简单技术报告框架

李东风

2019-07-04

Table of Contents

[1 问题介绍 1](#_Toc13129118)

[2 数据分析演示 2](#_Toc13129119)

[2.1 样例数据 2](#_Toc13129120)

[2.2 相关系数 3](#_Toc13129121)

[2.3 散点图矩阵 3](#_Toc13129122)

[3 总结 4](#_Toc13129123)

# 问题介绍

本子目录（Rmarkdown）是一个简单的R Markdown文件模板。 只要复制此子目录改为适当名称， 然后在复制得到的子目录中新建一个RStudio项目， 在RStudio项目中打开report.Rmd文件进行修改， 就可以将其输出为HTML、Word、PDF等格式。

这个模板是一个简单的R Markdown文件模板， 不使用bookdown的功能， 所以不支持公式、图表的自动编号和引用等高级功能。

但是，如果用到HTML输出和数学公式， 为了避免网络延迟使得公式显示困难以及断网情况下网页上公式的正常显示， 需要使得.Rmd源文件的上两层的目录中有MathJax目录， 本模板发布的打包文件中已经包含了此内容，只要不随便更改目录结构就没有问题。 这样一份MathJax本地库可以支持多个Rmarkdown和bookdown项目。 MathJax库总空间并不大但是有三万多个小文件， 很不利于项目备份与分发。

这个模板直接可以转换为HTML和Word格式， 如果操作系统中安装有LaTeX编译系统， 还可以通过LaTeX转换成PDF。 实际上， HTML文件可以在网络浏览器中通过打印到一个PDF打印机转换为PDF， Word可以在MS Office软件中另存为PDF。

公式测试：

可以用Rstudio的Knit快捷图标实现转换， 但有时这种功能不太可靠， 可以用如下命令进行转换：

## 转换为HTML：  
rmarkdown::render("report.Rmd", output\_format="html\_document", encoding="UTF-8")  
## 转换为Word：  
rmarkdown::render("report.Rmd", output\_format="word\_document", encoding="UTF-8")  
## 转换为PDF，借助于LaTeX：  
rmarkdown::render("report.Rmd", output\_format="pdf\_document", encoding="UTF-8")

如果有其他的Rmd文件，只要在命令中改成相应的Rmd文件名即可。

# 数据分析演示

## 样例数据

以R软件中带有的iris数据为例， 有3个品种的鸢尾花， 各抽取了50枚样品， 测量花瓣、花萼的长宽。

从数据中随机抽取10个观测，列表如下：

set.seed(1)  
d1 <- iris[sort(sample(150, size=10)), ]  
knitr::kable(d1)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Sepal.Length | Sepal.Width | Petal.Length | Petal.Width | Species |
| 7 | 4.6 | 3.4 | 1.4 | 0.3 | setosa |
| 14 | 4.3 | 3.0 | 1.1 | 0.1 | setosa |
| 21 | 5.4 | 3.4 | 1.7 | 0.2 | setosa |
| 43 | 4.4 | 3.2 | 1.3 | 0.2 | setosa |
| 51 | 7.0 | 3.2 | 4.7 | 1.4 | versicolor |
| 68 | 5.8 | 2.7 | 4.1 | 1.0 | versicolor |
| 74 | 6.1 | 2.8 | 4.7 | 1.2 | versicolor |
| 85 | 5.4 | 3.0 | 4.5 | 1.5 | versicolor |
| 106 | 7.6 | 3.0 | 6.6 | 2.1 | virginica |
| 129 | 6.4 | 2.8 | 5.6 | 2.1 | virginica |

其中knit::kable()用于显示数据框。

## 相关系数

对每种花， 分别计算四个测量值的相关系数:

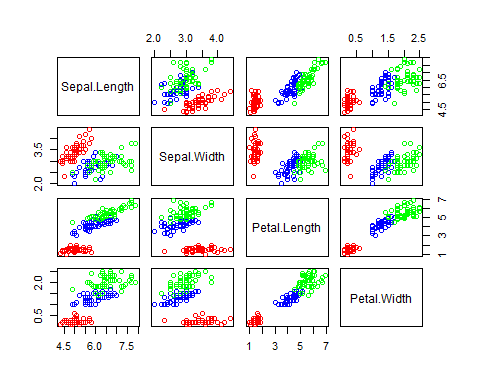
sps <- sort(unique(iris$Species))  
for(sp in sps){  
 d2 <- iris[iris$Species == sp,]  
 cat("Species: ", sp, "\n")  
 print(cor(d2[,1:4]), digits=2)  
}

## Species: setosa   
## Sepal.Length Sepal.Width Petal.Length Petal.Width  
## Sepal.Length 1.00 0.74 0.27 0.28  
## Sepal.Width 0.74 1.00 0.18 0.23  
## Petal.Length 0.27 0.18 1.00 0.33  
## Petal.Width 0.28 0.23 0.33 1.00  
## Species: versicolor   
## Sepal.Length Sepal.Width Petal.Length Petal.Width  
## Sepal.Length 1.00 0.53 0.75 0.55  
## Sepal.Width 0.53 1.00 0.56 0.66  
## Petal.Length 0.75 0.56 1.00 0.79  
## Petal.Width 0.55 0.66 0.79 1.00  
## Species: virginica   
## Sepal.Length Sepal.Width Petal.Length Petal.Width  
## Sepal.Length 1.00 0.46 0.86 0.28  
## Sepal.Width 0.46 1.00 0.40 0.54  
## Petal.Length 0.86 0.40 1.00 0.32  
## Petal.Width 0.28 0.54 0.32 1.00

## 散点图矩阵

作所有观测的四个测量值的散点图矩阵， 并用不同颜色区分品种：

cols <- c("red", "blue", "green")  
names(cols) <- sps  
plot(iris[,1:4], col=cols[as.character(iris$Species)])



# 总结

通过对分析结果的讨论， 给出对研究报告或者作业题的适当结论。