

# Bonnes Pratiques

Jean-Pierre Müller ([jean-pierre.mueller@unil.ch](mailto:jean-pierre.mueller@unil.ch))

22 novembre 2018

# Introduction

# Buts de cette formation:

1. Comprendre ce qu'est "RMarkdown" et pourquoi vous devez l'utiliser
2. Apprendre à utiliser "RMarkdown" dans différents contextes
3. Générer des fichiers de différents formats à l'aide de "RMarkdown"
4. Notebooks : Différencier "RMarkdown" des notebooks "Jupyter" et de l'environnement "Jupyter" et plus généralement de l'environnement Python.

# Reproductibilité — Réplicabilité

Patil Prasad, Peng Roger D. et Leek Jeffrey, « A statistical definition for reproducibility and replicability », bioRxiv, 29.07.2016, p. 066803. En ligne: <https://doi.org/10.1101/066803>, consulté le 14.10.2018.

## Reproductibilité

Une équipe indépendante analyse les mêmes données et doit retomber sur les mêmes résultats.

*« à partir des données, en effectuant les mêmes traitements numériques, je dois obtenir les mêmes résultats numériques. »*

## Réplicabilité

Une équipe indépendante produit de nouvelles données avec le même design et doit retomber sur des données (et résultats) semblables.

*« toutes choses égales par ailleurs, je dois aboutir aux mêmes conclusions. »*

# Reproductibilité — Réplicabilité

La « reproductibilité » n'est possible que si les logiciels utilisés et les données sont publics.

La « répliquabilité » peut être difficile à atteindre selon le domaine scientifique. Voir tableau 1 page 12 de : Leonelli, Sabina (2018) Re-Thinking Reproducibility as a Criterion for Research Quality. [Preprint] URL: [http://philsci-archive.pitt.edu/14352/1/Reproducibility\\_2018\\_SL.pdf](http://philsci-archive.pitt.edu/14352/1/Reproducibility_2018_SL.pdf)

**R, Rstudio, Rmarkdown**

# Une solution : R markdown

- R Markdown permet créer des documents qui contiendront toutes les informations sur vos analyses.
- Vos analyses sont associées à leurs résultats (tableaux, graphiques, ...) et sont enrobées de textes pour les expliquer.
- D'autres personnes peuvent donc comprendre ce que vous avez fait dans vos analyses et être certaines que vous les avez faites correctement.
- Vos fichiers de données, fichiers d'analyses R Markdown peuvent être partagés, ainsi que les fichiers qui en résultent.

# Une solution : R markdown

## Références

Xie Yihui, Allaire J. J. et Golemund Garrett, R Markdown: The Definitive Guide, 2018. En ligne: <https://bookdown.org/yihui/rmarkdown/>, consulté le 08.09.2018.

« R Markdown », <https://rmarkdown.rstudio.com/>, consulté le 14.10.2018.

« rmarkdown-cheatsheet-2.0.pdf », <https://www.rstudio.com/wp-content/uploads/2016/03/rmarkdown-cheatsheet-2.0.pdf>, consulté le 15.10.2018.



# R markdown : Préalables

Les logiciels suivants doivent être installés sur votre machine :

- R ( <https://cran.r-project.org/> )
- RStudio ( <https://www.rstudio.com/> )
- $T_E X - L_A T_E X$ 
  - ( <https://openclassrooms.com/fr/courses/1617396-redigez-des-documents-de-qualite-avec-latex/1617747-installer-latex> )
  - ou dans R :

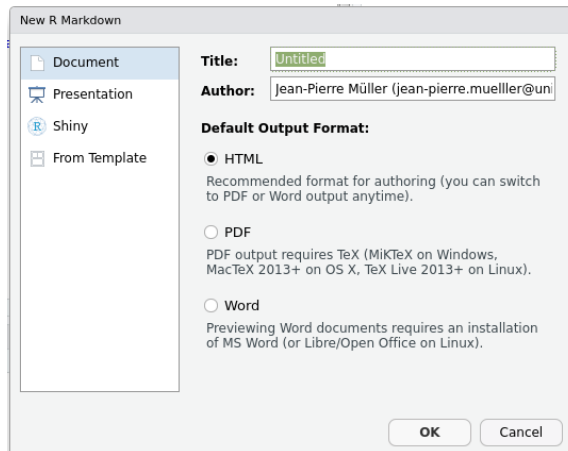
```
# install TinyTeX
install.packages("tinytex")
tinytex::install_tinytex()
```

# R markdown : les bases

Créer un document de base :

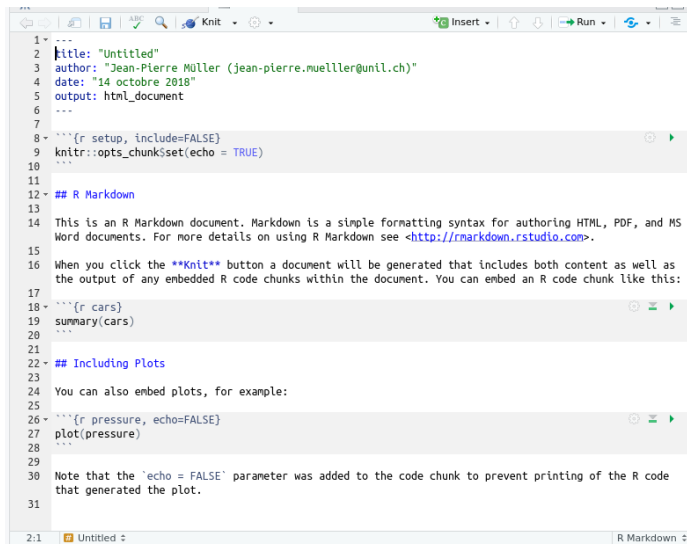
Dans RStudio, menu **File** -> **New File** -> **R Markdown**.

Saisir un titre, un auteur et valider le dialogue.



# R markdown : les bases (2)

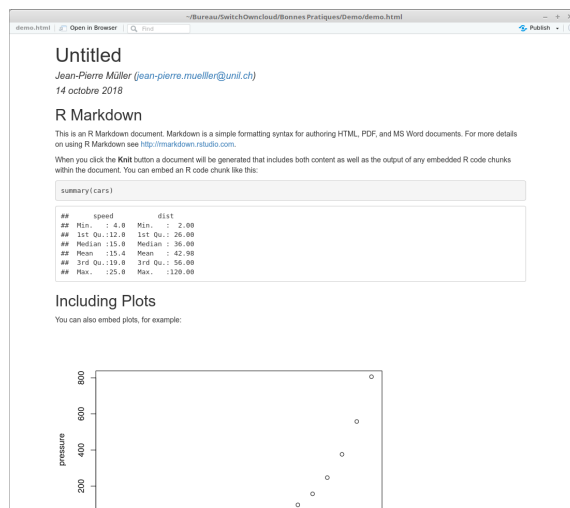
Un squelette de document est affiché



```
1 ---
2 title: "Untitled"
3 author: "Jean-Pierre Müller (jean-pierre.mueller@unil.ch)"
4 date: "14 octobre 2018"
5 output: html_document
6 ---
7
8 ```{r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = TRUE)
10 ```
11
12 ## R Markdown
13
14 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS
15 Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.
16
17 When you click the Knit button a document will be generated that includes both content as well as
18 the output of any embedded R code chunks within the document. You can embed an R code chunk like this:
19
20 ```{r cars}
21 summary(cars)
22 ```
23
24 ## Including Plots
25
26 You can also embed plots, for example:
27
28 ```{r pressure, echo=FALSE}
29 plot(pressure)
30 ```
31
32 Note that the 'echo = FALSE' parameter was added to the code chunk to prevent printing of the R code
33 that generated the plot.
```

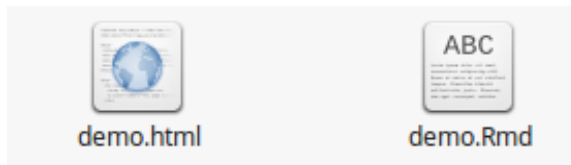
# R markdown : les bases (3)

Cliquer sur le bouton **Knit**, nommer le fichier: une fenêtre avec un document **.html** est affichée.



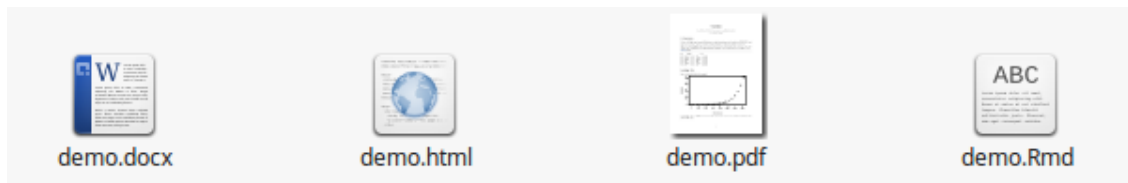
## R markdown : les bases (4)

Dans le dossier où le fichier a été sauvegardé, on a maintenant deux fichiers :



Sélectionner maintenant les articles du menu **Knit** : **Knit to PDF** et **Knit to Word**.

Un document PDF ou un document MSWord sont affichés. Et le dossier contient maintenant 4 fichiers:



# R markdown : les bases (5)

Le fichier du squelette contient maintenant dans la partie YAML metadata:

```
---  
title: "Untitled"  
author: "Jean-Pierre Müller (jean-pierre.mueller@unil.ch)"  
date: "14 octobre 2018"  
output:  
  word_document: default  
  pdf_document: default  
  html_document: default  
---
```

Vous pouvez supprimer les lignes "pdf\_..." et "word\_...".

# R markdown : mise en forme du texte

Texte normal

Texte normal

Pour forcer un retour à la ligne  
terminer la ligne par deux ou plus signes "espace".

Pour forcer un retour à la ligne  
terminer la ligne par deux ou plus signes "espace".

*\*Italiciser\** et mettre **\*\*en gras\*\***

*Italiciser* et mettre **en gras**

# R markdown : mise en forme du texte (2)

Dans un texte, ``du texte pré-formaté ou code source` ...`

Dans un texte, du texte pré-formaté ou code source ...

Exposant<sup>2</sup>

Exposant<sup>2</sup>

indice<sub>2</sub>

Indice<sub>2</sub>



# R markdown : mise en forme du texte (3)

~~Barré~~

~~Barré~~

caractères spéciaux : \\* \\_ \\

caractères spéciaux : \* \_ \

endash: --, emdash: ---

endash: –, emdash: —

# R markdown : mise en forme du texte (4)

```
> Beware of bugs in the above code; I have only proved it correct, not tried it.  
>  
> ---Donald Knuth
```

Beware of bugs in the above code; I have only proved it correct, not tried it.

—Donald Knuth

Ligne horizontale ou séparation dans une présentation  
\*\*\*

# R markdown : mise en forme du texte (5)

Définition

: MATH. Convention logique a priori.

**Définition**

MATH. Convention logique a priori.

<!--Commentaire-->

```\n

Bloc de texte pré-formaté

ou

code source

```\n

Bloc de texte pré-formaté

ou

code source

# R markdown : mise en forme du texte (6)

```
# Titre de Niveau 1
## Titre de Niveau 2
### Titre de Niveau 3
#### Titre de Niveau 4
##### Titre de Niveau 5
##### Titre de Niveau 6
```

Titre de Niveau 3

Titre de Niveau 4

Titre de Niveau 5

Titre de Niveau 6

# R markdown : mise en forme du texte (7)

- \* liste à puces
  - + élément 1
  - + élément 2
    - sous-niveau 2
    - sous-niveau 2
- \* liste à puces : 2ème puce (avec 2 espaces en fin de ligne)  
Suite (indenté de 4 espaces)
- \* liste à puces : 3ème puce

# R markdown : mise en forme du texte (8)

- liste à puces
  - élément 1
  - élément 2
    - sous-niveau 2
    - sous-niveau 2
- liste à puces : 2ème puce (avec 2 espaces en fin de ligne)  
Suite (indenté de 4 espaces)
- liste à puces : 3ème puce

# R markdown : mise en forme du texte (9)

1. Liste ordonnée
2. 2ème élément
  - i) sous-niveau 1
  - ii) sous-niveau 2
    - A. sous-sous-niveau 1
    - B. sous-sous-niveau 2

# R markdown : mise en forme du texte (10)

1. Liste ordonnée
2. 2ème élément
  1. sous-niveau 1
  2. sous-niveau 2
    - A. sous-sous-niveau 1
    - B. sous-sous-niveau 2



# R markdown : tableaux simples

| Right | Left | Default | Center |
|-------|------|---------|--------|
| 12    | 12   | 12      | 12     |
| 123   | 123  | 123     | 123    |
| 1     | 1    | 1       | 1      |

| Right | Left | Default | Center |
|-------|------|---------|--------|
| 12    | 12   | 12      | 12     |
| 123   | 123  | 123     | 123    |
| 1     | 1    | 1       | 1      |

---

# R markdown : liens

lien explicite : `<http://www.unil.ch>`  
[lien vers UNIL](http://www.unil.ch)

lien explicite : <http://www.unil.ch>  
[lien vers UNIL](#)

# R markdown : Images

Il est recommandé de créer un dossier contenant les images à insérer dans le document. Par exemple, pour les logos unil, créés à partir de <http://unilogo.unil.ch> :

```
![Légende : logo unil](unilogo_logos/  
Word-Excel-Powerpoint/unilogo_bleu_300dpi.png)
```

## R markdown : Images (2)



**UNIL** | Université de Lausanne

Faculté des SSP

Légende : logo unil

# R markdown : code R

## Morceaux de code R (code chunks)

Il existe deux modes pour les morceaux de code R:

```
```{r cars}  
summary(cars)  
```
```

et

La moyenne des vitesses est de  
`r mean(cars\$speed)` Mph.

## R markdown : code R (2)

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   :  2.00
##  1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##  Mean   :15.4    Mean   : 42.98
##  3rd Qu.:19.0    3rd Qu.: 56.00
##  Max.   :25.0    Max.   :120.00
```

La moyenne des vitesses est de 15.4 Mph.

# R markdown : code R (3)

## Instructions d'évaluation des Code Chunks

**eval=**

évaluer le bloc ?

**echo=**

afficher le code ?

**warning=**

afficher les alertes?

**error=**

afficher les erreurs?

**message=**

afficher les messages?

# R markdown : code R (4)

## Instructions d'évaluation des Code Chunks

**tidy=**

formater le code selon les normes ?

**results=**

forme des résultats, "hide", "asis", "markup".

**fig.width=, fig.height=**

dimensions (en pouces) des graphiques.

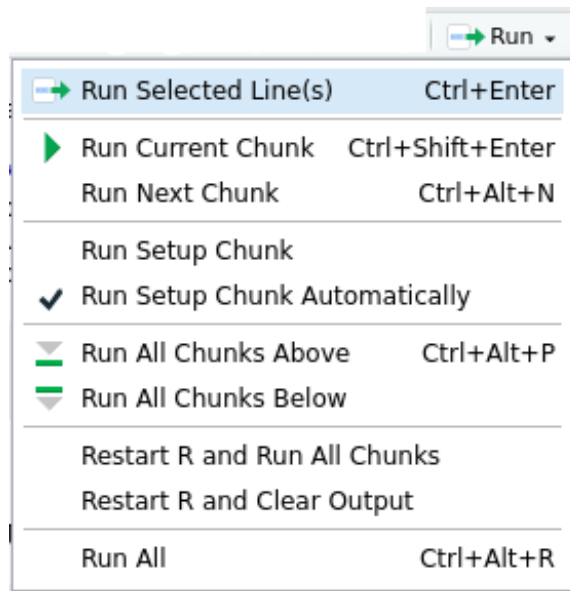
**fig.align=**

alignement de la figure, "left", "right", "center".



# R markdown : code R (5)

- Menu Global



- Boutons locaux



# R markdown : afficher des résultats

L'affichage des résultats par défaut est celui de R.

Il est possible d'obtenir des résultats de meilleure qualité en :

- effectuant les calculs dans un bloc avec l'instruction `echo=FALSE`.
- en affichant le résultat à l'aide d'un bloc spécifique.

L'affichage des résultats peut se faire en utilisant des fonctions des

```
library("knitr")  
library("pander")  
library("papeR")  
library("stargazer")
```

# R markdown : afficher des résultats

knitr::kable (exemple de la vignette)

Avec les instructions: `{r, echo=FALSE, results='asis'}`

```
kable(iris[1:4,], align=c('l', 'c', 'r', 'l', 'r'))
```

| Sepal.Length | Sepal.Width | Petal.Length | Petal.Width | Species |
|--------------|-------------|--------------|-------------|---------|
| 5.1          | 3.5         | 1.4          | 0.2         | setosa  |
| 4.9          | 3.0         | 1.4          | 0.2         | setosa  |
| 4.7          | 3.2         | 1.3          | 0.2         | setosa  |
| 4.6          | 3.1         | 1.5          | 0.2         | setosa  |

---

# R markdown : afficher des résultats

knitr::kable (exemple de la vignette)

Avec les instructions: `{r, echo=FALSE, results='asis'}`

```
kable(mtcars[1:4, 1:4], booktabs=TRUE)
```

|                | mpg  | cyl | disp | hp  |
|----------------|------|-----|------|-----|
| Mazda RX4      | 21.0 | 6   | 160  | 110 |
| Mazda RX4 Wag  | 21.0 | 6   | 160  | 110 |
| Datsun 710     | 22.8 | 4   | 108  | 93  |
| Hornet 4 Drive | 21.4 | 6   | 258  | 110 |

---

# R markdown : afficher des résultats

pander (exemple de la vignette)

Avec les instructions: {r, echo=FALSE, results='asis'}

```
library("pander")
t <- mtcars[1:3, 1:5]
emphasize.italics.cols(1)
emphasize.italics.rows(1)
emphasize.strong.cells(which(t > 20, arr.ind = TRUE))
pandoc.table(t)
```

# R markdown : afficher des résultats

|               | mpg  | cyl | disp | hp  | drat |
|---------------|------|-----|------|-----|------|
| Mazda RX4     | 21   | 6   | 160  | 110 | 3.9  |
| Mazda RX4 Wag | 21   | 6   | 160  | 110 | 3.9  |
| Datsun 710    | 22.8 | 4   | 108  | 93  | 3.85 |

---

# R markdown : afficher des résultats

papeR (exemple de la vignette)

Avec les instructions: {r, results='hide', message=FALSE}

```
library("papeR")  
library("nlme")  
(tab1 <- summarize(Orthodont, type = "numeric"))
```

Avec les instructions: {r, echo=FALSE, results='asis'}

```
kable(tab1)
```

# R markdown : afficher des résultats

|          | N   | Mean  | SD   | Min  | Q1 | Median | Q3 | Max  |
|----------|-----|-------|------|------|----|--------|----|------|
| distance | 108 | 24.02 | 2.93 | 16.5 | 22 | 23.75  | 26 | 31.5 |
| age      | 108 | 11.00 | 2.25 | 8.0  | 9  | 11.00  | 13 | 14.0 |

---



# R markdown : afficher des résultats

## papeR (exemple de la vignette)

Avec les instructions: `{r, results='hide', message=FALSE}`

```
library("papeR")  
library("nlme")  
(tab3 <- summarize(Orthodont, type = "factor",  
                   variables = "Sex",  
                   cumulative = TRUE))
```

Avec les instructions: `{r, echo=FALSE, results='asis'}`

```
kable(tab3)
```

# R markdown : afficher des résultats

|     | Level  | N  | %    | $\Sigma$ % |
|-----|--------|----|------|------------|
| Sex | Male   | 64 | 59.3 | 59.3       |
|     | Female | 44 | 40.7 | 100.0      |

---

# R markdown : afficher des résultats

## papeR (exemple de la vignette)

Avec les instructions: `{r, results='hide', message=FALSE}`

```
library("papeR")  
library("nlme")  
(tab4 <- summarize(Orthodont, type = "numeric",  
                    group = "Sex",  
                    test = FALSE))
```

Avec les instructions: `{r, echo=FALSE, results='asis'}`

```
kable(tab4)
```

# R markdown : afficher des résultats

|     |          | Sex    | N  | Mean  | SD   | Min  | Q1 | Median | Q3    | Max  |
|-----|----------|--------|----|-------|------|------|----|--------|-------|------|
| 1   | distance | Male   | 64 | 24.97 | 2.90 | 17.0 | 23 | 24.75  | 26.50 | 31.5 |
| 1.1 |          | Female | 44 | 22.65 | 2.40 | 16.5 | 21 | 22.75  | 24.25 | 28.0 |
| 2   | age      | Male   | 64 | 11.00 | 2.25 | 8.0  | 9  | 11.00  | 13.00 | 14.0 |
| 2.1 |          | Female | 44 | 11.00 | 2.26 | 8.0  | 9  | 11.00  | 13.00 | 14.0 |

---

# R markdown : afficher des résultats

## papeR (exemple de la vignette)

Avec les instructions: {r, results='hide', message=FALSE}

```
library("papeR")
library("nlme")
data(Orthodont, package = "nlme")
linmod <- lm(distance ~ age + Sex, data = Orthodont)
## Extract pretty summary
(pretty_lm <- prettify(summary(linmod)))
```

Avec les instructions: {r, echo=FALSE, results='asis'}

```
kable(pretty_lm)
```

# R markdown : afficher des résultats

|             | Estimate   | CI (lower) | CI (upper) | Std. Error | t value   | Pr(>  t ) |     |
|-------------|------------|------------|------------|------------|-----------|-----------|-----|
| (Intercept) | 17.7067130 | 15.5014071 | 19.9120189 | 1.1122095  | 15.920304 | <0.001    | *** |
| age         | 0.6601852  | 0.4663472  | 0.8540231  | 0.0977589  | 6.753194  | <0.001    | *** |
| Sex: Female | -2.3210227 | -3.2031499 | -1.4388955 | 0.4448862  | -5.217115 | <0.001    | *** |

---

# R markdown : afficher des résultats

stargazer (exemple de la vignette)

Avec les instructions: ``{r, echo = FALSE, message=FALSE, results='asis'}`

```
library("stargazer")  
stargazer(attitude, type="html")
```

| Statistic  | N   | Mean  | St. Dev. | Min | Max |
|------------|-----|-------|----------|-----|-----|
| rating     | 306 | 4.633 | 12.173   | 40  | 85  |
| complaints | 306 | 6.600 | 13.315   | 37  | 90  |
| privileges | 305 | 3.133 | 12.235   | 30  | 83  |
| learning   | 305 | 6.367 | 11.737   | 34  | 75  |
| raises     | 306 | 4.633 | 10.397   | 43  | 88  |
| critical   | 307 | 4.767 | 9.895    | 49  | 92  |
| advance    | 304 | 2.933 | 10.289   | 25  | 72  |

# R markdown : afficher des résultats

stargazer (exemple de la vignette)

Avec les instructions: ``{r, echo = FALSE, message=FALSE, results='asis'}`

```
library("stargazer")
## 2 OLS models
linear.1 <- lm(rating ~ complaints + privileges + learning + raises + critical,
data=attitude)
linear.2 <- lm(rating ~ complaints + privileges + learning, data=attitude)
## create an indicator dependent variable, and run a probit model
attitude$high.rating <- (attitude$rating > 70)
probit.model <- glm(high.rating ~ learning + critical + advance, data=attitude,
family = binomial(link = "probit"))
stargazer(linear.1, linear.2, probit.model, title="Results", align=TRUE, type="html")
```



# R markdown : afficher des résultats

| Results    |                            |                     |                              |
|------------|----------------------------|---------------------|------------------------------|
|            | <i>Dependent variable:</i> |                     |                              |
|            | rating<br><i>OLS</i>       |                     | high.rating<br><i>probit</i> |
|            | (1)                        | (2)                 | (3)                          |
| complaints | 0.692***<br>(0.149)        | 0.682***<br>(0.129) |                              |
| privileges | -0.104<br>(0.135)          | -0.103<br>(0.129)   |                              |
| learning   | 0.249<br>(0.160)           | 0.238*<br>(0.139)   | 0.164***<br>(0.053)          |
| raises     | -0.033<br>(0.202)          |                     |                              |
| critical   | 0.015<br>(0.147)           |                     | -0.001<br>(0.044)            |
| advance    |                            |                     | -0.062<br>(0.042)            |
| Constant   | 11.011<br>(11.704)         | 11.258<br>(7.318)   | -7.476**<br>(3.570)          |

# R markdown : afficher des résultats

Lors de la production de fichiers ".pdf" supprimer `type="html"` dans les paramètres de la fonction `stargazer()`.

# R markdown : afficher des résultats

rapport

<http://rapport-package.info/>

```
library("rapport")
```

# R markdown :

kableExtra

<http://haozhu233.github.io/kableExtra/>

« Create Awesome HTML Table with knitr::kable and kableExtra », [https://haozhu233.github.io/kableExtra/awesome\\_table\\_in\\_html.html](https://haozhu233.github.io/kableExtra/awesome_table_in_html.html), consulté le 08.09.2018.

```
library(kableExtra)
```

# R markdown : notes de bas de page

notes de bas de page, par exemple<sup>1</sup>

<sup>1</sup>: texte de la note

Ne marche pas dans une présentation!

notes de bas de page, par exemple<sup>1</sup>

# R markdown : équations mathématiques

Formule  $(x_i - \overline{x})^2$  dans un texte

Formule  $(x_i - \overline{x})^2$  dans un texte

Bloc mathématique centré:  $\frac{1}{n} \sum_{i=1}^n E(X_i)$

Bloc mathématique centré:

$$\frac{1}{n} \sum_{i=1}^n E(X_i)$$

# R markdown : équations math. (2)

Une feuille de triche simple...

« undergradmath.pdf », <http://joshua.smcvt.edu/undergradmath/undergradmath.pdf>, consulté le 17.10.2018.

Plus compliquée :

Lammarsch Marion, « LaTeX Cheat Sheet », p. 4. En ligne: [ftp://ftp.dante.de/tex-archive/info/latex-refsheet/LaTeX\\_RefSheet.pdf](ftp://ftp.dante.de/tex-archive/info/latex-refsheet/LaTeX_RefSheet.pdf), consulté le 17.10.2018.

# R markdown : citer

Télécharger un fichier .csl de <https://www.zotero.org/styles> et le placer dans le même dossier que le fichier R Markdown.

Ajouter dans l'entête YAML du fichier R Markdown:

```
bibliography: bibliographie.bib  
csl: fichier_telechargé.csl
```

Après le dernier paragraphe du document, ajouter,

```
# References
```

Exporter vos références au format bibtex dans un fichier `bibliographie.bib` et le placer dans le même dossier que le fichier R Markdown.



# R markdown : citer (2)

Exemple de fichier bibtex :

```
@article{wan_analytical_2012,  
  title = {Analytical {Sociology}: {A} {Bungean} {Appreciation}},  
  volume = {21},  
  issn = {0926-7220, 1573-1901},  
  shorttitle = {Analytical {Sociology}},  
  url = {http://link.springer.com/10.1007/s11191-011-9427-3},  
  doi = {10.1007/s11191-011-9427-3},  
  language = {en},  
  number = {10},  
  urldate = {2018-02-16},  
  journal = {Science \& Education},  
  author = {Wan, Poe Yu-ze},  
  month = oct,  
  year = {2012},  
  pages = {1545--1565}  
}
```

# R markdown : citer (3)

exemple [voir @wan\_analytical\_2012, pp. 1550-1560 ou  
@kuhn\_structure\_1996, ch. 4].

exemple [@kuhn\_structure\_1996; @lebart\_visualizations\_1998].

Kuhn [-@kuhn\_structure\_1996] ajoute...

@wan\_analytical\_2012 [p. 1563] démontre ...

# R markdown : pdf output et écrire en français

```
---  
title: "Untitled"  
lang: fr  
author: "Jean-Pierre Müller (jean-pierre.mueller@unil.ch)"  
date: "14 octobre 2018"  
geometry: "left=3cm,right=3cm,top=2cm,bottom=2cm"  
output:  
  pdf_document:  
    toc: true  
    number_sections: true  
documentclass: article  
classoption: landscape, a4paper  
bibliography: bibliographie.bib  
csl: sociological-methodology.csl  
---
```

# R markdown : gestion fine de tableaux et graphiques

- Tableaux:
  - créer un `data.frame` dans un bloc
  - Et l'afficher en suite dans un 2ème bloc.
- Graphiques, Images: il est possible d'utiliser

```
knitr::include_graphics(file)
```

# R markdown : calculs longs

Si un calcul est particulièrement long où intensif, il est pratique de ne pas l'inclure dans le fichier Markdown.

- préparer un fichier annexe réalisant le calcul.
- sauvegarder les résultats `save(resultats, file = "xy.RData")`
- relire les résultats dans l'entête du fichier Markdown avec `load(file=xy.RData)`
- diffuser aussi le fichier annexe réalisant le calcul

# R markdown : hasard

Les fonctions de R simulant des tirages aléatoires constituent un autre problème de reproductibilité. Utiliser :

```
set.seed(seed)
```

avant tout calcul faisant intervenir des tirages aléatoires.

ou

utiliser l'instruction d'évaluation des code chunks:

```
cache=TRUE
```

voir <https://yihui.name/knitr/options/> .

# R markdown : persistance logiciels

Les logiciels contiennent des bugs, qui sont corrigés, et donc de nouvelles versions des ces logiciels peuvent amener à des résultats différents de ceux obtenus avec des versions précédentes. Il est possible de corriger ce problème en conservant les versions utilisées lors de la création du fichier. Utiliser :

```
library(packrat)
```

# R markdown : présentations

Dans RStudio, menu **File -> New File -> R Markdown**.

- Sélectionner **Présentation** .
- Saisir un titre, un auteur et valider le dialogue.



# R markdown : présentations (2)

Touches de contrôle:

- 'f': fullscreen mode
- 'w': widescreen mode
- 'o': overview mode
- 'h': code highlight mode
- 'p': monter les notes du

# R markdown : poster

Bucior Ben, Generate Academic Posters in R Markdown and CSS, inspired by « reveal.js »: bbucior/drposter, 2018 [10 juillet 2016]. En ligne: <https://github.com/bbucior/drposter>, consulté le 16.10.2018.

```
devtools::install_github("bbucior/drposter",  
                          dep=FALSE)
```

- Menu  
File -> New File -> R Markdown -> From Template -> drposter Poster
- **knit** génère un fichier .html à imprimer à l'aide de **chrome** (uniquement!)

# R markdown : poster (2)

Autres solutions :

<https://gist.github.com/Pakillo/4854e5d760351206084f6be8abe476b2>

# Rmarkdown : Autres formats prédéfinis

Voir :

- Menu **File** -> **New File** -> **R Markdown** -> **From Template** ->

Packages R contenant des templates :

rticles, tint, resumer, linl, tufte, prettydoc...

Voir aussi : <http://jianghao.wang/post/2017-12-08-rmarkdown-templates/>

**Bookdown**

# Bookdown

Permet de générer un site web, des fichiers word, odt, ebook et pdf à partir de plusieurs fichiers .Rmd.

<https://bookdown.org/>

Xie Yihui, bookdown: Authoring Books and Technical Documents with R Markdown, 2018. En ligne: <https://bookdown.org/yihui/bookdown/>, consulté le 08.09.2018.

# Bookdown (2)

```
# stable version on CRAN  
install.packages("bookdown")  
# or development version on GitHub  
# devtools::install_github('rstudio/bookdown')
```

Télécharger et unzip-er :

<https://github.com/rstudio/bookdown-demo>

( ou exemple minimal : <https://github.com/yihui/bookdown-minimal> )

## Bookdown (3)

Fermer Rstudio et double-cliquer sur `bookdown-demo.Rproj` .

Le quadrant Environnement / History de RStudio affiche maintenant un onglet supplémentaire :



Pour générer le "livre", cliquer sur Build Book.  
La page web du livre va s'afficher.



# Bookdown (4)

Les principales caractéristiques d'un ouvrage peuvent être gérées dans cet environnement.

- Numérotation des figures et tableaux
- Index
- Renvois

La génération du pdf fait appel à de nombreuses fonctionnalités de xetex, xelatex, ...

# Bookdown (5)

Cela peut devenir compliqué selon ce que l'on veut obtenir. Il vous faudra peut-être utiliser:

```
output:  
  pdf_document:  
    extra_dependencies:  
      hyperref: ["unicode=true", "breaklinks=true"]  
      lmodern: null
```

ou, plus général :

```
output:  
  pdf_document:  
    includes:  
      in_header: preamble.tex  
      before_body: doc-prefix.tex  
      after_body: doc-suffix.tex
```

# Bookdown (6)

Liens :

Casas Pablo, « How to self-publish a book: Customizing Bookdown », Data Science Heroes Blog, 28.08.2018, <https://blog.datascienceheroes.com/how-to-self-publish-a-book-customizing-bookdown/>, consulté le 13.11.2018.

Casas Pablo, « How to self-publish a book on Amazon using Bookdown », Data Science Heroes Blog, 21.08.2018, <https://blog.datascienceheroes.com/how-to-self-publish-a-book/>, consulté le 13.11.2018.

**Notebooks**

# La guerre des notebooks

**Notebook** Equivalent du carnet d'expériences dans un laboratoire (par exemple : <http://www.cnrs.fr/infoslabos/cahier-laboratoire/> ) ou des notes de terrain pour l'ethnologue ( par exemple: <https://medium.com/media-ethnography/field-notes-and-participant-observation-in-ethnographic-studies-a-skill-summary-bb74e3881258> ), mais pour les travaux informatiques.

But annoncé : Reproductibilité.

Origine : ancienne, mais essor avec les développeurs et utilisateurs de python (= data scientists).

Très utilisé : Perkel Jeffrey M., « Why Jupyter is data scientists' computational notebook of choice », Nature 563, 30.10.2018, p. 145. En ligne: <https://doi.org/10.1038/d41586-018-07196-1>, consulté le 30.10.2018.

# La guerre des notebooks (2)

## L'attentat

<https://docs.google.com/presentation/d/1n2RIMdmv1p25Xy5thJUhkKGvjtV-dkAIsUXP-AL4ffl/>

## Quelques remarques

Différence de culture entre statisticiens et informaticiens.

<https://yihui.name/en/2018/09/notebook-war/>

<https://minimaxir.com/2017/06/r-notebooks/>

# R Notebooks

Un R notebook permet de générer un fichier `.nb.html` qui peut être ouvert dans un browser web (pour voir le résultat) ou dans RStudio (pour voir le code). C'est un compromis entre la séduction des fichiers `.html` et la possibilité de d'accéder aux coulisses du code.

Dans **RStudio**, menu **File** -> **New File** -> **R Notebook**.

Exécuter tous les morceaux de code R et cliquer sur le bouton **Preview**.

## R Notebooks (2)

Le fichier `.nb.html` est créé :

- vous pouvez l'ouvrir dans un browser web.
- vous pouvez l'ouvrir depuis RStudio; le fichier original `.Rmd` sera recréé et vous pourrez le sauver...

Il est possible aussi de générer le fichier `.Rmd` depuis le menu Code du fichier affiché dans un browser.

N.B. : Aucun code R n'est exécuté dans le fichier `.nb.html`.



# R Jupyter Notebook

Jupyter : acronyme de Julia, Python, et R.

Il faut installer Python pour l'utiliser :

<https://www.python.org/downloads/>

(sur linux utiliser le package manager de votre distribution)

# Jupyter install - Déconseillé (python seulement)

Avec python seulement et dans un terminal :

*# On Windows*

```
python -m pip install -U pip setuptools
```

*# On OS X or Linux*

```
pip install -U pip setuptools
```

et

*# Python 3*

```
pip3 install jupyter
```

Pour l'utiliser :

```
jupyter notebook
```

# anaconda et conda

De <https://geohackweek.github.io/Introductory/01-conda-tutorial/> :

**"What is Conda?** conda is an open source package and environment management system for any programming language"

**"What is Anaconda?** Anaconda is a distribution of conda. It is a data science platform that comes with a lot of packages."

Liens : <https://www.anaconda.com/> , <https://conda.io/docs/>

**Attention** : Anaconda installe une copie de tous ses composants dans son propre dossier.

# Installer Jupyter via anaconda

Installer anaconda :

<https://www.anaconda.com/download>

et suivre les instructions destinées à votre OS sur:

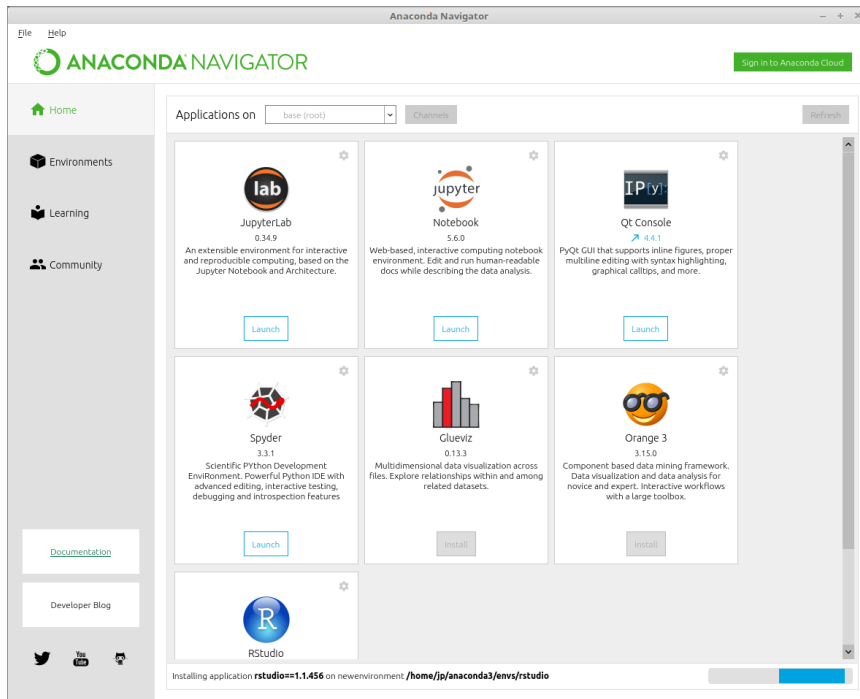
<http://docs.anaconda.com/anaconda/install/>

Lancer anaconda sur MacOS ou linux; dans un terminal saisir :

`anaconda-navigator`

Sous Windows, utiliser l'entrée du menu Démarrer . . . .

# Anaconda Navigator



Ouvrir Jupyter depuis anaconda : cliquer sur le bouton Launch de la tuille Jupyter.

# Utiliser Jupyter

Une page web <http://localhost:8888> s'ouvre :

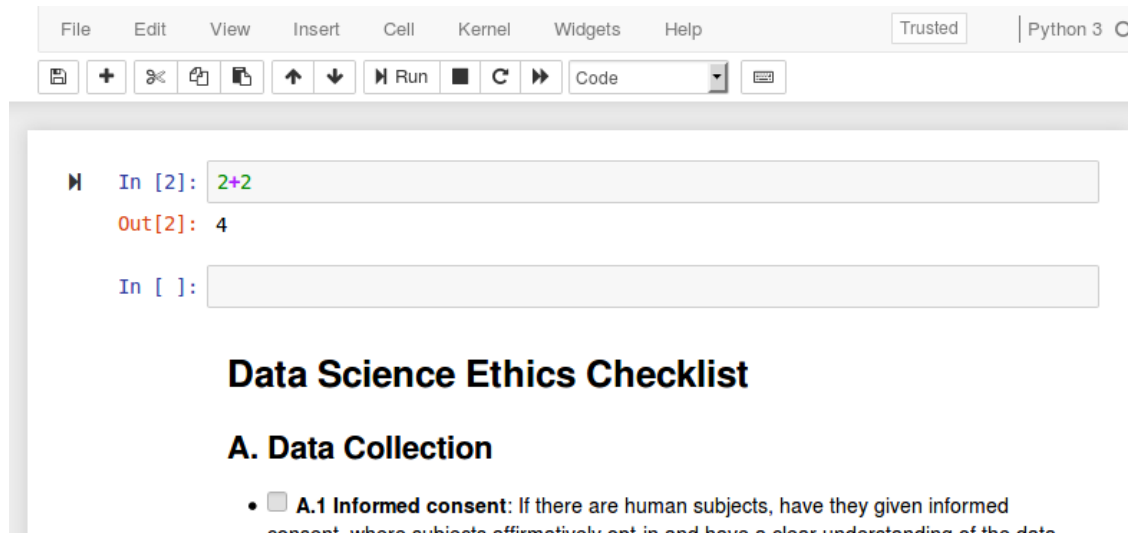


3 onglets: Files, Running, Clusters

Créer un nouveau fichier, menu **New**, article **Python 3**

# Utiliser Jupyter (2)

Nouvelle page web :



- Edition et interprétation du code par cellule
- Essentiellement deux types de cellules

# Exemple : Utiliser un notebook existant

Aller sur

<https://github.com/jupyter/jupyter/wiki/A-gallery-of-interesting-Jupyter-Notebooks>

et choisir un notebook qui vous intéresse, par exemple :

<http://nbviewer.jupyter.org/github/fbkarsdorp/python-course/blob/master/Chapter%201%20-%20Getting%20started.ipynb>

Le télécharger (icône Download Notebook), l'ouvrir depuis Jupyter, le compléter / l'exécuter.

Enregistrer une version pdf ou html : menu **File, Download as** .



# Installer le kernel R pour Jupyter

Déconseillé! Depuis un terminal, dans R :

```
install.packages(c('repr', 'IRdisplay', 'evaluate',  
                  'crayon', 'pbdZMQ', 'devtools',  
                  'uuid', 'digest'))  
devtools::install_github('IRkernel/IRkernel')
```

```
IRkernel::installspec()
```

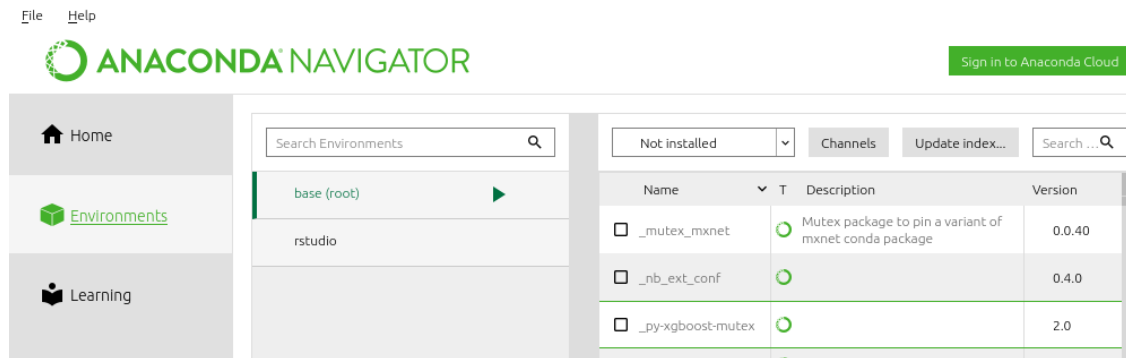
et dans le terminal

```
pip install rpy2
```

# Activer R dans Jupyter

Ou l'activer dans anaconda...

Dans Anaconda Navigator : Environments



Sélectionner dans la liste des packages à installer :

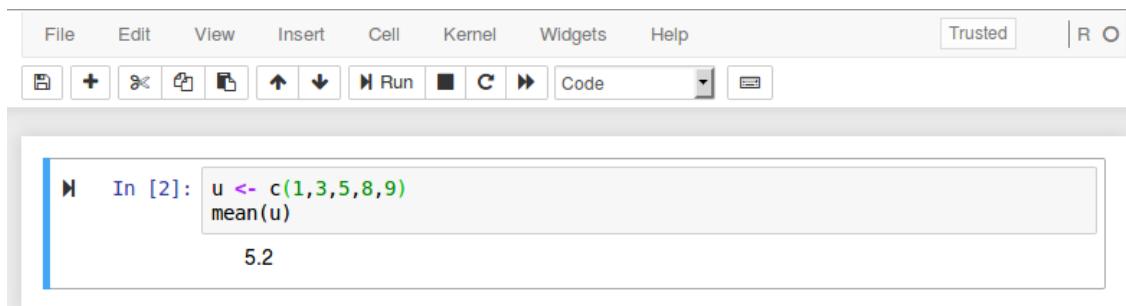
- rstudio, r-essentials, rpy2, r-irkernel
- et d'autres éléments r- . . . . .

Eventuellement créer un nouvel environnement rstudio.

# Utiliser R dans un Jupyter Notebook

Créer un Notebook R, dans Jupyter, dans l'onglet Files : menu New, article R .

Les cellules de code sont interprétées comme contenant des commandes R :



The screenshot shows the Jupyter Notebook interface. At the top is a menu bar with 'File', 'Edit', 'View', 'Insert', 'Cell', 'Kernel', 'Widgets', and 'Help'. To the right of the menu bar is a 'Trusted' status indicator and a language selector showing 'R'. Below the menu bar is a toolbar with icons for saving, adding a new cell, undo, redo, deleting a cell, moving a cell up/down, running the cell, and a dropdown menu currently set to 'Code'. The main area contains a single code cell with the prompt 'In [2]:'. The code inside the cell is `u <- c(1,3,5,8,9)` followed by `mean(u)` on the next line. The output of the cell, '5.2', is displayed below the code.

```
In [2]: u <- c(1,3,5,8,9)
        mean(u)

        5.2
```

# Utiliser R dans un Jupyter Python Notebook

Dans une cellule du Notebook, saisir :

```
# charger r magic  
%load_ext rpy2.ipython
```

```
# Exemple de code R  
%R require(ggplot2)  
%R y <- 2+2  
%R y
```

# Jupyter : commandes utiles

Quelques commandes utiles dans une cellule:

```
%lsmagic  
?%alias_magic  
?range
```

**Pour terminer**

# Dernières remarques : Structurer votre projet

```
library(adapr)
```

Gelfond Jonathan, Goros Martin, Hernandez Brian et al., « A System for an Accountable Data Analysis Process in R », The R Journal 10 (1), 2018, pp. 6-21.  
En ligne: <https://journal.r-project.org/archive/2018/RJ-2018-001/index.html>,  
consulté le 14.10.2018.

Utile pour un **DMP** sophistiqué!

# Dernières remarques : Données, anonymisation et protection des données

Vos données doivent être distribuées avec vos fichiers de production des résultats, articles, etc..

Attention de ne pas diffuser des données permettant d'identifier ne serait-ce qu'un seul individu ...

... Il est donc souvent difficile de diffuser un fichier de données brut sans traitement.

Solution: diffuser le fichier de nettoyage des données brutes et de restriction aux variables utilisées dans les analyses. Ainsi que le fichier des données pré-traitées obtenu.



# Dernières remarques : Diffuser vos analyses et publications

<https://www.dlcm.ch/> (l'UniL est partenaire)

Présentation:

[http://misc.www.switch.ch/media/2018/engines/9%20DLCM\\_IT\\_for\\_Research\\_Day\\_2018.pdf](http://misc.www.switch.ch/media/2018/engines/9%20DLCM_IT_for_Research_Day_2018.pdf)

git/SVN

<https://support.rstudio.com/hc/en-us/articles/200532077?version=1.1.447&mode=desktop>

FORS

<https://forsbase.unil.ch/>

**Happy computing ! ;-)**