Recherche Reproductibl

What

001-motivatir example

Getting Starte

KT-V

knitr

Example

UU2-minimum-

working-examp

--- -

005-beamer

presentation

006-sensitivity analysis-one-

parameter

parameters

008-large-

009-rmarkdov

010-rmarkdow presentation

Final Remark

# Avoir une présence en ligne Outils pour la diffusion rapide et reproductible de la recherche

Sahir Rai Bhatnagar<sup>1</sup>

May 14, 2019

<sup>&</sup>lt;sup>1</sup>https://github.com/sahirbhatnagar/raqc

### Remerciements

Recherche Reproductib

What? Why?

> 001-motivatir example

Getting Starte

HT-X

knitr

Example

002-minimumworking-example

003-model-outpu

005-beamerpresentation

006-sensitivity analysis-one-

007-sensitivityanalysis-many-

parameters 008-large-

009-rmarkdown

Final Remarks

- La comité organisateur
- Pierre Racine et Sophie Baillargeon
- Don Knuth (T<sub>E</sub>X)
- Friedrich Leisch (Sweave)
- Yihui Xie (knitr)
- Vous





## Avis #1

Reproductil
What?

001-motivati

Getting Starte

Getting Starte

RStudio

Example

002-minimum-

working-examp

003-model-out

005-heame

presentatio

analysis-one-

parameter

parameters

008-largedocuments

010-rmarkdow

Final Remark

 Ceci est une introduction au outils pour la recherche reproductible

## Avis #1

Recherche Reproductib

Why?

101-motivati example

Getting Starte

RStudio

002-minimum-

working-example

004-figures 005-beamer-

presentation

006-sensitivit analysis-one-

parameter 007-sensitivity

008-large-

009-rmarkdow

Final Remark

- Ceci est une introduction au outils pour la recherche reproductible
- Le niveau de cet atelier est "intermédiaire" et suppose des connaissances de base en R ainsi que de l'environnement RStudio

## Avis #1

Recherche Reproductibl

Why?

01-motivati xample

Getting Starte

RStudi

Examples 002-minimum-

working-example 003-model-outpu

004-ngures 005-beamer-

presentation

analysis-oneparameter

parameter 008-large-

009-rmarkdow

Final Remark

- Ceci est une introduction au outils pour la recherche reproductible
- Le niveau de cet atelier est "intermédiaire" et suppose des connaissances de base en R ainsi que de l'environnement RStudio
- N'hésitez pas à posez des questions

Avis #2



R Markdown v2







Reproductible

Why? 001-motivati example

Getting Starte

ETpX

knit:

Example

working-exame

003-model-out

004-figur

nresentation

006-sensitivit

analysis-oneparameter

analysis-man

008-large-

document

010-rmarkdow

Final Remark

Je n'ai aucune relation commerciale avec ces logiciels.

# Avis #3

Recherche Reproductible

What? Why? 001-motivating example

Getting Started

RStudio knitr

Examples 002-minimum-

working-example

004-figures

presentation

006-sensitivit analysis-one-

007-sensitivit analysis-man

008-large-

009-rmarkdow 010-rmarkdowi presentation

Final Remark

- Le matériel pour cet atelier est basé sur plusieurs ressources
- Voir ce lien pour une liste complète de références: https://github.com/sahirbhatnagar/raqc
- Une grande partie du contenu de ces diapositives est basée sur ces deux livres:





# Eat Your Own Dog Food

Recherche Reproductib

Why?

uu1-motivat example

Getting Start

INT V

knit

Exampl

002-minimum-

working-exampl

004-figures

005-beamer

presentation

006-sensitivity analysis-one-

parameter 007-sensitivity-

008-large-

documents

010-rmarkdow

Final Remark

- Ces diapositives sont reproductibles
- code:

https://github.com/sahirbhatnagar/raqc/tree/master/slides

# Le programme de l'atelier

Recherche Reproductib

Why? 001-motivatir

**Getting Starte** 

ETEX

knit

Example

working-example

004-figures 005-beamer-

005-beamerpresentation

006-sensitivity analysis-oneparameter

007-sensitivity analysis-many

008-largedocuments

009-rmarkdow

Final Remarks

- 8h30 à 10h00: Introduction aux raports reproductibles avec knitr et RMarkdown
- 10h00 à 10h30: Pause
- 10h30 à 12h00:
- 13h30 à 15h00: Créer un siteweb avec blogdown
- 15h00 à 15h30: Pause
- 15h30 à 17h:
- 17h: Fin de l'atelier

# What is Science Anyway?

Recherche Reproductib

Reproductib What?

001-motivat

Cotting Start

octting ottart

lonita

Evample

002-minimu

working-example

003-model-out

004-figur

005-beame presentation

006-sensitivit analysis-one

parameter

analysis-manyparameters

008-large-

document

010-rmarkdov

presentation

Final Remark

What is Science Anyway?

Recherche Reproductible

#### What? Why?

01-motivatii xamnle

**Getting Starte** 

ETeX

RStud knit:

Examples 002-minimumworking-example

working-example

004-figures

presentation

analysis-one parameter

analysis-mar parameters

008-large-

009-rmarkdow

Final Remarks

### 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

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

What?

Recherche Reproductibl

What?

001-motivati

**Getting Starte** 

IST.V

RStud knit

Example

002-minimumworking-eyampl

003-model-out

004-figures

005-beamer

presentation 006-sensitivi

parameter

parameters

008-largedocuments

010-rmarkdow

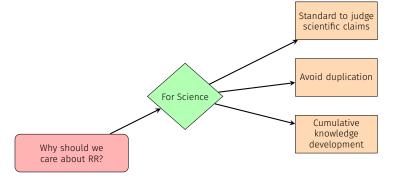
Final Remarks

# 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

Why?



Recherche Reproductib

Why?

001-motivatingexample

Getting Starte

FT-X

knit

Example

002-minimum-

003-model-outp

004-figures

presentation

006-sensitivityanalysis-one-

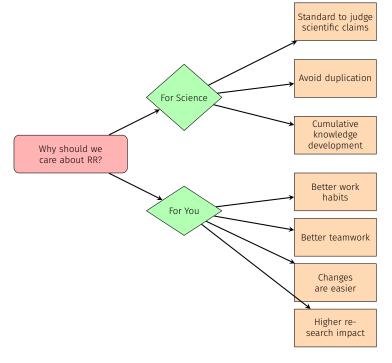
analysis-oneparameter

parameters

documents 009-rmarkdov

010-rmarkdown

Final Remarks



# A Motivating Example

Recherche Reproductibl

What? Why? 001-motivating-

example
Getting Starte

ET<sub>E</sub>X

RStudio

LAdilible

002-minimumworking-examn

002-model-out

004-figures

005-beame

presentation

analysis-one

parameter

parameters

008-largedocuments

010-rmarkdow

Final Remark

Demonstrate: 001-motivating-example

Survey: https://www.surveymonkey.com/s/CDVXW3C

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

Recherche Reproductible

Why? 001-motivatir

Getting Starte

ur v

RStudi

Examples

002-minimumworking-example

003-model-outp

005-beamer-

presentation

analysis-one-

007-sensitivit

008-large-

009-rmarkdow

Final Remark

### Free and Open Source Software

- RStudio: Creating, managing, compiling documents
- धा<sub>E</sub>X: Markup language for typesetting a document
- R: Statistical analysis language
- knitr: Integrate McXand R code. Based on Prof. Friedrich Leisch's Sweave

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

# Comparison

impossible to do effort and time consuption ET<sub>E</sub>X MS Word LaTeX document complexity and size

Figure 1: Comparison

- ETFX 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 MEX

Recherche Reproductibl What? Why?

Getting Started

RStudio

Examples 002-minimum-

003-model-outp 004-figures 005-beamer-

006-sensitivity analysis-oneparameter 007-sensitivity

008-largedocuments 009-rmarkdow

Final Romarks



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

# The Genius Behind ETEX

Figure 3: The T<sub>E</sub>X 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).

Recherche Reproductibl

Why?

example

detting Starte

ETEX

Examp

002-minimumworking-example

003-model-outpl 004-figures 005-beamerpresentation

006-sensitivity analysis-oneparameter 007-sensitivity-

008-largedocuments

009-rmarkdow 010-rmarkdow presentation

Final Remark

# Integrated Development Environment (IDE)

Recherche Reproductib

Reproductil

Why?

example

Getting Starte

we u

RStudio

#### Example

002-minimumworking-examp

003-model-outp

004-figure

004-ngures

presentatio

analysis-one

parameter 007-sensitivity

parameters

008-large-

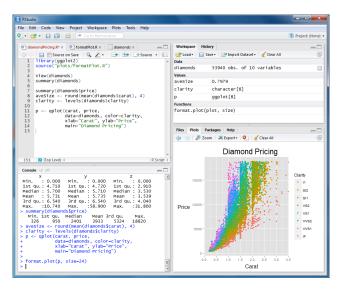
009-rmarkdow

010-rmarkdow presentation

Final Remark

RStudio

# Integrated Development Environment (IDE)



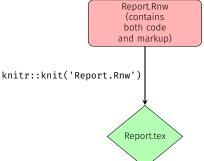
Final Remarks

Demonstrate: Explore RStudio

knitr

### What **knitr** does

LETEX example:



knitr

### What **knitr** does

LETEX example: Report.Rnw (contains both code and markup) knitr::knit('Report.Rnw') Report.tex

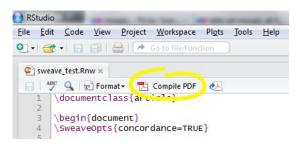
latex2pdf('Report.tex')

# Compiling a . Rnw document

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

knitr::knit2pdf()

or in RStudio:



Recherche Reproductible

What? Why?

example

Getting Starte

ETeX

knitr

002-minimumworking-example

working-example

005-beamer-

006-sensitivity analysis-one-

007-sensitivityanalysis-many-

008-largedocuments

010-rmarkdown presentation

Final Remark

Recherche Reproductible

Why? 001-motivatir

001-motivatir example

Getting Starte

ETeX

knitr

002-minimumworking-example

004-figures 005-beamer-

006-sensitivit analysis-one-

007-sensitivity analysis-many parameters

009-rmarkdowr

Final Remarks

# Incorporating R code

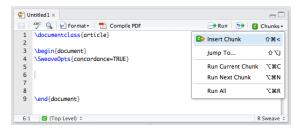
Insert R code in a Code Chunk starting with



and ending with



### In RStudio:



```
Outils pour la
diffusion rapide
et reproductible
de la recherche
```

# Example 1

```
Recherche
Reproductib
What?
```

001-motivatir

Getting Starte

ETEX

knitr

### Example

002-minimumworking-exampl

003-model-out

004-figures

005-beamer

presentation

analysis-one-

parameter 007-sensitivity

parameters 008-large-

009-rmarkdow 010-rmarkdow

Final Remark

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

@
```

### produces

```
library(magrittr)
rnorm(50) %>% mean
```

## [1] 0.12

```
Outils pour la
diffusion rapide
et reproductible
de la recherche
```

# Example 2

```
Recherche
Reproductibl
```

Why? 001-motivatin example

Getting Started

ETeX

knitr

KHIL

### Exampl

working-examp

003-model-out

005-beamer-

presentation 006-sensitivit

parameter

analysis-many parameters

008-largedocuments

009-rmarkdow 010-rmarkdow presentation

Final Remark

```
<<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
```

```
Outils pour la
diffusion rapide
et reproductible
de la recherche
```

# Example 2.2

```
Recherche
Reproductib
What?
```

Why? 001-motivati

Getting Starte

ETEX

DStudio

knitr

#### .....

002-minimum-

working-exam

003-model-ou

004-figures

presentation

006-sensitivity analysis-one-

007-sensitivityanalysis-many-

008-large-

009-rmarkdow 010-rmarkdow

Final Remark

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

### produces

```
## [1] 4
## [1] 5
## [1] 6
## [1] 7
## [1] 8
```

```
Outils pour la
diffusion rapide
et reproductible
de la recherche
```

# Example 2.3

knitr

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

### produces

Demonstrate: Try it yourself

Recherche Reproductib

What? Why? 001-motivatir

**Getting Starte** 

KT-Y

DIEV

knitr

#### Examp

002-minimumworking-exampl

003-model-out

004-figures 005-beamer-

presentation

analysis-oneparameter

007-sensitivityanalysis-manyparameters

008-largedocuments 009-rmarkdo

010-rmarkdown presentation

Final Remarks

# 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

knitr

# Include a Figure

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

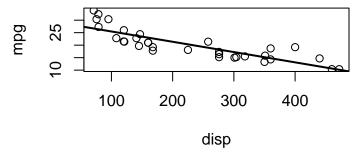


Figure 4: Linear regression

```
Outils pour la
diffusion rapide
et reproductible
de la recherche
```

Recherche Reproductib

What? Why?

001-motivatin

Getting Started

ETEX

RStudio

knitr

Exampl

002-minimumworking-examp

003-model-ou

005-beamer-

presentation 006-sensitivit

parameter 007-sensitivity

parameters

documents

010-rmarkdow

Final Remark

### Include a Table

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

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

# Minimum Working Example

Recherche Reproductib

Why?

001-motiva example

**Getting Start** 

detting starte

RStudio

KHIL

Example

#### 002-minimumworking-example

working-example

004-figures

005-beamer

presentatio

analysis-one

parameter 007-sensitivity

parameters

document

009-rmarkdov 010-rmarkdov

Final Romark

https://github.com/sahirbhatnagar/knitr-tutorial/tree/master/002-minimum-working-example

# Extracting output from Regression Models

Recherche Reproductib

Why?

001-motiva example

**Getting Starte** 

MT-V

RStudio

\_

002-minim

working-exam

003-model-output

004-figures

005-beame

presentation

analysis-one

parameter

parameters

documents

010-rmarkdow

Final Remark

https://github.com/sahirbhatnagar/knitr-tutorial/tree/master/003-model-output

# Figures

Recherche Reproductib

What?

001-motivat example

Getting Starte

ETeX

knitr

#### Exampl

working-example

003-model-outpi

#### 004-figures

005-beamerpresentation

006-sensitivi

analysis-one parameter

analysis-ma parameters

documents

009-rmarkdov 010-rmarkdov

Final Remark

https://github.com/sahirbhatnagar/knitr-tutorial/tree/master/004-figures

### **Beamer Presentations**

Recherche Reproducti

Why?

001-motivat example

**Getting Starte** 

were

RStudio

knit

Exampl

002-minimum-

002-modol-out

005 model out

004-ngu

005-beamerpresentation

006-sensitivit

analysis-oneparameter

analysis-many parameters

008-large-

009-rmarkdov

010-rmarkdow presentation

Final Remark

https://github.com/sahirbhatnagar/knitr-tutorial/tree/master/005-beamer-presentation

# Changing one Parameter in an Analysis

Recherche Reproductib

What:

001-motiva example

Getting Starte

MT-V

RStudio

knit

#### Example

002-minimum-

working-examp

004-figures

005-beamer

#### 006-sensitivityanalysis-oneparameter

007-sensitivityanalysis-many-

008-large-

009-rmarkdov

010-rmarkdow presentation

Final Remark

https://github.com/sahirbhatnagar/knitr-tutorial/tree/master/006-sensitivity-analysis-one-parameter

# Changing Many Parameters in an Analysis

Recherche Reproductib What?

001-motiva

example

Getting Starte

ETEX

knit

Example

002-minimum-

working-examp

003-model-out

004-figures

presentation

006-sensitivityanalysis-one-

007-sensitivityanalysis-many-

parameters 008-large-

009-rmarkdov

010-rmarkdow

Final Remark

https://github.com/sahirbhatnagar/knitrtutorial/tree/master/007-sensitivity-analysis-many-parameters

# Large Documents

Recherche Reproductib

What? Why?

> 001-motivat example

**Getting Starte** 

INT.V

RStudio

Evamnl

Lyampie

working-examp

003-model-out

004-figure

005-beame

006-sensitivit

analysis-one parameter

007-sensitivity analysis-many

#### 008-largedocuments

009-rmarkdow

Final Remark

https://github.com/sahirbhatnagar/knitr-tutorial/tree/master/008-large-documents

# **HTML** Reports

Recherche Reproducti

Why?

example

**Getting Starte** 

MT-V

RStudio

KHIL

Exampl

002-minimum-

working-examp

003-model-out

004-ngure

005-beame

006-sensitivit analysis-one-

parameter

analysis-many parameters

008-large-

009-rmarkdown

010-rmarkdowi

Final Remark

https://github.com/sahirbhatnagar/knitr-tutorial/tree/master/009-rmarkdown

### HTML Presentations

010-rmarkdownpresentation

https://github.com/sahirbhatnagar/knitrtutorial/tree/master/010-rmarkdown-presentation

Recherche Reproductib

Why?

001-motivatir example

Getting Starte

ET<sub>E</sub>X RStudio

Examples

002-minimumworking-examn

003-model-out

004-figures

005-beamerpresentation

006-sensitivit

analysis-oneparameter

analysis-ma parameters

008-large-

010-rmarkdow

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.

# Always Remember ...

Recherche Reproductib

Why?

001-motivat example

**Getting Star** 

INT V

lenita

Example

LXGIIIPIC

working-eyamr

working-exam

003-model-ou

004-figur

nrocontatio

006-sensitivi

analysis-one parameter

007-sensitivit

nno largo

document

009-rmarkdov

010-rmarkdow

Final Remarks

Reproducibility 
$$\propto \frac{1}{\text{copy paste}}$$

# Is the juice worth the squeeze?

Recherche Reproductib

What

001-motiva example

Getting Start

100.14

Land de

Evamnl

----

working-examp

003-model-out

004-figure

005-beame

006-sensitivit

analysis-one-

parameter 007 consistivity

parameter

documents

009-rmarkdov

010-rmarkdow presentation

Final Remarks

