



中国矿业大学 (北京)

China University of Mining & Technology, Beijing

R Markdown 应用之学位论文排版

.....

第十一届中国 R 会议（北京）@ 北大光华

黄湘云

2018-04-27



就是要把你解救出来

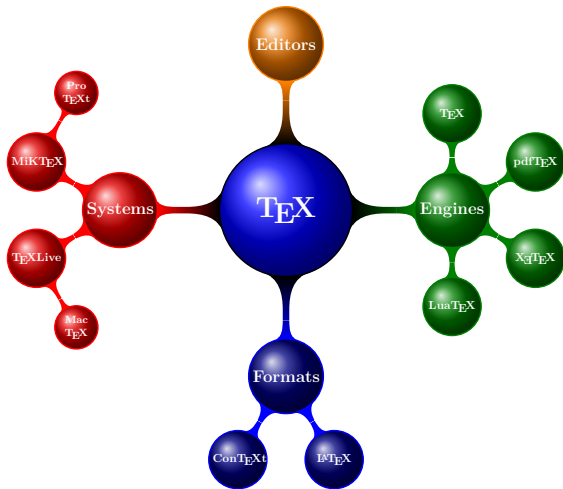




图 1: 来自 <https://www.packtpub.com/hardware-and-creative/latex-cookbook>





将 Word 和 LaTeX 模板改造成 R Markdown 模板

Word

-  交叉引用
-  字体、字号

LaTeX

-  公式无法预览
-  反斜杠太多

Pandoc's Markdown & R Markdown

什么是 Pandoc's Markdown

- ⇒ 支持转化加强版的 Markdown 文档，即 Pandoc's Markdown¹。-><http://pandoc.org/MANUAL.html#pandocs-markdown>
- ⇒ 原始的 Markdown 主要是输出 HTML 格式文档
- ⇒ Pandoc's Markdown 可以被 Pandoc 转化输出多种格式，如 HTML，PDF，EPUB，DOCX 等等

什么是 R Markdown

- ⇒ R Markdown 站在 Pandoc's Markdown 的肩膀上，手里攥着 Knitr，集成 C/C++，R，Python，Bash 等编程语言。
- ⇒ bookdown (Xie, 2016) 作为 R Markdown 的扩展，负责组织多个 Rmd 文件。
- ⇒ knitr 调各路编程语言的引擎跑源文档中的代码，然后将 Rmd 文档转化为 md，Pandoc 再根据 _output.yml 输出各种格式的文档。



举例子 LaTeX or Markdown

```
\documentclass{article}  
\begin{document}  
\section{section}  
\subsection{subsection}  
Hello, R Markdown.  
\end{document}
```

```
# section  
## subsection  
Hello, R Markdown.
```

```
pandoc demo.tex -o demo.pdf  
pandoc demo.md -o demo.pdf
```

更多例子请看 <https://bookdown.org/yihui/bookdown/>

回顾一下

- Pandoc: 文档转化器
- knitr: 哪里有代码, 点哪里
- bookdown: 分章节组织文档
- tinytex: 自动安装缺失的 TeX 包
- TinyTeX: 一个轻量级的发行版

万事俱备, 模板要怎么做

- 举例子划重点

<https://github.com/XiangyunHuang/master-thesis-template>

基于 R Markdown 的论文模板结构

原始的纯 LaTeX 论文模板根据 bookdown 的组织方式拆分为

1. Pandoc's LaTeX 模板
2. preamble.tex 论文样式设置
3. before_body.tex 封面、授权页、摘要等
4. after_body.tex 致谢、个人简介等
5. 主体部分就是 R Markdown 文件

总之一把刷

```
docker run --name rocker -d -p 8787:8787 -e ROOT=TRUE \  
-e USER=rstudio -e PASSWORD=rstudio rocker/geospatial
```

docker

安装依赖

```
sudo yum install -y yum-utils \  
    device-mapper-persistent-data lvm2
```

repo

```
sudo yum-config-manager --add-repo \  
    https://download.docker.com/linux/centos/docker-ce.repo
```

镜像站点

```
sudo sed -i 's+download.docker.com+mirrors.tuna.tsinghua.edu.cn/d  
    /etc/yum.repos.d/docker-ce.repo
```

安装docker

```
sudo yum install docker-ce
```

启动

```
sudo service docker start
```



docker

```
sudo docker run -d -p 8787:8787 -e ROOT=TRUE -e USER=xiangyun -e
文件(F) 编辑(E) 查看(V) 搜索(S) 终端(T) 帮助(H)
Cloud2016@xiangyun:/home/Cloud2016 $ ps aux | grep docker
Cloud20+ 28315 0.0 0.0 112676 996 pts/1 S+ 10:55 0:00 grep --color=auto --exclude-dir=.bzzr --exclude-dir
=CVS --exclude-dir=.git --exclude-dir=.hg --exclude-dir=.svn docker
Cloud2016@xiangyun:/home/Cloud2016 $ service docker stop
Redirecting to /bin/systemctl stop docker.service
Cloud2016@xiangyun:/home/Cloud2016 $ ps aux | grep docker
Cloud20+ 29545 0.0 0.0 112676 996 pts/1 S+ 10:56 0:00 grep --color=auto --exclude-dir=.bzzr --exclude-dir
=CVS --exclude-dir=.git --exclude-dir=.hg --exclude-dir=.svn docker
Cloud2016@xiangyun:/home/Cloud2016 $ sudo service docker start
[sudo] Cloud2016 的密码:
Redirecting to /bin/systemctl start docker.service
Cloud2016@xiangyun:/home/Cloud2016 $ ps aux | grep docker
root 29906 2.4 0.3 602888 25320 ? Ssl 10:56 0:00 /usr/bin/dockerd
root 29926 0.0 0.0 496664 7684 ? Ssl 10:56 0:00 docker-containerd -l unix:///var/run/docker/libcon
tainerd --shim docker-containerd-shim --runtime docker-runc
Cloud20+ 30113 0.0 0.0 112676 996 pts/1 S+ 10:56 0:00 grep --color=auto --exclude-dir=.bzzr --exclude-dir
=CVS --exclude-dir=.git --exclude-dir=.hg --exclude-dir=.svn docker
Cloud2016@xiangyun:/home/Cloud2016 $ docker-machine ls
zsh: docker-machine: 未找到命令...
zsh: command not found: docker-machine
Cloud2016@xiangyun:/home/Cloud2016 $ sudo docker run -d -p 8787:8787 -e ROOT=TRUE \
> -e USER=xiangyun -e PASSWORD=cloud \
> rocker/geospatial
Unable to find image 'rocker/geospatial:latest' locally
latest: Pulling from rocker/geospatial
3e17c6eae66c: Pull complete
ef4a96c34b16: Downloading [=====>] 58.89MB/185.4MB
d191e567e27c: Downloading [=====>] 60.54MB/119.8MB
cecb308e9eeb: Download complete
3dd31f307f4c: Download complete
4ec0353599ca: Downloading [=====>] 23.24MB/339.8MB
20e027fb82e6: Waiting
```

图 2: 安装 docker



Pandoc 提供的 LaTeX 模板

```
\PassOptionsToPackage{unicode=true}{hyperref} % options for packa
\PassOptionsToPackage{hyphens}{url}
$if(colorlinks)$
\PassOptionsToPackage{dvipsnames,svgnames*,xllnames*}{xcolor}
$endif$
\documentclass[$if(fontsize)$$fontsize$, $endif$$if(lang)$$babel-l
中间省略
$for(include-after)$
$include-after$
$endfor$
\end{document}
```

作用类似 LaTeX 的 class 文件 (*.cls, *.sty)

_output.yml

```
bookdown: pdf_book:
  includes:
    in_header: latex/preamble.tex
    before_body: latex/before_body.tex
    after_body: latex/after_body.tex
  keep_tex: yes
  dev: "cairo_pdf"
  latex_engine: xelatex
  citation_package: natbib
  pandoc_args: "--top-level-division=chapter"
  toc_depth: 4
  toc_unnumbered: yes
  toc_appendix: no
  toc_bib: yes
  template: pandoc-templates/default.latex
  quote_footer: ["\\begin{flushright}", "\\end{flushright}"]
```



index.Rmd

```
title: "空间广义线性混合效应模型及其应用"  
subtitle: "理学院2015级硕士学位论文"  
institute: "中国矿业大学（北京）"  
author: "黄湘云"  
date: "2018-04-27"  
site: bookdown::bookdown_site  
documentclass: ctexbook  
mainfont: Nimbus Roman No9 L  
geometry: [tmargin=3.0cm, bmargin=3.0cm, lmargin=3.0cm, rmargin=3.0cm]  
classoption: "hyperref, a4paper, UTF8, zihao = -4, linespread = 1"  
bibliography: [refer.bib, packages.bib]  
natbiboptions: super, square  
biblio-style: natbib  
link-citations: yes  
colorlinks: yes
```

R Markdown 的其他应用

- 可重复的数据分析报告
- 搭建个人博客 <https://github.com/rbind>
- 写书 <https://bookdown.org/>

数据分析报告

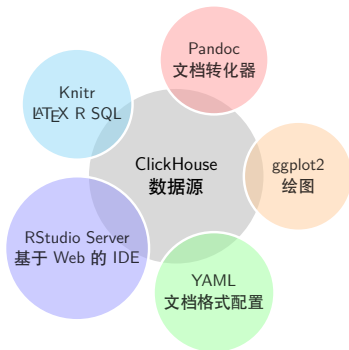
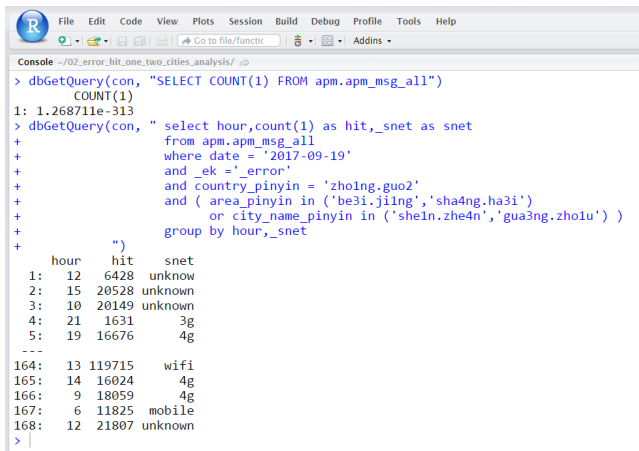


图 3: 日志数据分析报告

数据分析报告



The screenshot shows an R console window with the following content:

```
Console ~/02_error_hit_one_two_cities_analysis/

> dbGetQuery(con, "SELECT COUNT(1) FROM apm.apm_msg_all")
COUNT(1)
1: 1.268711e-313
> dbGetQuery(con, " select hour,count(1) as hit,_snet as snet
+                      from apm.apm_msg_all
+                      where date = '2017-09-19'
+                      and _ek = 'error'
+                      and country_pinyin = 'zho1ng.guo2'
+                      and ( area_pinyin in ('be3i.ji1ng','sha4ng.ha3i')
+                        or city_name_pinyin in ('she1n.zhe4n','gua3ng.zho1u') )
+                      group by hour,_snet
+                      ")
  hour    hit    snet
1:   12   6428 unknown
2:   15  20528 unknown
3:   10  20149 unknown
4:   21   1631      3g
5:   19  16676      4g
---
164:  13 119715     wifi
165:  14  16024      4g
166:   9  18059      4g
167:   6  11825  mobile
168:  12  21807 unknown
>
```

图 4: ClickHouse in R

Q & A

✉ 邮件: xiangyunfaith@outlook.com²

✉ 博客: <https://xiangyun.netlify.com/>

参考文献 I

Xie, Y. (2016). *bookdown: Authoring Books and Technical Documents with R Markdown*. Chapman and Hall/CRC, Boca Raton, Florida. ISBN 978-1138700109.