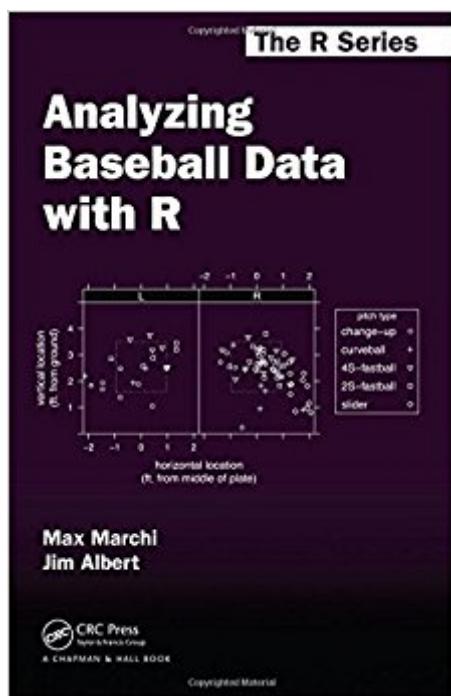


The book was found

# Analyzing Baseball Data With R (Chapman & Hall/CRC The R Series)



## Synopsis

With its flexible capabilities and open-source platform, R has become a major tool for analyzing detailed, high-quality baseball data. *Analyzing Baseball Data with R* provides an introduction to R for sabermetricians, baseball enthusiasts, and students interested in exploring the rich sources of baseball data. It equips readers with the necessary skills and software tools to perform all of the analysis steps, from gathering the datasets and entering them in a convenient format to visualizing the data via graphs to performing a statistical analysis. The authors first present an overview of publicly available baseball datasets and a gentle introduction to the type of data structures and exploratory and data management capabilities of R. They also cover the traditional graphics functions in the base package and introduce more sophisticated graphical displays available through the lattice and ggplot2 packages. Much of the book illustrates the use of R through popular sabermetrics topics, including the Pythagorean formula, runs expectancy, career trajectories, simulation of games and seasons, patterns of streaky behavior of players, and fielding measures. Each chapter contains exercises that encourage readers to perform their own analyses using R. All of the datasets and R code used in the text are available online. This book helps readers answer questions about baseball teams, players, and strategy using large, publically available datasets. It offers detailed instructions on downloading the datasets and putting them into formats that simplify data exploration and analysis. Through the book's various examples, readers will learn about modern sabermetrics and be able to conduct their own baseball analyses.

## Book Information

Series: Chapman & Hall/CRC The R Series (Book 14)

Paperback: 352 pages

Publisher: Chapman and Hall/CRC; 1 edition (October 31, 2013)

Language: English

ISBN-10: 1466570229

ISBN-13: 978-1466570221

Product Dimensions: 1 x 6 x 9 inches

Shipping Weight: 1 pounds (View shipping rates and policies)

Average Customer Review: 4.4 out of 5 stars 29 customer reviews

Best Sellers Rank: #84,210 in Books (See Top 100 in Books) #22 in Books > Sports & Outdoors > Coaching > Baseball #412 in Books > Textbooks > Science & Mathematics > Mathematics > Statistics #547 in Books > Science & Math > Mathematics > Applied > Probability & Statistics

## Customer Reviews

"There are some great resources out there for learning R and for learning how to analyze baseball data with it. In fact, a few pretty smart people wrote a fantastic book on the subject, coincidentally titled *Analyzing Baseball Data with R*. I can't say enough about this book as a reference, both for baseball analysis and for R. Go and buy it." •Bill Petti, *The Hardball Times*, September 2015 "The authors present a potpourri of well-conceived case-studies that give insight into both the game's complexity and R's simplicity. Virtually no previous knowledge of statistical theory and software is required to master the data analyses and to follow the explications in this book. The authors' style of writing is pleasurable and bespeaks their passion for the game. Narratives and R commands are so smoothly intermingled that the source code hardly disturbs the flow of reading, and a wealth of graphs break up the grey. A great asset of the book is that it encourages the reader to learn the ropes of sabermetrics by actually running the example analyses on one's own computer." •Journal of the Royal Statistical Society, Series A, 2015 "If you are interested in statistics, especially baseball statistics, you will find this book fascinating and very useful. It provides many details, websites, and useful descriptions for using the R programming environment. This is not only a book on statistics; there are many references to famous player statistics, making this a very enjoyable book to read. And even if you don't like baseball but still find statistics very exciting, then this book provides a great introduction to R that can be used for any other type of statistical data set." •IEEE Insulation Magazine, November/December 2014 "I have spent most of the past decade working in baseball as a statistical analyst for the New York Mets. This type of employment can be highly valued, especially among quantitatively inclined college students who are coincidentally passionate baseball fans. It is from these students from whom I am most frequently asked, 'what book would you recommend for someone who wants to get started in sabermetrics?' Invariably, my response has been [Jim Albert and Jay Bennett's] *Curve Ball*. I have a new response. I always felt that *Curve Ball* was the best place for a budding sabermetrician to start. However, it later dawned on me that while *Curve Ball* provided a sound framework for thinking probabilistically about baseball, I devoted a huge proportion of my time at work to computer programming. In their new book, Albert and Max Marchi, a native Italian who now works for the Cleveland Indians, have closed the loop by offering the aspiring sabermetrician a blueprint. The reader who digests this book alongside her keyboard will emerge as a practicing sabermetrician •having knowledge of the key ideas in sabermetric theory, a historical understanding of from whence those ideas came, and the practical ability to compute with baseball data. It is a sabermetric workshop in paperback." •Ben S. Baumer, *International Statistical Review*

Max Marchi is a baseball analyst with the Cleveland Indians. He was previously a statistician at the Emilia-Romagna Regional Health Agency. He has been a regular contributor to The Hardball Times and Baseball Prospectus websites and has consulted for MLB clubs. Jim Albert is a professor of statistics at Bowling Green State University. He has authored or coauthored several books and is the editor of the Journal of Quantitative Analysis of Sports. His interests include Bayesian modeling, statistics education, and the application of statistical thinking in sports.

Great book though it suffers from typos in a few areas and the issue seems to get worse the farther into the text one gets. As an example, in Chapter 8 naming of objects isn't consistent in between text passages as well as in across the book and downloadable scripts. That said, I suppose one could say that is the nature of data analysis. A good amount of analytical work and even more debugging/data correction. I still consider this a great resource though it could use some polish.

I now know how to write rudimentary code to analyze baseball data. I tried to use other books on R, and translate them to baseball; however, they were convoluted. This is the Holy Grail for anyone who wants to learn R and how it relates to baseball.

Probably unlike most people who will buy this book, I am more well versed in R than I am Sabermetrics. I bought this book to teach me more about baseball statistics, and I figured it would be worth it considering Jim Albert's involvement. I have taken formal classes in R in graduate school, and let me tell you this book was a dynamite review and I think even better at teaching basic coding and packages than some of the books solely dedicated to R out there. It also directed me to some databases I was unfamiliar with and where to find specific datasets, which is amazing. This is a must buy for those looking to conduct statistical analyses on their favorite team while utilizing free software. Looking forward to next season.

This book is excellent and I recommend it highly if you're looking to (a) improve your fluency with R and/or (b) learn how to work with the standard baseball datasets (Retrosheet, Lahman, etc.). I'm only halfway through and already my R skills have improved greatly.

Great book

I now know how to write rudimentary code to analyze baseball data. I tried to use other books on R, and translate them to baseball; however, they were convoluted. This is the Holy Grail for anyone who wants to learn R and how it relates to baseball.

Tons of very valuable info applied to baseball research. Being new to R I found it easy to get started and tinkering with the projects in the book, however, I would have liked a more detailed intro to some of the language... Not really a big deal due to pretty much everything being available online.

After opening the book and reading 5 pages I already learned enough to get started and hit the ground running. Thanks

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

Analyzing Baseball Data with R (Chapman & Hall/CRC The R Series) Data Classification: Algorithms and Applications (Chapman & Hall/CRC Data Mining and Knowledge Discovery Series) Statistics and Data Analysis for Microarrays Using R and Bioconductor, Second Edition (Chapman & Hall/CRC Mathematical and Computational Biology) Computational Statistics Handbook with MATLAB, Third Edition (Chapman & Hall/CRC Computer Science & Data Analysis) RNA-seq Data Analysis: A Practical Approach (Chapman & Hall/CRC Mathematical and Computational Biology) Modern Data Science with R (Chapman & Hall/CRC Texts in Statistical Science) Modelling Survival Data in Medical Research, Third Edition (Chapman & Hall/CRC Texts in Statistical Science) Baseball Defense Mastery: Fundamentals, Concepts & Drills For Defensive Prowess (Baseball Defense, Baseball Book, Baseball Coaching, Baseball Drills, Outfield, Infield) Baseball Hitting Mastery: Art of the Line Drive Swing (Baseball Book, Baseball Hitting Mechanics, Baseball Hitting Drills, Baseball Swing) Big Data For Business: Your Comprehensive Guide to Understand Data Science, Data Analytics and Data Mining to Boost More Growth and Improve Business - Data Analytics Book, Series 2 Data Analytics: What Every Business Must Know About Big Data And Data Science (Data Analytics for Business, Predictive Analysis, Big Data Book 1) Data Analytics: Applicable Data Analysis to Advance Any Business Using the Power of Data Driven Analytics (Big Data Analytics, Data Science, Business Intelligence Book 6) Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data Introduction to Scientific Programming and Simulation Using R (Chapman & Hall/CRC The R Series) Statistical Computing with R (Chapman & Hall/CRC The R Series) Variational Methods in Image Processing (Chapman & Hall/CRC Mathematical and Computational Imaging Sciences Series) Introduction to Stochastic Processes

(Chapman & Hall/CRC Probability Series) Introduction to Scientific Programming and Simulation Using R, Second Edition (Chapman & Hall/CRC The R Series) Using R for Numerical Analysis in Science and Engineering (Chapman & Hall/CRC The R Series) Introduction to Modern Cryptography, Second Edition (Chapman & Hall/CRC Cryptography and Network Security Series)

[Contact Us](#)

[DMCA](#)

[Privacy](#)

[FAQ & Help](#)