

01-R Markdown 历史和工作原理

梁昊

02/04/2021

- 1 R Markdown 的历史
- 2 R Markdown 的工作原理
- 3 R Markdown 文档的结构
- 4 R Markdown 学术应用场景
- 5 Rstudio 的安装、设置和包管理
- 6 课后作业

Section 1

R Markdown 的历史

时间节点

R Markdown 始于 2012 年，配合 knitr package 进行格式转化

- 但那时候功能非常有限，尤其表格、数学公式、引用等都不完善

2014 年 rmarkdown package 诞生了

经过 6 年多的发展，R Markdown 已经形成完整生态，功能越发强大

- Demo
- gallery

R Markdown 关键人物

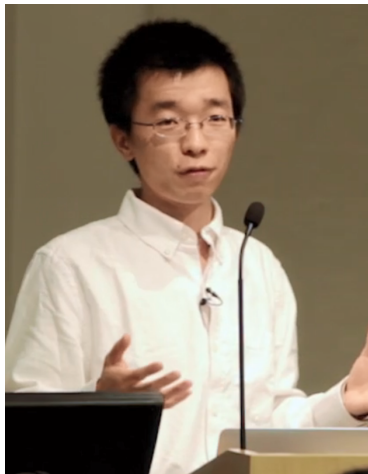


图 1: Yihui Xie 谢益辉



图 2: Hadley Wickham

Section 2

R Markdown 的工作原理

R Markdown 站在巨人的肩膀上

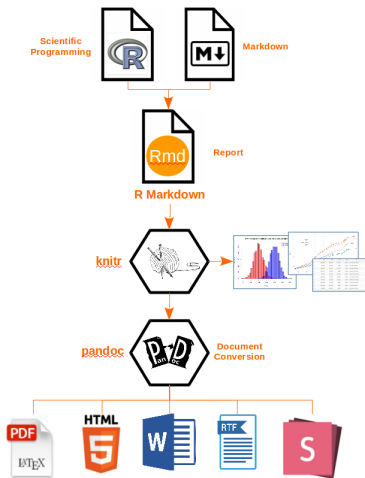


图 3: R Markdown Workflow

R Markdown 主要依赖两个包来运行和输出 (knitr and Pandoc)

- R Markdown stands on the shoulders of knitr and Pandoc.

安装rmarkdown以后, 会自动集成 knitr 和 Pandoc

R Markdown 基础包安装

html 和 pdf 是 R Markdown 输出最基础的格式，html 相对简单

pdf 格式的输出需要 TeX，TeX Live、MiKTeX 等 (~5Gb 大小)TeX 工具包的安装是繁重而痛苦

tinytex 告别了繁琐的安装

```
install.packages('rmarkdown')
install.packages('tinytex')
tinytex::tlmgr_repo('https://mirrors.tuna.tsinghua.edu.cn/CTAN/') #国内镜像
tinytex::install_tinytex() # install TinyTeX
tinytex::is_tinytex() # 安装成功返回TURE，必须检测
```


点击 knit 按钮时发生了什么-1



图 4: Rmd 文档点击 knit 按钮时

点击 knit 按钮时发生了什么-2

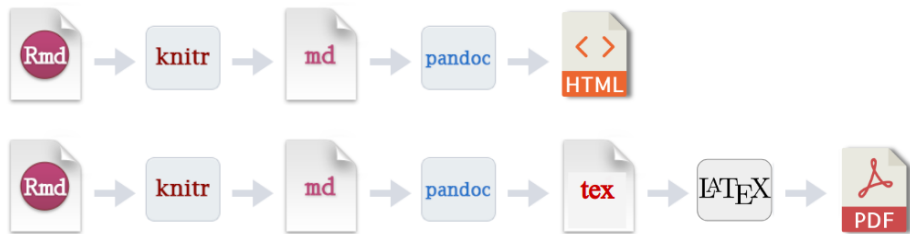


图 5: 点击按钮时执行了 `rmarkdown::render()` 命令, pdf 生成需要 TeX

Section 3

R Markdown 文档的结构

R Markdown 主要由三个部分组成

YAML
header

formatted
text

code
chunk

```

1  ----
2  title: "BI5302 Variance Heterogeneity Practical"
3  author: "Alex Douglas"
4  date: "17/10/2019"
5  output:
6    pdf_document: default
7    html_document: default
8  fontsize: 11pt
9  ----
10
11 Setup global options for knitr package. Normally I would not display these but I leave them here for your
12 information. The arguments ``width.cutoff`` and ``tidy = TRUE`` keeps the displayed code within the code
13 boxes (see what happens if you omit this).
14
15 ---
16
17 - Benthic Biodiversity experiment
18 These data were obtained from a mesocosm experiment which aimed to examine the effect of benthic polychaete
19 ("Nereis diversicolor") biomass on sediment nutrient release (NH4-, NO3- and PO4-). At the start of the
20 experiment replicate mesocosms were filled with
21 Import all the packages required for this exercise:
22
23 ---[r import data]
24 nereis <- read.table("/Users/nhy163/Documents/Alex/tmp/Nereis2.txt", header = TRUE)
25 nereis$fbiomass <- factor(nereis$biomass)
26 str(nereis)
27
28 3. How many replicates are there for each biomass and nutrient combination?
  
```

图 6: 来源:https://alexnd106.github.io/intro2R/Rmarkdown_intro.html

Section 4

R Markdown 学术应用场景

主要应用场景及对应的 R 包

- 写论文 ~ `rmarkdown/articles/pagedown`

写书 ~ `bookdown`
- 写毕业论文 ~ `thesisdown`

写学术墙报 ~ `posterdown`
- 写简历 ~ `vitae`

学术网站 ... ~ `blogdown/distill` ...

Section 5

Rstudio 的安装、设置和包管理

Rstudio 的安装和设置

本教程 Rstudio 版本必须为 1.4，R 的版本必须为 4.0+，建议都安装最新版本

建议安装时，错误提示等均选择英文，不安装中文提示包

Rstudio CRAN 镜像选为国内，速度快

Rstudio Python 引擎设定

Rstudio pdf 输出引擎设定

包管理和需要安装的包

关于包管理，建议采用命令行

- 详见我的R 包管理终极教程

必须要安装的包 (之前已安装了 rmarkdown 和 tinytex):

```
install.packages(c("tidyverse", "pacman", "rticles",  
  "pagedown", "bookdown", "reticulate", "  
  kableExtra"))
```

加载包 (library)

`library` 为默认加载命令，只能加载一个包

`pacman` 安装好以后，可以一次加载多个包，强烈推荐

关于 “::”，该命令为在需要用某个函数时临时直接加载包，不用事先 `library`，这样可以减轻系统负担。比如 `pacman::p_load` 命令就是在未加载 `pacman` 这个包时，使用了其 `p_load` 命令。

```
library(tidyverse)
```

```
library(reticulate)
```

```
library(kableExtra)
```

```
# OR
```

```
pacman::p_load(tidyverse, reticulate, kableExtra) #
```

如果其中有包未安装，则该条命令会安装并加载它

Section 6

课后作业

课后作业

安装好 rmarkdown 和 tinytex 包

- 新建一个 R Markdown notebook 文档，只输入英文
- 分别生成 html 和 pdf 文件