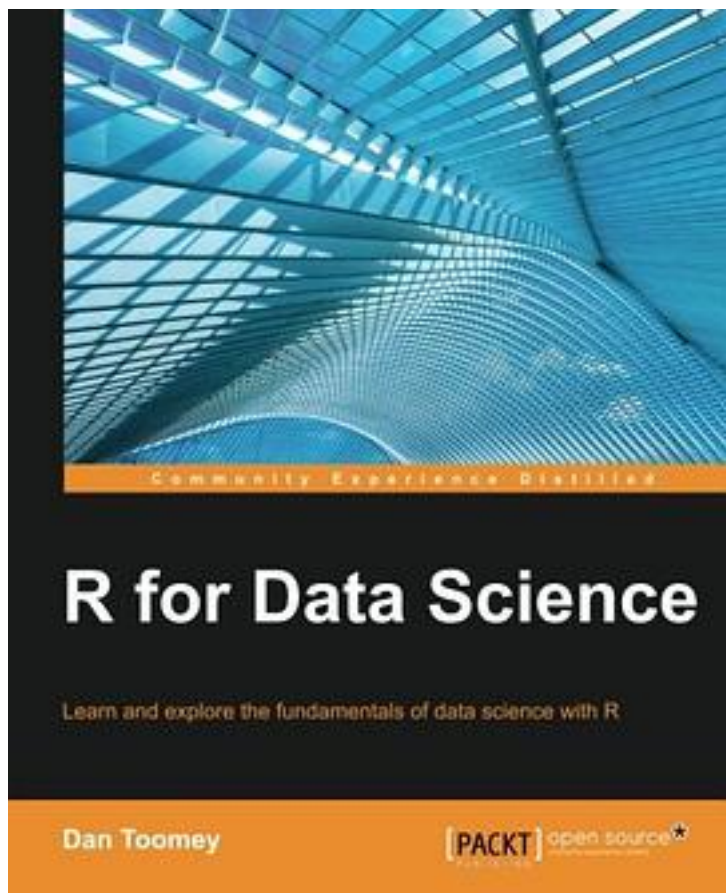


R for Data Science



[R for Data Science_ 下载链接1](#)

著者:Hadley Wickham

出版者:O'Reilly Media

出版时间:2016-12-25

装帧:Paperback

isbn:9781491910399

<http://r4ds.had.co.nz/>

作者介绍:

Hadley Wickham is an Assistant Professor and the Dobelman Family Junior Chair in Statistics at Rice University. He is an active member of the R community, has written and contributed to over 30 R packages, and won the John Chambers Award for Statistical Computing for his work developing tools for data reshaping and visualization. His research focuses on how to make data analysis better, faster and easier, with a particular emphasis on the use of visualization to better understand data and models.

Garrett Golemund is a statistician, teacher and R developer who currently works for RStudio. He sees data analysis as a largely untapped fountain of value for both industry and science. Garrett received his Ph.D at Rice University in Hadley Wickham's lab, where his research traced the origins of data analysis as a cognitive process and identified how attentional and epistemological concerns guide every data analysis.

Garrett is passionate about helping people avoid the frustration and unnecessary learning he went through while mastering data analysis. Even before he finished his dissertation, he started teaching corporate training in R and data analysis for Revolutions Analytics. He's taught at Google, eBay, Axciom and many other companies, and is currently developing a training curriculum for RStudio that will make useful know-how even more accessible.

Outside of teaching, Garrett spends time doing clinical trials research, legal research, and financial analysis. He also develops R software, he's co-authored the lubridate R package--which provides methods to parse, manipulate, and do arithmetic with date-times--and wrote the ggsubplot package, which extends the ggplot2 package.

目录: Chapter 1 Data Visualization with ggplot2

Chapter 2 Workflow: Basics

Chapter 3 Data Transformation with dplyr

Chapter 4 Workflow: Scripts

Chapter 5 Exploratory Data Analysis

Chapter 6 Workflow: Projects

Chapter 7 Tibbles with tibble

Chapter 8 Data Import with readr

Chapter 9 Tidy Data with tidyr

Chapter 10 Relational Data with dplyr

Chapter 11 Strings with stringr

Chapter 12 Factors with forcats

Chapter 13 Dates and Times with lubridate

Chapter 14 Pipes with magrittr

Chapter 15 Functions

Chapter 16 Vectors

Chapter 17 Iteration with purrr

Chapter 18 Model Basics with modelr

Chapter 19 Model Building

Chapter 20 Many Models with purrr and broom

Chapter 21 R Markdown

Chapter 22 Graphics for Communication with ggplot2

Chapter 23 R Markdown Formats

Chapter 24 R Markdown Workflow

• • • • • ([收起](#))

标签

R

数据科学

数据分析

编程

统计

Programming

R语言

统计学

评论

入门书。Tidyverse 的哲学可看作 R 的新（现代）范式，目的就是取其精华、去其糟粕。请 pandas/matplotlib/sklearn 诸君继续努力，因为我被迫用你们时经常怀念 tidyverse 的某些优雅与效率

非常好。之前一直因为R混乱的语法而弃用很多次，但是pandas和jupyter notebook又不够灵活，勉强用python做了一段时间的数据分析。这次偶然接触了tidyverse套件，瞬间觉得R非常清晰好用。另外这本的理念很好，完全摒弃了其他R工具书从头到尾讲统计的陋习。

基本上是hadley在安利他的各种包tidyr, dplyr, purrr, ggplot2, hadley的东西有更高的抽象，用起来还是比较舒服的。只不过base包的东西已经很熟，表示老狗学新把戏的意愿不是很高。本书在手机上看完，囧rz..

Hadley Wickham的数据哲学

花了一年的时间，跟同事一起把这本书以上课时间讲完了。受益匪浅。是一本非常应用的书。入门R的同学可以看，这是tidyverse体系的R。容易上手。也够用。如果是14年之前接触R的，也建议读一下，真的很方便。但是如果对R使用和时间要求比较高，那还是要从传统的R入手，这本书应用性强，理论性太浅。

读得不太认真，有需要再复习吧。R毕竟不是专门做统计的，大部分场合下当然没有sta ta之类直观，纵有万般好，也不必过分推崇。我决定实用主义地穿插着用各种软件，虽然姿势难看，但是省劲儿啊……

辗转于matlab/python+pandas/R后，终于发现了R配合tidyverse配合ggplot2配合RStudio，才是最契合交互式代码+图像快速探索数据，形成数据分析思路的终极利器。

tidyverse太美啦！

遇到Tidyverse大概是这个夏天遇到的最爽的事了

大概因为有点pandas和sklearn的基础，一刷非常顺畅~暑假跟着R for finance类的教材初学R真是不堪回首=。=#选对入门书太重要！

手头常备。

大牛的作品，果然不一般，清楚简明易懂。

刷新对R认识的作品

Hadley
Wickham不仅一人支撑起了R的半壁江山，还精心教育社区成员数据分析的正确姿势。

入门书籍，tidyverse占了大头，基本上也是很主流的包比如ggplot2还有tibble这种，重点是第四部分的函数建模部分，主要是了解lm (x~y) 的写法。至于第五章讲交互的意义就真的很迷了，真的不是拿来凑数的？

我十分十分十分不理解这本书的编排。第3章 ggplot，第5章 dplyr，这都 OK；第4、6、8章讲 R 和 RStudio 基础，这是什么鬼套路？！学完 tidyverse 收工

刷新对R的三观

真的讨厌r

这书可以翻翻，里面的很多package并不常用，像什么tidy很少用的。作为参考书是可以的，但是入门这本书并不好

Hadley出品，必属精品。技术和数据哲学融为一体，R数据科学圣经。

[R for Data Science_下载链接1](#)

书评

这本书的定位是 data science 入门书，特点是使用了 tidyverse 的一套哲学。整体思路可借用书中的一张图来说明：首先明确几点原则：1. 工具不是重点，创造价值才是目的。具体到数据科学，表现形式往往是提供解决方案或者做出某种决策。至于使用什么语言，采用什么工具，不本...

学R是在2012年在Jenny Bryan的stats 540的课上，Jenny也是本书几个章节的作者之一。那时，课上画图用的是lattice，这个包只包括了一些常见的画图函数，有些图没有，经常还得去找其他的包，理解不同函数的输入参数。同一个事情，不同包的控制参数名称完全不一样，经常找不到或...

[R for Data Science_下载链接1](#)