

Week1 Rmarkdown and Reproducibility

2023年12月27日 17:03

阅读R for Data Science Chapter 21 RMarkdown

Introduction

rmarkdown满足3种需求:

- 1.决定者, 关注结果
- 2.合作者, 关注结果和过程
- 3.日常工作记录

手头常备工具书:

R Markdown Cheat Sheet

R Markdown Reference Guide

R Markdown Basics

接下来将介绍text, code和YAML header

Text formatting with Markdown

这些格式忘了可以翻阅Help中的Markdown Quick Reference

Text

换行后空一行或两个空格后换行实现换行

内容 斜体

内容 加粗

~内容~ 下标

^内容^ 上标

标题

一级标题

二级标题

三级标题

列表

*加空格 项目符号, 换行后TAB加+和空格实现二级项目符号

1.

2.

3.实现列出条目

***或---是横线

>

>后面是引用

```
| Col1 | Col2 | Col3 |
| :---: | :---: | :---: |
| value1 | value2 | value3 |
| value4 | value5 | value6 |
| value7 | value8 | value9 |
```

可以画出表格, 其中的.:可以使内容居中

Code chunks(代码块)

创建代码块有三种方式

1. Ctrl+Alt+I
2. 点击'Insert'按钮，在Code>Insert chunk中
3. 键入``{r}和``

运行代码块中的代码可以使用快捷键
Ctrl+Shift+Enter

1.Chunk name

代码块可以取名字通过``{r by-name}

这样做有三个优势：

1. 可以通过脚本编辑器左下角的代码导航快速找到对应代码块
2. 代码块做的图可以有有用的名字，方便后续使用
3. 建立缓存快网络，避免每次运行都执行

2.Chunk options

选项跟在代码块名字后面，与名字用逗号分隔，多个选项间也用逗号分隔
可以用options自定义块输出下面介绍最常用的一些选项，更多可参考

<http://yihui.name/knitr/options/>

1. eval = FALSE 不输出结果，只输出代码
2. include = FALSE运行代码但是不输出结果到报告上。将此用于不希望使报告混乱的setup代码
3. echo = FALSE 不将代码输出到报告上，但把结果输出到报告上
4. message = FALSE或warning = FALSE不输出提示和警告信息
5. results = 'hide'不输出结果到报告上；fig.show='hide'不输出图到报告上
6. error = TRUE即使报错，依然输出到报告上

3.Table

通常输出的数据框是这样的

```
mtcars[1:5, ]
```

#>	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
#> Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
#> Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
#> Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1
#> Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
#> Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	0	0	3	2

可以使用函数

```
knitr::kable(  
mtcars[1:5,],  
caption = 'A knitr kable'  
)
```

来使输出表格变为

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	0	0	3	2

4.Caching(缓存)

选项设置为`cache=TRUE`,可以保存运行结果，当有修改时再重新运行
当代码块依赖于其它代码块时可以用`dependson='依赖代码块名字'`进行链接

5.Global options

可以用`knitr::opts_chunk$set()`进行全局设置

例如:

```
```{r setup, include=FALSE}
knitr::opts_chunk$set(warning=FALSE)
```

可以全局设定不输出提示信息，那些选项同上

#### YAML header

可以通过调整YAML标头的参数来控制许多其它整个文档的设置。

#### document parameters(文档参数)

可以查阅

[http://rmarkdown.rstudio.com/developer\\_parameterized\\_reports.html#parameter\\_user\\_interfaces](http://rmarkdown.rstudio.com/developer_parameterized_reports.html#parameter_user_interfaces)

来看到更多文档参数

例子:

```

title: My Document
output: html_document
params:
 year: 2018
 region: Europe
 printcode: TRUE
 data: file.csv

```