

R + LATEX

Une très brève introduction

Objectifs & Philosophie

Objectifs & Philosophie

- ▶ Installation: RStudio + R + L^AT_EX

Objectifs & Philosophie

- ▶ Installation: RStudio + \mathbb{R} + \LaTeX
- ▶ Résumé de la journée: \mathbb{R} et \LaTeX
 1. Le monde de l'open source
 2. Quelques bases de programmation \mathbb{R}
 3. Faire des graphiques en \mathbb{R} avec \LaTeX

Objectifs & Philosophie

- ▶ Installation: RStudio + \mathbb{R} + \LaTeX
- ▶ Résumé de la journée: \mathbb{R} et \LaTeX
 1. Le monde de l'open source
 2. Quelques bases de programmation \mathbb{R}
 3. Faire des graphiques en \mathbb{R} avec \LaTeX
- ▶ Philosophie: Les outils *avant* la méthode

Installation

2 Languages, 1 Software



Installation

2 Languages, 1 Software



1. R : www.cran.rstudio.com

Installation

2 Languages, 1 Software



1. R : www.cran.rstudio.com
2. L^AT_EX: www.latex-project.org/get/

Installation

2 Languages, 1 Software



1. R : www.cran.rstudio.com
2. L^AT_EX: www.latex-project.org/get/
3. RStudio: www.rstudio.com/products/rstudio/download/

Installation

2 Languages, 1 Software

The screenshot shows the RStudio interface with the following details:

- Project:** (None)
- Environment:** Shows "Environment is empty".
- Files:** Displays several files: JfMorin.tex*, Class3.tex, beamerthemeCement_WorkshopR.sty, CreateMaps.R, Class2.tex, and mapWorld.
- R Script:** The main pane contains R code for mapping:

```
## MAP THE ADDITIVE SCALE
library(RColorBrewer)
library(mapprojtools)
library(ggplot2)
library(rworldmap)

# Replace Badly coded countries in df
DataRegion[DataRegion == "Russian Federation"] <- "Russia"
DataRegion[DataRegion == "United Kingdom"] <- "UK"
DataRegion[DataRegion == "Congo"] <- "Republic of Congo"
DataRegion[DataRegion == "DR Congo"] <- "Democratic Republic of the Congo"

# 
mapWorld = map_data(map='world')
mapWorld = merge(Data, mapWorld, by='region', all.y=TRUE)
mapWorld = mapWorld[order(mapWorld$order), ] # <---

# Fix missing code in World
mapWorld$region[mapWorld$region == "Denmark"] <- "Greenland"

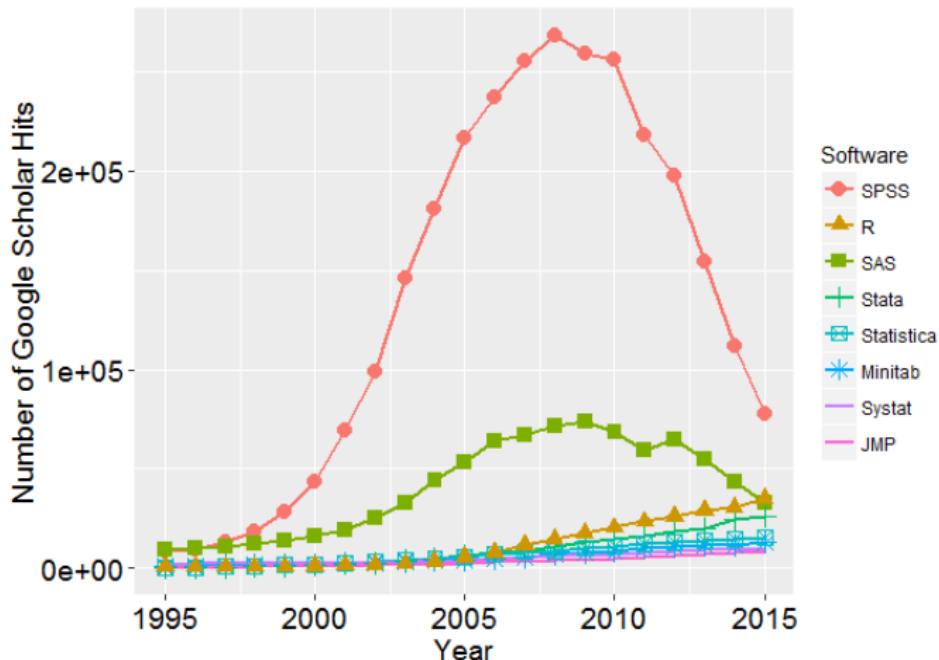

```
- Console:** Shows the command `./CreateMaps.R` being run, with the output:

```
>
>
>
>
>
>
>
> 1+1
[1] 2
> |
```

Le monde de l'Open Source

Why R ?

Why R ?



Pourquoi R ? Les raisons de l'aimer

1. Gratisssss
2. Disponible pour tous les systèmes d'exploitation
3. Graphiques + L^AT_EX
4. Popularité + Packages
5. *Open source*: Développer par et pour les chercheurs

Pourquoi R ? Les raisons de détester

1. Programmer du code = Courbe d'apprentissage raide
2. Développement éclectique. Par moment... chaotique

Why L^AT_EX?

Why L^AT_EX? Reasons to Love

- ▶ Bibliography: BIBTEX
- ▶ Table of content, tables, etc.
- ▶ Deals automatically with stuff like tables, figures, etc.
- ▶ Pretty templates
- ▶ Code + *Open source* = A large expert community on the web

Why L^AT_EX? Reasons to Hate

- ▶ Tough to learn... I mean very tough. But the basics are easy
- ▶ Incompatible with MS Word
- ▶ ~~No spell checker~~
- ▶ No “Change trackers” and stuff like that
- ▶ Final document available only after code compilation
- ▶ Some journals do not accept L^AT_EX submission... others strongly encourage it

L^AT_EX: A Nice Table

Table 1. Length of Bananas and Apples

Quantile	Bananas	Apples
0%	59	44
50%	69	64
100%	77	71

L^AT_EX: The Nice Table Code

```
\begin{table}
  \centering
  \caption{Length of Bananas and Apples}
  \begin{tabular}{lrr}
    Quantile & Bananas & Apples \\ \hline
    0\%      & 59      & 44 \\
    50\%     & 69      & 64 \\
    100\%    & 77      & 71 \\
  \end{tabular}
  \label{tab:bananasapples}
\end{table}
```

L^AT_EX: It's a joke, right?!



LATEX: Nope



LATEX

Tableau 1: Tests des hypothèses

	Vote pour le NPD						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Évaluation du chef NPD					3.87*** (0.22)	3.81*** (0.24)	3.17*** (0.52)
Droite idéologique			-2.86*** (0.46)	-3.24*** (0.53)			-2.66*** (0.57)
Québec	0.69*** (0.09)	0.61*** (0.16)		0.92** (0.34)		0.56** (0.17)	0.93** (0.35)
Femme		0.05 (0.09)		-0.08 (0.19)		-0.03 (0.10)	-0.08 (0.20)
Francophone		-0.02 (0.17)		-0.37 (0.35)		-0.29 (0.18)	-0.63 (0.37)
allophone		-0.17 (0.15)		-0.38 (0.34)		-0.18 (0.17)	-0.22 (0.36)
Moins de 34 ans		-0.03 (0.15)		-0.17 (0.34)		-0.13 (0.16)	-0.26 (0.36)
Plus de 55 ans		-0.23* (0.10)		-0.33 (0.21)		-0.24* (0.11)	-0.23 (0.22)
Haut revenu		-0.33** (0.12)		-0.36 (0.24)		-0.30* (0.13)	-0.32 (0.25)
Faible revenu		0.30* (0.15)		0.33 (0.31)		0.40* (0.17)	0.49 (0.33)
Pas de diplôme secondaire		-0.23 (0.15)		0.04 (0.36)		-0.12 (0.17)	0.03 (0.38)
Diplôme universitaire		0.13 (0.10)		-0.61** (0.21)		-0.12 (0.11)	-0.79*** (0.22)
-constante	-1.05*** (0.05)	-0.86*** (0.11)	0.34 (0.20)	0.96** (0.35)	-3.17*** (0.15)	-2.95*** (0.19)	-1.21* (0.51)
N	2,745	2,464	655	610	2,636	2,381	602
Log Likelihood	-1,650.11	-1,487.30	-383.02	-346.16	-1,412.88	-1,276.31	-317.77
AIC	3,304.22	2,996.60	770.04	716.31	2,829.77	2,576.62	661.54

Source : Étude électorale canadienne, 2011.

Note : Régression logistique binomiale.

*p<0.05 ; **p<0.01 ; ***p<0.001

L^AT_EX: Code (Part 1)

<...>

```
1 \documentclass{article}
2 \usepackage[utf8]{inputenc}
3 \usepackage[T1]{fontenc}
4 \usepackage{cjk}
5 \usepackage{amsmath}
6 \usepackage{amssymb}
7 \usepackage{amsthm}
8 \usepackage{mathbbm}
9 \usepackage{bm}
10 \usepackage{color}
11 \usepackage{listings}
12 \usepackage{fancyvrb}
13 \usepackage{framed}
14 \usepackage{tikz}
15 \usepackage{tikz-qtree}
16 \usepackage{tikz-qtree-compat}
17 \usepackage{tikz-qtree}
18 \usepackage{tikz-qtree}
19 \usepackage{tikz-qtree}
20 \usepackage{tikz-qtree}
21 \usepackage{tikz-qtree}
22 \usepackage{tikz-qtree}
23 \usepackage{tikz-qtree}
```

L^AT_EX: Code (Part 2)

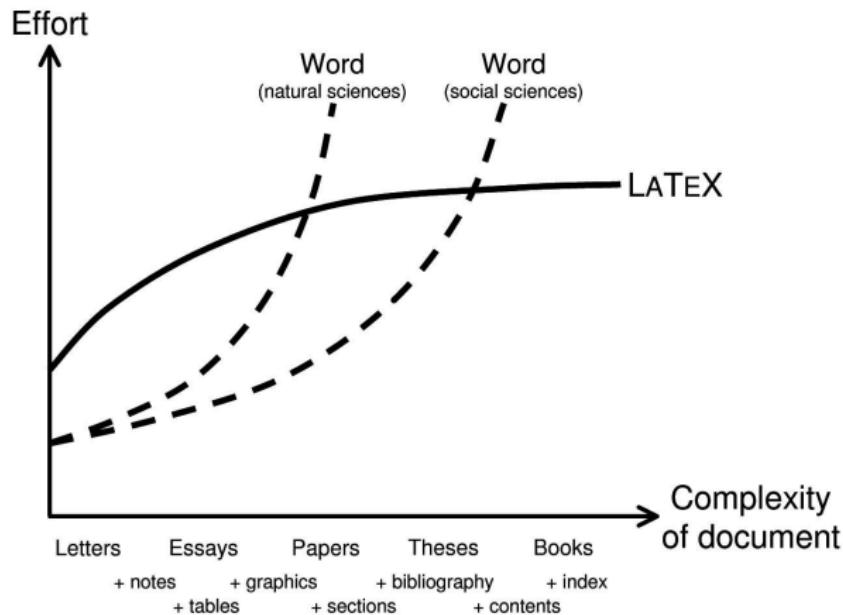
<...>

```
1 \documentclass{article}
2 \usepackage[utf8]{inputenc}
3 \usepackage[T1]{fontenc}
4 \usepackage{amsmath}
5 \usepackage{amsfonts}
6 \usepackage{amssymb}
7 \usepackage{pstricks}
8 \usepackage{color}
9 \usepackage{listings}
10 \usepackage{float}
11 \usepackage{fancyvrb}
12 \usepackage{array}
13 \usepackage{longtable}
14 \usepackage{booktabs}
15 \usepackage{url}
16 \usepackage{hyperref}
17 \usepackage{mathbbm}
18 \usepackage{bm}
19 \usepackage{mathtt}
20 \usepackage{listings}
21 \usepackage{mathtools}
22 \usepackage{color}
23 \usepackage{listings}
```

LATEX



LATEX



R + LATEX

<...>

```
1 \begin{document} \begin{center} \textbf{R} \textbf{+} \textbf{LATEX} \end{center} \end{document}
```

R + LATEX

Tableau 1: Tests des hypothèses

	Vote pour le NPD						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Évaluation du chef NPD					3.87*** (0.22)	3.81*** (0.24)	3.17*** (0.52)
Droite idéologique			-2.86*** (0.46)	-3.24*** (0.53)			-2.66*** (0.57)
Québec	0.69*** (0.09)	0.61*** (0.16)		0.92** (0.34)		0.56** (0.17)	0.93** (0.35)
Femme		0.05 (0.09)		-0.08 (0.19)		-0.03 (0.10)	-0.08 (0.20)
Francophone		-0.02 (0.17)		-0.37 (0.35)		-0.29 (0.18)	-0.63 (0.37)
allophone		-0.17 (0.15)		-0.38 (0.34)		-0.18 (0.17)	-0.22 (0.36)
Moins de 34 ans		-0.03 (0.15)		-0.17 (0.34)		-0.13 (0.16)	-0.26 (0.36)
Plus de 55 ans		-0.23* (0.10)		-0.33 (0.21)		-0.24* (0.11)	-0.23 (0.22)
Haut revenu		-0.33** (0.12)		-0.36 (0.24)		-0.30* (0.13)	-0.32 (0.25)
Faible revenu		0.30* (0.15)		0.33 (0.31)		0.40* (0.17)	0.49 (0.33)
Pas de diplôme secondaire		-0.23 (0.15)		0.04 (0.36)		-0.12 (0.17)	0.03 (0.38)
Diplôme universitaire		0.13 (0.10)		-0.61** (0.21)		-0.12 (0.11)	-0.79*** (0.22)
-constante	-1.05*** (0.05)	-0.86*** (0.11)	0.34 (0.20)	0.96** (0.35)	-3.17*** (0.15)	-2.95*** (0.19)	-1.21* (0.51)
N	2,745	2,464	655	610	2,636	2,381	602
Log Likelihood	-1,650.11	-1,487.30	-383.02	-346.16	-1,412.88	-1,276.31	-317.77
AIC	3,304.22	2,996.60	770.04	716.31	2,829.77	2,576.62	661.54

Source : Étude électorale canadienne, 2011.

Note : Régression logistique binomiale.

*p<0.05 ; **p<0.01 ; ***p<0.001

R Programming Bases

R = Programming Language

- ▶ Calculation operators
- ▶ Assignment operators
- ▶ Logical operators
- ▶ Control instructions

Calculation Operators

- +
- -
- /
- %%

Logical Operators

- ==
- !=
- >=
- <=
- <
- >
- &
- |
- %in%

Control Instructions

- ▶ if... else
- ▶ for loop

R Data Structure

Data Structure

- ▶ Constants

Data Structure

- ▶ Constants
- ▶ Vectors

Data Structure

- ▶ Constants
- ▶ Vectors
- ▶ Data frames

Constants

<...>

```
1 //constant  
2 //constant  
3 //constant
```

Vecteurs

<...>

```
1 Векторът еднодимензионална структура данни, която съдържа
2 последователност от елементи и има определени
3 методи за работа с тях.
```

Data Frames

<...>

```
1 | Data frame is a two-dimensional, column-oriented data structure, similar to a table or matrix.
```

Constantes

Vecteurs

Data frames

Constantes

Vecteurs

Data frames

1 

Constantes

1 

Vecteurs



Data frames

Constantes

1 

Vecteurs

1	
2	
3	
4	
5	
6	

Data frames

	v1	v2	v3	v4
1				
2				
3				
4				
5				
6				

aFruit <- “banana”

Constantes



Vecteurs

1	
2	
3	
4	
5	
6	

Data frames

	v1	v2	v3	v4
1				
2				
3				
4				
5				
6				

fruits[1] <- “banana”

Constantes



Vecteurs

1	
2	
3	
4	
5	
6	

Data frames

	v1	v2	v3	v4
1				
2				
3				
4				
5				
6				

Data[1,1] <- “banana”

Constantes



Vecteurs

1	A yellow banana icon with a brown stem, enclosed within a thin black rectangular border.
2	
3	
4	
5	
6	

Data frames

	v1	v2	v3	v4
1	A yellow banana icon with a brown stem, enclosed within a thin black rectangular border.			
2				
3				
4				
5				
6				

Data\$fruits[1] <- “banana”

Constantes

1	
---	-----------------------------------------------------------------------------------

Vecteurs

1	
2	
3	
4	
5	
6	

Data frames

	v1	v2	v3	v4
1				
2				
3				
4				
5				
6				

Functions

R Base Functions

- ▶ `length()`
- ▶ `min()`
- ▶ `max()`
- ▶ `sum()`
- ▶ `median()`
- ▶ `mean()`

R Function: mean()

<...>

```
1 mean(...)
```

Creating a Function in R

R Function: meanGirls()



R Function: meanGirls()

<...>

```
1 function(x) {  
2   mean(x)  
3 }  
4
```

R Function: meanGirlsPlus()

<...>

```
1 #> Function that adds 1 to the mean of a vector
2 #> Returns the mean plus one
3 #> 
4 #> @param x
5 #>   A vector of numbers
6 #> @return The mean plus one
7 #>
8 #> Examples
9 #>
```

Then?

Then? More R Functions...

```
meanGirls()  
meanGirlsPlus()  
meanBoys()  
meanBoysPlus()
```

Then? A \mathbb{R} Package

MeanSexPak

```
meanGirls()  
meanGirlsPlus()  
meanBoys()  
meanBoysPlus()
```

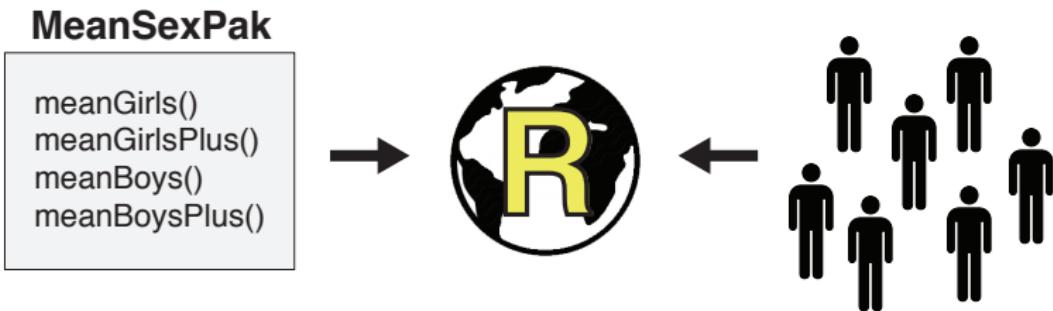
Then? Package Publication

MeanSexPak

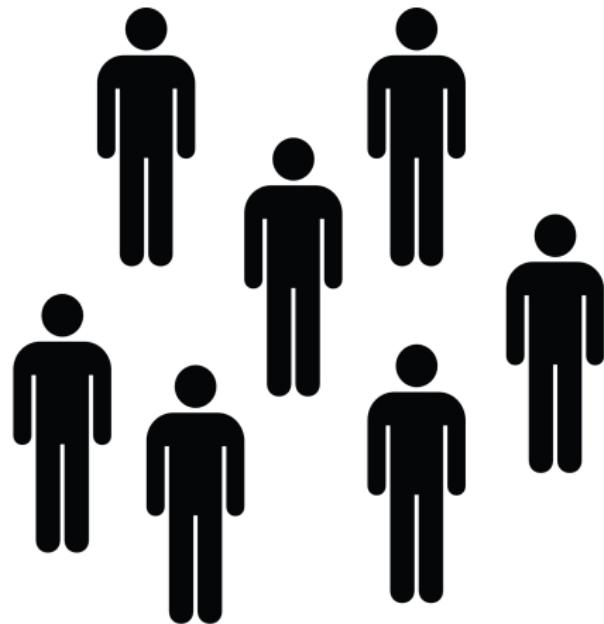
```
meanGirls()  
meanGirlsPlus()  
meanBoys()  
meanBoysPlus()
```



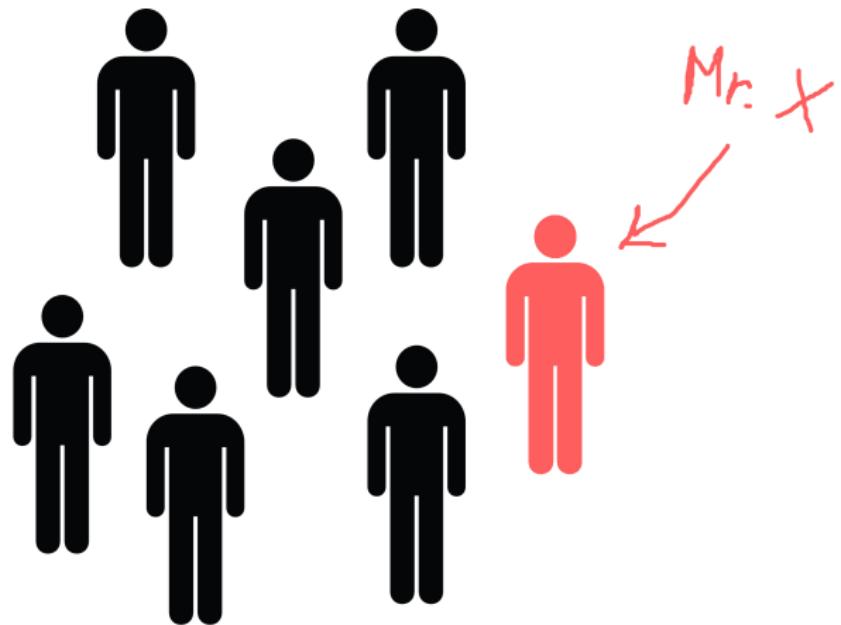
Then? Diffusion to the Community



R Community



Mr. X



Then? Mr. X Installs the Package

<...>

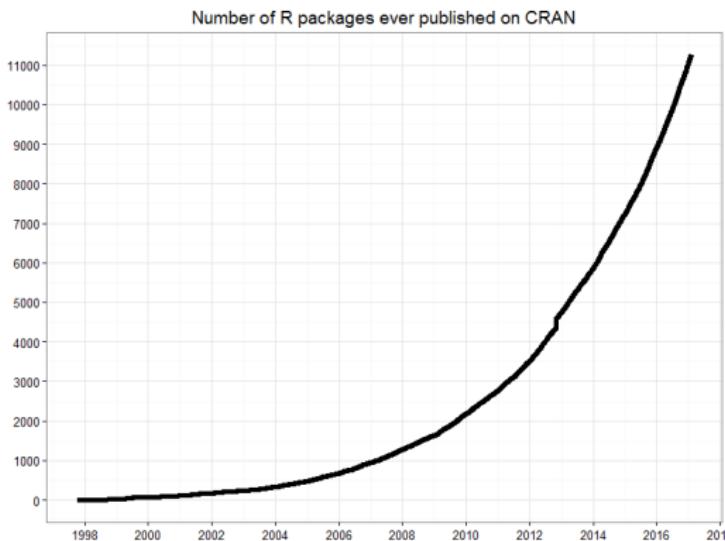
```
1 [REDACTED]
```

Ensuite? Mr. X Uses the Package

<...>

```
1 package ensuite
2
3 # This file contains code that runs when the package
4 # is loaded into R. It is called the "initialization" code.
```

Number of R Packages



**Enough blabla...
Let's Code!**

**Et maintenant,
comment on apprend ?**

Contre vents et marées: apprendre R et savoir naviguer malgré les intempéries

Et maintenant, comment on apprend ?

Plan de la présentation

- ▶ Apprentissage
 - ▶ Présentation de DataCamp
 - ▶ Choix de cours : serpents et échelles
 - ▶ DataCamp vs. les autres sites
 - ▶ Livres, manuels et autres
 - ▶ Opportunités d'apprentissage
- ▶ Navigage
 - ▶ Stack Overflow
 - ▶ Slack
 - ▶ Comment régler ses problèmes en programmation ?

Présentation de DataCamp



DataCamp, c'est un site internet où apprendre R, Python, Git, SQL, etc. Mais, c'est beaucoup plus que cela !

- ▶ 217 instructeurs provenant de plusieurs disciplines
- ▶ 254 cours offerts, allant de débutants à avancés
- ▶ Exercices théoriques et pratique, accompagnés de vidéos
- ▶ Une équipe dynamique qui développe constamment du nouveau contenu

Présentation de DataCamp

Et maintenant, comment on apprend ?