

Ggplot2 Elegant Graphics For Data Analysis

Master the syntax for working with R's plotting functions in graphics and stats in this easy reference to formatting plots. The approach in Visualizing Data in R 4 toward the application of formatting in ggplot() will follow the structure of the formatting used by the plotting functions in graphics and stats. This book will take advantage of the new features added to R 4 where appropriate including a refreshed color palette for charts, Cairo graphics with more fonts/symbols, and improved performance from grid graphics including ggplot 2 rendering speed. Visualizing Data in R 4 starts with an introduction and then is split into two parts and six appendices. Part I covers the function plot() and the ancillary functions you can use with plot(). You'll also see the functions par() and layout(), providing for multiple plots on a page. Part II goes over the basics of using the functions qplot() and ggplot() in the package ggplot2. The default plots generated by the functions qplot() and ggplot() give more sophisticated-looking plots than the default plots done by plot() and are easier to use, but the function plot() is more flexible. Both plot() and ggplot() allow for many layers to a plot. The six appendices will cover plots for contingency tables, plots for continuous variables, plots for data with a limited number of values, functions that generate multiple plots, plots for time series analysis, and some miscellaneous plots. Some of the functions that will be in the appendices include functions that generate histograms, bar charts, pie charts, box plots, and heatmaps. What You Will Learn Use R to create informative graphics Master plot(), qplot(), and ggplot() Discover the canned graphics functions in stats and graphics Format plots generated by plot() and ggplot() Who This Book Is For Those in data science who use R. Some prior experience with R or data science is recommended.

Geocomputation with R is for people who want to analyze, visualize and model geographic data with open source software. It is based on R, a statistical programming language that has powerful data processing, visualization, and geospatial capabilities. The book equips you with the knowledge and skills to tackle a wide range of issues manifested in geographic data, including those with scientific, societal, and environmental implications. This book will interest people from many backgrounds, especially Geographic Information Systems (GIS) users interested in applying their domain-specific knowledge in a powerful open source language for data science, and R users interested in extending their skills to handle spatial data. The book is divided into three parts: (I) Foundations, aimed at getting you up-to-speed with geographic data in R, (II) extensions, which covers advanced techniques, and (III) applications to real-world problems. The chapters cover progressively more advanced topics, with early chapters providing strong foundations on which the later chapters build. Part I describes the nature of spatial datasets in R and methods for manipulating them. It also covers geographic data import/export and transforming coordinate reference systems. Part II represents methods that build on these foundations. It covers advanced

map making (including web mapping), "bridges" to GIS, sharing reproducible code, and how to do cross-validation in the presence of spatial autocorrelation. Part III applies the knowledge gained to tackle real-world problems, including representing and modeling transport systems, finding optimal locations for stores or services, and ecological modeling. Exercises at the end of each chapter give you the skills needed to tackle a range of geospatial problems. Solutions for each chapter and supplementary materials providing extended examples are available at <https://geocompr.github.io/geocompkg/articles/>. Dr. Robin Lovelace is a University Academic Fellow at the University of Leeds, where he has taught R for geographic research over many years, with a focus on transport systems. Dr. Jakub Nowosad is an Assistant Professor in the Department of Geoinformation at the Adam Mickiewicz University in Poznan, where his focus is on the analysis of large datasets to understand environmental processes. Dr. Jannes Muenchow is a Postdoctoral Researcher in the GIScience Department at the University of Jena, where he develops and teaches a range of geographic methods, with a focus on ecological modeling, statistical geocomputing, and predictive mapping. All three are active developers and work on a number of R packages, including `stplanr`, `sabre`, and `RQGIS`.

Discover best practices for data analysis and software development in R and start on the path to becoming a fully-fledged data scientist. This book teaches you techniques for both data manipulation and visualization and shows you the best way for developing new software packages for R. *Beginning Data Science in R* details how data science is a combination of statistics, computational science, and machine learning. You'll see how to efficiently structure and mine data to extract useful patterns and build mathematical models. This requires computational methods and programming, and R is an ideal programming language for this. This book is based on a number of lecture notes for classes the author has taught on data science and statistical programming using the R programming language. Modern data analysis requires computational skills and usually a minimum of programming. *What You Will Learn* Perform data science and analytics using statistics and the R programming language Visualize and explore data, including working with large data sets found in big data Build an R package Test and check your code Practice version control Profile and optimize your code Who This Book Is For Those with some data science or analytics background, but not necessarily experience with the R programming language. *Applied Spatial Data Analysis with R*, second edition, is divided into two basic parts, the first presenting R packages, functions, classes and methods for handling spatial data. This part is of interest to users who need to access and visualise spatial data. Data import and export for many file formats for spatial data are covered in detail, as is the interface between R and the open source GRASS GIS and the handling of spatio-temporal data. The second part showcases more specialised kinds of spatial data analysis, including spatial point pattern analysis, interpolation and geostatistics, areal data analysis and disease

mapping. The coverage of methods of spatial data analysis ranges from standard techniques to new developments, and the examples used are largely taken from the spatial statistics literature. All the examples can be run using R contributed packages available from the CRAN website, with code and additional data sets from the book's own website. Compared to the first edition, the second edition covers the more systematic approach towards handling spatial data in R, as well as a number of important and widely used CRAN packages that have appeared since the first edition. This book will be of interest to researchers who intend to use R to handle, visualise, and analyse spatial data. It will also be of interest to spatial data analysts who do not use R, but who are interested in practical aspects of implementing software for spatial data analysis. It is a suitable companion book for introductory spatial statistics courses and for applied methods courses in a wide range of subjects using spatial data, including human and physical geography, geographical information science and geoinformatics, the environmental sciences, ecology, public health and disease control, economics, public administration and political science. The book has a website where complete code examples, data sets, and other support material may be found: <http://www.asdar-book.org>. The authors have taken part in writing and maintaining software for spatial data handling and analysis with R in concert since 2003.

See How Graphics Reveal Information Graphical Data Analysis with R shows you what information you can gain from graphical displays. The book focuses on why you draw graphics to display data and which graphics to draw (and uses R to do so). All the datasets are available in R or one of its packages and the R code is available at rosuda.org/GDA. Graphical data analysis is useful for data cleaning, exploring data structure, detecting outliers and unusual groups, identifying trends and clusters, spotting local patterns, evaluating modelling output, and presenting results. This book guides you in choosing graphics and understanding what information you can glean from them. It can be used as a primary text in a graphical data analysis course or as a supplement in a statistics course. Colour graphics are used throughout.

GGPlot2 is a powerful and a flexible R package for producing elegant graphics piece by piece. It has become a popular package for data visualization. This book presents the essentials of ggplot2 to easily create beautiful graphics in R. Key features: - Covers the most important graphic functions- Short, self-contained chapters with practical examples.

Turn your R code into packages that others can easily download and use. This practical book shows you how to bundle reusable R functions, sample data, and documentation together by applying author Hadley Wickham's package development philosophy. In the process, you'll work with devtools, roxygen, and testthat, a set of R packages that automate common development tasks.

Devtools encapsulates best practices that Hadley has learned from years of working with this programming language. Ideal for developers, data scientists,

and programmers with various backgrounds, this book starts you with the basics and shows you how to improve your package writing over time. You'll learn to focus on what you want your package to do, rather than think about package structure. Learn about the most useful components of an R package, including vignettes and unit tests Automate anything you can, taking advantage of the years of development experience embodied in devtools Get tips on good style, such as organizing functions into files Streamline your development process with devtools Learn the best way to submit your package to the Comprehensive R Archive Network (CRAN) Learn from a well-respected member of the R community who created 30 R packages, including ggplot2, dplyr, and tidyr "Advanced Engineering Mathematics" is written for the students of all engineering disciplines. Topics such as Partial Differentiation, Differential Equations, Complex Numbers, Statistics, Probability, Fuzzy Sets and Linear Programming which are an important part of all major universities have been well-explained. Filled with examples and in-text exercises, the book successfully helps the student to practice and retain the understanding of otherwise difficult concepts.

Written by the author of the lattice system, this book describes lattice in considerable depth, beginning with the essentials and systematically delving into specific low levels details as necessary. No prior experience with lattice is required to read the book, although basic familiarity with R is assumed. The book contains close to 150 figures produced with lattice. Many of the examples emphasize principles of good graphical design; almost all use real data sets that are publicly available in various R packages. All code and figures in the book are also available online, along with supplementary material covering more advanced topics.

Much of the data available today is unstructured and text-heavy, making it challenging for analysts to apply their usual data wrangling and visualization tools. With this practical book, you'll explore text-mining techniques with tidytext, a package that authors Julia Silge and David Robinson developed using the tidy principles behind R packages like ggraph and dplyr. You'll learn how tidytext and other tidy tools in R can make text analysis easier and more effective. The authors demonstrate how treating text as data frames enables you to manipulate, summarize, and visualize characteristics of text. You'll also learn how to integrate natural language processing (NLP) into effective workflows. Practical code examples and data explorations will help you generate real insights from literature, news, and social media. Learn how to apply the tidy text format to NLP Use sentiment analysis to mine the emotional content of text Identify a document's most important terms with frequency measurements Explore relationships and connections between words with the ggraph and widyr packages Convert back and forth between R's tidy and non-tidy text formats Use topic modeling to classify document collections into natural groups Examine case studies that compare Twitter archives, dig into NASA metadata, and analyze thousands of Usenet messages This third edition of Paul Murrell's classic book on using R for graphics represents a major update, with a complete overhaul in focus and scope. It focuses primarily on the two core graphics packages in R - graphics and grid - and has a new section on integrating graphics. This section includes three new chapters: importing external images in to R; integrating the graphics and grid systems; and advanced SVG graphics.

The emphasis in this third edition is on having the ability to produce detailed and customised graphics in a wide variety of formats, on being able to share and reuse those graphics, and on being able to integrate graphics from multiple systems. This book is aimed at all levels of R users. For people who are new to R, this book provides an overview of the graphics facilities, which is useful for understanding what to expect from R's graphics functions and how to modify or add to the output they produce. For intermediate-level R users, this book provides all of the information necessary to perform sophisticated customizations of plots produced in R. For advanced R users, this book contains vital information for producing coherent, reusable, and extensible graphics functions.

A far-reaching course in practical advanced statistics for biologists using R/Bioconductor, data exploration, and simulation.

Advanced R helps you understand how R works at a fundamental level. It is designed for R programmers who want to deepen their understanding of the language, and programmers experienced in other languages who want to understand what makes R different and special. This book will teach you the foundations of R; three fundamental programming paradigms (functional, object-oriented, and metaprogramming); and powerful techniques for debugging and optimising your code. By reading this book, you will learn:

- The difference between an object and its name, and why the distinction is important
- The important vector data structures, how they fit together, and how you can pull them apart using subsetting
- The fine details of functions and environments
- The condition system, which powers messages, warnings, and errors
- The powerful functional programming paradigm, which can replace many for loops
- The three most important OO systems: S3, S4, and R6
- The tidy eval toolkit for metaprogramming, which allows you to manipulate code and control evaluation
- Effective debugging techniques that you can deploy, regardless of how your code is run
- How to find and remove performance bottlenecks

The second edition is a comprehensive update: New foundational chapters: "Names and values," "Control flow," and "Conditions" comprehensive coverage of object oriented programming with chapters on S3, S4, R6, and how to choose between them Much deeper coverage of metaprogramming, including the new tidy evaluation framework use of new package like rlang (<http://rlang.r-lib.org>), which provides a clean interface to low-level operations, and purrr (<http://purrr.tidyverse.org/>) for functional programming Use of color in code chunks and figures

Hadley Wickham is Chief Scientist at RStudio, an Adjunct Professor at Stanford University and the University of Auckland, and a member of the R Foundation. He is the lead developer of the tidyverse, a collection of R packages, including ggplot2 and dplyr, designed to support data science. He is also the author of R for Data Science (with Garrett Golemund), R Packages, and ggplot2: Elegant Graphics for Data Analysis.

Forecasting is required in many situations. Stocking an inventory may require forecasts of demand months in advance. Telecommunication routing requires traffic forecasts a few minutes ahead. Whatever the circumstances or time horizons involved, forecasting is an important aid in effective and efficient planning. This textbook provides a comprehensive introduction to forecasting methods and presents enough information about each method for readers to use them sensibly.

Tell your story and show it with data, using free and easy-to-learn tools on the web. This introductory book teaches you how to design interactive charts and customized

maps for your website, beginning with simple drag-and-drop tools such as Google Sheets, Datawrapper, and Tableau Public. You'll also gradually learn how to edit open source code templates like Chart.js, Highcharts, and Leaflet on GitHub. Hands-On Data Visualization takes you step-by-step through tutorials, real-world examples, and online resources. This practical guide is ideal for students, nonprofit organizations, small business owners, local governments, journalists, academics, and anyone who wants to take data out of spreadsheets and turn it into lively interactive stories. No coding experience is required. Build interactive charts and maps and embed them in your website Understand the principles for designing effective charts and maps Learn key data visualization concepts to help you choose the right tools Convert and transform tabular and spatial data to tell your data story Edit and host Chart.js, Highcharts, and Leaflet map code templates on GitHub Learn how to detect bias in charts and maps produced by others

John Chambers turns his attention to R, the enormously successful open-source system based on the S language. His book guides the reader through programming with R, beginning with simple interactive use and progressing by gradual stages, starting with simple functions. More advanced programming techniques can be added as needed, allowing users to grow into software contributors, benefiting their careers and the community. R packages provide a powerful mechanism for contributions to be organized and communicated. This is the only advanced programming book on R, written by the author of the S language from which R evolved.

R in Action, Third Edition teaches you to use the R language, including the popular tidyverse packages, through hands-on examples relevant to scientific, technical, and business developers. Built specifically for statistical computing and graphics, the R language, along with its amazing collection of libraries and tools, is one of the most powerful tools you can use to tackle data analysis for business, research, and other data-intensive domains. This revised and expanded third edition of R in Action covers the new tidyverse approach to data analysis and R's state-of-the-art graphing capabilities with the ggplot2 package. R in Action, Third Edition teaches you to use the R language, including the popular tidyverse packages, through hands-on examples relevant to scientific, technical, and business developers. Focusing on practical solutions to real-world data challenges, R expert Rob Kabacoff takes you on a crash course in statistics, from dealing with messy and incomplete data to creating stunning visualisations. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications.

This book offers solutions to all 284 exercises in Advanced R, Second Edition. All the solutions have been carefully documented and made to be as clear and accessible as possible. Working through the exercises and their solutions will give you a deeper understanding of a variety of programming challenges, many of which are relevant to everyday work. This will expand your set of tools on a technical and conceptual level. You will be able to transfer many of the specific programming schemes directly and will discover far more elegant solutions to everyday problems. Features: When R creates copies, and how it affects memory usage and code performance Everything you could ever want to know about functions The differences between calling and exiting handlers How to employ functional programming to solve modular tasks The motivation, mechanics, usage, and limitations of R's highly pragmatic S3 OO system The R6 OO system, which is more like OO programming in other languages The rules that R uses to parse and evaluate expressions How to use metaprogramming to generate HTML or LaTeX with elegant R code How to identify and resolve performance bottlenecks Professionelle elektronische Ausgabe erhältlich direkt bei [Page 6/12](http://diestel-graph-</p></div><div data-bbox=)

theory.com/german/Profi.html Detailliert und klar, sowie stets mit Blick auf das Wesentliche, führt dieses Buch in die Graphentheorie ein. Zu jedem Themenkomplex stellt es sorgfältig die Grundlagen dar und beweist dann ein oder zwei tiefere typische Sätze, oftmals ergänzt durch eine informelle Diskussion ihrer tragenden Ideen. Es vermittelt so exemplarisch die wichtigsten Methoden der heutigen Graphentheorie, einschließlich moderner Techniken wie Regularitätslemma, Zufallsgraphen, Baumzerlegungen und Minoren. Aus den Besprechungen: "Eine hervorragende und mit größter Sorgfalt geschriebene Einführung in die moderne Graphentheorie, die sich in den Kanon der prägenden Lehrbücher einreihen wird. Vorbehaltlos zu empfehlen." DMV-Jahresbericht "Ein Höhepunkt ist das Kapitel zur Minorentheorie von Robertson und Seymour: mit Abstand die beste in der Literatur zu findende Darstellung." Mathematika „Das Buch wurde enthusiastisch aufgenommen – und hat es allemal verdient. Eine meisterhaft klare Darlegung der modernen Graphentheorie." ICA Bulletin "Fantastisch gelungen ... ein verdammt gutes Buch." MAA Reviews "Tief, klar, wunderbar. Ein anspruchsvolles Buch aus dem Herzen der Graphentheorie, voll von Tiefe und Integrität." SIAM Review

"Practical recipes for visualizing data"--Cover.

ggplot2Elegant Graphics for Data AnalysisSpringer Science & Business Media
Master the Shiny web framework—and take your R skills to a whole new level. By letting you move beyond static reports, Shiny helps you create fully interactive web apps for data analyses. Users will be able to jump between datasets, explore different subsets or facets of the data, run models with parameter values of their choosing, customize visualizations, and much more. Hadley Wickham from RStudio shows data scientists, data analysts, statisticians, and scientific researchers with no knowledge of HTML, CSS, or JavaScript how to create rich web apps from R. This in-depth guide provides a learning path that you can follow with confidence, as you go from a Shiny beginner to an expert developer who can write large, complex apps that are maintainable and performant. Get started: Discover how the major pieces of a Shiny app fit together Put Shiny in action: Explore Shiny functionality with a focus on code samples, example apps, and useful techniques Master reactivity: Go deep into the theory and practice of reactive programming and examine reactive graph components Apply best practices: Examine useful techniques for making your Shiny apps work well in production An accessible primer on how to create effective graphics from data This book provides students and researchers a hands-on introduction to the principles and practice of data visualization. It explains what makes some graphs succeed while others fail, how to make high-quality figures from data using powerful and reproducible methods, and how to think about data visualization in an honest and effective way. Data Visualization builds the reader's expertise in ggplot2, a versatile visualization library for the R programming language. Through a series of worked examples, this accessible primer then demonstrates how to create plots piece by piece, beginning with summaries of single variables and moving on to more complex graphics. Topics include plotting continuous and categorical variables; layering information on graphics; producing effective "small multiple" plots; grouping, summarizing, and transforming data for plotting; creating maps; working with the output of statistical models; and refining plots to make them more comprehensible. Effective graphics are essential to communicating ideas and a great way to better understand data. This book provides the practical skills students and practitioners need to visualize quantitative data and get the most out of their research findings. Provides hands-on instruction using R and ggplot2 Shows how the "tidyverse" of data analysis tools makes working with R easier and more consistent Includes a library of data sets, code, and functions

Written for statisticians, computer scientists, geographers, research and applied scientists, and others interested in visualizing data, this book presents a unique foundation for producing almost every quantitative graphic found in scientific journals, newspapers, statistical packages,

and data visualization systems. It was designed for a distributed computing environment, with special attention given to conserving computer code and system resources. While the tangible result of this work is a Java production graphics library, the text focuses on the deep structures involved in producing quantitative graphics from data. It investigates the rules that underlie pie charts, bar charts, scatterplots, function plots, maps, mosaics, and radar charts. These rules are abstracted from the work of Bertin, Cleveland, Kosslyn, MacEachren, Pinker, Tufte, Tukey, Tobler, and other theorists of quantitative graphics.

The richly illustrated Interactive Web-Based Data Visualization with R, plotly, and shiny focuses on the process of programming interactive web graphics for multidimensional data analysis. It is written for the data analyst who wants to leverage the capabilities of interactive web graphics without having to learn web programming. Through many R code examples, you will learn how to tap the extensive functionality of these tools to enhance the presentation and exploration of data. By mastering these concepts and tools, you will impress your colleagues with your ability to quickly generate more informative, engaging, and reproducible interactive graphics using free and open source software that you can share over email, export to pdf, and more. Key Features: Convert static ggplot2 graphics to an interactive web-based form Link, animate, and arrange multiple plots in standalone HTML from R Embed, modify, and respond to plotly graphics in a shiny app Learn best practices for visualizing continuous, discrete, and multivariate data Learn numerous ways to visualize geo-spatial data This book makes heavy use of plotly for graphical rendering, but you will also learn about other R packages that support different phases of a data science workflow, such as tidyr, dplyr, and tidyverse. Along the way, you will gain insight into best practices for visualization of high-dimensional data, statistical graphics, and graphical perception. The printed book is complemented by an interactive website where readers can view movies demonstrating the examples and interact with graphics.

"This book introduces you to R, RStudio, and the tidyverse, a collection of R packages designed to work together to make data science fast, fluent, and fun. Suitable for readers with no previous programming experience"--

Data visualization is one of the most important part of data science. Many books and courses present a catalogue of graphics but they don't teach you which charts to use according to the type of the data. In this book, we start by presenting the key graphic systems and packages available in R, including R base graphs, lattice and ggplot2 plotting systems. Next, we provide more than 200 practical examples to create great graphics for the right data using either the ggplot2 package and extensions or the traditional R graphics. With this book, you 'll learn: - How to quickly create beautiful graphics using ggplot2 packages - How to properly customize and annotate the plots - Type of graphics for visualizing categorical and continuous variables - How to add automatically p-values to box plots, bar plots and alternatives - How to add marginal density plots and correlation coefficients to scatter plots - Key methods for analyzing and visualizing multivariate data - R functions and packages for plotting time series data - How to combine multiple plots on one page to create production-quality figures.

Join the revolution ignited by the ground-breaking R system! Starting with an introduction to R, covering standard regression methods, then presenting more advanced topics, this book guides users through the practical and powerful tools that the R system provides. The emphasis is on hands-on analysis, graphical display and interpretation of data. The many worked examples, taken from real-world research, are accompanied by commentary on what is done and why. A website provides computer code and data sets, allowing readers to reproduce all analyses. Updates and solutions to selected exercises are also available.

Assuming only basic statistical knowledge, the book is ideal for research scientists, final-year undergraduate or graduate level students of applied statistics, and practising statisticians. It is both for learning and for reference. This revised edition reflects changes in R since 2003 and

has new material on survival analysis, random coefficient models, and the handling of high-dimensional data.

Statistical Rethinking: A Bayesian Course with Examples in R and Stan builds readers' knowledge of and confidence in statistical modeling. Reflecting the need for even minor programming in today's model-based statistics, the book pushes readers to perform step-by-step calculations that are usually automated. This unique computational approach ensures that readers understand enough of the details to make reasonable choices and interpretations in their own modeling work. The text presents generalized linear multilevel models from a Bayesian perspective, relying on a simple logical interpretation of Bayesian probability and maximum entropy. It covers from the basics of regression to multilevel models. The author also discusses measurement error, missing data, and Gaussian process models for spatial and network autocorrelation. By using complete R code examples throughout, this book provides a practical foundation for performing statistical inference. Designed for both PhD students and seasoned professionals in the natural and social sciences, it prepares them for more advanced or specialized statistical modeling. Web Resource The book is accompanied by an R package (rethinking) that is available on the author's website and GitHub. The two core functions (map and map2stan) of this package allow a variety of statistical models to be constructed from standard model formulas.

This book is about using interactive and dynamic plots on a computer screen as part of data exploration and modeling, both alone and as a partner with static graphics and non-graphical computational methods. The area of interactive and dynamic data visualization emerged within statistics as part of research on exploratory data analysis in the late 1960s, and it remains an active subject of research today, as its use in practice continues to grow. It now makes substantial contributions within computer science as well, as part of the growing fields of information visualization and data mining, especially visual data mining. The material in this book includes:

- An introduction to data visualization, explaining how it differs from other types of visualization.
- A description of four toolbox of interactive and dynamic graphical methods.
- An approach for exploring missing values in data.
- An explanation of the use of these tools in cluster analysis and supervised classification.
- An overview of additional material available on the web.
- A description of the data used in the analyses and exercises.

The book's examples use the software R and GGobi. R (Ihaka & Gentleman 1996, R Development Core Team 2006) is a free software environment for statistical computing and graphics; it is most often used from the command line, provides a wide variety of statistical methods, and includes high-quality static graphics. R arose in the Statistics Department of the University of Auckland and is now developed and maintained by a global collaborative effort.

Summary R in Action, Second Edition presents both the R language and the examples that make it so useful for business developers. Focusing on practical solutions, the book offers a crash course in statistics and covers elegant methods for dealing with messy and incomplete data that are difficult to analyze using traditional methods. You'll also master R's extensive graphical capabilities for exploring and presenting data visually. And this expanded second edition includes new chapters on time series analysis, cluster analysis, and classification methodologies, including decision trees, random forests, and support vector machines. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Business pros and researchers thrive on data, and R speaks the language of data analysis. R is a powerful programming language for statistical computing. Unlike general-purpose tools, R provides thousands of modules for solving just about any data-crunching or presentation challenge you're likely to face. R runs on all important platforms and is used by thousands of major corporations and institutions worldwide. About the Book R in Action, Second Edition teaches you how to use the R language by presenting examples relevant to scientific, technical, and business developers.

Focusing on practical solutions, the book offers a crash course in statistics, including elegant methods for dealing with messy and incomplete data. You'll also master R's extensive graphical capabilities for exploring and presenting data visually. And this expanded second edition includes new chapters on forecasting, data mining, and dynamic report writing. What's Inside Complete R language tutorial Using R to manage, analyze, and visualize data Techniques for debugging programs and creating packages OOP in R Over 160 graphs About the Author Dr. Rob Kabacoff is a seasoned researcher and teacher who specializes in data analysis. He also maintains the popular Quick-R website at statmethods.net. Table of Contents PART 1 GETTING STARTED Introduction to R Creating a dataset Getting started with graphs Basic data management Advanced data management PART 2 BASIC METHODS Basic graphs Basic statistics PART 3 INTERMEDIATE METHODS Regression Analysis of variance Power analysis Intermediate graphs Resampling statistics and bootstrapping PART 4 ADVANCED METHODS Generalized linear models Principal components and factor analysis Time series Cluster analysis Classification Advanced methods for missing data PART 5 EXPANDING YOUR SKILLS Advanced graphics with ggplot2 Advanced programming Creating a package Creating dynamic reports Advanced graphics with the lattice package available online only from manning.com/kabacoff2

Do you want to use R to tell stories? This book was written for you—whether you already know some R or have never coded before. Most R texts focus only on programming or statistical theory. Practical R for Mass Communication and Journalism gives you ideas, tools, and techniques for incorporating data and visualizations into your narratives. You'll see step by step how to: Analyze airport flight delays, restaurant inspections, and election results Map bank locations, median incomes, and new voting districts Compare campaign contributions to final election results Extract data from PDFs Whip messy data into shape for analysis Scrape data from a website Create graphics ranging from simple, static charts to interactive visualizations for the Web If you work or plan to work in a newsroom, government office, non-profit policy organization, or PR office, Practical R for Mass Communication and Journalism will help you use R in your world. This book has a companion website with code, links to additional resources, and searchable tables by function and task. Sharon Machlis is the author of Computerworld's Beginner's Guide to R, host of InfoWorld's Do More With R video screencast series, admin for the R for Journalists Google Group, and is well known among Twitter users who follow the #rstats hashtag. She is Director of Editorial Data and Analytics at IDG Communications (parent company of Computerworld, InfoWorld, PC World and Macworld, among others) and a frequent speaker at data journalism and R conferences.

Provides both rich theory and powerful applications Figures are accompanied by code required to produce them Full color figures

R is revolutionizing the world of statistical computing. Powerful, flexible, and best of all free, R is now the program of choice for tens of thousands of statisticians. Destined to become an instant classic, R Graphics presents the first complete, authoritative exposition on the R graphical system. Paul Murrell, widely known as the leading expert

o ggplot2 is one of the most sophisticated and advanced packages of R and its use is constantly growing in the community of R programmers. This book is the perfect starting point for your journey in learning about one of the most refined and widely used plotting tools—ggplot2. You will start by taking a tour of the most relevant packages available in R to create plots, such as graphics, lattice, and ggplot2. Following this, you will take a deep dive into the concepts of ggplot2 by covering different kinds of plots that will help you gear up for your first hands-on project. After this first taste of the capabilities of ggplot2, you will explore in detail the visualization approach implemented

in this package, the so-called grammar of graphics, providing you with the basics to understand the backstage concepts in the plotting functions. Finally, you will take a look at some specialized and sophisticated applications of ggplot2, such as how to realize a complex scatterplot matrix, heatmaps, and how to make spatial visualization graphs. This guide for practicing statisticians, data scientists, and R users and programmers will teach the essentials of preprocessing: data leveraging the R programming language to easily and quickly turn noisy data into usable pieces of information. Data wrangling, which is also commonly referred to as data munging, transformation, manipulation, janitor work, etc., can be a painstakingly laborious process. Roughly 80% of data analysis is spent on cleaning and preparing data; however, being a prerequisite to the rest of the data analysis workflow (visualization, analysis, reporting), it is essential that one become fluent and efficient in data wrangling techniques. This book will guide the user through the data wrangling process via a step-by-step tutorial approach and provide a solid foundation for working with data in R. The author's goal is to teach the user how to easily wrangle data in order to spend more time on understanding the content of the data. By the end of the book, the user will have learned: How to work with different types of data such as numerics, characters, regular expressions, factors, and dates The difference between different data structures and how to create, add additional components to, and subset each data structure How to acquire and parse data from locations previously inaccessible How to develop functions and use loop control structures to reduce code redundancy How to use pipe operators to simplify code and make it more readable How to reshape the layout of data and manipulate, summarize, and join data sets

Statistical Inference via Data Science: A ModernDive into R and the Tidyverse provides a pathway for learning about statistical inference using data science tools widely used in industry, academia, and government. It introduces the tidyverse suite of R packages, including the ggplot2 package for data visualization, and the dplyr package for data wrangling. After equipping readers with just enough of these data science tools to perform effective exploratory data analyses, the book covers traditional introductory statistics topics like confidence intervals, hypothesis testing, and multiple regression modeling, while focusing on visualization throughout. Features: ? Assumes minimal prerequisites, notably, no prior calculus nor coding experience ? Motivates theory using real-world data, including all domestic flights leaving New York City in 2013, the Gapminder project, and the data journalism website, FiveThirtyEight.com ? Centers on simulation-based approaches to statistical inference rather than mathematical formulas ? Uses the infer package for "tidy" and transparent statistical inference to construct confidence intervals and conduct hypothesis tests via the bootstrap and permutation methods ? Provides all code and output embedded directly in the text; also available in the online version at moderndive.com This book is intended for individuals who would like to simultaneously start developing their data science toolbox and start learning about the inferential and modeling tools used in much of modern-day research. The book can be used in methods and data science courses and first courses in statistics, at both the undergraduate and graduate levels.

With more than 200 practical recipes, this book helps you perform data analysis with R quickly and efficiently. The R language provides everything you need to do statistical work, but its structure can be difficult to master. This collection of concise, task-oriented

recipes makes you productive with R immediately, with solutions ranging from basic tasks to input and output, general statistics, graphics, and linear regression. Each recipe addresses a specific problem, with a discussion that explains the solution and offers insight into how it works. If you're a beginner, R Cookbook will help get you started. If you're an experienced data programmer, it will jog your memory and expand your horizons. You'll get the job done faster and learn more about R in the process.

Create vectors, handle variables, and perform other basic functions
Input and output data
Tackle data structures such as matrices, lists, factors, and data frames
Work with probability, probability distributions, and random variables
Calculate statistics and confidence intervals, and perform statistical tests
Create a variety of graphic displays
Build statistical models with linear regressions and analysis of variance (ANOVA)
Explore advanced statistical techniques, such as finding clusters in your data

"Wonderfully readable, R Cookbook serves not only as a solutions manual of sorts, but as a truly enjoyable way to explore the R language—one practical example at a time."—Jeffrey Ryan, software consultant and R package author

An Essential Reference for Intermediate and Advanced R Programmers
Advanced R presents useful tools and techniques for attacking many types of R programming problems, helping you avoid mistakes and dead ends. With more than ten years of experience programming in R, the author illustrates the elegance, beauty, and flexibility at the heart of R. The book develops the necessary skills to produce quality code that can be used in a variety of circumstances. You will learn:

- The fundamentals of R, including standard data types and functions
- Functional programming as a useful framework for solving wide classes of problems
- The positives and negatives of metaprogramming
- How to write fast, memory-efficient code

This book not only helps current R users become R programmers but also shows existing programmers what's special about R. Intermediate R programmers can dive deeper into R and learn new strategies for solving diverse problems while programmers from other languages can learn the details of R and understand why R works the way it does.

[Copyright: 25c49e9293e071f1ce6a78f3b01b6926](#)