第三章:可重复性报告

王敏杰

2020年7月26日

四川师范大学

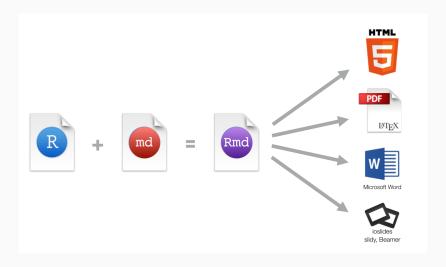
为什么要做可重复性报告

交流-理解-重复

- 需要**展示和分享**我们的数据分析结果给同行、老板或者 老师
- 为了让老板能快速地的理解我们的分析思路和方法, 最好的方法,就是将分析背景、分析过程、分析结果 以及图表等形成报告
- 让读者能重复和验证我们的结果,确保结论的真实可信

因此,本章将介绍用 Rmarkdown 生成分析报告(可重复性报告)

什么是 Rmarkdown



章节

```
# 第一章 (注意 "#" 与 " 第一章" 之间有空格)
## 第一节 (同上, "##" 与 " 第一节" 之间有空格)
### 第一小节 (同上, "###" 与 " 第一小节" 之间有空格)
```

■ 正文

This is a sentence. ... 这是正文...

■ 序列

Now a list begins:

- no importance
- again
- repeat

A numbered list:

- 1. first
- 2. second

■ 其他标记

```
__bold__
_italic_
~~strike through~~
```

创建 RMarkdown

创建 RMarkdown

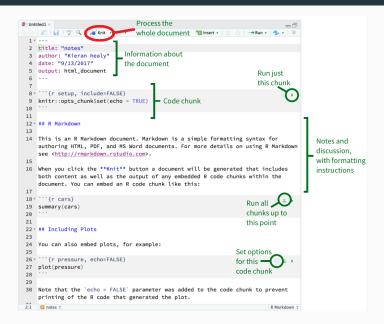
```
install.packages("rmarkdown")
```

Rstudio 中创建: File -> New File -> R Markdown.

基本构成(图中绿色括号地方)

- metadata
- text
- code

创建 RMarkdown



生成 html 文档

希望 html 文档有章节号、目录或者更好显示表格,可以修改头文件(用下面的内容替换 Rmarkdown 的头文件)

```
title: Habits
author: John Doe
date: "2020-07-26"
output:
  html document:
    df print: paged
    toc: yes
    number sections: yes
```

生成 pdf 文档

优雅的 pdf 文档

- pdf 文档可以插入漂亮的矢量图和优雅的数学公式,所以备受同学们的喜欢。
- 但往往我们写中文的时候,编译不成功,解决方案就 是使用 tinytex,可以看这个视频。

Rmarkdown 使用方法

插入公式

我相信你已经熟悉了 latex 语法, 那么我们在 Rmarkdwon 里输入 \$\$\frac{\sum (\bar{x} - x_i)^2}{n-1}\$\$, 那么实际输出:

$$\frac{\sum (\bar{x} - x_i)^2}{n - 1}$$

插入公式

也可以使用 latex 的等式环境,比如

```
$$
\Theta = \begin{pmatrix}\alpha & \beta\\
\gamma & \delta
\end{pmatrix}
$$
```

输出

$$\Theta = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

插入图片

```
```{r, out.width='35%', fig.align='center', fig.cap='this is caption'}
knitr::include_graphics("images/R_logo.png")
...
```



图 1: this is caption

#### 运行代码

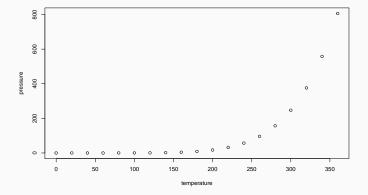
## 表格

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa

需要更优美的表格, 可参考这里

# 生成图片

```
```{r}
plot(pressure)
...
```



把这段代码复制到你的 Rmarkdown 文档试试

```
```{r, out.width = '85%', fig.showtext = TRUE}
library(tidyverse)
library(nycflights13)
library(showtext)
showtext_auto()
flights %>%
 group_by(dest) %>%
 summarize(
 count = n().
 dist = mean(distance, na.rm = TRUE),
 delay = mean(arr delay, na.rm = TRUE)
) %>%
 dplyr::filter(delay > 0, count > 20, dest != "HNL") %>%
 ggplot(mapping = aes(x = dist, y = delay)) +
 geom_point(aes(size = count), alpha = 1 / 3) +
 geom_smooth(se = FALSE) +
 ggtitle(" 这是我的标题")
```

#### 延伸阅读

- Markdown tutorial
   https://www.markdowntutorial.com (10 分钟学完)
- LaTeX tutorial https://www.latex-tutorial.com/quick-start/
- Rmarkdown 介绍
  https://bookdown.org/yihui/rmarkdown/
- Rmarkdown 手册 https: //bookdown.org/yihui/rmarkdown-cookbook/