Reproducible

What? Why? 001-motivating-

etting Starte

LATEX

knitr

Details

Code Chu

Hooks

Child

Docume

Custom

- .

002-minimumworking-example 003-modeloutput 004-beamer-

005-simulations 006-sensitivityanalysis

R Markdown

Final Remark

Reproducible Research

An Introduction to knitr

Sahir Bhatnagar¹

May 28, 2014

¹sahir.bhatnagar@mail.mcgill.ca

Reproducible

What? Why? 001-motivatingexample

Getting Star LAT_EX

RStudio knitr

Detail

Hooks Child Document Custom

Exercises

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown

Final Remark

Acknowledgements

- Dr. Erica Moodie
- Maxime Turgeon, Kevin McGregor, Greg Voisin
- You





Reproducible

What? Why? 001-motivating

LAT_EX RStudio

RStudio knitr

Detai

Code

Child

Custom

Custom Environn

Exercises

LACI CIBCS

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

Introduction

Final Remark

Disclaimer











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

Reproducible

What? Why? 001-motivatingexample

Getting Sta LATEX RStudio

Details

Code Ch Hooks Child

Documen Custom Environm

Exercises

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

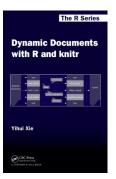
R Markdown Introduction

Final Remark

Disclaimer

 Material for this tutorial comes from many sources. For a complete list see: https://github.com/sahirbhatnagar/knitr-tutorial

• Alot of the content in these slides are based on these two books





Reproducible

What? Why? 001-motivating-

etting Start

RStudi

D-4-11

Detail

Code Chu Hooks

Docume

Custom

Evercises

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown Introduction

Final Remark

Eat Your Own Dog Food

- These slides are reproducible
- Source code: https://github.com/sahirbhatnagar/knitr-tutorial

What is Science Anyway?

Reproducible

What?

Why? 001-motivating-

etting Starte

LATEX

RStudio

D . . .

Details

Code Chun

Hooks

Child

Docume

Custom

Custom

Exercises

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown

Final Remark

What? example

LAT_EX

004-beamer-006-sensitivityanalysis

What is Science Anyway?

According to the American Physical Society:

Science is the systematic enterprise of gathering knowledge about the universe and organizing and condensing that knowledge into testable laws and theories. The success and credibility of science are anchored in the willingness of scientists to expose their ideas and results to independent testing and replication by other scientists



What?

001-motivating-

LAT_EX

004-beamer-006-sensitivityanalysis

RR: A Minimum Standard to Verify Scientific **Findings**

What?

LAT_EX

004-beamer-006-sensitivityanalysis

RR: A Minimum Standard to Verify Scientific **Findings**

Reproducible Research (RR) in Computational Sciences

The data and the code used to make a finding are available and they are sufficient for an independent researcher to recreate the finding

RR: Intro to knitr

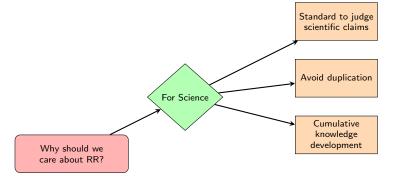
Why? 001-motivating-

LAT_EX

knitr

Custom

002-minimum-003-model-004-beamer-006-sensitivityanalysis



Reproducible

What?

Why? 001-motivatingexample

Getting Starte

RStud knitr

Detail

Code Chunk Hooks Child Documents

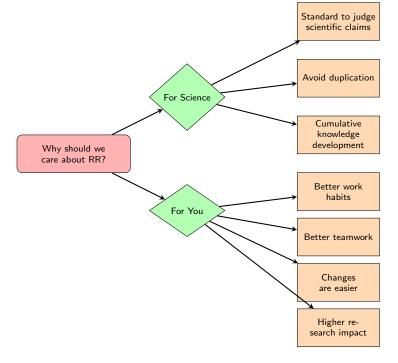
Environmen

Evercice

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown

Final Remark



RR: Intro to knitr

001-motivatingexample

LAT_EX

004-beamer-006-sensitivityanalysis

A Motivating Example

Demonstrate: 001-motivating-example

Reproducible

What? Why? 001-motivatin

Getting Started

IATEX RStudio

D . .

Code Chunk Hooks Child

Custom

Evereione

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown

Final Remark

Tools for Reproducible Research²

Free and Open Source Software

- RStudio: Creating, managing, compiling documents
- LATEX: Markup language for typesetting a document
- R: Statistical analysis language
- knitr: Integrate Lagrand R code. Based on Prof. Friedrich Leisch's Sweave

²http://onepager.togaware.com/

RR: Intro to knitr

IAT_EX

006-sensitivityanalysis

Comparison

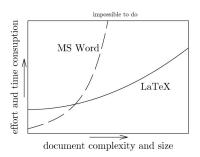


Figure 1: Comparison

- LATEX has a greater learning curve
- Many tasks are very tedious or impossible (most cases) to do in MS Word or Libre Office

The Philosophy behind LATEX

Research
What?
Why?
001-motivating
example

Getting Star LATEX

RStudi knitr

Code Chun Hooks Child Documents Custom

Exercises

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown Introduction

nal Damari



Figure 2: Adam Smith, author of *The Wealth of Nations* (1776), in which he conceptualizes the notion of the division of labour

Division of Labour

Composition and logical structuring of text is the author's specific contribution to the production of a printed text. Matters such as the choice of the font family, should section headings be in bold face or small capitals? Should they be flush left or centered? Should the text be justified or not? Should the notes appear at the foot of the page or at the end? Should the text be set in one column or two? and so on, is the typesetter's business

Reproducibl

What? Why? 001-motivating example

Getting Star

LATEX

RStudio knitr

Detai

Code Chu Hooks Child Documen Custom

Environmen

002-minimumworking-exampl 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

Introduction

Final Remark

The Genius Behind LATEX



Figure 3: The TEX project was started in 1978 by Donald Knuth (Stanford). He planned for 6 months, but it took him nearly 10 years to complete. Coined the term "Literate programming": mixture of code and text segments that are "human" readable. Recipient of the Turing Award (1974) and the Kyoto Prize (1996).



Reproducible

What? Why?

001-motivatingexample

etting Start

IAT_EX RStudio

knit

D-4-11

Detail

Code Chui

Child

Custom

Custom

Exercises

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown

Final Remark

Integrated Development Environment (IDE)

Reproducible

What? Why? 001-motivating

LAT_EX

RStudio

Detaile

Code Chu Hooks

Child Document Custom

Exercises

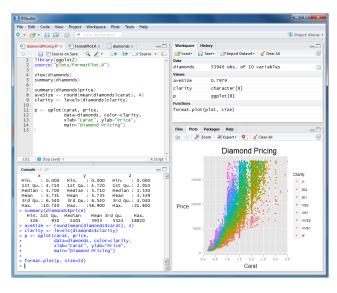
002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivity-

analysis

R Markdowr
Introduction

Final Remark

Integrated Development Environment (IDE)



Demonstrate: Explore RStudio, projects and .Rprofile

Reproducible

What? Why? 001-motivating-

Getting Started

RStudi

knitr

D 1 1

Detail

Code Chu Hooks

Child Docume

Custom

Environment

Exercise

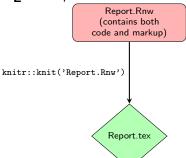
002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown

Final Remark

What knitr does

LATEX example:



Reproducible

What? Why? 001-motivatingexample

LATEX

KStu

knitr

Detail

Code Chu Hooks

Docume

Custom

_

Exercise

002-minimumworking-example 003-modeloutput 004-beamerpresentation

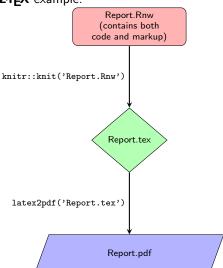
006-sensitivityanalysis

Introduction

Final Remark

What knitr does

LEX example:





knitr

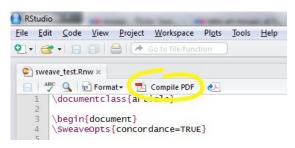
006-sensitivityanalysis

Compiling a .Rnw document

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

knitr::knit2pdf()

or in RStudio:





LAT_EX

knitr

006-sensitivityanalysis

Incorporating R code

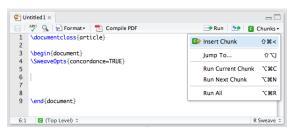
• Insert R code in a Code Chunk starting with

<< >>=

and ending with

@

In RStudio:



```
RR: Intro to
   knitr
```

LAT_EX

knitr

Hooks

002-minimum-003-model-006-sensitivityanalysis

Example 1

```
<<example-code-chunk-name, echo=TRUE>>=
library(magrittr)
rnorm(50) %>% mean
```

produces

```
library(magrittr)
rnorm(50) %>% mean
## [1] 0.031
```

```
RR: Intro to
   knitr
```

LAT_EX

knitr

Hooks

006-sensitivityanalysis

Example 2

```
<<example-code-chunk-name2, echo=TRUE, tidy=TRUE>>=
for(i in 1:5){ (i+3) %>% print}
0
```

produces

```
for (i in 1:5) {
    (i + 3) %>% print
   [1] 4
   [1] 8
```

```
RR: Intro to
   knitr
```

LAT_EX

knitr

Hooks

003-model-006-sensitivityanalysis

Example 2.2

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

produces

[1] 8

Reproducible

What? Why? 001-motivating-

Getting Started

RStudio

knitr

Details

Code Chu Hooks

Child Documen

Custom Environm

Exercises

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown Introduction

Final Remark

Example 2.3

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

produces

RR: Intro to knitr

knitr

006-sensitivityanalysis

R output within the text

- Include R output within the text
- We can do that with "S-expressions" using the command $\mathbb{Sexpr}{\ldots}$

Example:

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

produces

The iris dataset has 150 rows and 5 columns

```
RR: Intro to
   knitr
```

LAT_EX

knitr

006-sensitivityanalysis

Include a Figure

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

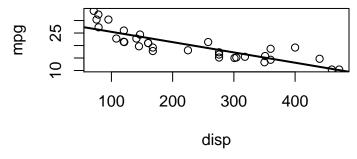


Figure 4: Linear regression

```
RR: Intro to
   knitr
```

LAT_EX

knitr

Hooks

003-model-

006-sensitivity-

analysis

Include a Table

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

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



Reproducible

What? Why? 001-motivating-

tting Starte

LATEX RStudio

Details

Code Chunks

Hooks

Custom

Environme

Exercise

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown Introduction

Final Remark

A selection of knitr code chunk options

content...

Reproducible

What? Why? 001-motivating-

etting Starte

IAT_EX RStudi

knitr Details

Cala

Code Chunks

Child

Custom

Environment

Exercise

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown Introduction

Final Remark

Set global chunk options

content...

RR: Intro to knitr

Reproducible

What? Why? 001-motivating-

tting Starte

LATEX RStudio

Details

Code Chunks

Hooks

Docume

Environment

Exercise

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown

Final Remark

Option Aliases

see page 109 yihui

Reproducible

What? Why? 001-motivating-

tting Starte

LATEX RStudio

knitr Details

Code Chunks

Hooks

Docume

Custom Environme

Exercises

Exercise

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown

Final Remark

Option Templates

see page 110 yihui



Chunk References

Personal

What? Why? 001-motivating-

tting Starte

LAT_EX RStudio

Details

Code Chunks

Hooks

Docume Custom

Environment

Exercise

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown

Final Remark

see page 79 yihui

RR: Intro to knitr

Code in Appendix

Research

What? Why? 001-motivating-

tting Starte

LATEX RStudio knitr

Details

Code Chunks

Hooks

Docume Custom

Environment

Exercise

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown

Final Remark

see page 110 yihui

RR: Intro to knitr

001-motivating-

LAT_EX

Hooks

004-beamer-006-sensitivityanalysis

A selection of knitr code chunk options

content...

Reproducible

Research

What? Why? 001-motivating-

Cotting Stort

Getting Starf

RStud

knitr

Details

Code Chui

Hooks Child

Documents

Custom Environmer

Exercises

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown

Final Remark

A selection of knitr code chunk options

see 83

RR: Intro to knitr

001-motivating-

LAT_EX

knitr

Custom Environments

003-model-004-beamer-006-sensitivity-

Example Environment

see 120

Reproducible

What? Why?

Why? 001-motivating-

etting Started

IAT_EX RStudio

knitr

Detail

Code Chun

Child

Docume

Custom

Exercises

002-minimumworking-example

003-modeloutput 004-beamerpresentation 005-simulations

R Markdown

Final Remark

Reproducible

What? Why?

001-motivating-

etting Started

LAT_EX RStudio

knitr

Details

Code Chu

Hooks

Documen

Custom

002-minimum-

working-exampl 003-model-

output 004-beamer-

presentation 005-simulations

R Markdown

Final Remark

Reproducible

What? Why?

001-motivatingexample

etting Started

LAT_EX RStudio

knitr

Details

Code Chui

HOOKS

Documen

Environment

Eversions

002-minimumworking-exampl 003-model-

004-beamerpresentation

005-simulations 006-sensitivity-

R Markdown

Final Remark

Reproducible

What? Why? 001-motivating-

C-11:-- C1-11-1

LAT_EX RStudio

knitr

Details

Code Chur

Child

Custom

_

002-minimumworking-example 003-modeloutput 004-beamer-

005-simulations 006-sensitivity-

R Markdown

Final Remark

Reproducible

What? Why?

001-motivatingexample

etting Started

LAT_EX RStudio

knitr

Detail

Code Chur

Child

Documen

Environment

_ .

002-minimumworking-example 003-modeloutput 004-beamerpresentation

006-sensitivityanalysis

K Markdown

Final Remark

Reproducible

What?

Why? 001-motivating-

etting Started

LAT_EX RStudio

knitr

Details

Code Chui

Child

Document

Environmen

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations

Introduction

Final Remarks

Reproducible

What? Why? 001-motivating

Getting Starte

RStudio

Details

Code Chunk Hooks Child

Custom

Environments

Exercise

002-minimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis

R Markdown

Final Remarks

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

Jeffrey T. Leek^{a,1} and Roger D. Peng^b

^aAssociate Professor of Biostatistics and Oncology and ^bAssociate 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.