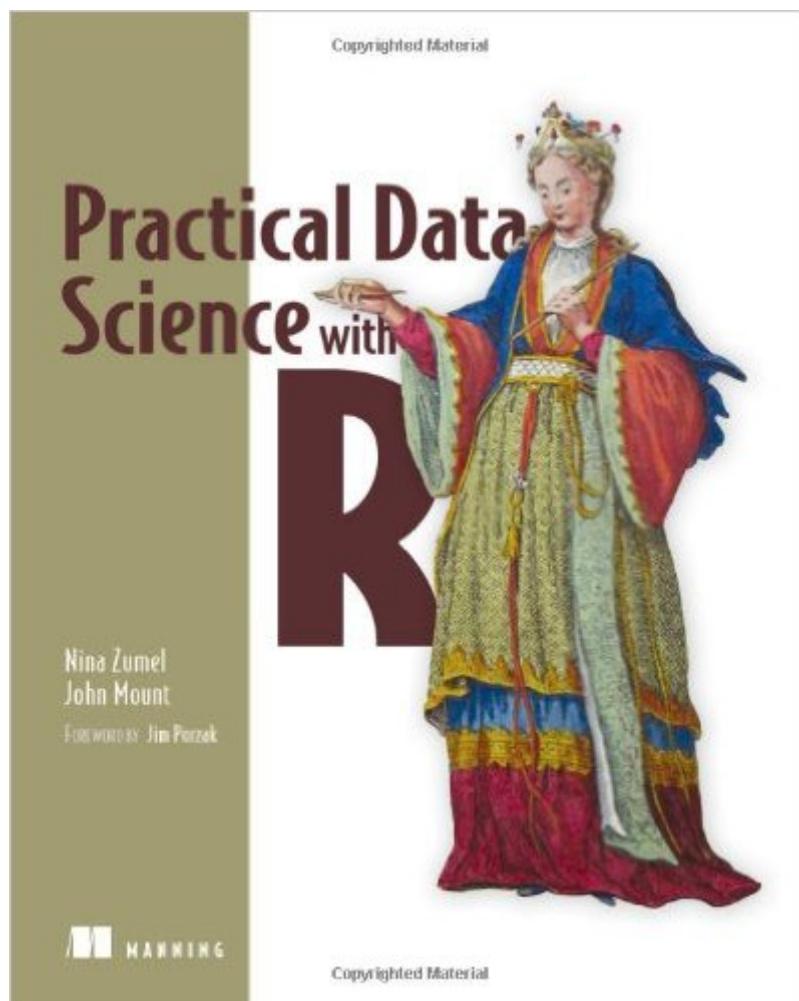


The book was found

# Practical Data Science With R



## Synopsis

Summary Practical Data Science with R lives up to its name. It explains basic principles without the theoretical mumbo-jumbo and jumps right to the real use cases you'll face as you collect, curate, and analyze the data crucial to the success of your business. You'll apply the R programming language and statistical analysis techniques to carefully explained examples based in marketing, business intelligence, and decision support. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications.

**About the Book**

Business analysts and developers are increasingly collecting, curating, analyzing, and reporting on crucial business data. The R language and its associated tools provide a straightforward way to tackle day-to-day data science tasks without a lot of academic theory or advanced mathematics. Practical Data Science with R shows you how to apply the R programming language and useful statistical techniques to everyday business situations. Using examples from marketing, business intelligence, and decision support, it shows you how to design experiments (such as A/B tests), build predictive models, and present results to audiences of all levels. This book is accessible to readers without a background in data science. Some familiarity with basic statistics, R, or another scripting language is assumed.

What's Inside

- Data science for the business professional
- Statistical analysis using the R language
- Project lifecycle, from planning to delivery
- Numerous instantly familiar use cases
- Keys to effective data presentations

**About the Authors**

Nina Zumel and John Mount are cofounders of a San Francisco-based data science consulting firm. Both hold PhDs from Carnegie Mellon and blog on statistics, probability, and computer science at [win-vector.com](http://win-vector.com).

**Table of Contents**

**PART 1**

**INTRODUCTION TO DATA SCIENCE**

- The data science process
- Loading data into R
- Exploring data
- Managing data

**PART 2**

**MODELING METHODS**

- Choosing and evaluating models
- Memorization methods
- Linear and logistic regression
- Unsupervised methods
- Exploring advanced methods

**PART 3**

**DELIVERING RESULTS**

- Documentation and deployment
- Producing effective presentations

## Book Information

Paperback: 389 pages

Publisher: Manning; 1st edition (April 13, 2014)

Language: English

ISBN-10: 1617291560

ISBN-13: 978-1617291562

Product Dimensions: 7.3 x 1 x 9.1 inches

Shipping Weight: 1.6 pounds ([View shipping rates and policies](#))

Average Customer Review: 4.3 out of 5 stars [See all reviews](#) (31 customer reviews)

Best Sellers Rank: #83,509 in Books (See Top 100 in Books) #17 in Books > Textbooks > Computer Science > Artificial Intelligence #20 in Books > Computers & Technology > Programming > Software Design, Testing & Engineering > Tools #37 in Books > Computers & Technology > Programming > Software Design, Testing & Engineering > Testing

## Customer Reviews

tl;dr: A well rounded, occasionally high-level introductory text that will leave you feeling prepared to participate in the Data Science conversation at work, from earliest planning to presentation and maintenance.

Details: Was excited to see this book coming to publication. I'm a fan of practical, non-academic approaches to subjects and prefer working from concrete examples to abstract principles (rather than the other way around). I think this is both the most difficult and most needed type of resources that can be put into print. This book handles the task ok; it falls a bit short on practical, concrete, use cases as it alternates between working with hands on datasets and shotgun coverage of principles and techniques at a higher level. I'd have much preferred sticking with single data-sets for longer (say, a couple chapters per data set), but didn't feel cheated out of hands on work.

Pros:- Easy access to the datasets via Github; good documentation on where to find others- Key Takeaways provided at end of chapter are good summaries of overall information provided.- A good focus on not just data analysis, but the process as a whole; very Agile like, practical, and non-dogmatic.- Battle tested advice: You can tell some of the advice comes from hard-fought battles - ex: Why not use the sample() function instead of manually creating a sample column? Because with a sample column, you can repeatably sample the same data (e.g. all columns < 2) for repeatable output and for regression testing (avoiding introducing bugs).- Builds your analyst vocabulary, increasing your all-important google-fu skills. Not knowing what to Google is, imho, the single hardest problem when learning a new set of problems / api's.

I've had to hire recent graduates with degrees in machine learning, operations research and even "data science." One of the problems with such people: they don't know anything practical. They probably know the basics of regression and some classification routines, as learned in their coursework. They've probably worked on one or many data science like problems, using machine learning techniques or regression or what not. Many of them have never done a SQL query, or done the dirty business of data cleaning which takes up most of the data scientist's time. They'll always have gaps in their education; maybe they wrote a dissertation on an application of trees or deep

learning, and have never used any of the other myriad tools available to the data scientist. None of them have ever done data science for money, and so none of them know about practical things like git or what the process looks like in an industrial setting. It is for these people that this book appears to be written. In an ideal world, all larval data scientists would be taught a course based on this book, or at least go through it themselves. It is also useful to experienced practitioners, as it covers many things, and can be a good practical reference to keep around. The book is ordered as a data science project would be ordered, from start to finish; so, as you proceed down an engagement, reviewing the chapters in order will be helpful. Ch1 describes the job of the data scientist, the workflow, and the characters you run into on a project. Ch2 outlines some of the tools used to get at the data, including the authors tool, "SQL Screwdriver.

[Download to continue reading...](#)

Data Analytics: Practical Data Analysis and Statistical Guide to Transform and Evolve Any Business. Leveraging the Power of Data Analytics, Data ... (Hacking Freedom and Data Driven) (Volume 2) Data Analytics: What Every Business Must Know About Big Data And Data Science (Data Analytics for Business, Predictive Analysis, Big Data) Analytics: Data Science, Data Analysis and Predictive Analytics for Business (Algorithms, Business Intelligence, Statistical Analysis, Decision Analysis, Business Analytics, Data Mining, Big Data) Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data (Data-Centric Systems and Applications) Practical Data Management for Risk Data Aggregation and BCBS 239 Compliance Second Edition Swift: Programming, Master's Handbook; A TRUE Beginner's Guide! Problem Solving, Code, Data Science, Data Structures & Algorithms (Code like a PRO in ... engineering, r programming, iOS development) Ruby: Programming, Master's Handbook: A TRUE Beginner's Guide! Problem Solving, Code, Data Science, Data Structures & Algorithms (Code like a PRO in ... web design, tech, perl, ajax, swift, python,) Java Programming: Master's Handbook: A TRUE Beginner's Guide! Problem Solving, Code, Data Science, Data Structures & Algorithms (Code like a PRO in ... web design, tech, perl, ajax, swift, python) Php: Programming, Master's Handbook: A TRUE Beginner's Guide! Problem Solving, Code, Data Science, Data Structures & Algorithms (Code like a PRO in ... engineering, r programming, iOS development,) Python: Programming, Master's Handbook; A TRUE Beginner's Guide! Problem Solving, Code, Data Science, Data Structures & Algorithms (Code like a PRO ... engineering, r programming, iOS development) Agile Data Science: Building Data Analytics Applications with Hadoop Analytics: Data Science, Data Analysis and Predictive

Analytics for Business Data Smart: Using Data Science to Transform Information into Insight A collection of Advanced Data Science and Machine Learning Interview Questions Solved in Python and Spark (II): Hands-on Big Data and Machine ... Programming Interview Questions) (Volume 7) Practical Data Science with R Deep Learning Step by Step with Python: A Very Gentle Introduction to Deep Neural Networks for Practical Data Science Big Data and Social Science: A Practical Guide to Methods and Tools (Chapman & Hall/CRC Statistics in the Social and Behavioral Sciences) Mathematical Problems in Data Science: Theoretical and Practical Methods

[Dmca](#)