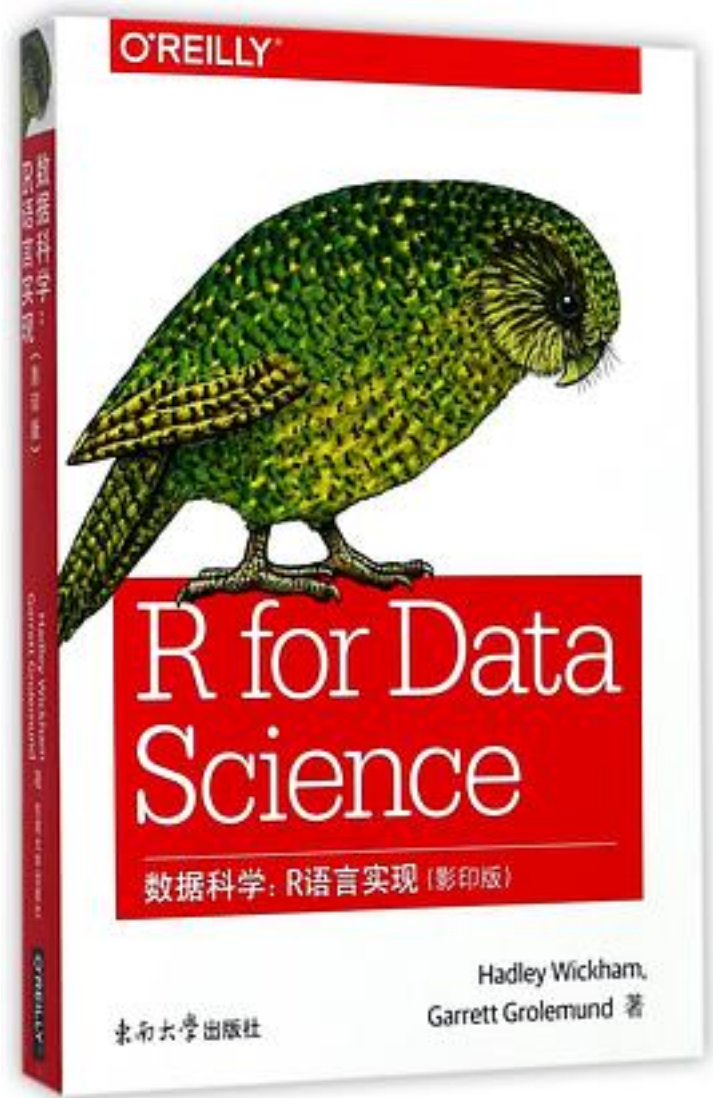


数据科学--R语言实现(影印版)(英文版)



[数据科学--R语言实现\(影印版\)\(英文版\) 下载链接1](#)

著者:Hadley Wickham

出版者:东南大学出版社

出版时间:2017-10-1

装帧:

isbn:9787564173531

学习如何利用R语言洞察、知晓、理解原始数据。

《数据科学：R语言实现（影印版 英文版）》介绍了R、RStudio以及tidyverse，后者是一组相互配合工作的R包，能够使数据科学快速、流畅、富有乐趣。

《数据科学：R语言实现（影印版 英文版）》旨在帮助你尽快地上手数据科学相关的工作，并不要求读者具备编程经验。

《数据科学：R语言实现（影印版 英文版）》Hadley Wickham和Garrett Grolemund将一步步指导你对数据进行导入、提炼、探索以及建模并发布成果。除了处理数据所需的基本工具，你还将会对数据科学的周期拥有一个完整的、宏观的理解。

作者介绍:

Hadley Wickham是RStudio的首席科学家以及R基金会成员。他构建了一套使数据科学变得更加快捷、富有乐趣的工具。可以通过其个人网站了解更多的信息：<http://hadley.nz>。

Garrett Grolemund是一名统计学家、教师以及RStudio的硕士生导师。他还是《Hands-On Programming with R》(O'Reilly)一书的作者。Garrett的很多授课视频可以在oreilly.com/safari上找到。

目录: Preface

Part I. Explore

1. Data Visualization with ggplot2

Introduction

First Steps

Aesthetic Mappings

Common Problems

Facets

Geometric Objects

Statistical Transformations

Position Adjustments

Coordinate Systems

The Layered Grammar of Graphics

2. Workflow: Basics

Coding Basics

What's in a Name?

Calling Functions

3. Data Transformation with dplyr

Introduction

Filter Rows with filter ()

Arrange Rows with arrange ()

Select Columns with select ()

Add New Variables with mutate ()

Grouped Summaries with summarize ()

Grouped Mutates (and Filters)

4. Workflow: Scripts

Running Code
RStudio Diagnostics
5. Exploratory Data Analysis
Introduction
Questions
Variation
Missing Values
Covariation
Patterns and Models
ggplot2 Calls
Learning More
6. Workflow: Projects
What Is Real?
Where Does Your Analysis Live?
Paths and Directories
RStudio Projects
Summary
Part II. Wrangle
7. Tibbles with tibble
Introduction
Creating Tibbles
Tibbles Versus data.frame
Interacting with Older Code
8. Data Import with readr
Introduction
Getting Started
Parsing a Vector
Parsing a File
Writing to a File
Other Types of Data
9. Tidy Data with tidyr
Introduction
Tidy Data
Spreading and Gathering
Separating and Pull
Missing Values
Case Study
Nontidy Data
10. Relational Data with dplyr
Introduction
nycflights13
Keys
Mutating Joins
Filtering Joins
Join Problems
Set Operations
11. Strings with stringr
Introduction
String Basics
Matching Patterns with Regular Expressions
Tools
Other Types of Pattern
Other Uses of Regular Expressions
stringi

12. Factors with forcats

Introduction

Creating Factors

General Social Survey

Modifying Factor Order

Modifying Factor Levels

13. Dates and Times with lubridate

Introduction

Creating Date/Times

Date-Time Components

Time Spans

Time Zones

Part III. Program

14. Pipes with magrittr

Introduction

Piping Alternatives

When Not to Use the Pipe

Other Tools from magrittr

15. Functions

Introduction

When Should You Write a Function?

Functions Are for Humans and Computers

Conditional Execution

Function Arguments

Return Values

Environment

16. Vectors

Introduction

Vector Basics

Important Types of Atomic Vector

Using Atomic Vectors

Recursive Vectors (Lists)

Attributes

Augmented Vectors

17. Iteration with purrr

Introduction

For Loops

For Loop Variations

For Loops Versus Functionals

The Map Functions

Dealing with Failure

Mapping over Multiple Arguments

Walk

Other Patterns of For Loops

Part IV. Model

18. Model Basics with modelr

Introduction

A Simple Model

Visualizing Models

Formulas and Model Families

Missing Values

Other Model Families

19. Model Building

Introduction

Why Are Low-Quality Diamonds More Expensive?

What Affects the Number of Daily Flights?

Learning More About Models

20. Many Models with purrr and broom

Introduction

gapminder

List-Columns

Creating List-Columns

Simplifying List-Columns

Making Tidy Data with broom

Part V. Communicate

21. R Markdown

Introduction

R Markdown Basics

Text Formatting with Markdown

Code Chunks

Troubleshooting

YAML Header

Learning More

22. Graphics for Communication with ggplot2

Introduction

Label

Annotations

Scales

Zooming

Themes

Saving Your Plots

Learning More

23. R Markdown Formats

Introduction

Output Options

Documents

Notebooks

Presentations

Dashboards

Interactivity

Websites

Other Formats

Learning More

24. R Markdown Workflow

Index

• • • • • ([收起](#))

[数据科学--R语言实现\(影印版\)\(英文版\) 下载链接1](#)

标签

R

研究

评论

[数据科学--R语言实现\(影印版\)\(英文版\) 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...

有人说这本书就是Hadley安利他的Tidyverse各种包，书名不如就叫《Tidyverse for Data Science》，没办法，谁让Tidyverse就是优雅好用呢。举个例子，我看完以后就记得一堆动词性函数（Verbs），比如用filter()筛选符合条件的行（observations），用select()筛选需要的列（variab...

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...

你们打五星的人，是水军吗。书的英文版的确很棒。中文版删了那么多东西你们不知道

吗??? 长表和短表那一章全被删除, 因子那一章也被删减。简直恶心到家了。图灵社区怎么翻译质量这么烂 书本身英文版值4星, 除了介绍tidyverse之外对base R和其他高效率的包如data.table介绍基本...

[数据科学--R语言实现\(影印版\)\(英文版\)_下载链接1](#)