Recherche Reproductible (RR)

Quoi?

001-exemple

Getting Starte

RStudio

knitr

Examples

002-minimum

working-example

004 figures

004-ngures

nresentation

006-sensitivity

analysis-oneparameter

analysis-many

parameters

000 -----

009-rmarkdowr

.....

Final Remarks

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

Sahir Rai Bhatnagar¹

May 14, 2019

¹https://github.com/sahirbhatnagar/raqc

Remerciements

Recherche Reproductible (RR)

Pourquoi? 001-exemp

Getting Starte

RStudio knitr

Examples 002-minimumworking-example

004-figures

presentation 006-sensitivity analysis-oneparameter

007-sensitivityanalysis-manyparameters 008-largedocuments

009-rmarkdown

Final Remark

- La comité organisateur
- Pierre Racine et Sophie Baillargeon
- Don Knuth (T_EX)
- Friedrich Leisch (Sweave)
- Yihui Xie (knitr)
- Vous





Avis #1

Recherche Reproductibl (RR)

Pourquoi? 001-exempl

Getting Starte

ETEX

RStudi knitr

002-minimum-

working-example 003-model-outpu

004-figures

005-beame

006-sensitivit analysis-one-

007-sensitivity-

analysis-manyparameters

documents

009-rmarkdow

Final Remarks

 Ceci est une introduction au outils pour la recherche reproductible

Avis #1

Recherche Reproductible (RR)

Pourquoi? 001-exemp

Getting Starte

RStudio knitr

Examples 002-minimumworking-examp 003-model-out 004-figures

presentation 006-sensitivity analysis-one-

007-sensitivityanalysis-manyparameters

009-rmarkdowi

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 Reproductible (RR)

001-exemplejustificatif

Getting Starte

knitr Examples

002-minimumworking-example 003-model-outpu

005-beamerpresentation 006-sensitivity

007-sensitivityanalysis-manyparameters

009-rmarkdowi

Final Remarks

- 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











Je n'ai aucune relation commerciale avec ces logiciels.

Avis #3

Recherche Reproductible (RR) Ouoi?

Getting Starte

knitr Examples 002-minimum-

003-model-outpu

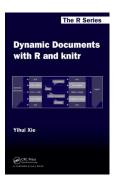
005-beamerpresentation

parameter 007-sensitivityanalysis-manyparameters

documents 009-rmarkdown 010-rmarkdown

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

Ces diapositives sont reproductibles

Voir ragc-slides.Rnw: https://github.com/sahirbhatnagar/raqc/tree/master/slides

Le programme de l'atelier

Reproductibl (RR) Quoi? Pourquoi?

Getting Starte

ETEX

kniti

Evami

002-minimumworking-exampl 003-model-outp

004-figures

005-beamerpresentation 006-sensitivity analysis-one-

007-sensitivityanalysis-manyparameters 008-largedocuments

009-rmarkdown

Final Remark

- 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

C'est quoi la science?

Recherche Reproductib (RR)

Quoi?

Pourquoi? 001-exemple-

Gotting Starte

INT V

RStud

knit

Examples

non minim

working-example

003-model-outp

004-ngures

005-beamer-

006-sensitivit

parameter

analysis-ma

008-large-

009-rmarkdov

010-rmarkdow

Final Remarks

C'est quoi la science?

Recherche Reproductible (RR)

Quoi?

Pourquoi? 001-exemp

Getting Starte

IST-Y

RStudio

knitr

002-minimum-

003-model-output

004-figures

presentation 006-sensitivi

006-sensitivity analysis-oneparameter

007-sensitivity analysis-many parameters

parameters

009-rmarkdow

Final Romark

Final Remarks

Selon l'American Physical Society:

La science est l'entreprise systématique consistant à rassembler des connaissances sur l'univers et à les organiser et les condenser en lois et **théories vérifiables**...

Recherche Reproductible

Quoi?

ourquoi?

Getting Starte

Getting Starte

RStudio knitr

Examples 002-minimumworking-example

003-model-outp 004-figures

004-figures

presentation 006-sensitivity analysis-one-

007-sensitivity analysis-many parameters

009-rmarkdow 010-rmarkdow

Final Remark

C'est quoi la science?

Selon l'American Physical Society:

La science est l'entreprise systématique consistant à rassembler des connaissances sur l'univers et à les organiser et les condenser en lois et **théories vérifiables**...

Le succès et la crédibilité de la science sont fondés sur la volonté des scientifiques d'exposer leurs idées et leurs résultats à des tests indépendants et à leur reproduction par d'autres scientifiques.

Recherche Reproductible

Quoi?

Pourquoi? 001-exempl

Getting Starte

KT-Y

RStu

KHIL

Exampl

002-minimu

working-example 003-model-outp

004-figures

005-beamer-

presentation

006-sensitivit analysis-oneparameter

007-sensitivityanalysis-many-

parameters

008-largedocuments

009-rmarkdow

presentation

Final Remarks

RR: Une norme minimale pour vérifier les résultats scientifiques

Recherche Reproductible (RR)

Quoi?

ourquoi? 101-exemp

Getting Starte

NT-X

RStudi

Examples

002-minimumworking-exampl

003-model-outpu 004-figures

005-beamerpresentation 006-sensitivity

analysis-oneparameter 007-sensitivity

analysis-many parameters

009-rmarkdow

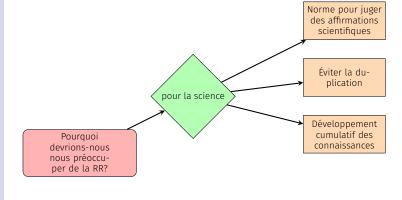
Final Remark

RR: Une norme minimale pour vérifier les résultats scientifiques

Recherche reproductible (RR) dans la science des données

Les données et le code utilisés pour effectuer une constatation sont disponibles et suffisent à un chercheur indépendant pour recréer la constatation.

Pourquoi?



Recherche Reproductible (RR)

Quoi

Pourquoi? 001-exemp

Cotting Start

Getting Start

DStudi

knit

Exampl

002-minimun

working-example

004 figures

004-figures

005-beamerpresentation

006-sensitivity analysis-one-

007-sensitivity-

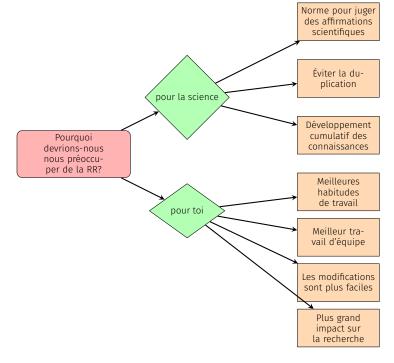
analysis-manyparameters

parameters nns-large-

009-rmarkdow

010-rmarkdown

Final Remarks



Un exemple justificatif

Recherche Reproductib (RR)

Quoi?

001-exemplejustificatif

Getting Starte

IST-V

RStudio

RStude

Examples 002-minimum

working-example

004-figures

005-beamer

006-sensitivit

analysis-oneparameter

analysis-many

parameters 009 Jarge

009-rmarkdow

010-rmarkdow

Final Remarks

Demonstrate: 001-exemple-justificatif

Outils pour la diffusion rapide de la recherche

Tools for Reproducible Research²

Free and Open Source Software

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

²http://onepager.togaware.com/

ET_EX

Comparison

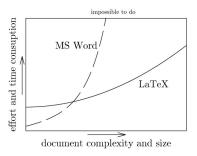


Figure 1: Comparison

- LTFX 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 ETEX

Recherche Reproductible RR) Quoi?

Getting Starte

RStudio

Examples 002-minimumworking-exampl 003-model-outp

005-beamerpresentation 006-sensitivity analysis-oneparameter

007-sensitivity analysis-many parameters 008-largedocuments 009-rmarkdow

documents 009-rmarkdow 010-rmarkdow presentation

Final Remark



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_EX 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).

Reproductibl (RR)

Quoi?

001-exemp

Getting Starte

RStudi

Examples
002-minimumworking-examp
003-model-out
004-figures
005-beamer-

presentation 006-sensitivity analysis-oneparameter

007-sensitivity analysis-many parameters 008-largedocuments 009-rmarkdow

Final Rema

Integrated Development Environment (IDE)

Recherche Reproductib (RR)

Quoi

001-exempl

Getting Starte

detting starte

ETEX

RStudio

knitr

Examples

working-exan

003-model-outp

004-figures

007 1150103

UU5-beamer-

006-sensitivit analysis-one-

007-sensitivity

analysis-many parameters

parameters

009-rmarkdow

009-rmarkdowi

F ... I B I

rillat Reillarks

Recherche Reproductible (RR)

Quoi?

Pourquoi? 001-exemple

Getting Starte

RStudio

Fxample

002-minin working-e

004-figures 005-beamer-

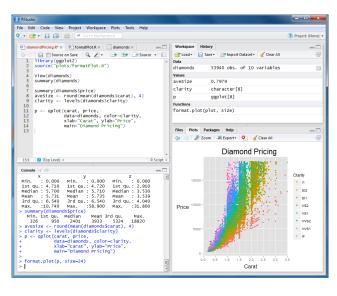
006-sensitivit analysis-oneparameter

007-sensitivityanalysis-manyparameters

documents 009-rmarkdow 010-rmarkdowi

Final Remarks

Integrated Development Environment (IDE)



Demonstrate: Explore RStudio

Recherche Reproductible (RR)

Quoi

001-exem

Gotting Start

were a

RStudio

knitr

002-minimu

working-examp

004-figures

005-hean

presentat

006-sensitivit analysis-onenarameter

007-sensitivityanalysis-many-

anatysis-many parameters

000 rmarkdow

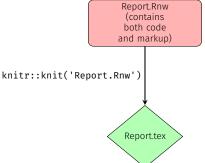
009-rmarkdowr

p. ----

presentation

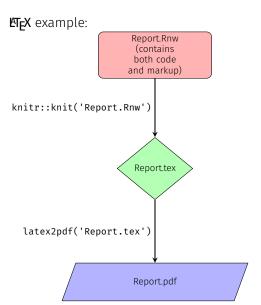
What knitr does

 \LaTeX example:



knitr

What **knitr** does

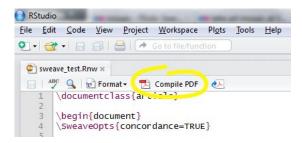


Compiling a . Rnw document

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

knitr::knit2pdf()

or in **RStudio**:



Recherche Reproductible (RR)

Quoi? Pourquoi? 001-exemple

Getting Starte

ETEX

knitr

knitr

Examples 002-minimum

003-model-outpu

005-beamerpresentation

analysis-oneparameter 007-sensitivityanalysis-many-

analysis-manyparameters 008-largedocuments

009-rmarkdown

Final Remark

Recherche Reproductible (RR)

Quoi? Pourquoi? 001-exemple

Getting Starte

KT_EX

RStudio

knitr

Examples 002-minimumworking-examp

003-model-outp 004-figures 005-beamerpresentation

analysis-oneparameter 007-sensitivityanalysis-manyparameters

documents 009-rmarkdow 010-rmarkdown presentation

Final Remarks

Incorporating R code

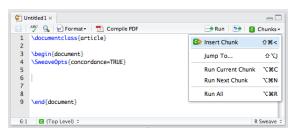
Insert R code in a Code Chunk starting with



and ending with



In **RStudio**:



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

de la recherc

Recherche Reproductible (RR)

Quoi?

Pourquoi? 201-exemple Justificatif

Getting Starte

MT-Y

piEv

knitr

Examples 002-minimum-

working-example

004-figures

005-beam presentati

006-sensitivit analysis-one-

007-sensitivityanalysis-manyparameters

documents 009-rmarkdow

010-rmarkdow

Final Remark

Example 1

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

0
```

produces

[1] 0.12

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

manyrs

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

Recherche Reproductible

Quoi? Pourquoi? 001-exemple

Getting Started

ETicX

FIEV

11.5000

knitr

Examples 002-minimum-

003-model-output 004-figures

005-beamerpresentation

006-sensitivity analysis-oneparameter

analysis-manyparameters

documents 009-rmarkdow

010-rmarkdow presentation

Final Remark

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

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

Example 2.2

```
Recherche
Reproductibl
(RR)
```

Pourquoi?

Getting Starte

detting starte

ELEX

knitr

Examples

002-minimum

working-example

nna-figures

004-ligures

presentation

analysis-oneparameter

007-sensitivity analysis-many parameters

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

```
Recherche
Reproductib
(RR)
Quoi?
```

justincatir

Getting Start

RStudio

RStudio knitr

Examples 002-minimum

working-example 003-model-outpu

004-figures

presentation 006-sensitivi analysis-one

parameter 007-sensitivity analysis-many

analysis-many parameters 008-large-

009-rmarkdow

010-rmarkdow

Final Remark

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

a
```

produces

Demonstrate: Try it yourself

Recherche Reproductibl (RR) Quoi?

Pourquoi? 001-exemplejustificatif

Getting Starte

RStudio

knitr

Examples

002-minimumworking-exampl

003-model-outs

004-figures 005-beamerpresentation

006-sensitivityanalysis-oneparameter

analysis-many parameters 008-large-

009-rmarkdow

Final Remark

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

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

Recherche
Reproductible
(RR)
Quoi?
Pourquoi?

Getting Started

ETeX

11.51.01.0

knitr

```
Examples
002-minimum-
working-exampl
```

003-model-output 004-figures

005-beamerpresentation

006-sensitivity analysis-oneparameter 007-sensitivity-

analysis-manyparameters 008-large-

009-rmarkdow

Final Remark

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)

0
```

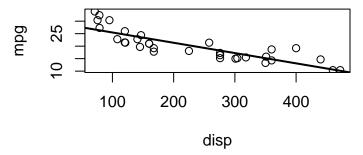


Figure 4: Linear regression

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

Recherche Reproductible (RR)

Quoi?

Pourquoi? 001-exemple

Getting Starte

ETEX

knitr

Examples

002-minimumworking-examp

003-model-output 004-figures 005-beamer-

presentation 006-sensitivity analysis-oneparameter

007-sensitivityanalysis-manyparameters 008-large-

009-rmarkdow 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 (RR)

Quoi?

001-exempl

Getting Starte

INT.V

RStudio

knitr

002-minimum-

002-minimumworking-example

003-model-output

004-figures

005-beamerpresentation

006-sensitivity analysis-oneparameter

007-sensitivityanalysis-many-

anatysis-manyparameters

009-rmarkdowi

010-rmarkdow

E ... I B I

rillat Reillarks

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

Extracting output from Regression Models

Reproductib (RR)

Quoi? Pourquoi

001-exemple

Getting Starte

ETEX

RStud

Examples

002-minimumworking-exampl

003-model-output

005 moder output

005-beamer-

006-sensitivity analysis-one-

007-sensitivityanalysis-many-

analysis-manyparameters

000 nmarkdown

010-rmarkdown

rillat Kelliai K

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

Figures

Recherche Reproductib (RR)

Quoi?

001-exempl

Getting Starte

were

RStudi

knit

002-minimum

working-example

004-figures

004-ngures

presentatio

006-sensitivit analysis-one-

007-sensitivity analysis-many

parameters

009-rmarkdow

010-rmarkdow

Final Remarks

rillat Reillarks

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

Beamer Presentations

Recherche Reproductib (RR)

Quoi?

001-exempl

Getting Starte

INT V

DC+nq

knit

Examr

002-minimum-

working-example

004-figures

005-beamer-

presentation

006-sensitivity analysis-one-

007-sensitivityanalysis-many-

parameters

009-rmarkdowr

009-rmarkdow

i illat Kelliaik.

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

Changing one Parameter in an Analysis

Recherche Reproductibl (RR)

Quoi? Pourquoi

rourquor: 001-exempl iustificatif

Getting Starte

ETEX

RStud

Evamn

002-minimumworking-example

003-model-output

004-figures

005-beamer

006-sensitivityanalysis-oneparameter

007-sensitivityanalysis-many-

analysis-manyparameters

009-rmarkdowi

009-rmarkdowr

Einal Domarks

Final Remarks

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

Changing Many Parameters in an Analysis

Recherche Reproductibl (RR)

Quoi? Pourquoi?

001-exempl justificatif

Getting Starte

ETEX

RStud

Examples 002-minimum

002-minimumworking-example

003-model-outpu 004-figures

005-beamer-

006-sensitivit analysis-one-

007-sensitivityanalysis-manyparameters

008-largedocument

009-rmarkdow

Final Remarks

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

Large Documents

Recherche Reproductib (RR)

Quoi?

001-exempl

Getting Starte

we w

DOM: N

RStu

Knit

002-minimum-

working-example

003-modet-outpu

005-beame

presentation

analysis-oneparameter

analysis-ma

008-large-

009-rmarkdow

010-rmarkdownpresentation

Final Remark

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

HTML Reports

Recherche Reproductib (RR)

Pourquoi

001-exemple

Getting Starte

KT-Y

RStud

knitr

Examples 002-minimum-

002-minimumworking-example

003-model-outpu

005-beame

presentation

006-sensitivity analysis-oneparameter

007-sensitivit

008-large-

009-rmarkdown

010-rmarkdown

Final Remarks

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

HTML Presentations

Recherche Reproductib (RR)

Quoi?

001-exempl

Getting Starte

detting starte

BIEX

RStud

Evamo

002-minimum-

working-example 003-model-outpu

004-figures

005-beamer

006-sensitivit

parameter 007-sensitivity-

analysis-manyparameters

008-large-

009-rmarkdow

010-rmarkdownpresentation

Final Remarks

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

Reproductible (RR)

Quoi?

Pourquoi? 001-exempl iustificatif

Getting Started

RStudio

Examples 002-minimum-

working-example 003-model-output

004-figures

presentation

006-sensitivity analysis-oneparameter

007-sensitivityanalysis-many-

parameters

009-rmarkdown

009-rmarkdow

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.

Always Remember ...

Recherche Reproductibl (RR)

Quoi?

001-exempl

Getting Starte

KT-Y

RStudi

kniti

Example

002-minimun

working-exampl

003-model-out

007 hanns

005-beame

presentation

analysis-one

007-sensitivity

analysis-many parameters

nne-largo-

009-rmarkdov

009-rmarkdow

presentation

Final Remarks

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

Is the juice worth the squeeze?

Recherche Reproductib (RR)

Quoi?

001-exempl

Getting Starte

IT V

RStudio

knit

Example

002-minimum

working-example

004-figures

005-beame

procentatio

006-sensitivity analysis-one-

007-sensitivity-

analysis-manyparameters

documents

009-rmarkdow

.....

Final Remarks

