Utiliser R Markdown pour créer des documents dynamiques

par Kevin Cazelles

01 September 2020

Le package rmarkdown permet de créer des documents dynamiques qui intègrent des morceaux de code R et ce qu’ils génèrent (dont figures et tableaux). Ce document est une introduction à R Markdown qui passe en revue différents aspects techniques nécessaires à la bonne compréhension et l’utilisation du package rmarkdown. Le code source de ce document est en lui-même un exemple d’application du package. Le lecteur trouvera également un ensemble de références disponibles en ligne pour aller plus loin.

# Avant-propos

Ce document est une introduction à R Markdown conçu pour répondre aux objectifs suivants:

1. comprendre ce qu’est R Markdown;
2. comprendre les liens entre R, Markdown et Pandoc;
3. apprendre la syntaxe Pandoc Markdown;
4. apprendre à créer des documents dynamiques avec le package rmarkdown.

Ce document a lui-même été édité avec R Markdown et peut-être converti en une page internet (au format HTML, .html), au format Word (.docx) ou encore en document portable (.pdf). Bien que dans “R Markdown”, il y ait “R” et “Markdown”, ce document n’est pas une introduction à R (<http://www.r-project.org>). Le contenu est accessible aux utilisateurs débutants et pour les lecteurs intéressés qui ne connaissent pas encore R, une courte introduction à R est nécessaire. Ils pourront consulter l’une des différentes introductions disponibles sur le site du CRAN: <https://cran.r-project.org/other-docs.html>.

Ce document est déployé automatiquement par GitHub Actions (voir <https://github.com/KevCaz/Rmarkdowndocfr/actions>), lors de la dernière compilation, la version 2.3 du package rmarkdown a été utilisée. De plus amples informations relatives à l’environnement dans lequel le document a été compilé sont indiquées ci-dessous.

sessionInfo()

## R version 4.0.2 (2020-06-22)  
## Platform: x86\_64-pc-linux-gnu (64-bit)  
## Running under: Ubuntu 20.04 LTS  
##   
## Matrix products: default  
## BLAS/LAPACK: /usr/lib/x86\_64-linux-gnu/openblas-openmp/libopenblasp-r0.3.8.so  
##   
## locale:  
## [1] LC\_CTYPE=en\_US.UTF-8 LC\_NUMERIC=C   
## [3] LC\_TIME=en\_US.UTF-8 LC\_COLLATE=en\_US.UTF-8   
## [5] LC\_MONETARY=en\_US.UTF-8 LC\_MESSAGES=C   
## [7] LC\_PAPER=en\_US.UTF-8 LC\_NAME=C   
## [9] LC\_ADDRESS=C LC\_TELEPHONE=C   
## [11] LC\_MEASUREMENT=en\_US.UTF-8 LC\_IDENTIFICATION=C   
##   
## attached base packages:  
## [1] stats graphics grDevices utils datasets methods base   
##   
## loaded via a namespace (and not attached):  
## [1] compiler\_4.0.2 magrittr\_1.5 tools\_4.0.2 htmltools\_0.5.0  
## [5] yaml\_2.2.1 stringi\_1.4.6 rmarkdown\_2.3 knitr\_1.29   
## [9] stringr\_1.4.0 xfun\_0.16 digest\_0.6.25 rlang\_0.4.7   
## [13] evaluate\_0.14

rmarkdown::pandoc\_version()

## [1] '2.7.3'

# Introduction

Dans de nombreux milieux professionnels, à des fins de communication diverses, sont produits régulièrement des documents intégrant des analyses de données (tableaux, figures, tests statistiques, etc.). Pour créer de tels documents, il faut être en mesure de manipuler des données les analyser et créer des figures pour les intégrer dans le document final. [R](http://www.r-project.org) est un langage de programmation qui répond à ces besoins avec un grand nombre de package qui permettent de manipuler et traiter un spectre très large de données et de les visualiser efficacement. Ce langage offre également la possibilité d’intégrer code et les produits du code (résultats de tests, tableaux, figures, etc.) directement dans un document qui est alors qualifié de dynamique. Le package le plus utilisé pour créer des documents dynamiques est la package rmarkdown l’intégration de R dans un document écrit avec Markdown et qui peut être converti en de nombreux formats de document (dont PDF, word, HTML). C’est ce package qui est l’objet de ce document, bien qu’il nous faudra en mentionner d’autre ainsi que d’autres langages et outils pour comprendre ce package, en commençant par Markdown.

Écrire un document en Markdown, quésaco? Markdown est un langage de balisage léger, c’est-à-dire un langage dans lequel on peut utiliser des ensembles de caractères spécifiques (des balises) pour délimiter une zone de texte pour laquelle un formatage associé (e.g. **text en gras**) est appliqué. Markdown est aujourd’hui très répandu sur Internet, à tel point que vous pourriez une des différentes syntaxes existantes sans me le savoir. Sans être exhaustif, il est utile de donner un peu plus de détails sur ce point pour bien comprendre ce quel Markdown on utilise avec R Markdown. La syntaxe originale de Markdown est le fruit du travail de John Gruber[[1]](#footnote-25), programmeur, bloggeur et baladodiffuseur de Philadelphie en collaboration avec Aaron Swartz (lui même connu pour avoir participer a la création de Creative Commons et son tragique destin qui fut l’objet d’un film). Sur le site de John Gruber, “daringfireball”, Markdown est décrit depuis décembre 2004 et on peut même y télécharger la version 1.0.1 (voir <http://daringfireball.net/projects/markdown>). L’idée de départ est simple et élégante : produire un langage léger qui simplifie les balises HTML (utilisé par tout les sites Internet). L’idée n’est pas tant de remplacer le HTML mais plutôt d’en augmenter l’efficacité d’écriture et de fait, il est beaucoup plus rapide d’écrire en Markdown qui couvre les opérations de formatage les plus courantes (listes, hyperliens, etc.). Notons qu’il existe d’autres langages qui répondent aux me objectif, par example ReStructuredText[[2]](#footnote-28).

Après la publication de Markdown, John Gruber a cessé de travailler sur Markdown[[3]](#footnote-30) et d’autres développeurs, sans doute séduits par le langage, ont proposé différentes additions syntaxiques. Il s’agissait surtout de lever certaines limitations tout en préservant l’esprit d’origine. Ci-dessous, en voici une liste non exhaustive de différentes variantes Markdown:

* [GitHub Flavored Markdown (GFM)](https://help.github.com/articles/github-flavored-markdown/)
* [Kramdown](http://kramdown.gettalong.org/syntax.HTML#math-blocks)
* [Markdown Extra](https://michelf.ca/projects/php-markdown/extra/)
* [Multi Markdow](https://github.com/fletcher/MultiMarkdown/wiki/MultiMarkdown-Syntax-Guide#math-support)
* [Pandoc Markdown](http://pandoc.org/README.HTML#pandocs-markdown)

Fort heureusement, depuis 2014, CommonMark (<https://commonmark.org/>) propose une spécification (norme technique) pour Markdown de plus en plus utilisée[[4]](#footnote-38), ce qui signifie qu’en allant d’un outil à l’autre qui utilise cette spécification, il n’y a pas de questions à se poser quant à savoir ce qui marche ou non en terme de syntaxe (un problème parfois frustrant quand on utilise plusieurs outils qui utilisent différentes syntaxes Markdown), il suffit de se reporter à la spécification!

Il est important de souligner que R Markdown (voir <http://rmarkdown.rstudio.com>)[[5]](#footnote-41) utilise [la variante Markdown de Pandoc](https://rmarkdown.rstudio.com/lesson-8.html)[[6]](#footnote-43). Une précision relative à Pandoc (<http://www.pandoc.org>) s’impose: il s’agit, comme l’indique le site internet, d’un “convertisseur de document universel”, en une ligne de commande Pandoc convertit un document dans un format donné en un document d’un autre format. Par exemple, Pandoc permet de passer d’un fichier .tex (LaTeX) à un fichier .docx (Word)! La variante Markdown de Pandoc a été pensé pour rester fidèle à l’esprit originel de Markdown tout en incluant davantage d’éléments de document communs différent format de documents[[7]](#footnote-45). Pour rendre le document dynamique, R Markdown ajoute notamment les fonctionnalités de l’excellent package [knitr](http://cran.r-project.org/web/packages/knitr/index.html) qui intègre du code R et ce qu’il produit. Le document dynamique en question pourra être converti dans un grand nombre de format grâce à l’utilisation de [Pandoc](http://www.pandoc.org/installing.html). Les intérêts de R Markdown sont nombreux :

1. utiliser une syntaxe épurée pour distinguer la mise en page, le texte, le formatage du texte et le code R,
2. utiliser les fonctionnalités du package knitr pour créer morceaux de codes, tableaux et figures dynamiquement,
3. utiliser le format de sérialisation YAML pour personnaliser la mise en page des documents produits grâce à la fonction render(), elle-même appelée en quelques clics avec [RStudio](http://www.rstudio.com),
4. obtenir ainsi un document avec un contenu R dynamique en différents formats dont HTML (fichier .html), PDF (.pdf) et Word (.docx).

Il existe une documentation abondante relative à R Markdown, malheureusement presque exclusivement écrite en anglais. Pour les lecteurs capables de lire la la langue de Shakespeare, je recommande [le site officiel de R Markdown](http://rmarkdown.rstudio.com) sur lequel vous trouverez, entre autres:

1. un condensé d’utilisation sous forme de *Cheat Sheet*, voir <http://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf>;
2. un guide de référence disponible sur le site de RStudio <http://www.rstudio.com/wp-content/uploads/2015/03/rmarkdown-reference.pdf>.

Il y a également sur le site <https://bookdown.org> :

1. un livre en anglais très complet sur le sujet, par l’un des grands architectes de R Markdown, voir <https://bookdown.org/yihui/rmarkdown/>;
2. [un livre d’astuces (*cook book*)](https://bookdown.org/yihui/rmarkdown-cookbook/) en cours d’écriture.

Une section complète du site de l’auteur principal de knitr, <http://yihui.name/knitr/>, permet de bien comprendre l’intégration des morceaux de code et des résultats de ce code, contenu qui fait l’objet d’un livre de ce même auteur (Xie, 2017). Je tiens également à signaler la [présentation de Mansun Kuo sur Rpubs](http://rpubs.com/mansun_kuo/24330). Pour apprendre la syntaxe Pandoc Markdown, vous pouvez vous reportez à la très complète page internet écrite en français par Jean-Daniel Bonjour, <http://enacit1.epfl.ch/markdown-pandoc>.

La syntaxe [Pandoc Markdown](http://rmarkdown.rstudio.com/authoring_pandoc_markdown.html) est, à mon sens, très complète, mais elle ne permet pas une personnalisation totale de la mise en page de votre document. Il existe des moyens simples pour aller plus loin avec R Markdown, détaillé dans le livre d’astuce mentionné ci-dessus. Si toutefois vous souhaitez aller encore plus loin dans la mise en forme de votre document, vous pouvez utiliser une des trois approches suggérez ci-dessous. Néanmoins, avant d’envisager l’un de ces recours, il est très important de se demander si cette personnalisation est nécessaire.

1. vous pouvez travailler sur le document dans le format qui vous intéresse après l’avoir générer, c’est une solution qui a une limite évident: à chaque fois que le document sera compilé, les modifications post-compilation seront perdues,
2. vous pouvez utiliser un *template* personnel[[8]](#footnote-59), un fichier qui gère partie la mise en forme de votre document) associé à votre document,
3. vous pouvez toujours composé en un langage donné, p. ex. [HTML](http://www.w3.org/html/) ou [Latex](http://www.latex-project.org), mais les commandes non valides pour un format donné, seront ignorées.

Avant de rentrer dans le vif du sujet, je souhaiterais insister sur le fait que cette introduction est axée sur la création de document en format Word, PDF et HTML pour de nombreux usages (rapport, article, etc.), mais que les possibilités offertes par rmarkdown sont plus vastes (voir le site dédié <http://rmarkdown.rstudio.com>), entre autres

* pour les documents :
  + [PDF](https://en.wikipedia.org/wiki/Portable_Document_Format)
  + [HTML](http://rmarkdown.rstudio.com/html_document_format.html)
  + [Word](http://rmarkdown.rstudio.com/word_document_format.html)
  + [Tufte handout (Pdf)](http://rmarkdown.rstudio.com/tufte_handout_format.html)
  + [Package Vignette (Html)](http://rmarkdown.rstudio.com/package_vignette_format.html)
* pour les présentations :
  + [Beamer](http://rmarkdown.rstudio.com/beamer_presentation_format.html)
  + [ioslide](http://rmarkdown.rstudio.com/ioslides_presentation_format.html)
  + [Slidy](http://rmarkdown.rstudio.com/slidy_presentation_format.html)

De plus, il existe des packages qui se basent sur rmarkdown et en étendent les applications. Par exemple, blogdown vous permet d’écrire des postes de blogue avec R Markdown dans un site permet généré par Hugo (<https://gohugo.io/>). De me, grâce à bookdown] (<https://bookdown.org/yihui/bookdown/>), vous pouvez créer de long documents HTML structuré comme un [Gitbook](https://www.gitbook.com). Ainsi, les fonctionnalités décrites dans ce document permettent d’utiliser une très grande variété de médias pour partager les fruits de votre labeur.

# Utiliser un fichier R Markdown

## Organisation générale

Un fichier R Markdown (dont l’extension est .rmd ou .Rmd), c’est :

1. un document de texte plein qui utilise la variante syntaxique Pandoc de Markdown;
2. des blocs de code particuliers qui, en plus de pouvoir être utilisé pour présenter du code, le code va pouvoir être exécuter et le résultat du code va pouvoir être intégré dans le document. Ces blocs commencent et finissent trois accents graves (*backtick* ou *backquote* en anglais): ` et les trois accents graves ouvrant le bloc sont suivis d’une accolade qui commence par r ou R, par exemple

* ```{R name, option1, option2}  
  # code R à exécuter  
  ```
* ou encore
* ```{r option1}  
  # code R à exécuter  
  ```

1. enfin, un *Front Matter*[[9]](#footnote-76) YAML optionnel (au début du fichier, dans un bloc de trois tirets : ---) qui permet de donner des indications sur le ou les documents à générer.

Au début, il peut-être un peu difficile de se retrouver un tel fichier car on y croise différents langages, notamment

1. un langage de programmation pour les analyses, R;
2. un langage de balisage pour l’écriture du document, Markdown;
3. un langage de sérialisation de données pour spécifier les sorties du documents, YAML.

On peut ajouter qu’avec la variante syntaxique Pandoc de Markdown les symboles mathématiques TeX[[10]](#footnote-77) sont utilisés pour rendre facile l’écriture, entre autres, des équations[[11]](#footnote-79), en un sens c’est un quatrième langage à connaître!

## Créer un fichier R Markdown

Créer un fichier R Markdown, c’est simplement créer un fichier dont l’extension est .Rmd ou .rmd, ce qui peut être vais avec n’importe quelle éditeur de texte, ou de code, ou même en ligne de commande. Par example, étant moi-même utilisateur du terminal et travaillant dans un environment Linux, j’entre dans ce dernier la commande:

$ echo "---\nauthor: Kevin Cazelles\n---" > mondoc.Rmd

et le fichier est créé! Pour les utilisateurs de R Studio, cela peut se faire en 2 clics, comme illustré sur les figures qui suivent. L’intérêt d’utiliser R Studio pour cette opération est que le fichier ainsi créé contient des indications relatives à l’utilisation du fichier en question.



Première étape, utilisez l’icône de création de nouveaux fichiers.



Seconde étape, choisissez le format de sortie désiré



Un document contenant différentes instructions et examples est généré.

## Spécifier les documents à obtenir avec YAML

Le site officiel <https://yaml.org/> indique:

[*Traduction*]

YAML est un standard de sérialisation de données pour tous langages de programmation adopté à l’humain.

Autrement dit, dans un fichier au format YAML (extension .yaml ou .yml), on entre des données qui pourront être utilisées dans par de nombreux langages de programmation et ce fichier est facile à lire par l’humain. Pour que le contenu du fichier soit effectivement utilisables dans un langage de programmation donné, il faut que le code ad hoc soit disponible. Pour les utilisateurs de R, c’est le package yaml[[12]](#footnote-86) qui permet de lire de tels fichiers et de les importer sous forme de listes.

Pour comprendre l’essentiel du fonctionnement du YAML, je vous recommande le tutoriel à l’URL suivante : <https://sweetohm.net/article/introduction-yaml.html> en vous concentrant sur la syntaxe et nous sur comment utiliser Python pour le faire. De manière générale:

nomduchamp: données

et on saute une ligne pour ajouter un autre champ:

nomduchamp: données  
nomduchamp2: données2

Un champ données peut contenir une chaîne de caractère, une date, des chiffres ou encore une liste:

nomduchamp: [élément1, élément2]

ou encore

nomduchamp:  
 - élément1  
 - élément2

Noté que l’indentation est utilisée pour signifier la hiérarchie entre les différents éléments et que les commentaires sont introduits par un “#”.

nomduchamp:  
 souchamp1: élément1  
# ce qui suit sont les élément dans le 2eme souschamp  
 souchamp2:  
 souchamp2a: élément2a  
 souchamp2b:  
 - élément2ba  
 élément2ba

Aussi, voici un exemple qui, bien que non-exhaustif, montre comment un fichier est structuré et comment l’important dans R (le fichier utilisé est disponible dans le répertoire GitHub <https://github.com/KevCaz/Rmarkdowndocfr/blob/master/aux/data_01.yaml>):

author: Kevin Cazelles  
date: 2020-06-06  
description: "Il s'agit d'un exemple, utilisez les guillemets pour les longs textes avec ponctuation!"  
# Les différents variables  
var1: [C, C++, Dart, Julia, Python, R, Rust]  
var2: y  
var3:  
 var3a:  
 x: 1  
 "y": 2  
 var3b:  
 - a  
 - b  
 var3c: |  
 utiliser plusieurs  
 lignes et garder les sauts de lignes

Pour l’importer avec le package yaml, j’utilise les commandes suivantes:

library(yaml)  
yaml.load\_file("aux/data\_01.yaml")

## $author  
## [1] "Kevin Cazelles"  
##   
## $date  
## [1] "2020-06-06"  
##   
## $description  
## [1] "Il s'agit d'un exemple, utilisez les guillemets pour les longs textes avec ponctuation!"  
##   
## $var1  
## [1] "C" "C++" "Dart" "Julia" "Python" "R" "Rust"   
##   
## $var2  
## [1] TRUE  
##   
## $var3  
## $var3$var3a  
## $var3$var3a$x  
## [1] 1  
##   
## $var3$var3a$y  
## [1] 2  
##   
##   
## $var3$var3b  
## [1] "a" "b"  
##   
## $var3$var3c  
## [1] "utiliser plusieurs\nlignes et garder les sauts de lignes"

Les chaînes de caractères y, yes, true, TRUE ainsi que n, no, false, FALSE seront toutes interprétées comme variables booléennes (respectivement TRUE ou FALSE dans R). C’est une fonctionnalité bien pratique mais à double tranchant: par exemple si y est utilisé comme nom de champ, le nom du champ sera interprété comme TRUE et pour palier ce comportement, il faut ajouter des guillemets (voir l’exemple ci-dessus).

Les remarques précédentes sont générales et vous pourriez être amené à utiliser YAML dans un autre contexte que l’édition de document avec R Markdown, par exemple pour stocker des données. Revenons à l’utilisation de YAML dans un document R Markdown. Ce qu’il faut bien garder à l’esprit, c’est que les options utilisées servent à spécifier différents propriétés des documents à générer à partir du fichier .Rmd (p-ex le titre, la date, les polices de caractères utilisées, ajout d’une table des matières, etc.) grâce à des *templates* utilisés par Pandoc. Ainsi, les options disponibles dépendent des templates utilisés qui sont spécifiques à un format donné, les options par défaut varient ainsi d’un format à l’autre. Pour un aperçu assez complet des options YAML utilisables dans fichier R Markdown, rendez-vous à la dernière page du [guide de référence](http://www.rstudio.com/wp-content/uploads/2015/03/rmarkdown-reference.pdf). Notez qu’il est possible de créer ses propres templates et donc d’ajouter autant d’options que désiré.

À titre d’exemple, ce document utilise un grand nombre d’options en voici une reproduction quasi-identique (sans le résumé complet et sans la majorité des commentaires que vous pouvez trouverez dans le fichier source de ce document) :

---  
title: "Utiliser R Markdown pour créer des documents dynamiques"  
date: "01 September 2020"  
author: par Kevin Cazelles  
lang: fr  
abstract: "Le package [...] pour aller plus loin."  
fontfamily: fourier  
linestretch: 1  
fontsize: 10pt  
lof: yes  
output:  
 html\_document:  
 toc: yes  
 toc\_float: yes  
 highlight: default  
 theme: flatly  
 include:  
 after\_body: aux/footer.html  
 PDF\_document:  
 highlight: tango  
 toc: yes  
 toc\_depth: 3  
 fig\_caption: yes  
 keep\_tex: yes  
 latex\_engine: PDFlatex #xelatex  
 number\_section: yes  
 includes:  
 before\_body: aux/license.tex  
 word\_document:  
 #toc: yes  
 fig\_caption: yes  
 highlight: pygments  
 md\_document:  
 variant: markdown\_strict  
bibliography: aux/mybiblio.bib  
csl: aux/journal-of-theoretical-biology.csl  
header-includes:  
 - \usepackage{fancyhdr}  
 - \pagestyle{fancy}  
 - \fancyfoot[CO,CE]{Documents dynamiques avec R Markdown}  
 - \fancyfoot[R]{\thepage}  
---

Comme expliqué précédemment, l’indentation est utilisée pour regrouper les options par format. Ainsi, les lignes entre “pdf\_document:” et “word\_document:”, spécifient différent aspect relative à la génération au format PDF. Par exemple, pour ajouter une table des matière, on ajoute toc: yes, “toc” signifiant *table of content* (table des matières) et pour en contrôler la profondeur (le niveau maximal de sous-titre affiché dans la table), on utilise “toc\_depth”. De même, “tof” (*table of figure*) permet d’ajouter une table des figures. Aussi, l’option “number\_section:true” permet d’obtenir une numérotation des différentes parties.

## Éditer le contenu du fichier R Markdown

Pour cela, on utilise simplement du texte plein avec la syntaxe Pandoc Markdown (détaillée dans la section suivante), la différence avec un document Markdown est qu’il est possible d’exécuter du code R et d’utiliser les résultat du code, ce qui rend le document dynamique (voir la section dédiée).

## Obtenir le document final (ou les documents finaux)

Une fois le fichier est prêt (ou qu’on souhaite voir le résultat) et que le package rmarkdown est installé[^4], on utilise la fonction render() du package en lui indiquant le chemin de du fichier .Rmd. Si on est un utilisateur de RStudio, on peut cliquer sur Knit dans la barre de menu contextuelle associée au fichier ( voir la documentation associée, ?render). Dans la suite du document, la function render() est utilisée:

render("ex\_Rmardown.rmd", "all")

L’argument “all” permet d’obtenir tous les fichiers pour lesquels il existe une spécification dans le fichier YAML. Les formats utilisés ici sont :

* [PDF](https://en.wikipedia.org/wiki/Portable_Document_Format)
* [HTML](http://rmarkdown.rstudio.com/html_document_format.html)
* [Word](http://rmarkdown.rstudio.com/word_document_format.html)

Les documents au format PDF requièrent [l’installation de Latex](https://www.tug.org/texlive/quickinstall.html) étant donné que pour produire un PDF à partir d’un fichier “.rmd”, un premier fichier Latex est produit. Pour le lecteur qui désire conserver le fichier Latex utilisé pour produire le PDF, il suffit d’ajouter l’option “keep\_tex: true” dans les options relatives au format PDF dans le [Le fichier YAML][]. Pour le Markdown, un fichier “.md” est généré suivant la variante de Markdown précisée dans le YAML par “variant: markdown” (voir la section suivante). Le format “Tufte handout” est une mise en page qui inclue des marges larges où sont insérées les illustrations (dont les figures) et que nous devons au chercheur Edward Tufte.

# Le variante Pandoc de Markdown

Dans cette partie, je détaille les éléments de formatage du texte proposés par la syntaxe Pandoc de Markdown. L’ensemble est très bien présenté sur le [site de référence](http://rmarkdown.rstudio.com/authoring_pandoc_markdown.html)[[13]](#footnote-94) et très bien résumé à la première page du [guide de référence](http://www.rstudio.com/wp-content/uploads/2015/03/rmarkdown-reference.pdf). Pour une source en français, j’ai trouvé un excellent tour d’horizon nommé : “Élaboration et conversion de documents avec Markdown et Pandoc” (<http://enacit1.epfl.ch/markdown-pandoc/>) écrit par Jean-Daniel Bonjour. Notez que certains symboles sont réservés au formatage du texte. Cependant, quand leur affichage est requis, on les fait précéder du caractère d’échappement qui est, pour Markdown, l’antislash : “\”. Par exemple, j’entre :

\\ \& \# \$ \[

Pour obtenir : \ & # $ [. J’ajoute que certains sauts de ligne sont obligatoires pour obtenir la mise en page désirée (pour les notes de bas de page par exemple).

## Décoration du texte

* Pour écrire du *texte en italique*, vous avez deux possibilités :

\*le texte à mettre en italique\*  
\_le texte à mettre en italique\_

* Pour écrire du **texte en gras**, encore deux possibilités :

\*\*le texte à mettre en gras\*\*  
\_\_le texte à mettre en gras\_\_

* Pour écrire du ***texte en gras et italique***, utilisez :

\*\*le \_texte en italique et en gras\_\*\*  
```md  
  
- Pour obtenir un ~~texte rayé~~, entrez&nbsp;:  
  
```md  
~~texte rayé~~

* Pour écrire un élément en exposant, utilisez :

^texte en exposant^

* Pour écrire un élément en indice, tapez :

~texte en indice~

Notez qu’il n’y a pas de balises pour le soulignement du texte. De manière générale, quand un élément de mise en page manque dans la syntaxe, il est toujours possible d’utiliser des commandes d’un langage. Par exemple, pour souligner dans un document qui sera produit en HTML, je peux utiliser <u>texte souligné</u> mais cela ne me permettra pas d’avoir un texte souligné en Word ou en PDF. De même que si j’utilise \underline{texte souligné} le texte sera souligné en PDF, mais pas en HTML ni en Word. Procéder de la sorte n’est pas toujours souhaité car le document R Markdown perd en généralité, en ce sens où il ne pourra pas être correctement généré dans tous les formats. Cela n’est cependant pas nécessairement un problème, par exemple, si vous souhaiter obtenir le document en un seul format, ce fonctionnement devient un atout puisque vous pouvez utiliser toute la gamme de mise en forme offert le langage en question.

## Les titres

Le plus simple est d’utiliser un nombre croissant de # (ATX heading) pour descendre dans l’arborescence des titres:

# Un titre d'ordre 1  
## Un titre d'ordre 2  
### Un titre d'ordre 3

Il est aussi possible d’utiliser une série de “=” en dessous des titre de premier niveau et une ligne de “-” en dessous des titres de niveau 2. Cette option a la qualité de permettre de repérer facilement les titres dans le code source.

Un titre d'ordre 1  
==================  
Un titre d'ordre 2  
------------------  
### Un titre d'ordre 3

## Les listes

Les listes sont très intuitives en Markdown, alors qu’elles requièrent des balises un peu lourdes aussi bien en Latex qu’en HTML. Dans les exemples donnés, il faut toujours séparer le texte principal de la liste par des sauts de ligne. Aussi, il y a une différence de format

### Listes non numérotées

Pour obtenir une liste non numérotée j’entre :

\* objet 1,  
\* objet 2,  
\* objet 3.

ou bien :

+ objet 1,  
+ objet 2,  
+ objet 3.

ou encore :

- objet 1,  
- objet 2,  
- objet 3.

et même :

+ objet 1,  
\* objet 2,  
- objet 3.

Dans tous les cas, cela donne

* objet 1,
* objet 2,
* objet 3.

Et si j’utilise on ajoute un espace entre les éléments de la liste alors le rendu change un peu, la liste est plus aérée, par exemple en HTML, une balise paragraphe, <p> </p> est ajoutée, ainsi

\* objet 1,  
  
\* objet 2,  
  
\* objet 3.

devient:

* objet 1,
* objet 2,
* objet 3.

En utilisant une indentation de 4 espaces (ou une tabulation), on peut obtenir des listes hiérarchisées, ainsi:

- objet 1,  
 + machin 1  
 - chose 1  
 - chose 2  
 + machin 2  
- objet 2,  
- objet 3.

donne:

* objet 1,
  + machin 1
    - chose 1
    - chose 2
  + machin 2
* objet 2,
* objet 3.

En ajoutant au moins deux espaces à la fin des éléments d’une liste, chaque élément de la liste est formaté comme un paragraphe  :

* objet 1,
* objet 2,
* objet 3.

Pour alterner des listes avec du texte ou du code, il faut utiliser des sauts de lignes avec l’indentation adéquate. Ainsi, avec les lignes suivantes :

- élément 1&nbsp;:  
  
 Un petit texte qui pourrait expliciter ce qu'est l'élément 1.  
  
- machin 2:  
  
 for (i in 1:2) print(i)

j’obtiens :

* élément 1 :
* Un petit texte qui pourrait expliciter ce qu’est l’élément 1.
* machin 2:
* for (i in 1:2) print(i)

### Listes numérotées

Pour une liste numérotée, c’est aussi très simple, en entrant par exemple si je rentre :

1. machin 1,  
2. machin 2,  
3. machin 3.

J’obtiens:

1. machin 1,
2. machin 2,
3. machin 3.

Si les nombres ne sont pas écrits manière ordonnée, cela ne changera pas le résultat. Néanmoins, le premier nombre détermine le point de la liste. En écrivant :

3. machin 1,  
3. machin 2,  
3. machin 3,  
5. machin 4.

j’obtiens :

1. machin 1,
2. machin 2,
3. machin 3,
4. machin 4.

Pour ne pas se soucier des numéros, il existe un style par défaut :

#. machin 1,  
#. machin 2,  
#. machin 3.

on retrouve bien la première liste numérotée :

1. machin 1,
2. machin 2,
3. machin 3.

Plusieurs styles de numérotation sont disponibles, p. ex. :

#) élément 1  
 #) élément 2  
 #) élément 3  
 (1) truc 1  
 (2) truc 2  
 (5) truc 3  
 i. machin 1  
 i. machin 2  
 i. machin 3

nous donne :

1. élément 1
2. élément 2
3. élément 3
4. truc 1
5. truc 2
6. truc 3
   1. machin 1
   2. machin 2
   3. machin 3

Nous avons aussi la possibilité de mélanger les niveaux numérotés et les niveaux non-numérotés :

1. machin 1,  
 1. machin 1.1,  
 2. machin 1.2,  
2. machin 2,  
 - machin 2.1,  
 - machin 2.2,  
 - machin 3,  
3. machin 4,  
4. machin 5.

ce qui donne :

1. machin 1,
   1. machin 1.1,
   2. machin 1.2,
2. machin 2,
   * machin 2.1,
   * machin 2.2,

* machin 3,

1. machin 4,
2. machin 5.

Enfin, il possible de mettre manuellement fin à une liste en introduisant un commentaire entre les listes à séparer :

(1) truc 1  
 (2) truc 2  
 (3) truc 2b  
  
 <!-- end -->  
  
 (1) truc 3  
 (2) truc 4

ces lignes sont rendues ainsi :

1. truc 1
2. truc 2
3. truc 2b
4. truc 3
5. truc 4

## Blocs de citation

Pour utiliser un bloc de citation (la balise “blockquote” en HTML), il suffit d’utiliser “>” avant la citation. Ainsi les lignes suivantes :

> la citation est ajoutée comme ceci, elle nous donne une indentation adéquate  
pour une mise en page agréable dont le style peut être facilement travailler  
en HTML grâce au CSS.

deviennent :

La citation est ajoutée comme ceci, elle nous donne une indentation adéquate pour une mise en page agréable dont le style peut être facilement travailler en HTML grâce au CSS.

Pour ajouter une hiérarchie dans les citations, on entre :

> La citation de départ  
>  
>> une hiérarchie dans la citation

ce qui donne :

La citation de départ

une hiérarchie dans la citation

## Blocs de code

Pour l’ensemble des exemples qui suivent, que je présente sous forme de lignes de code, j’utilise un environnement simple qui utilise une police d’écriture à chasse fixe et affiche tous les caractères tels qu’ils sont entrés (les balises ne sont pas interprétées).

Il est possible d’ajouter des morceaux de code colorés selon la nature des éléments de la syntaxe d’un langage de programmation donné. Les morceaux de codes peuvent être écrits entre série de tildes (~) et une accolade est nécessaire pour préciser le langage[[14]](#footnote-104). Pour présenter un morceau de code C, par exemple :

~~~~~~~~~~~~~~~{.c}  
// Commentaire  
int c,d,g ;  
c = 10;  
d = 4;  
int func(int a,int b) { return a\*b; }  
g = func(c,d)  
printf("%i",g)  
~~~~~~~~~~~~~~~

devient :

// Commentaire  
int c,d,g;  
c = 10;  
d = 4;  
int func(int a,int b) { return a\*b; }  
g = func(c,d);  
printf("%i",g);

Plus simplement, on peut utiliser des blocs de trois apostrophes inversés (*backquote* en anglais) et enlever l’accolade et le point pour préciser le langage et ainsi, avec le code ci-dessous

```c  
// Commentaire  
int c,d,g ;  
c = 10;  
d = 4;  
int func(int a,int b) { return a\*b; }  
g = func(c,d)  
printf("%i",g)  
```

on obtient bien la même bloc de code que précédemment.

// Commentaire  
int c,d,g ;  
c = 10;  
d = 4;  
int func(int a,int b) { return a\*b; }  
g = func(c,d)  
printf("%i",g)

Dans le document, j’utilise justement ces blocs de code pour montrer comment utiliser Markdown. Une autre façon simple pour faire un bloc de code est d’ajouter simplement 4 espaces au début de chaque ligne. C’est d’ailleurs ce que j’utilise quand je montre l’utilisation des apostrophes inversés ci-dessus!

Un des grands intérêts de R Markdown est d’avoir des blocs de codes R qui peuvent être exécutés par R! Je détaille cette fonctionnalité plus bas.

## Les barres horizontales

Une ligne barre horizontale peut être ajoutée comme suit :

-------------

ou encore :

\* \* \* \* \*

## Mathématiques

### Symboles mathématiques

Pour utiliser les symboles mathématiques dans le texte, les commandes associées doivent être placées entre deux “$”. Bien sur, il faut connaître les combinaisons de caractère associées aux différents symboles. Ce sont les même que celles proposées par [Latex](http://latex-project.org/intro.html) et qui seront utilisées par [MathJax](http://www.mathjax.org) (par défaut) pour générer les expressions mathématiques dans le fichier HTML. Pour quelques exemples, [regarder ce site](http://www.suluclac.com/Wiki+MathJax+Syntax), pour quelque choses de plus complet, jetez un œil [ici](https://en.wikibooks.org/wiki/LaTeX/Mathematics). Voici tout de même quelques exemples :

* quelques lettres grecques :
* $\alpha$, $\beta$, $\delta$, $\lambda$, $\pi$, $\phi$, $\omega$, $\varpi$, $\vartheta$
* pour obtenir : , , , , , , , ,
* quelques symboles mathématiques :
* $\sum$, $\prod$, $\int$, $\infty$, $\lim$
* pour obtenir : , , , ,
* quelques combinaisons :
* $\mu \in\mathbb{R}$, $\lim\_{x \rightarrow 3} f(x)$
* pour obtenir : ,

### Équations

Pour faire des équations, il faut placer l’équation entre double $$. Créons une première équation :

$$\frac{vache}{oiseau} = \frac{2\pi}{l}$$

Pour utiliser des équations numérotées, il faut ajouter “(@label)” avant l’équation, je ré-utilise l’équation précédente :

(@eq1) $$\frac{vache}{oiseau} = \frac{2\pi}{l}$$

Je rajoute une seconde équation qui utilise les balises Latex pour créer un système :

(@eq2) $$\begin{array}{ccc}  
x^2+y^2 &=& z^2 \\  
xy &=& z  
\end{array}$$

La référence aux équations se fait en utilisant “(@label)” dans le texte. Ainsi, en écrivant “(@eq1)”, j’appelle l’équation (1). De même je peux faire référence à l’équation (2) en utilisant “(@eq2)”[[15]](#footnote-114). Les références seront gérées correctement, mais le rendu visuel n’est pas nécessairement celui désiré. Ils existent différentes solutions pour palier ceci. Pour les documents HTML, il est possible d’utiliser le méchanisme de rendu HTML de bookdown (voir <https://bookdown.org/yihui/bookdown/markdown-extensions-by-bookdown.html>), ce qui revient à utiliser un autre type de format de sortie, mais qui reste du HTML. Sinon, il est possible d’utiliser le filtre [pandoc-eqnos](https://github.com/tomduck/pandoc-eqnos) (noter qu’il y a des filtres équivalents pour les tables et les figures). Enfin, si vous voulez simplement un PDF, vous pouvez travailler comme dans un document LaTeX (voir <https://www.overleaf.com/learn/latex/Cross%20referencing%20sections,%20equations%20and%20floats>).

## Les images

Pour insérer une image, deux solutions nous sont offertes :

1. la combinaison dite “inline” : !+[nom de l’image (inclus dans la légende)]+(adresse)+{#label}. Par exemple,

* ![Le logo de R](./images/Rlogo.png)



Le logo de R

1. la combinaison dite “reference” : !+[nom de l’image (inclus dans la légende)]+[id] et ailleurs dans le document le détail [id]+:+adresse. Par exemple

* [img2]: ./images/Markdown.png  
    
   ![Le logo sur la page du Markdown original][img2]

Pour faire des références aux figures, comme pour les équations, vous pouvez utiliser un filtre dédié, [pandoc-fignos](https://github.com/tomduck/pandoc-fignos), ou, si vous souhaitez simplement avoir les références dans un document PDF’ vous pouvez peut injecter une balise \\label{lelabel} et utiliser \\ref{lelabel}. Notez qu’avec R Markdown, on peut très facilement inclure des graphiques produits avec R et meme n’importe quelle image (voir plus bas).

## Les tables

La création facilitée de tableaux est l’une des extensions bien utile de Pandoc Markdown. Il existe plusieurs extensions pour faire des tableaux. Pour avoir des détails sur le sujet, je vous recommande la [page écrite par Jean-Daniel Bonjour](http://enacit1.epfl.ch/markdown-pandoc/#tableaux). Je présente ci-dessous un exemple de tableau de style “pipe table”:

| Aligné à gauche | Aligné au centre | Par défaut | Aligné à droite  
 | :------- | :-------: | ------ | -------:  
 | truc 1.1 | truc 2.1 | \*\*\_truc 3.1\_\*\* | truc 4.1  
 | truc 1.2 | truc 2.2 | ~~truc 3.2~~ | truc 4.2  
 | truc 1.3 | truc 2.3 | \*truc 3.3\* | truc 4.1  
  
 : La légende associé au tableau.

La légende associé au tableau.

|  |  |  |  |
| --- | --- | --- | --- |
| Aligné à gauche | Aligné au centre | Par défaut | Aligné à droite |
| truc 1.1 | truc 2.1 | ***truc 3.1*** | truc 4.1 |
| truc 1.2 | truc 2.2 | ~~truc 3.2~~ | truc 4.2 |
| truc 1.3 | truc 2.3 | *truc 3.3* | truc 4.1 |

Si vous préférez les obtenir grâce à une interface de type “WYSIWYG” (*What you see is what you get*), vous pouvez utiliser [ce générateur de table](http://www.tablesgenerator.com/markdown_tables). Pour les références, vos avec les mêmes options que pour les figures et les équations, le filtre dédié étant [pandoc-tablenos](https://github.com/tomduck/pandoc-tablenos). Notez qu’avec R Markdown, vous pouvez facilement faire des tables à partir de data frames. Je reviens sur là-dessus plus bas.

## Références

### Liens hypertextes

Les liens hypertextes sont utilisés sous la forme [groupe de mots sur lequel cliquer]+(adresse du lien). Pour créer un lien vers la page Markdown de Wikipedia, j’utilise :

[Markdown](https://fr.wikipedia.org/wiki/Markdown)

et voila le lien vers la page [Markdown](https://fr.wikipedia.org/wiki/Markdown) de Wikipedia.

### Notes de bas de page

On peut produire une note de bas de page en plaçant la balise “[^id]” là où la note doit être insérée. Pour préciser le texte qui y est associé, on utilise: “[^id]:texte associé” où l’on souhaite dans le document, il faut simplement que les notes de bas pages (rassemblées ou non) soit séparé du reste du texte par un saut de ligne. Personnellement, je préfère rassembler les notes d’une section à la fin de la section en question. Par exemple :

Un bout de texte avec une note[^note1] et une autre [^note2].  
  
[^note1]: à la fin d'une section par exemple.  
[^note2]: ou encore, à la fin du document.

Un bout de texte avec une note[[16]](#footnote-130) et une autre[[17]](#footnote-131).

Notez qu’il faut séparer le texte des notes de bas de pages du texte principal d’un saut de ligne.

### Références à une section

La référence à une section se fait à l’aide de deux crochets. Dans le premier crochet, on trouve le texte associé au lien et dans le second, le nom de la partie à laquelle on fait référence :

Référence à la [section sur les liens hypertextes][Liens hypertextes]

Référence à la [section sur les liens hypertextes](#liens-hypertextes)

On peut également utiliser le nom de la section dans le premier crochet et rien dans le second :

Rappelez vous la section [Liens hypertextes][]

Rappelez vous la section [Liens hypertextes](#liens-hypertextes)

### Références bibliographiques

Un des points forts de Pandoc est la possibilité de gérer de manière très efficace votre bibliographie grace au processeur de citation [pandoc-citeproc](https://github.com/jgm/pandoc-citeproc). Un grand nombre de fichiers de bibliographie sont bien gérés[[18]](#footnote-135), dont les fichiers bibtex. Pour plus de renseignements pour utiliser pandoc-citeproc dans un document R Markdown, visitez la [page du site de R Markdown consacrée](http://rmarkdown.rstudio.com/authoring_bibliographies_and_citations.html).

L’exemple qui suit est basé sur le fichier bibtex (mybiblio.bib) disponible dans le dossier “aux” qui est donnée avec ce document et que je donne ci-dessous :

@article{Lande1979,  
 author = {Lande, R},  
 journal = {Evolution (N. Y).},  
 number = {1},  
 pages = {402--416},  
 title = {Quantitative Genetic Analysis of Multivariate Evolution,  
 Applied to Brain : Body Size Allometry Russell Lande},  
 volume = {33},  
 year = {1979}  
}  
@article{Oreskes1994,  
 author = {Oreskes, Naomi and Shrader-Frechette, Kristin and Belitz, Kenneth},  
 journal = {Science (80-. ).},  
 number = {5147},  
 pages = {641--646},  
 title = {Verification, Validation, and Confirmation of Numerical  
 Models in the Earth Sciences},  
 volume = {263},  
 year = {1994}  
}  
@article{Knauff2014,  
 author = {Knauff, Markus and Nejasmic, Jelica},  
 issn = {1932-6203},  
 journal = {PLoS One},  
 number = {12},  
 pages = {e115069},  
 title = {{An Efficiency Comparison of Document Preparation Systems  
 Used in Academic Research and Development}},  
 volume = {9},  
 year = {2014}  
}

Ce fichier est spécifié dans le YAML : “bibliography: mybiblio.bib” (voir la section “[Le fichier YAML][]”). Le style de citation peut être spécifié, entre autres, à l’aide d’un fichier CSL, lui aussi spécifié dans le YAML et fourni dans le dossier “aux”  : “csl: ./aux/journal-of-theoretical-biology.csl”. Dans le texte, on peut appeler une référence en utilisant la forme sans parenthèse ou avec parenthèse. Pour la forme sans parenthèse, il suffit d’utiliser “@id” où “id” est l’identifiant donné par le fichier “mybiblio.bib”. Pour une référence sans parenthèse :

1. Selon @Oreskes1994, la modélisation [...].

1. Selon Oreskes et al. (1994), la modélisation […].

Pour les citations entre parenthèses, les citations sont insérées entre crochets et séparées par des points-virgules :

2. Dans la littérature [...] [@Oreskes1994] bien que [...] [@Lande1979; @Knauff2014].

1. Dans la littérature […] (Oreskes et al., 1994) bien que […] (Knauff and Nejasmic, 2014; Lande, 1979).

On peut facilement ajouter du texte dans la parenthèse

3. Dans la littérature, [...] [voir @Oreskes1994] bien que [...] [@Lande1979 entre autres].

1. Dans la littérature, […] (voir Oreskes et al., 1994) bien que […] (Lande, 1979 entre autres).

Il est important de savoir que la liste des références est mise à la fin du document, nous ajoutons donc #Références (ou # bibliographie) à la fin du document pour lui donner un titre. Une autre remarque pour vous dire qu’il existe de nombreux fichiers “csl” (acronyme pour [*Citation Style Langage*](https://en.wikipedia.org/wiki/Citation_Style_Language)) sur le [site de Zotero](https://www.zotero.org/styles).

# Intégration de R dans le document

L’intérêt du package rmarkdown est d’étendre la syntaxe Pandoc Markdown avec les fonctionnalités du [package knitr](http://cran.r-project.org/web/packages/knitr/index.html) pour insérer non seulement du code R mais aussi les sorties associées (sorties console et figures). Nous obtenons ainsi un document dynamique en ce sens que si les données associées et/ou le code R changent, le document évolue aussi. Cela permet, entre autres, de créer des rapports automatisés.

## Les blocs de code R

Il y a deux manières d’insérer des sorties R dans le document: 1. directement dans le texte (“inline”); 2. en utilisant un bloc de code dédié.

Pour inclure une sortie texte directement dans un paragraphe, on utilise  : `r expression`. Ainsi, par exemple, il est possible d’insérer l’heure et la date à R en utilisant la fonction *Sys.time()*, l’utilisation de `r Sys.time()`  nous donne r Sys.time().

Le reste de cette section se concentre sur les bloc de code R (*code chunks*) qui s’utilisent comme les [blocs de code](#blocs-de-code) détaillés précédemment, mais, à la différence de ces-derniers, il est possible d’identifier le bloc de code (idbloc ci-dessous) et varier différents paramètres d’affichage du code source et des sorties (l’ensemble des paramètres sont disponible à l’URL suivante <https://yihui.org/knitr/options/>). Typiquement, l’utilisation d’un tel bloc de code ressemble à ceci :

```{r, idbloc, param1 = val1, param2 = val2}  
ligne de code 1  
ligne de code 2  
...  
ligne de code n  
```

Notons que la casse n’est pas important pour le R de l’accolade et qu’il n’est pas obligatoire d’identifier les blocs de code, il est donc possible d’utiliser

```{R param1 = val1, param2 = val2}  
ligne de code 1  
ligne de code 2  
...  
ligne de code n  
```

Si l’identification du bloc de code n’est pas obligatoire, il est en revanche important de ne pas nommer deux blocs de code R avec le me identifiant, cela posera problème au moment de générer le code (un message d’erreur sera retourné).

Les commentaires sont introduits, comme dans R, sous la forme de lignes de codes commençant par un #. Débutons avec un exemple simple qui inclut un commentaire et une addition :

```{r, addition}  
# une addition avec R.  
2+3  
```

Notez que “addition” qui suit “r” dans l’accolade est l’identifiant du morceau de code (la virgule entre les deux premiers éléments est facultative), qui permet notamment de facilement identifier les lignes précédentes me donnent :

# une addition avec R.  
2+3

## [1] 5

J’obtiens ainsi le code de R dans un environnement adéquate (voir la coloration du code) avec la sortie console associée, en l’occurrence, le résultat de l’addition. Je présente dans la suite un certain nombre d’options qui donnent une large flexibilité dans la création des morceaux de code. Pour avoir accès à plus de précisions, reportez-vous au [guide de référence](http://www.rstudio.com/wp-content/uploads/2015/03/rmarkdown-reference.pdf) ou au [site de knitr](http://yihui.name/knitr/options/).

## Modifier l’affichage du code source et des sorties associées

Grâce à différents paramètres, il est possible de sélectionner finement ce qui est affiché, aussi bien pour le code que pour les sorties associées. Cela permet à l’utilisateur de mettre en avant certaines parties du code, mais aussi de choisir finement les sorties R (figures, tables, etc) à ajouter dans le document.

Dans cette section, j’ai choisi de détailler quelques paramètres clefs, mais il y en a d’autres (voir <https://yihui.org/knitr/hooks/>) qui permettent d’aller plus loin. Il est cependant important de noter que tout n’est pas nécessairement faisable avec ces paramètres. Ainsi, il est possible de sélectionner différentes coloration syntaxiques avec le paramètre “highlight” dans [l’en-tête YAML](#X5408e78486463c18291aa7ba9e4e59516c78e04)). Aussi, pour changer la police de font des blocs de code dans un document HTML, il faudra modifier certaines propriétés CSS du document.

### Modifier l’affichage du code source avec le paramètre echo

Prenons le bloc de code suivant (volontairement non identifié).

``{r}  
# une addition de variables avec R  
a <- 2  
b <- 3  
a + b  
```

# une addition de variables avec R  
a <- 2  
b <- 3  
a + b

## [1] 5

Pour enlever le code (et uniquement le code), il suffit d’utiliser echo = FALSE, (echo = TRUE par défaut). Ainsi, le bloc de code sanscode

``{r sanscode, echo = FALSE}  
# une addition de variables avec R  
a <- 2  
b <- 3  
a + b  
```

donne:

## [1] 5

Le paramètre echo peut aussi être utilisé pour choisir les lignes à montrer. Pour cela, on utilise un vecteur indiquant les positions des lignes à montrer. Par exemple, pour montrer uniquement les lignes 1 et 4 on utilise echo de la sorte

``{r code14, echo = c(1, 4)}  
# une addition de variables avec R  
a <- 2  
b <- 3  
a + b  
```

ce qui donne:

# une addition de variables avec R  
a + b

## [1] 5

On peut aussi choisir les lignes à ne pas montrer en utilisant des valeurs négatives. Ainsi, je peux enlever les trois premières lignes du bloc de code:

``{r codesans13, echo = -c(1:3)}  
# une addition de variables avec R  
a <- 2  
b <- 3  
a + b  
```

ce qui donne:

a + b

## [1] 5

### Modifier l’affichage des sorties

#### Le paramètre comment

Les symboles devant les sorties console de R peuvent être modifiées en utilisant le paramètre comment. Avec le bloc comment1

``{r comment1, comment = "$ "}  
# une addition de variables avec R  
a <- 2  
b <- 3  
a + b  
```

on obtient

# une addition de variables avec R  
a <- 2  
b <- 3  
a + b

$ [1] 5

De me, comment2  :

``{r comment2, comment="R>>"}  
# une addition de variables avec R  
a <- 2  
b <- 3  
a + b  
```

il nous donne:

# une addition de variables avec R  
a <- 2  
b <- 3  
a + b

R>> [1] 5

#### Le paramètre results

Le paramètre results permet de choisir comment les sorties d’un bloc de code R vont être traitées. Par défaut, c’est le mode “markup” qui est utilisé:

``{r markup, results = 'markup'}  
# une division avec R  
cat("Exemple de division avec R: 1/998.001 = ", 1/998.001)  
```

# une division avec R  
cat("Exemple de division avec R: 1/998.001 = ", 1/998.001)

## Exemple de division avec R: 1/998.001 = 0.001002003

* Avec result='asis', les résultats sont affichées comme un paragraphe du document texte principal :
* ``{r asis, results = 'asis'}  
   # une division avec R  
   cat("Exemple de division avec R: 1/998.001 = ", 1/998.001)  
   ```

# une division avec R  
cat("Exemple de division avec R: 1/998.001 = ", 1/998.001)

Exemple de division avec R: 1/998.001 = 0.001002003

* Avec result='hide', les sorties console ne sont pas ajoutées :
* ``{r hide, results='hide'}  
   # une division avec R  
   cat("Exemple de division avec R: 1/998.001 = ", 1/998.001)  
   ```

# une division avec R  
cat("Exemple de division avec R: 1/998.001 = ", 1/998.001)

* results='hold' permet d’afficher *toutes* les sorties après le morceau de code. Sans cette option, les sorties sont ajoutées au fur et à mesure de l’exécution du code et donc le bloc est interrompu. Par exemple, avec le code suivant’ :
* ``{r} a <- 2 print(a) b <- 3 print(b) 2 + 3 ``` on obtient :

a <- 2  
print(a)

## [1] 2

b <- 3  
print(b)

## [1] 3

2 + 3

## [1] 5

alors qu’avec results='hold',

``{r hold, results = 'hold'}  
a <- 2  
print(a)  
b <- 3  
print(b)  
2 + 3  
```

on a :

a <- 2  
print(a)  
b <- b  
print(b)  
2 + 3

## [1] 2  
## [1] 3  
## [1] 5

### Choisir les messages retournés

Il y a trois type de messages retournés par R:

1. les messages : une simple indication (voir la fonction message());
2. les avertissements qui soulignent que quelque chose est peut-être problématiques, mais le code est exécutable (voir la fonction warning());
3. les erreurs qui indiquent que quelques chose ne marche pas dans l’exécution du code.

Considérons le bloc suivant qui inclut un message, un avertissement, une erreur (en commentaire) et une addition.

``{r}  
message("Ceci est un message")  
warning("Ceci est un avertissement")  
# stop("PCeci est un problème")  
2 + 3  
```

message("Ceci est un message")

## Ceci est un message

warning("Ceci est un avertissement")

## Warning: Ceci est un avertissement

# stop("PCeci est un problème")  
2 + 3

## [1] 5

il est possible de supprimer le message en utilisant message = FALSE.

``{r sansmessage, message = FALSE}  
message("Ceci est un message")  
warning("Ceci est un avertissement")  
# stop("Ceci est un problème")  
2 + 3  
```

message("Ceci est un message")  
warning("Ceci est un avertissement")

## Warning: Ceci est un avertissement

# stop("Ceci est un problème")  
2 + 3

## [1] 5

De la même manière il est possible de supprimer l’avertissement avec warning = FALSE.

``{r sansavertissment, warning = FALSE}  
message("Ceci est un message")  
warning("Ceci est un avertissement")  
# stop("Ceci est un problème")  
2 + 3  
```

message("Ceci est un message")

## Ceci est un message

warning("Ceci est un avertissement")  
# stop("Ceci est un problème")  
2 + 3

## [1] 5

Notez que si vous utilisez results = 'hide' les avertissements seront tout de même ajoutés et cette remarque est valable pour les messages.

``{r sansavertissment2, warning = FALSE, results = 'hide'}  
message("Ceci est un message")  
warning("Ceci est un avertissement")  
# stop("Ceci est un problème")  
2 + 3  
```

message("Ceci est un message")

## Ceci est un message

warning("Ceci est un avertissement")  
# stop("Ceci est un problème")  
2 + 3

Enfin, par défaut, si une erreur advient dans votre code, le document ne sera pas généré (c’est d’ailleurs la raison pour laquelle la ligne stop("Ceci est un problème") est commentée précédemment). Il est cependant parfois souhaitable de montrer une erreur (à des fins pédagogiques, par exemple). Pour ce faire, il faut utiliser error = TRUE:

``{r erreur, error = TRUE}  
message("Ceci est un message")  
warning("Ceci est un avertissement")  
stop("Ceci est un problème")  
2 + 3  
```

message("Ceci est un message")

## Ceci est un message

warning("Ceci est un avertissement")

## Warning: Ceci est un avertissement

stop("Ceci est un problème")

## Error in eval(expr, envir, enclos): Ceci est un problème

2 + 3

## [1] 5

### Modifier le mode d’evaluation du code

Par défaut, avec rmarkdown, dans un bloc de code R, le code est exécuté par R et les sorties sont ajoutées dans le document. Ci-dessus, nous avons vu comment modifier ce qui est présenté dans le code source et changer certain aspect de la sortie. Il est également possible de modifier l’exécution du code.

#### Enlever toutes sorties

Dans certains cas, il peut s’avérer utile de montrer le code source sans l’exécuter. Pour empêcher l’évaluation du code, on utilise eval = FALSE. Ainsi, avec

``{r eval, eval = FALSE}  
install.packages(`rmarkdown`)  
```

l’installation du package rmarkdown n’est pas exécutée mais bien affichée :

install.packages(`rmarkdown`)

On peut aussi utiliser un [blocs de code](#blocs-de-code) sans accolade

``r  
install.packages(`rmarkdown`)  
```

install.packages(`rmarkdown`)

Une des raisons de préférer eval = FALSE à un bloc sans accolade est la possibilité de suspendre l’évaluation d’un bloc de code tout en conservant les autres paramètres dans l’accolade, ce qui est très pratique lorsqu’on développe du contenu avec R Markdown.

#### Exécuter le code silencieusement

Un autre cas relativement commun avec utilise R Markdown est de vouloir exécuter de manière silencieuse. En d’autres termes, on ne veut ni présenter le code ni aucune sortie, mais le code doit avoir été exécuté. Dans un rapport dynamique, par exemple, on est souvent amener à exécuter un script ou charger des fonctions qui serviront plus bas à créer des sorties qui seront intégrer dans le rapport. Avec ce qui a été vu précédemment, on sait qu’en utilisant les options echo = FALSE, results = hide, message = FALSE et warning = FALSE ensemble dans l’accolade, on est capable de d’arriver à ce résultat. Il existe cependant une manière plus rapide d’arriver à ce résultat en utilisant include = FALSE qui exécute le code silencieusement. Ci-dessous je créé une fonction qui fait une addition simple et j’assigne le résultat à la variable res

``{r include, include = FALSE}  
mon\_addition <- function(a, b) {  
 return(a + b)  
}  
res <- mon\_addition(2, 3)  
```

Ci-dessous j’utilise ce bloc de code …

… et, comme attendu, ni code ni sortie ne sont retournés, mais pourtant le code a bien été exécuté, et donc si je demande l’affichage de res

print(res)

## [1] 5

j’obtiens bien 5 et je peux également utiliser ma fonction mon\_addition() :

mon\_addition(12, 30)

## [1] 42

## Tables, figures et widget

Les remarques de la section précédente sont valides pour tout type de sortie. Les tables et les figures créées avec R sont des sorties très souvent utilisées qu’il est bon de savoir personnaliser davantage.

### Tables

Avec le package knitr, il est possible d’intégrer dans le texte une table créée sous R. Commençons par créer un objet data.frame :

var1 <- 20 \* runif(12)  
tab1 <- data.frame(  
 experience = paste0("traitement\_", rep(1:3, each = 4)),  
 replicat = rep(letters[1:4], 3),  
 var1 = var1,  
 var2 = var1 + rnorm(12)  
 )

La fonction *kable()* du package knitr nous permet d’obtenir notre object *tab1* en différents formats (dont “latex”, “html”, “markdown”). Ici, en utilisant “markdown” par défaut on aura une table qui sera convertie dans le format désiré par Pandoc. De plus, la fonction *kable()* inclut le paramètre caption pour ajouter une légende.

library(knitr)  
kable(tab1, caption = "Table créée à partir de \*df1\*")

Table créée à partir de *df1*

|  |  |  |  |
| --- | --- | --- | --- |
| experience | replicat | var1 | var2 |
| traitement\_1 | a | 4.692345 | 3.108282 |
| traitement\_1 | b | 19.656305 | 19.081249 |
| traitement\_1 | c | 4.736799 | 4.702092 |
| traitement\_1 | d | 15.068817 | 15.571055 |
| traitement\_2 | a | 19.327255 | 16.997823 |
| traitement\_2 | b | 17.938904 | 18.725079 |
| traitement\_2 | c | 18.229453 | 18.200675 |
| traitement\_2 | d | 6.657505 | 7.228780 |
| traitement\_3 | a | 19.763822 | 19.603095 |
| traitement\_3 | b | 6.528706 | 6.725647 |
| traitement\_3 | c | 19.966825 | 18.675096 |
| traitement\_3 | d | 13.505030 | 13.566587 |

Il y a bien d’autre paramètres dans la fonction *kable()*, par exemple, il est possible de choisir le nombre de chiffres utilisés utiliser après la virgule ou encore d’inclure les numéros de lignes.

kable(tab1, caption = "Table créée à partir de \*df1\*", digits = 3, row.names = TRUE)

Table créée à partir de *df1*

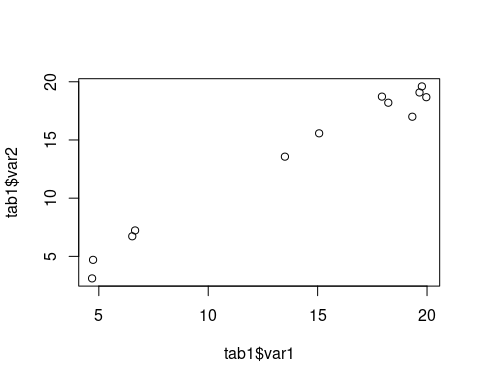
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | experience | replicat | var1 | var2 |
| 1 | traitement\_1 | a | 4.692 | 3.108 |
| 2 | traitement\_1 | b | 19.656 | 19.081 |
| 3 | traitement\_1 | c | 4.737 | 4.702 |
| 4 | traitement\_1 | d | 15.069 | 15.571 |
| 5 | traitement\_2 | a | 19.327 | 16.998 |
| 6 | traitement\_2 | b | 17.939 | 18.725 |
| 7 | traitement\_2 | c | 18.229 | 18.201 |
| 8 | traitement\_2 | d | 6.658 | 7.229 |
| 9 | traitement\_3 | a | 19.764 | 19.603 |
| 10 | traitement\_3 | b | 6.529 | 6.726 |
| 11 | traitement\_3 | c | 19.967 | 18.675 |
| 12 | traitement\_3 | d | 13.505 | 13.567 |

Pour en apprendre davantage, reportez-vous à la documentation de cette fonction. Le package [kableExtra](https://cran.r-project.org/web/packages/kableExtra/vignettes/awesome_table_in_html.html) offre de très nombreuses fonctionnalités pour créer des tableaux plus complexes, je vous recommande chaudement d’y jeter un coup d’œil.

### Les figures produites avec R

Avec le package rmarkdown, il est très facile d’insérer les figures produites avec R dans un document, le plus simple est de regarder ce que nous génère l’appel à un plot simple.

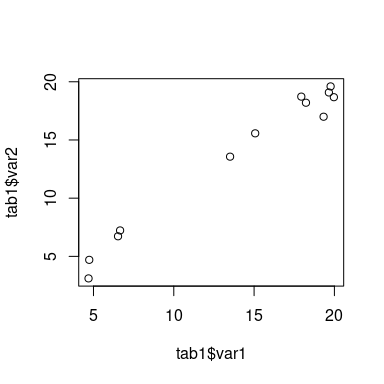
plot(tab1$var1, tab1$var2)



J’obtiens la figure demandée qui est d’assez grande dimension et numérotée “Figure 3”. Pour formater une figure obtenue avec R, il existe des options pour contrôler, entre autres, sa taille, son alignement et sa légende (taille et alignement ne sont pas supportés pour la sortie Word). Pour visualiser l’ensemble des options disponibles, je vous invite à regarder la page 3 du [guide de référence](http://www.rstudio.com/wp-content/uploads/2015/03/rmarkdown-reference.pdf). Je vais changer la taille des figures, l’alignement et ajouter une légende

``{r, fig.cap="une figure avec une légende",  
fig.height=4, fig.width=4, fig.align='left'}  
plot(tab1$var1,tab1$var2)  
```

plot(tab1$var1, tab1$var2)



une figure avec une légende

On peut alors profiter des facilités graphiques de R pour créer des figures élégantes. Je peux, par exemple, personnaliser la figure précédente avec quelques lignes supplémentaires.

par(bty="l", font=2)  
plot(tab1$var1, tab1$var2, pch=15:19, col=rep(c(8,4,6),each=5), xlab="Mon axe des abscisses",  
 ylab="Mon axe des ordonnées")  
 legend("bottomright", letters[1:5], pch=15:19, bty="n")

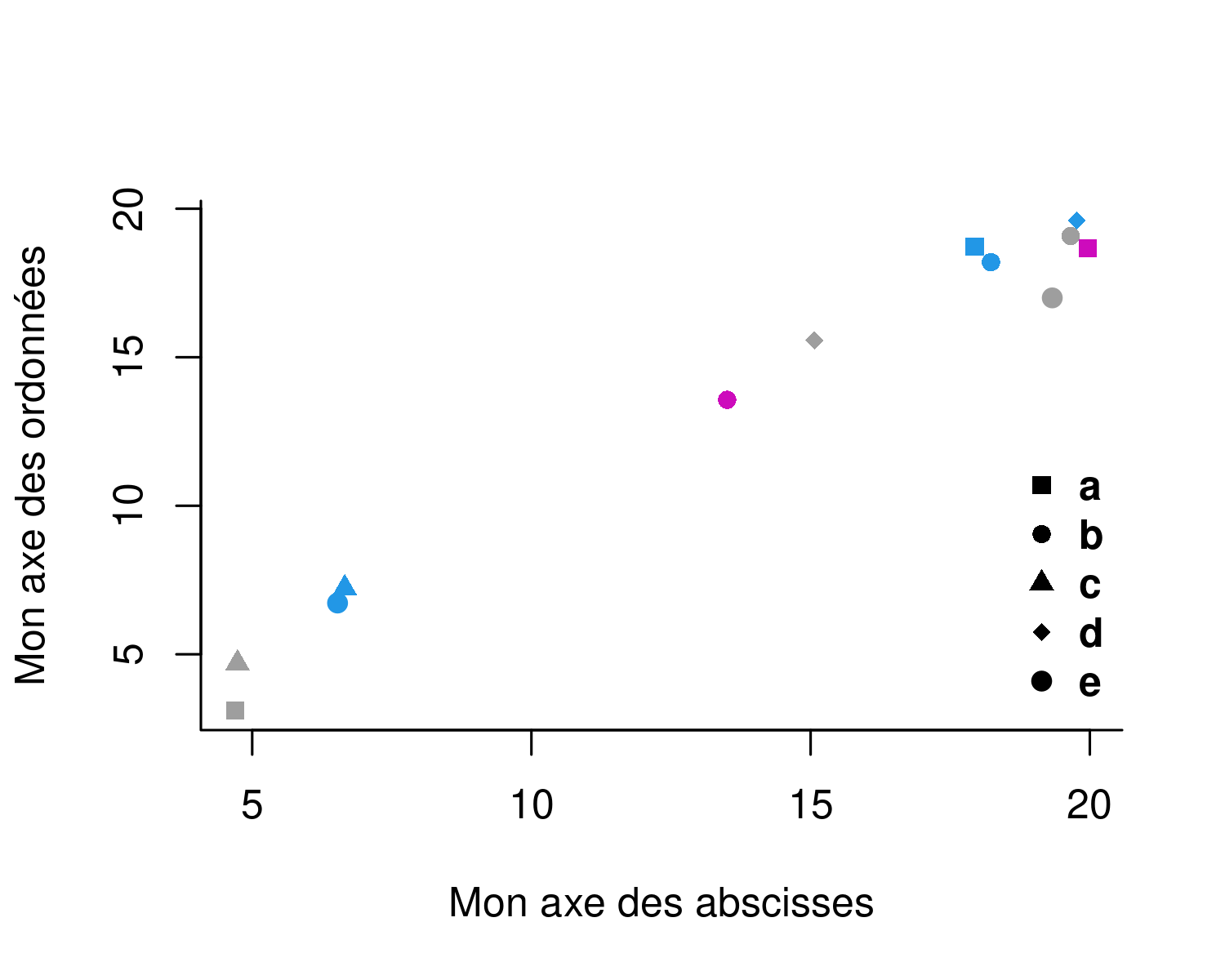
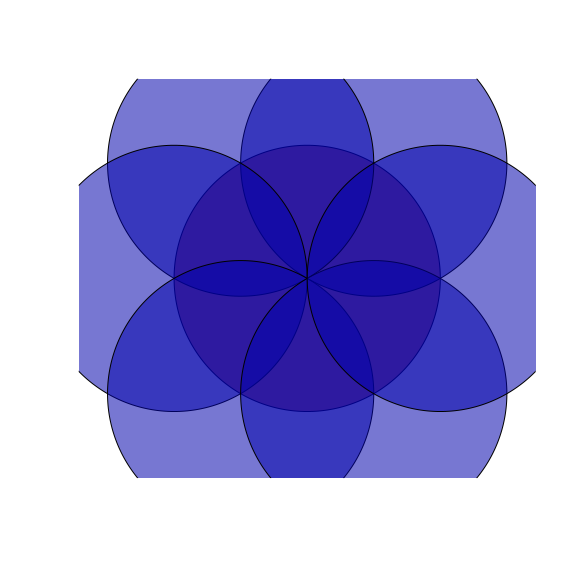


Figure associé à la table 1 […]. Le gris pour l’expérience 1, bleu pour la seconde et violet pour la troisième.

J’ai ajouté à la fin de la légende “\label{figcol}”, je peux donc l’appeler en utilisant “\ref{figcol}” mai ce n’est valable que pour la sortie PDF. De mon point de vue, on peut considérée qu’on a une extension graphique en ligne de commande un peu moins verbeuse que [PGF/TikZ](http://sourceforge.net/projects/pgf/).

kcircle <- function(rad=1,centre=c(0,0),from=0,to=2\*pi,dt=0.001,...){  
 pt <- seq(from, to, dt)  
 polygon(centre[1]+rad\*cos(pt),centre[2]+rad\*sin(pt),...)  
}  
plot(c(-1.5,1.5), c(-1.5,1.5), asp=1, type="n", ann=FALSE,  
 axes=FALSE, xaxs="i", yaxs="i")  
kcircle(col="#AA000088")  
for (i in 1:6) kcircle(1,c(cos(i\*pi/3),sin(i\*pi/3)), col="#0000AA88")



Ma rosace

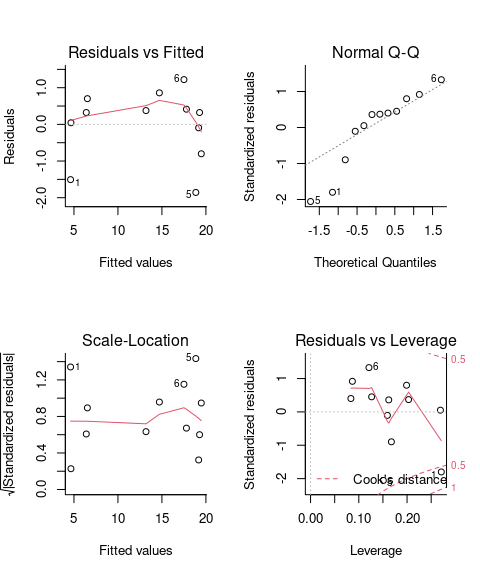
### Une application, modèle linéaire dynamique

Je souhaite faire un modèle linéaire avec la variable explicative *var1* et la variable à expliquer *var2* de notre *tab1*. Nous créons alors un petit document sachant que les données évolueront. Je vais faire une inclusion donc utiliser du code R dans mon document.

mod1 <- lm(var2 ~ var1, data = tab1)  
summary(mod1)

##   
## Call:  
## lm(formula = var2 ~ var1, data = tab1)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.8574 -0.2712 0.3232 0.4855 1.2207   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.04986 0.70237 0.071 0.945   
## var1 0.97300 0.04645 20.949 1.36e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.9807 on 10 degrees of freedom  
## Multiple R-squared: 0.9777, Adjusted R-squared: 0.9755   
## F-statistic: 438.9 on 1 and 10 DF, p-value: 1.363e-09

par(bty = "l", mfrow = c(2, 2))  
plot(mod1)



Mon modèle linéaire

# Fonctionnalités avancées

!!WIP!! !!WIP!! !!WIP!! !!WIP!! !!WIP!!

!!WIP!! —- À venir —- !!WIP!!

!!WIP!! !!WIP!! !!WIP!! !!WIP!! !!WIP!!

Dans cette section, je fais une mention rapide de quelques fonctionnalités que je qualifie d’avancées, au sens ou leur utilisation est peut-être moisn répandue et leurs intérêt mois évident pour l’utilisateur débutant.

## Changer les réglages par défaut

## Utiliser le cache

## Extraire le code R

## Utiliser du code depuis un script externe

## Utiliser les hooks

# Références

Knauff, M., Nejasmic, J., 2014. An Efficiency Comparison of Document Preparation Systems Used in Academic Research and Development. PLoS One 9, e115069.

Lande, R., 1979. Quantitative Genetic Analysis of Multivariate Evolution , Applied to Brain : Body Size Allometry Russell Lande. Evolution (N. Y). 33, 402–416.

Oreskes, N., Shrader-Frechette, K., Belitz, K., 1994. Verification, validation, and confirmation of numerical models in the earth sciences. Science (80-. ). 263, 641–646.

Xie, Y., 2017. Dynamic documents with R and knitr. CRC Press, Boca Raton, Florida.

1. <https://en.wikipedia.org/wiki/John_Gruber>, consulté le 1er juin 2020. [↑](#footnote-ref-25)
2. <https://fr.wikipedia.org/wiki/ReStructuredText>, consulté le 1er juin 2020. [↑](#footnote-ref-28)
3. <https://blog.codinghorror.com/responsible-open-source-code-parenting/>, consulté le 1er juin 2020. [↑](#footnote-ref-30)
4. Par example Goldmark, <https://github.com/yuin/goldmard>, un parser Markdown écrit en Go et utilisé par Hugo (un générateur de site très populaire), est compatible avec Common Mark. [↑](#footnote-ref-38)
5. La version 1 n’utilise pas Pandoc. [↑](#footnote-ref-41)
6. Pandoc est d’ailleurs capable de gérer différentes variantes de Markdown. [↑](#footnote-ref-43)
7. <https://pandoc.org/MANUAL.html#pandocs-markdown>, 1er juin 2020. [↑](#footnote-ref-45)
8. voir <https://pandoc.org/MANUAL.html#templates>, consulté le 1er juin 2020. [↑](#footnote-ref-59)
9. du contenu préliminaire. [↑](#footnote-ref-76)
10. voir ><https://fr.wikibooks.org/wiki/LaTeX/%C3%89crire_des_math%C3%A9matiques>>, consulté le 2 juin 2020. [↑](#footnote-ref-77)
11. Voir la section “Math” à l’URL suivant ><https://pandoc.org/MANUAL.html#pandocs-markdown>>, consulté le 2 juin 2020. [↑](#footnote-ref-79)
12. <https://cran.r-project.org/web/packages/yaml/index.html>, consulté le 8 juin 2020. [↑](#footnote-ref-86)
13. Le site de documentation est totalement écrit en Markdown et reprend ce que nous pouvons lire sur le [site de Pandoc](http://pandoc.org/demo/example9/pandocs-markdown.html) et que je reprend mais en français! [↑](#footnote-ref-94)
14. Voir la section *Fenced code blocks* du manuel de référence <https://pandoc.org/MANUAL.html>. [↑](#footnote-ref-104)
15. voir <https://tex.stackexchange.com/questions/111868/pandoc-how-can-i-get-numbered-latex-equations-to-show-up-in-both-pdf-and-html-o>) [↑](#footnote-ref-114)
16. la première [↑](#footnote-ref-130)
17. la seconde [↑](#footnote-ref-131)
18. Les logiciels de gestion de bibliographie génèrent ces fichiers. [↑](#footnote-ref-135)