Reproducible

What? Why? 001-motivating-

etting Starte

LAT<sub>E</sub>X

knitr

Details

Code Chui

Hooks

Child

Docume

Custom

Evereione

002-minimumworking-exampl 003-modeloutput

output 004-beamerpresentation

006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Final Remarks

### Reproducible Research

An Introduction to knitr

Sahir Bhatnagar<sup>1</sup>

May 28, 2014

<sup>&</sup>lt;sup>1</sup>sahir.bhatnagar@mail.mcgill.ca

Reproducible

What? Why? 001-motivating

Getting Start LAT<sub>E</sub>X

RStudio knitr

Details

Code Chunl Hooks Child Documents

Custom Environm

#### Exercises

002-mnimumworking-example 003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis-one-

R Markdown

Circl Domesto

### Acknowledgements

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





Reproducible

What? Why? 001-motivati

Getting Start

RStudi

Detail

Code Chun Hooks

Documer

Environment

Evercise

002-minimumworking-example 003-modeloutput 004-beamerpresentation

005-simulations 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-manyparameters

R Markdown Introduction

Final Remarks

## Disclaimer #1

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

Reproducible Research

Why? Why? 001-motivating example

IAT<sub>E</sub>X RStudio

RStudi knitr

Details

Code

Hooks Child

Custom

Custom Environm

Exercises

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

005-simulations 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown Introduction

Final Remarks

## Disclaimer #2



R Markdown v2







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<sub>E</sub>X RStudio

knitr

Detail

Code Chunks Hooks Child Documents Custom

Evercises

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

R Markdown Introduction

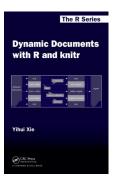
Final Remarks

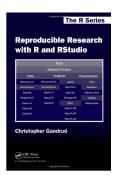
## Disclaimer #3

 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-

atting Start

IAT<sub>E</sub>X RStud

Dotoil

Details

Code Chu Hooks

Child Documer

Custom

Exercises

002-minimumworking-example 003-modeloutput 004-beamerpresentation

005-simulations 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-

etting Starte

LATEX

RStud

Detail

Details

Code Chun

Hooks

Docume

Custom

Custom

Exercises

002-minimumworking-example 003-modeloutput 004-beamerpresentation

005-simulations 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Final Remarks

# What?

006-sensitivityanalysis-one-

# 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<sub>E</sub>X

004-beamer-

006-sensitivityanalysis-one-

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



### What?

example

LAT<sub>E</sub>X

006-sensitivityanalysis-one-

# 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-motivating-

Getting Started

LATEX

knitr

Details

Codo

Hooks

Docume

Custom

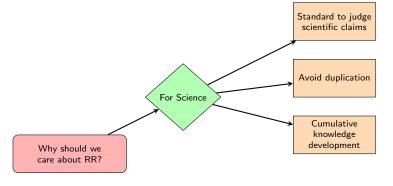
Exercises

002-minimumworking-example 003-modeloutput 004-beamerpresentation

005-simulations 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown Introduction

Final Remarks



Reproducible

What?

Why? 001-motivatingexample

Getting Starte

RStud knitr

Detail

Code Chunk Hooks Child

Documer Custom

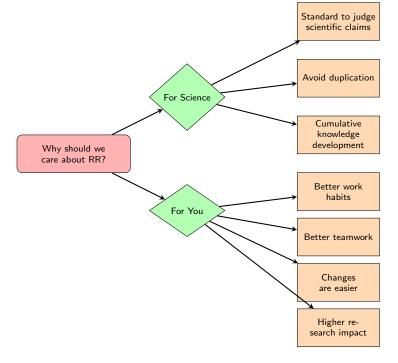
\_ .

002-minimumworking-exam 003-model-

005-simulations 006-sensitivityanalysis-oneparameter

R Markdown

Final Remarks



RR: Intro to knitr

A Motivating Example

001-motivatingexample

LAT<sub>E</sub>X

003-model-004-beamer-

006-sensitivityanalysis-one-

Demonstrate: 001-motivating-example

Reproducible

What? Why? 001-motivatin

#### Getting Started

IAT<sub>E</sub>X RStudio

Dotoil

Code Chunks Hooks Child

Custom

Exercises

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

005-simulations 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Final Remarks

# Tools for Reproducible Research<sup>2</sup>

### Free and Open Source Software

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

<sup>&</sup>lt;sup>2</sup>http://onepager.togaware.com/

Reproducible

What? Why? 001-motivat

etting Starte

#### IATEX

kni+

knit

Detai

Code

Child Docume

Custom Environme

Evercises

Exercises

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

005-simulations 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-manyparameters

R Markdown Introduction

Circl Describe

### Comparison

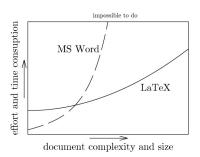


Figure 1 : Comparison

- 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

Getting Start

RStudio

kniti

Code Chun Hooks Child Documents

Child Document Custom Environme

Exercises

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

R Markdown Introduction A STATE OF THE PARTY OF THE PAR

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

producible

What? Why? 001-motivating

Getting Sta

LATEX

RStudio

Detai

Code Chu Hooks Child Documen

Custom Environme

Exercises

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

R Markdown

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)

Reproducible

What? Why?

001-motivatingexample

etting Start

IAT<sub>E</sub>X RStudio

len i to

\_

Detail

Code Chui

HOOKS

Docume

Custom

Custom

Exercises

002-minimumworking-example 003-modeloutput 004-beamerpresentation

006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Final Remark

Reproducible

What? Why? 001-motivating

example
Getting Starte

IAT<sub>E</sub>X RStudio

knitr

Details

Hooks Child

Documen Custom Environm

Exercises

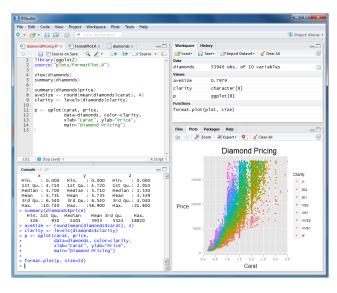
002-minimumworking-example 003-modeloutput 004-beamerpresentation

006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-manyparameters

R Markdown Introduction

Final Remarks

# Integrated Development Environment (IDE)



Demonstrate: Explore RStudio, projects and .Rprofile

Reproducible

What? Why? 001-motivating-

Getting Starte

LAT<sub>E</sub>X

knitr

Detail

Code Chui Hooks

Docum

Custom

#### Exercise

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

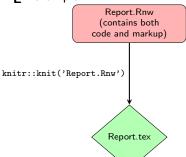
006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Final Remarks

### What knitr does

### **LEX** example:



RR: Intro to knitr

001-motivating-

LAT<sub>E</sub>X

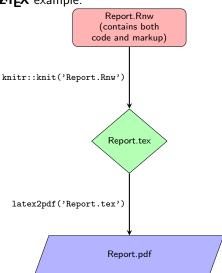
knitr

Hooks

006-sensitivityanalysis-one-

### What knitr does

**LATEX** example:



RR: Intro to knitr

001-motivating-

knitr

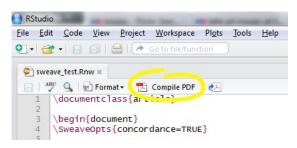
006-sensitivityanalysis-one-

### Compiling a .Rnw document

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

knitr::knit2pdf()

or in RStudio:





Reproducible

What? Why? 001-motivatin

example
Getting Started

LAT<sub>E</sub>X RStud

knitr

Kniti

Details Code Chu

Hooks Child Documents

Documents Custom Environme

Exercises

003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown Introduction

Final Remarks

### Incorporating R code

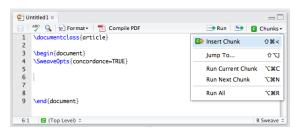
• Insert R code in a Code Chunk starting with



and ending with



#### In RStudio:



```
RR: Intro to
   knitr
```

001-motivating-

LAT<sub>E</sub>X

knitr

Hooks

006-sensitivityanalysis-one-

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
```

Reproducible

What? Why? 001-motivating-

example
Getting Started

LAT<sub>E</sub>X RStudi

knitr

.....

Details Code (

Hooks Child Documents

Custom

Exercises

002-minir

working-example 003-modeloutput 004-beamerpresentation

006-sensitivityanalysis-oneparameter 007-sensitivity-

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<sub>E</sub>X

knitr

Hooks

003-model-

006-sensitivityanalysis-one-

### Example 2.2

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

#### produces

```
RR: Intro to
```

Reproducible

What? Why? 001-motivating-

Getting Started

RStudio

knitr

Details

Code Chun Hooks

Child Documen Custom

Livironniene

002-minimumworking-examp

003-modeloutput 004-beamerpresentation 005-simulations 006-sensitivity-

analysis-oneparameter 007-sensitivityanalysis-manyparameters

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

Reproducible

What? Why? 001-motivating example

RStud

Irnita

knitr

Dotoil

Code Chui Hooks

Custom

Environn

Exercises

002-minimumworking-example 003-modeloutput 004-beamerpresentation

005-simulations 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown Introduction

### 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

```
RR: Intro to
   knitr
```

001-motivating-

LAT<sub>E</sub>X

knitr

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)
@
```

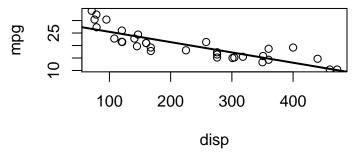


Figure 4: Linear regression

```
RR: Intro to
```

Reproducible

What? Why? 001-motivating-

Getting Starte

LAT<sub>E</sub>X

knitr

knitr

Details

Code Chu Hooks

Child Documer

Custom Environm

Exercises

002-minimumworking-example 003-modeloutput

presentation 005-simulations 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown Introductio

Final Remarks

### Include a Table

```
<<table.ex, results='asis'>>=
library(xtable)
iris[1:5,1:5] %>%
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

#### RR: Intro to knitr

001-motivating-

LAT<sub>E</sub>X

Code Chunks

Hooks

003-model-004-beamer-

006-sensitivityanalysis-one-

## A selection of knitr code chunk options

content...

Reproducible

What? Why? 001-motivating-

Cotting Start

LATEX RStudio

D . . .

Details

Code Chunks

Child

Custom

Environn

Exercises

002-minimumworking-exampl 003-modeloutput 004-beamer-

004-beamerpresentation 005-simulations 006-sensitivityanalysis-oneparameter

analysis-man parameters R Markdown

Final Remarks

## Set global chunk options

content...

### RR: Intro to knitr

# Option Aliases

Reproducible

What? Why? 001-motivating-

C-44:-- C4--4

LAT<sub>E</sub>X RStudi

knitr

Details

Code Chunks

Hooks

Docume

Custom Environn

Exercise

002-minimumworking-example 003-modeloutput 004-beamerpresentation

006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Final Remarks

see page 109 yihui

### RR: Intro to knitr

Reproducible

What? Why? 001-motivating-

. . . .

Getting Star

RStuc

Details

.

Code Chunks

Hooks

Docume

Custom Environn

Evercises

002-minimu

working-example 003-modeloutput 004-beamerpresentation

006-sensitivityanalysis-oneparameter 007-sensitivity-

R Markdown

Final Remarks

### **Option Templates**

see page 110 yihui

### Chunk References

Research

What? Why? 001-motivating-

Catting Start

Getting Start

knitr

Details

Code Chunks

Hooks

Child

Custom

Custom

\_ .

#### Exercise

002-minimumworking-example 003-modeloutput 004-beamerpresentation

006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Final Remarks

see page 79 yihui

### RR: Intro to knitr

## Code in Appendix

Reproducible

What? Why? 001-motivating-

tting Starte

IAT<sub>E</sub>X RStud

D-4-11

Details

Code Chunks

Hooks

Docume

Custom

\_ .

#### Exercises

working-example 003-modeloutput 004-beamerpresentation

006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Circl Demode

see page 110 yihui

Reproducible

What? Why? 001-motivating-

tting Starte

IAT<sub>E</sub>X RStudi

Knitr

Details

U--I--

Hooks

Docume

Custom

Exercises

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

006-sensitivity analysis-oneparameter 007-sensitivity analysis-many

R Markdown

Final Remarks

# A selection of knitr code chunk options

content...

# A selection of knitr code chunk options

Reproducible

What? Why? 001-motivating-

Cetting Starte

LATEX

RStud

knitr

Details

Code Chur

Hooks Child

Documents

Custom

\_\_\_\_\_

Exercises

002-minimumworking-example 003-modeloutput 004-beamerpresentation

006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Final Remark

see 83

# **Example Environment**

Reproducible

What? Why? 001-motivating-

example

etting Start

LAT<sub>E</sub>X

knitr

Detail

Code Chu

Hooks

Child

Custom

Environments

Exercises

working-example 003-modeloutput 004-beamerpresentation

006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

R Markdown

Final Remark

see 120

Reproducible

What?

Why? 001-motivating-

etting Started

LATEX

knitr

Detail

Code Chur

Hooks

Child

Custom

Custom

Exercises

LACI CISCS

#### 002-minimumworking-example 003-model-

004-beamerpresentation

005-simulations 006-sensitivityanalysis-one-

parameter 007-sensitivi

R Markdown

Final Remarks

Reproducible

What?

Why? 001-motivating-

etting Started

LATEX

knitr

Detail

Code Chu

Hooks

Child

Docume

Custom

Exercises

working-examp

#### 003-modeloutput

004-beamer-

005-simulations 006-sensitivity-

analysis-oneparameter 007-sensitivit

R Markdown

E: 15

Reproducible

What?

Why? 001-motivatingexample

etting Started

LATEX

knitr

Detail

Code Chui

Hooks

Documer

Custom

Liiviioiiiiic

Exercises

002-minimumworking-exampl 003-model-

#### 004-beamerpresentation

005-simulations 006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-many-

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-beamer-

#### 005-simulations

006-sensitivityanalysis-oneparameter 007-sensitivityanalysis-manyparameters

R Markdown Introduction

Final Remarks

Reproducible

What?

Why? 001-motivating-

etting Started

LATEX

knitr

Detail

Code Chur

Hooks

Child

Document

Ellallollille

Evercises

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

006-sensitivityanalysis-oneparameter 007-sensitivity-

R Markdown

Final Remarks

Reproducible

What?

Why? 001-motivating-

etting Starte

LATEX

knitr

Dotail

Code Chui

Hooks

Child

Documen

Environme

Evercises

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

005-simulations 006-sensitivityanalysis-one-

parameter 007-sensitivityanalysis-manyparameters

R Markdown

Final Remarks

Reproducible

What?

Why? 001-motivating-

etting Started

LATEX

knitr

Detail

Code Chui

Hooks

Child

Documen

LIIVIIOIIIII

Exercises

working-example 003-modeloutput 004-beamer-

005-simulations 006-sensitivityanalysis-oneparameter 007-sensitivity-

D Markdown

Introduction

Final Remarks

Reproducible

What? Why? 001-motivating

Getting Starte

RStudio

Details

Code Chunk Hooks Child

Custom

Environment

Exercise

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

006-sensitivity analysis-oneparameter 007-sensitivity analysis-many

R Markdown

Final Remarks

# 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.

Reproducible

What? Why? 001-motivating-

LAT<sub>E</sub>X RStudio

knitr

Detaile

Code Chur

Hooks

Child

Documer

Custom

E....

LXCI CI3C3

working-example 003-modeloutput 004-beamer-

005-simulati

006-sensitivityanalysis-oneparameter 007-sensitivity-

R Markdown

Final Remarks

