

R Markdown

Alison PATOU
Patou.alison@gmail.com



Programme

- Introduction
- R Markdown
- Projets en groupe



A savoir

- Tout est sur mon GitHub :
<https://github.com/apatou>
- (L'essentiel à retenir du cours, les dataset, exercices, ...)
- Merci de m'envoyer à chaque fin de séance vos TPs : patou.alison@gmail.com
- 1 projet à la dernière séance

1

Introduction

Comment optimiser et
présenter son code ?



Notebook

Il existe sous Python, différents notebooks afin de centraliser code et résultats : Jupyter est souvent cité en référence.

Quid de R ?

Notebook

Qu'est ce que c'est ?

Un notebook contient le code du développeur mais aussi les différentes étapes d'analyses, les visualisations, les commentaires et des découpages grâce à des titres et sous-titres pour une meilleure lisibilité.

Cela permet notamment une meilleure :

- Efficacité
- Interactivité
- Collaborativité
- Reproductibilité
- Automatisation
- ...

Notebook

Quand l'utiliser ?

Le notebook s'utilise durant les différentes étapes d'analyse :

- **Nettoyage des données** : faire le tri entre les données importantes et celles qui ne le sont pas dans l'analyse des ensembles de mégadonnées
- **Modélisation statistique** : méthode mathématique permettant d'établir la probabilité de répartition d'une caractéristique particulière
- **Création et mise en œuvre de modèles d'apprentissage automatique** : étude, programmation et apprentissage de modèles
- **Visualisation de données** : représentation graphique de données pour faire apparaître des structures, des tendances, des relations, etc.

R Markdown

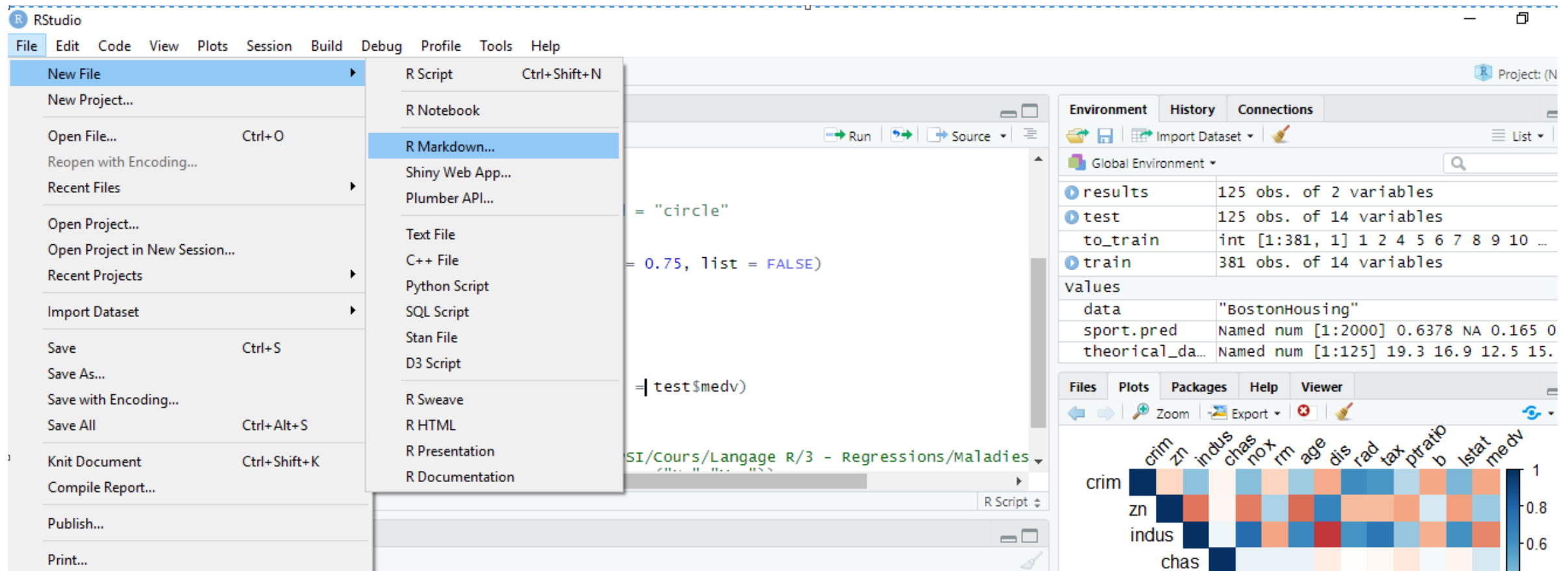
R Markdown

- R Markdown est un package qui, combiné à R Studio, permet la création rapidement d'un notebook de travail et ainsi créer une page Web.
- Le package Markdown fonctionne avec le package ***knitr***.

Le package ***knitr*** permet de créer des tableaux mais aussi des graphiques et ainsi permettre une reproductibilité des travaux de manière efficace et simple.

R Markdown

Ouvrir R Markdown :



The screenshot shows the RStudio interface. The 'File' menu is open, and 'R Markdown...' is selected. The background shows a code editor with R code and a console window.

File Menu Options:

- New File
 - R Script (Ctrl+Shift+N)
 - R Notebook
 - R Markdown...**
 - Shiny Web App...
 - Plumber API...
- New Project...
- Open File... (Ctrl+O)
- Reopen with Encoding...
- Recent Files
- Open Project...
- Open Project in New Session...
- Recent Projects
- Import Dataset
- Save (Ctrl+S)
- Save As...
- Save with Encoding...
- Save All (Ctrl+Alt+S)
- Knit Document (Ctrl+Shift+K)
- Compile Report...
- Publish...
- Print...

Code Editor Content:

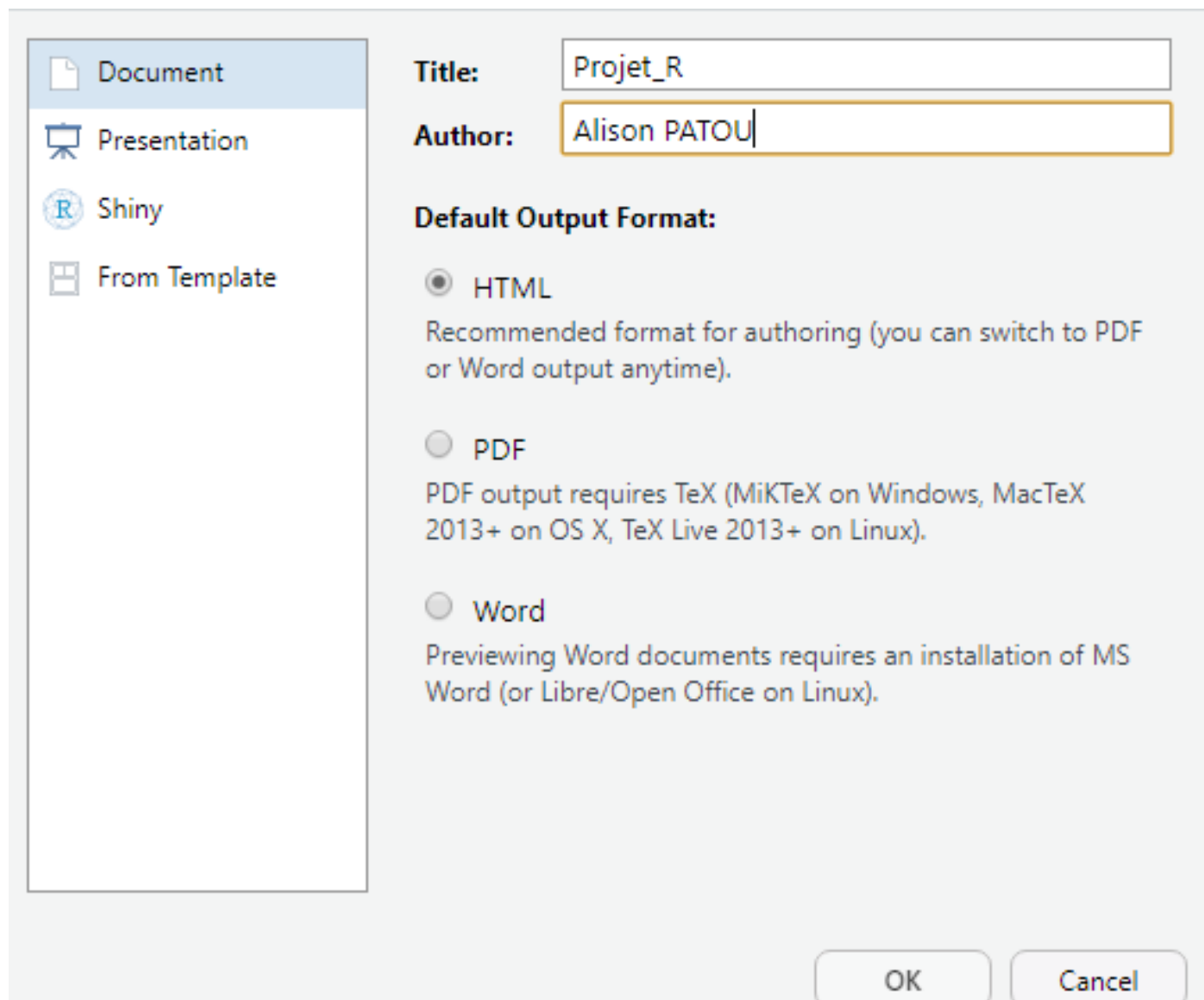
```
= "circle"  
  
= 0.75, list = FALSE)  
  
= test$medv)  
  
SI/Cours/Langage R/3 - Regressions/Maladies
```

Environment Panel:

Object	Details
results	125 obs. of 2 variables
test	125 obs. of 14 variables
to_train	int [1:381, 1] 1 2 4 5 6 7 8 9 10 ...
train	381 obs. of 14 variables

Viewer Panel:

Heatmap showing the relationship between variables: crim, zn, indus, chas, nox, rm, age, dis, rad, tax, ptratio, b, lstat, medv. The color scale ranges from 0.6 (dark blue) to 1.0 (dark red).



Document

Presentation

Shiny

From Template

Title:

Author:

Default Output Format:

☒ **HTML**
Recommended format for authoring (you can switch to PDF or Word output anytime).

☐ **PDF**
PDF output requires TeX (MiKTeX on Windows, MacTeX 2013+ on OS X, TeX Live 2013+ on Linux).

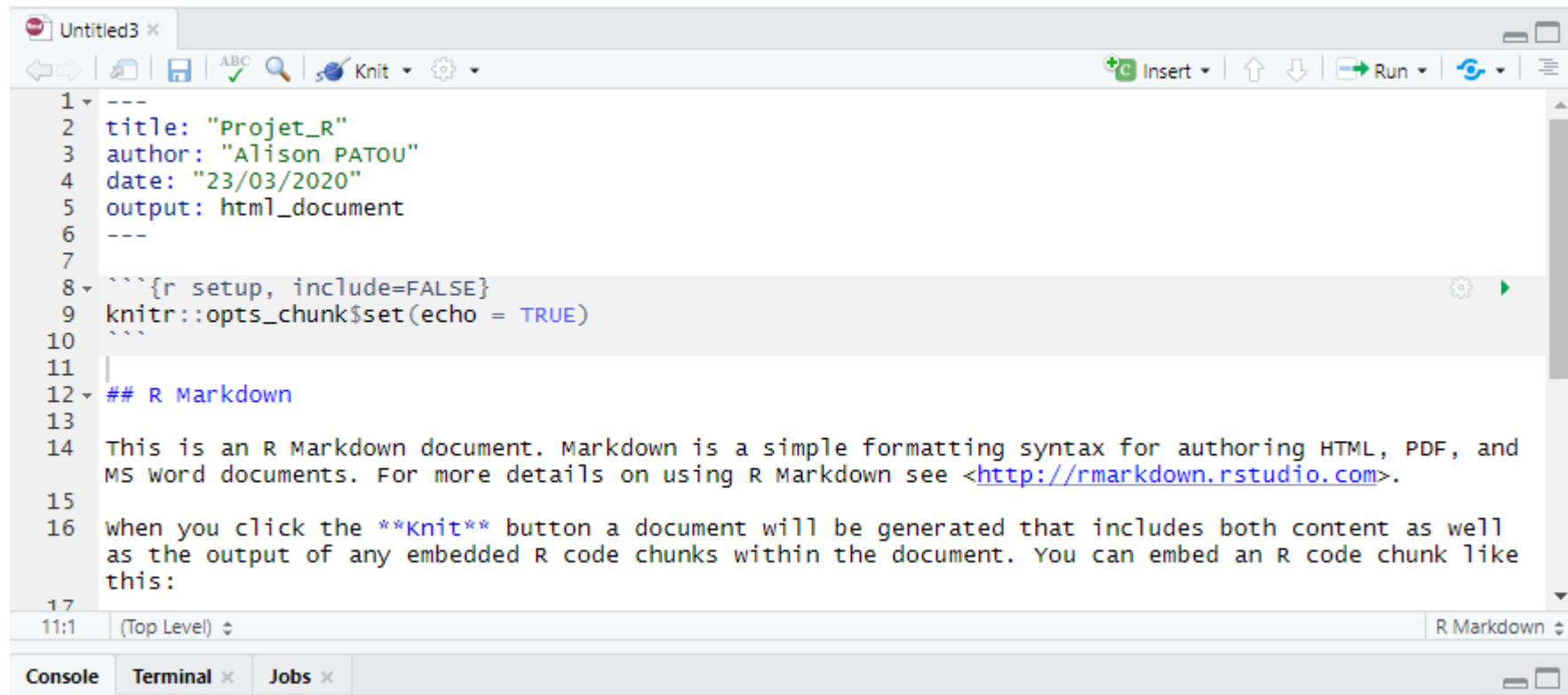
☐ **Word**
Previewing Word documents requires an installation of MS Word (or Libre/Open Office on Linux).

OK Cancel

R Markdown

- **3 fichiers en « sortie » possibles :**
- HTML : le plus commun, sera lu via un navigateur
- PDF : pour des rapports mais requiere tex
- Word

Structure R Markdown



The image shows a screenshot of an R Markdown document in a text editor. The document is titled 'Untitled3'. The code is as follows:

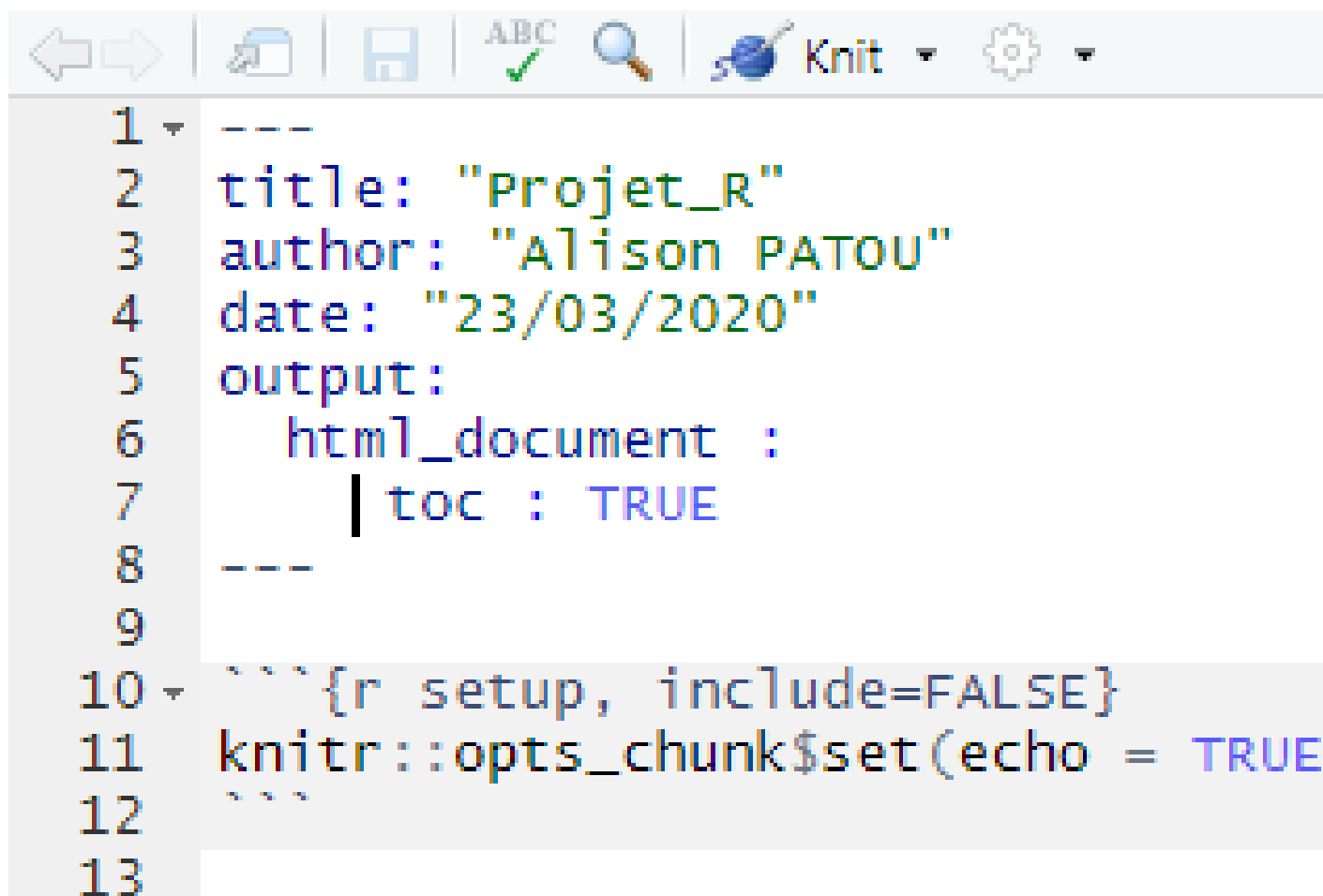
```
1 ---
2 title: "Projet_R"
3 author: "Alison PATOU"
4 date: "23/03/2020"
5 output: html_document
6 ---
7
8 ```{r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = TRUE)
10 ```
11
12 ## R Markdown
13
14 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and
15 MS 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
18 as the output of any embedded R code chunks within the document. You can embed an R code chunk like
19 this:
```

The editor interface includes a toolbar with icons for undo, redo, save, search, and knit. The status bar at the bottom shows '11:1 (Top Level)' and 'R Markdown'.

Header (présentation) du document

Corps du document

Header

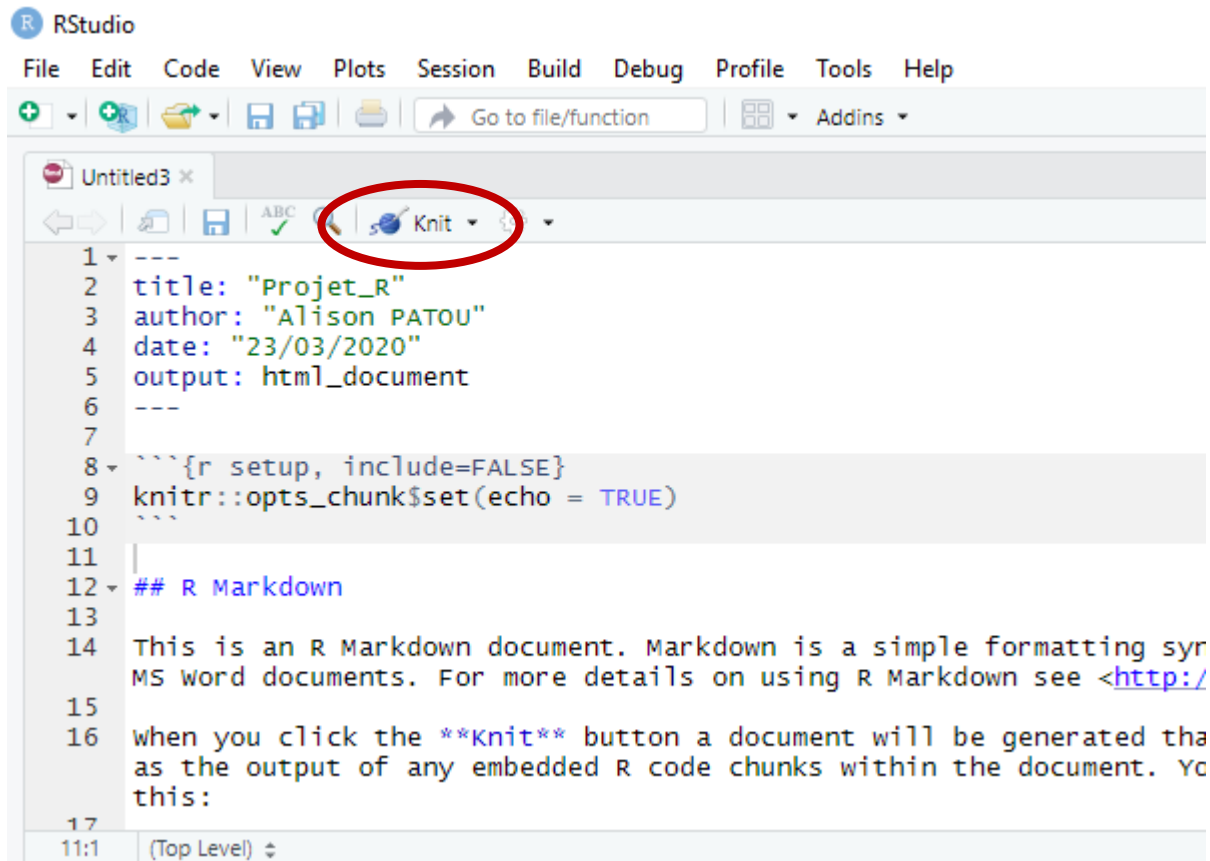


The image shows a screenshot of the RStudio editor interface. At the top is a toolbar with icons for navigation (back, forward), file operations (open, save), a text editor (ABC with a checkmark), a search icon, and the Knitr logo. Below the toolbar, a code editor displays a YAML header for a Knitr document. The code is as follows:

```
1 ---
2 title: "Projet_R"
3 author: "Alison PATOU"
4 date: "23/03/2020"
5 output:
6   html_document :
7     | toc : TRUE
8 ---
9
10 ```{r setup, include=FALSE}
11 knitr::opts_chunk$set(echo = TRUE)
12 ```
13
```

- Plusieurs customisations sont possible concernant le Header (numérotation des titres, insertion logo, ...)
- Toc : TRUE
 - > permet l'affichage de votre structure de document (titre, sous-titre, ...)

Exécution

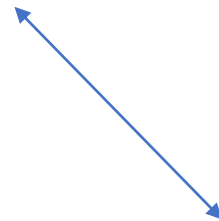


Le bouton knit va tricoter (exécuter)
votre code pour créer votre notebook

Exécution

Vos deux fichiers ont été créés (le .Rmd et le .html)

Markdown_cours	23/03/2020 11:15	Document HTML	629 Ko
Markdown_cours.Rmd	23/03/2020 11:15	Fichier RMD	1 Ko



RStudio

File Edit Code View Plots Console Environment

~/Personnel/EPISI/Cours/Langage R/4 - Rmarkdown/Markdown_cours.html

Markdown_cours.html Open in Browser Find

Projet_R

Alison PATOU
23/03/2020

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:


```
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
```

Including Plots

You can also embed plots, for example:

Mise en forme

Syntax	Becomes
Plain text	Plain text
End a line with two spaces to start a new paragraph.	End a line with two spaces to start a new paragraph.
<code>*italics*</code> and <code>_italics_</code>	<i>italics and italics</i>
<code>**bold**</code> and <code>__bold__</code>	bold and bold
<code>superscript^2^</code>	^{superscript} 2
<code>--strikethrough--</code>	strikethrough
<code>[link](www.rstudio.com)</code>	link
<code># Header 1</code>	<h2>Header 1</h2>
<code>## Header 2</code>	<h3>Header 2</h3>
<code>### Header 3</code>	<h4>Header 3</h4>
<code>#### Header 4</code>	<h5>Header 4</h5>
<code>##### Header 5</code>	<h6>Header 5</h6>
<code>##### Header 6</code>	<h7>Header 6</h7>
<code>endash: --</code>	endash: —
<code>emdash: ---</code>	emdash: —
<code>ellipsis: ...</code>	ellipsis: ...
<code>inline equation: $\\$A = \pi * r^{2}\\$</code>	inline equation: $A = \pi * r^2$
<code>image: </code>	image: 

NB : La Cheat Sheet R Markdown est sur mon github (<https://github.com/apatou>)

Code

L'avantage de Rmarkdown c'est que vous pouvez continuer d'écrire votre code comme dans un script R classique

Votre Code Chunk (corps de texte) va se trouver sous cette zone :

```
7
8 {r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = TRUE)
10
11
```

Chunk options		
option	default value	description
Code evaluation		
child	NULL	A character vector of filenames. Knitr will knit the files and place them into the main document.
code	NULL	Set to R code. Knitr will replace the code in the chunk with the code in the code option.
engine	'R'	Knitr will evaluate the chunk in the named language, e.g. <code>engine = 'python'</code> . Run <code>names(knitr::knit_engines\$get())</code> to see supported languages.
eval	TRUE	If FALSE, knitr will not run the code in the code chunk.
include	TRUE	If FALSE, knitr will run the chunk but not include the chunk in the final document.
purrr	TRUE	If FALSE, knitr will not include the chunk when running <code>purrr()</code> to extract the source code.
Results		
collapse	FALSE	If TRUE, knitr will collapse all the source and output blocks created by the chunk into a single block.
echo	TRUE	If FALSE, knitr will not display the code in the code chunk above it's results in the final document.
results	'markup'	If 'hide', knitr will not display the code's results in the final document. If 'hold', knitr will delay displaying all output pieces until the end of the chunk. If 'asis', knitr will pass through results without reformatting them (useful if results return raw HTML, etc.)
error	TRUE	If FALSE, knitr will not display any error messages generated by the code.
message	TRUE	If FALSE, knitr will not display any messages generated by the code.
warning	TRUE	If FALSE, knitr will not display any warning messages generated by the code.
Code Decoration		
comment	"##"	A character string. Knitr will append the string to the start of each line of results in the final document.
highlight	TRUE	If TRUE, knitr will highlight the source code in the final output.
prompt	FALSE	If TRUE, knitr will add > to the start of each line of code displayed in the final document.
strip.white	TRUE	If TRUE, knitr will remove white spaces that appear at the beginning or end of a code chunk.
tidy	FALSE	If TRUE, knitr will tidy code chunks for display with the <code>tidy_source()</code> function in the <code>formatR</code> package.

Code

The screenshot shows the RStudio interface with an R Markdown document open. The document is titled 'Projet_R' and is authored by 'Alison PATOU' on '23/03/2020'. The document content includes a title, author, date, and output format, followed by a setup block, a title block, an import block, a dataset import block, and a visualization block. The visualization block contains a plot of 'datasetfleur\$Valeurs'.

```
1 ---
2 title: "Projet_R"
3 author: "Alison PATOU"
4 date: "23/03/2020"
5 output: html_document
6 ---
7
8 ```{r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = TRUE)
10 ```
11
12 ## R Markdown
13
14 ### Import de la librairie
15 ```{r}
16 #Charger la librairie data.table
17 library(data.table)
18 ```
19
20 ### Import dataset
21 ```{r}
22 datasetfleur = fread("C:/Users/patou/Documents/Personnel/EPST/Cours/Langage R/2 - 3
23 ce/datasetfleur.csv")
24 ```
25
26 ### Visualisation
27 ```{r}
28 plot(datasetfleur$Valeurs)
29 ```
30 Globalement les valeurs semblent assez hom
```

The right pane shows the rendered HTML output of the document, which includes the title, author, date, and the first part of the content. The plot is not yet rendered.

- ## Titre
- ### Sous titre
- Le code R qui va être exécuté doit se trouver entre :
 - ``{r}
 - VOTRE CODE
 - ```

Pour aller plus loin

- Possibilité de customiser le rendu HTML avec du CSS (markdown.css)
- Ne pas hésiter à mettre des visualisations/graphiques dynamiques. Plusieurs packages en proposent :
 - Ggplot/ggplot2
 - Plotly
 - rAmChart



Merci