Petal Length",  
 main = "Sepal vs. Petal Length in Fisher's Iris data")  
```

* Options
* figure.width
* figure.height
* fig.align [center, left, right]
* The width and height are in inch!

## R-Markdown Syntax (plots)

qplot(Sepal.Length, Petal.Length, data = iris, color = Species,  
 xlab = "Sepal Length", ylab = "Petal Length",  
 main = "Sepal vs. Petal Length in Fisher's Iris data")

--- bg:#EEE

## Your turn (plots)

1. Create 3rd level header: e.g "Plots"
2. Create a chunk that is not displayed! (hint: include)

* load the ggplot2 package in that chunk

1. Create another chunk (use qplot)

* data = mtcars
* x = mpg (miles per US galon)
* y = wt (weight in lb (pound))
* if you like map cylinders (colour = cyl) to color

1. Knit to HTML
2. Play around with the alignment and size options!

* center the plot!
* resize till it looks correct

Failed? Your rescue!

--- bg:#EEE

## Your turn (plots)

The preparation chunk:

```{r preparation, include = F}  
require(ggplot2)  
```

Needs option: include = FALSE

* If startup messages appear use:

suppressPackageStartupMessages(library(gdata))

--- bg:#EEE

## Your turn (plots)

The plot chunk:

```{r plotchunk, fig.align = 'center', fig.width = 6, fig.height = 4}  
qplot(mpg, wt, data=mtcars, colour=cyl)  
```

## R-Markdown Syntax (tables)

Knitr can create markdown tables for you (Amazing!)

```{r results = 'asis'}  
kable(head(iris[,1:3]), format = "markdown")   
```

* requires chunk option: results = 'asis'
* Use the kable command
* kable parameters:
* align = c("c", "l", "r")
* digits = 3
* row.names = T/F

## R-Markdown Syntax (tables)

* Knitr can create markdown tables for you!
* needs chunk option: results = 'asis'

kable(head(iris[,1:3]), format = "markdown")

|  |  |  |
| --- | --- | --- |
| Sepal.Length | Sepal.Width | Petal.Length |
| 5.1 | 3.5 | 1.4 |
| 4.9 | 3.0 | 1.4 |
| 4.7 | 3.2 | 1.3 |
| 4.6 | 3.1 | 1.5 |
| 5.0 | 3.6 | 1.4 |
| 5.4 | 3.9 | 1.7 |

--- bg:#EEE

## Your turn (tables)

1. Create 3rd level header: e.g "Automagical Tables"
2. Create a chunk that is not shown! (hint: include)

* prepare 7 first rows of the mtcars dataset in a variable (hint: head)

1. Create another chunk

* use kable to display the 7 rows (hint: results = 'asis')
* do not forget to set kable option (format = "markdown")
* truncate the digits to 1 (digits = 1)

1. Knit to HTML

Failed? Your rescue!

--- bg:#EEE

## Your turn (tables)

The preparation chunk:

```{r preparation, include = F}  
subset\_of\_mtcars = head(mtcars, 7)  
```

* sets include = F

The table chunk:

```{r themagictable, results = 'asis'}  
kable(subset\_of\_mtcars, format = "markdown", digits = 1)  
```

* as output is already markdown we set: results = 'asis'
* truncate: digits = 1

--- bg:#EEE

## Your turn (tables)

The table:

kable(subset\_of\_mtcars, format = "markdown", digits = 1)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | mpg | cyl | disp | hp | drat | wt | qsec | vs | am | gear | carb |
| Mazda RX4 | 21.0 | 6 | 160 | 110 | 3.9 | 2.6 | 16.5 | 0 | 1 | 4 | 4 |
| Mazda RX4 Wag | 21.0 | 6 | 160 | 110 | 3.9 | 2.9 | 17.0 | 0 | 1 | 4 | 4 |
| Datsun 710 | 22.8 | 4 | 108 | 93 | 3.8 | 2.3 | 18.6 | 1 | 1 | 4 | 1 |
| Hornet 4 Drive | 21.4 | 6 | 258 | 110 | 3.1 | 3.2 | 19.4 | 1 | 0 | 3 | 1 |
| Hornet Sportabout | 18.7 | 8 | 360 | 175 | 3.1 | 3.4 | 17.0 | 0 | 0 | 3 | 2 |
| Valiant | 18.1 | 6 | 225 | 105 | 2.8 | 3.5 | 20.2 | 1 | 0 | 3 | 1 |
| Duster 360 | 14.3 | 8 | 360 | 245 | 3.2 | 3.6 | 15.8 | 0 | 0 | 3 | 4 |

## Where can I use it?

* GitHub (Software Development Platform)
* Version management based on Git
* Github Pages (used for this presentation!)

## Where can I use it?

* RStudio and RPubs
* Howto?
* Knit: .Rmd -> .md -> .html
* In preview windows click publish (needs account!)

## Where else it is used?

* Leanpub (Publish your books!)
* Advanced Programming in R (Hadley Wickham, knitr, mardown, pandoc etc..)

Leanpub Advanced R Programming

## Need a word document to share?

* You can use the conversion in R-Studio (word, open-office)

---  
output\_format: html\_document  
---

---  
output\_format: word\_document  
---

* As a way to create
* publishable manuscripts
* documents to share with your supervisor
* Try a word conversion now (requires word, open-office)!

Attachment.Rmd

## Wrapup

* Markdown
* write for the web (easy to learn)
* but has its limitations (e.g. figure size, references, citations)
* Scholary markdown, <http://bit.ly/11y0p23>

What happens if you ask for too much power from Markdown?  
– *Yihui Xie*

Blog-Post (Yihui)

## Wrapup

* conversion from MD works great in most cases (flexible)
* MD -> PDF
* MD -> Word
* MD -> HTML
* More complex documents (print, PDF)
* LaTeX is better
* Supports you with (references, citation, formulas, ...)
* But is much harder to learn

Repro. reporting (full course with latex)

--- bg:#EEE

## Reproduce this slides!

* download the slides (button below!)
* unzip the archive
* open the folder
* open in Rstudio: index.Rmd
* NOTE: Do not change anything in the file!
* knit it to HTML!
* do you get what you expected?

Get this presentation

--- bg:#EEE

## Reproduce this slides!

* install devtools:

install.packages("devtools")  
require(devtools)

* install slidify:

install\_github('slidify', 'ramnathv')  
install\_github('slidifyLibraries', 'ramnathv')  
require(slidify)

* knit to HTML again

<h1>Thanks for your attention!</h1>

Get this presentation

--- bg:#EEE

## No root no install rights ;-p

require(knitr)  
require(rmarkdown)  
require(ggplot2)

* Create a folder on your desktop (mylibrary)
* Add that folder to you library paths in R

.libPaths( c( .libPaths(), "c:/users/username/desktop/mylibrary"))

* Install the requirements

install.packages(c("knitr", "rmakrdown", "ggplot2"), lib = "c:/users/username/desktop/mylibrary")