Reproducible

What? Why? 001-motivating-

etting Starte

LATEX

knitr

Details

- Cturis

Code Chur

Hooks

Docum

Custom

Environm

Exercises

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter

R Markdown

Final Remarks

Reproducible Research

An Introduction to knitr

Sahir Bhatnagar¹

May 28, 2014

¹https://github.com/sahirbhatnagar/knitr-tutorial

Acknowledgements

Reproducible

What? Why? 001-motivating example

Getting Star LAT_EX

RStudio knitr

Detail

Hooks
Child
Documents
Custom

Custom

Exercises

working-exampl 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown Introduction

Final Remarks

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





Reproducible

What? Why? 001-motivatin

Getting Start

RStudi

Detail

Code Chunl Hooks Child

Documen Custom

Environment

Exerc

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown Introduction

Einal Damarka

Disclaimer #1

- Feel free to Ask questions
- Interrupt me often
- You don't need to raise your hand to speak

Reproducible Research

What? Why? 001-motivatingexample

LAT_EX RStudio

RStudio knitr

Detail

Code Chui Hooks

Documer

Custom Environm

Exercises

002-minimumworking-exampl 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivity-

R Markdown Introduction

Final Remarks

Disclaimer #2











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

Reproducible

What? Why? 001-motivating example

Getting Sta LAT_EX RStudio

knitr

Code Chunks Hooks Child Documents

Evercises

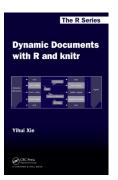
002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

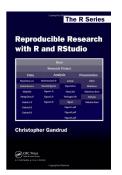
R Markdown Introduction

Disclaimer #3

 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 Starte

IAT_EX RStud

Dotoil

Detail

Code Chu Hooks

Child Documer

Custom Environm

Exercises

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown Introduction

Final Remarks

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-

atting Starte

LATEX

RStuc

Detail

Details

Code Chun

Hooks

Child

Docume

Custom

Environm

Exercises

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Final Remarks

What?

006-sensitivity-

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?

VVhy? 001-motivatingexample

LATEX RStudio

RStudi knitr

Detail

Detail

Code Chunks
Hooks
Child
Documents
Custom

Exercises

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown Introduction

Final Remark

RR: A Minimum Standard to Verify Scientific Findings

What?

example

LAT_EX

006-sensitivity-

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

Reproducible

\//ha+?

Why?

001-motivatingexample

etting Starte

LATEX

knitr

Detail

Code Chui

Docume

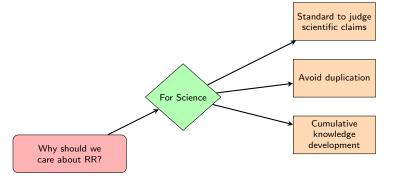
Custom

Exercises

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter

R Markdown

Final Remarks



Reproducible

What?

Why? 001-motivatingexample

Getting Starte

RStud knitr

D

Code

Hooks

Documen

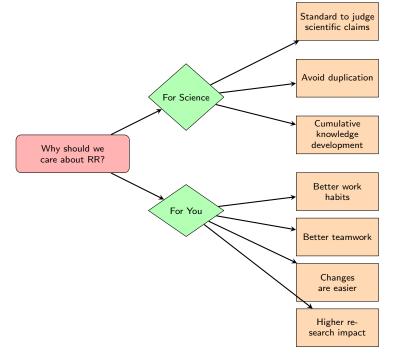
Liiviioiiiileii

002-minimur

working-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter

R Markdown

Final Remarks



A Motivating Example

001-motivatingexample

LAT_EX

003-model-004-figures 006-sensitivityanalysis-one-

Demonstrate: 001-motivating-example

What? Why? 001-motivating

Getting Started

LAT_EX RStudio

D . .

Code Chunk Hooks Child

Custom

Exercises

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter

R Markdown Introduction

Final Remarks

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 ATEXand R code. Based on Prof. Friedrich Leisch's Sweave

²http://onepager.togaware.com/

IAT_EX

004-figures 006-sensitivity-

Comparison

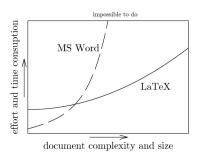


Figure 1 : Comparison

- 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 LAT_EX

RStud knitr

Details

Hooks
Child
Document
Custom
Environme

Exercises

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

Introduction

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

. ...

What? Why? 001-motivating

Getting Star

LATEX

RStudio

Detai

Code Chu Hooks Child Document Custom

Exercises

002-minimumworking-exampl 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-manyparameters

R Markdown Introduction

Einal Damark

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).



Integrated Development Environment (IDE)

001-motivating-

LAT_EX

RStudio

003-model-004-figures 006-sensitivityanalysis-one-

Reproducible

What? Why? 001-motivating

LATEX

RStudio

knitr

Code Chi

Child Documer Custom

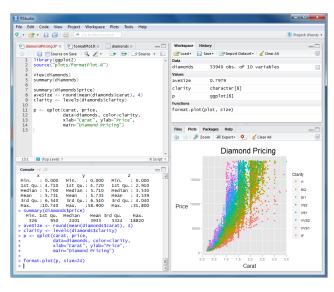
Exercises

002-minimumworking-exampl 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-one-

R Markdown Introduction

inal Remarks

Integrated Development Environment (IDE)



Demonstrate: Explore RStudio, projects and .Rprofile

Reproducible

What? Why? 001-motivating-

Getting Starte

RStudio

knitr

D . . .

Detail

Code Chui Hooks

Docume

Custom

Evercise

002-minimumworking-exampl 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivity-

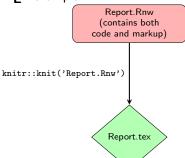
analysis-man parameters

analysis-one-

Final Remarks

What knitr does

LEX example:



001-motivating-

LAT_EX

knitr

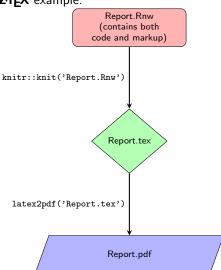
Hooks

004-figures 006-sensitivity-

analysis-one-

What knitr does

LATEX example:



001-motivating-

knitr

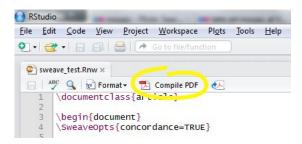
006-sensitivity-

Compiling a .Rnw document

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

knitr::knit2pdf()

or in RStudio:





What? Why? 001-motivating

Getting Started

LAT_EX RStudio

knitr

knitr

Details

Code Chunl Hooks Child Documents

Document Custom Environme

Exercises

working-exampl 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown Introduction Incorporating R code

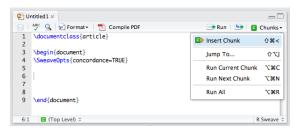
• Insert R code in a Code Chunk starting with

<< >>=

and ending with

@

In RStudio:



```
RR: Intro to knitr
```

What? Why? 001-motivating-

Getting Started

KStudio

knitr

Detail

Code Chun Hooks

Document Custom

Eversions

002-minimumworking-exampl 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter

R Markdown Introduction

Final Remarks

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
```

What? Why? 001-motivatingexample

Getting Started

RStudi

knitr

.....

Code Ch

Child Documen

Custom Environme

Exercises

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

presentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown Introduction

Final Remark

Example 2

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

produces

```
for (i in 1:5) {
     (i + 3) %>% print
}
## [1] 4
## [1] 5
## [1] 6
## [1] 7
## [1] 8
```

```
RR: Intro to
   knitr
```

001-motivating-

LAT_EX

knitr

Hooks

003-model-004-figures 006-sensitivityanalysis-one-

Example 2.2

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

produces

[1] 8

```
RR: Intro to
```

Research

What? Why? 001-motivating-

Getting Starte

RStudio

knitr

D-4-11

Code Chun

Hooks Child

Custom Environm

Evereine

002-minimumworking-exampl 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

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

knitr

006-sensitivityanalysis-one-

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
```

001-motivating-

LAT_EX

knitr

004-figures 006-sensitivityanalysis-one-

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)
@
```

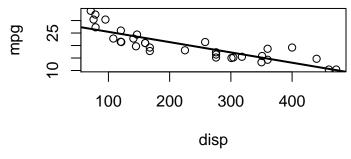


Figure 4: Linear regression

```
RR: Intro to
   knitr
```

001-motivating-

LAT_EX

knitr

Hooks

004-figures 006-sensitivityanalysis-one-

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-

etting Starte

LAT_EX RStudio

knitr

Details

Code Chunks

Hooks

Child

Custom

Environm

Evercises

Lxercises

working-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Final Remarks

A selection of knitr code chunk options

content...

001-motivating-

LAT_EX

Code Chunks

006-sensitivityanalysis-one-

Set global chunk options

content...

Option Aliases

Reproducible

What? Why? 001-motivating-

tting Starte

IAT_EX RStud

knitr

Details

Code Chunks

Child

Custom

Environment

Exercise

002-mninumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown Introduction

Final Remarks

see page 109 yihui

Reproducible

What? Why?

001-motivatingexample

tting Starte

IAT_EX RStud

knitr

Details

Code Chunks

Hooks

Custom

Custom Environm

Evercises

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivity-

R Markdown Introduction

Final Remarks

Option Templates

see page 110 yihui

Chunk References

Reproducible

What? Why? 001-motivating-

tting Starte

RStud

Details

Code Chunks

Hooks

Docume

Custom Environm

Evereieee

Exercise

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown Introduction

Final Remarks

see page 79 yihui

Code in Appendix

Reproducible

What? Why? 001-motivating-

tting Starte

IAT_EX RStud

Details

Code Chunks

Code Cit

Child

Custom

Environme

Evereieee

Exercise

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Final Remarks

see page 110 yihui

RR: Intro to knitr

001-motivating-

LAT_EX

Hooks

003-model-004-figures 006-sensitivityanalysis-one-

A selection of knitr code chunk options

content...

A selection of knitr code chunk options

Reproducible

What? Why? 001-motivating-

etting Starte

RStudi

D . . .

Detail

Hooks

Child Documents

Custom

Environment

Exercises

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Final Remarks

see 83

Example Environment

Percent

What? Why? 001-motivating-

tting Start

IAT_EX RStudi

knit

Detail

Code Chur

Child

Custom

Environments

Exercises

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Circl Domesto

see 120

Reproducible

What?

Why? 001-motivating-

etting Started

LAT_EX

knitr

Detail

Code Chur

Hooks

Child

Documer

Custom

E.....

Exercises

002-minimumworking-example

003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivity-

R Markdown

Final Remarks

Reproducible

What? Why?

001-motivating-

etting Started

IAT_EX

knitr

Detail

Code Chui

Hooks

Child

Custom

002-minimum-

003-modeloutput

004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown Introduction

Final Romarke

Reproducible

What?

Why? 001-motivating-

etting Started

LAT_EX RStud

knitr

Detail

Code Chui

Child

Document

Environme

Exercises

working-examp 003-modeloutput

004-figures

005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown Introduction

Final Remarks

Reproducible

What?

Why? 001-motivating-

etting Started

IAT_EX

knitr

Detail

Code Chur

Hooks

Child

Custom

Evercises

002-minimumworking-example 003-modeloutput

005-beamer-

presentation 006-sensitivityanalysis-one-

analysis-oneparameter 007-sensitivity analysis-many

R Markdown

Final Remarks

Reproducible

What?

Why? 001-motivating-

etting Started

LATEX

knitr

Detail

Code Chui

Child

Documen

Custom

_ ____

002-minimumworking-example 003-modeloutput 004-figures

006-sensitivityanalysis-oneparameter

parameter 007-sensitivityanalysis-many-

R Markdown

Final Remarks

Reproducible

What?

Why? 001-motivatingexample

etting Starte

LATEX

knitr

Dotail

Code Chui

Hooks

Child

Custom

Environmen

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

006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

parameters R Markdown

Introduction

Final Remarks

Reproducible

What?

Why? 001-motivating-

etting Started

LATEX

knitr

Detail

Code Chui

Hooks

Documen

Custom

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter

D Markdown

Introduction

Final Remarks

Reproducible

What? Why? 001-motivating

Getting Starte

RStudio knitr

Details

Code Chunk Hooks Child

Custom

_ .

Exercise

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivity-

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.

Reproducible

What? Why? 001-motivating-

etting Started

LAT_EX RStudio

knitr

Detaile

Code Chur

Hooks

Child

Docume

Custom

Evereieee

002-minimumworking-example 003-modeloutput 004-figures 005-beamerpresentation 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Introduction

