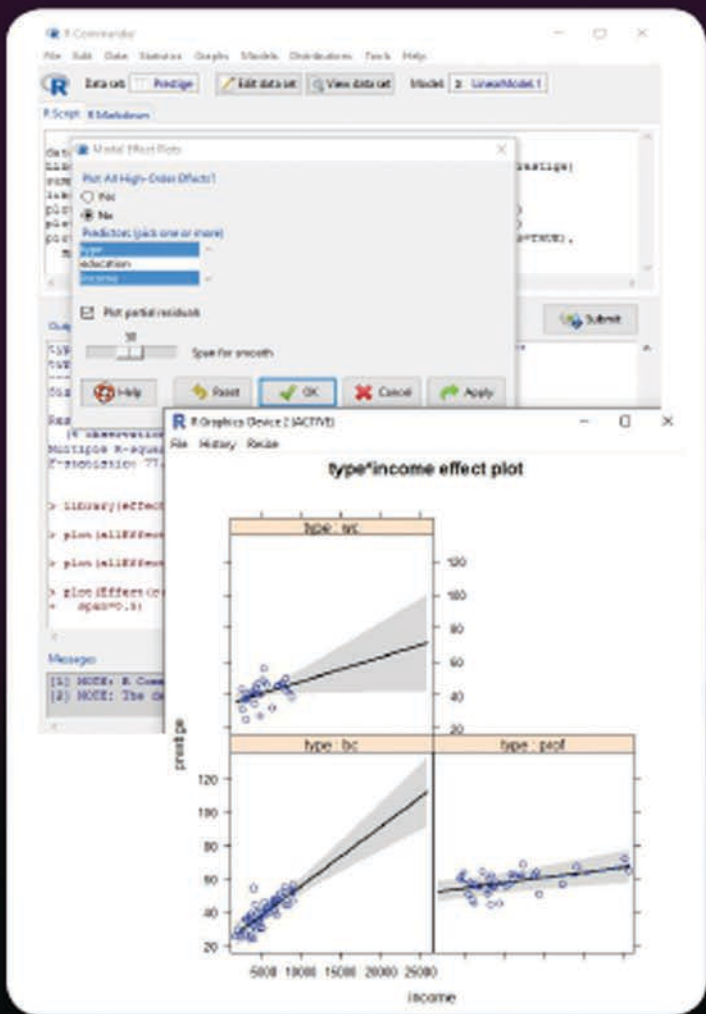


## The R Series

# Using the R Commander A Point-and-Click Interface for R



# John Fox



**CRC Press**  
Taylor & Francis Group

A CHAPMAN &amp; HALL BOOK

# **Using the R Commander**

## **A Point-and-Click Interface for R**

# **Chapman & Hall/CRC**

## **The R Series**

### **Series Editors**

**John M. Chambers**  
Department of Statistics  
Stanford University  
Stanford, California, USA

**Torsten Hothorn**  
Division of Biostatistics  
University of Zurich  
Switzerland

**Duncan Temple Lang**  
Department of Statistics  
University of California, Davis  
Davis, California, USA

**Hadley Wickham**  
RStudio  
Boston, Massachusetts, USA

### **Aims and Scope**

This book series reflects the recent rapid growth in the development and application of R, the programming language and software environment for statistical computing and graphics. R is now widely used in academic research, education, and industry. It is constantly growing, with new versions of the core software released regularly and more than 7,000 packages available. It is difficult for the documentation to keep pace with the expansion of the software, and this vital book series provides a forum for the publication of books covering many aspects of the development and application of R.

The scope of the series is wide, covering three main threads:

- Applications of R to specific disciplines such as biology, epidemiology, genetics, engineering, finance, and the social sciences.
- Using R for the study of topics of statistical methodology, such as linear and mixed modeling, time series, Bayesian methods, and missing data.
- The development of R, including programming, building packages, and graphics.

The books will appeal to programmers and developers of R software, as well as applied statisticians and data analysts in many fields. The books will feature detailed worked examples and R code fully integrated into the text, ensuring their usefulness to researchers, practitioners and students.

## Published Titles

**Spatial Microsimulation with R**, *Robin Lovelace and Morgane Dumont*

**Statistics in Toxicology Using R**, *Ludwig A. Hothorn*

**Stated Preference Methods Using R**, *Hideo Aizaki, Tomoaki Nakatani, and Kazuo Sato*

**Using R for Numerical Analysis in Science and Engineering**,  
*Victor A. Bloomfield*

**Event History Analysis with R**, *Göran Broström*

**Computational Actuarial Science with R**, *Arthur Charpentier*

**Statistical Computing in C++ and R**, *Randall L. Eubank and Ana Kupresanin*

**Basics of Matrix Algebra for Statistics with R**, *Nick Fieller*

**Reproducible Research with R and RStudio, Second Edition**,  
*Christopher Gandrud*

**R and MATLAB®**, *David E. Hiebeler*

**Nonparametric Statistical Methods Using R**, *John Kloke and Joseph McKean*

**Displaying Time Series, Spatial, and Space-Time Data with R**,  
*Oscar Perpiñán Lamigueiro*

**Programming Graphical User Interfaces with R**, *Michael F. Lawrence and John Verzani*

**Analyzing Sensory Data with R**, *Sébastien Lê and Theirry Worch*

**Parallel Computing for Data Science: With Examples in R, C++ and CUDA**, *Norman Matloff*

**Analyzing Baseball Data with R**, *Max Marchi and Jim Albert*

**Growth Curve Analysis and Visualization Using R**, *Daniel Mirman*

**R Graphics, Second Edition**, *Paul Murrell*

**Introductory Fisheries Analyses with R**, *Derek H. Ogle*

**Data Science in R: A Case Studies Approach to Computational Reasoning and Problem Solving**, *Deborah Nolan and Duncan Temple Lang*

**Multiple Factor Analysis by Example Using R**, *Jérôme Pagès*

**Customer and Business Analytics: Applied Data Mining for Business Decision Making Using R**, *Daniel S. Putler and Robert E. Krider*

**Implementing Reproducible Research**, *Victoria Stodden, Friedrich Leisch, and Roger D. Peng*

**Graphical Data Analysis with R**, *Antony Unwin*

**Using R for Introductory Statistics, Second Edition**, *John Verzani*

**Advanced R**, *Hadley Wickham*

**Dynamic Documents with R and knitr, Second Edition**, *Yihui Xie*

# **Using the R Commander A Point-and-Click Interface for R**

**John Fox**

McMaster University  
Hamilton, Ontario, Canada



**CRC Press**

Taylor & Francis Group

Boca Raton London New York

---

CRC Press is an imprint of the  
Taylor & Francis Group, an **informa** business

A CHAPMAN & HALL BOOK

CRC Press  
Taylor & Francis Group  
6000 Broken Sound Parkway NW, Suite 300  
Boca Raton, FL 33487-2742

© 2017 by Taylor & Francis Group, LLC  
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works

Printed on acid-free paper  
Version Date: 20160524

International Standard Book Number-13: 978-1-4987-4190-3 (Paperback)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access [www.copyright.com](http://www.copyright.com) (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

**Trademark Notice:** Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

**Visit the Taylor & Francis Web site at**  
**<http://www.taylorandfrancis.com>**

**and the CRC Press Web site at**  
**<http://www.crcpress.com>**

*To the memory of my mentor and friend,  
Mel Guyer.*





**Taylor & Francis**

Taylor & Francis Group

<http://taylorandfrancis.com>

---

# Contents

---

Preface	xiii
<b>1 Introducing R and the R Commander</b>	<b>1</b>
1.1 What Are R and the R Commander?	1
1.2 A Brief History of R and the R Commander	2
1.3 Chapter Synopses	4
1.4 What's Next?	5
1.5 Web Site	6
1.6 Typographical Conventions	6
<b>2 Installing R and the R Commander</b>	<b>9</b>
2.1 Acquiring and Installing R and the R Commander	9
2.2 Installing R and the R Commander under Microsoft Windows	10
2.2.1 Troubleshooting	11
2.3 Installing R and the R Commander under Mac OS X	13
2.3.1 Installing X11 (XQuartz) for Mac OS X	13
2.3.2 Installing the <b>Rcmdr</b> Package and Its Dependencies	15
2.3.3 Preventing R from Napping	15
2.3.4 Troubleshooting	15
2.4 Installing R and the R Commander on Linux and Unix Computers	16
2.4.1 Troubleshooting	17
2.5 Installing Optional Auxiliary Software: Pandoc and L <sup>A</sup> T <sub>E</sub> X	17
<b>3 A Quick Tour of the R Commander</b>	<b>19</b>
3.1 Starting the R Commander	19
3.2 The R Commander Interface	20
3.3 Reading Data into the R Commander	23
3.4 Examining and Recoding Variables	27
3.5 Making a Contingency Table	35
3.6 Creating a Report	38
3.6.1 Creating a Report by Cutting and Pasting	38
3.6.2 Creating a Report as a Dynamic Document	38
3.6.2.1 Using Markdown: The Basics	42
3.6.2.2 Adding a Little L <sup>A</sup> T <sub>E</sub> X Math to an R Markdown Document*	46
3.7 Editing Commands*	46
3.8 Terminating the R Commander Session	50
3.9 Customizing the R Commander*	50
3.9.1 Using the <i>Commander Options</i> Dialog	50
3.9.2 Setting <b>Rcmdr</b> Package Options	52
3.9.3 Managing Colors in the R Commander	52
3.9.4 Saving R Commander Options	53

<b>4</b>	<b>Data Input and Data Management</b>	<b>55</b>
4.1	Introduction to Data Management . . . . .	55
4.2	Data Input . . . . .	55
4.2.1	Entering Data into the R Commander Data Editor . . . . .	56
4.2.2	Reading Data from Plain-Text Files . . . . .	58
4.2.3	Importing Data from Spreadsheets and Other Sources . . . . .	61
4.2.4	Accessing Data Sets in R Packages . . . . .	63
4.3	Saving and Exporting Data from the R Commander . . . . .	66
4.4	Modifying Variables . . . . .	68
4.4.1	Recoding Variables . . . . .	68
4.4.2	Computing New Variables . . . . .	69
4.4.2.1	Complicated Expressions in Computing New Variables* . . . . .	70
4.4.3	Other Operations on Variables . . . . .	72
4.5	Manipulating Data Sets . . . . .	75
4.5.1	Special Operations on Data Sets* . . . . .	75
4.5.2	Subsetting Cases . . . . .	76
4.5.3	Merging Data Sets* . . . . .	78
<b>5</b>	<b>Summarizing and Graphing Data</b>	<b>81</b>
5.1	Simple Numerical Summaries . . . . .	81
5.2	Contingency Tables . . . . .	85
5.3	Graphing Distributions of Variables . . . . .	88
5.3.1	Graphing Numerical Data . . . . .	88
5.3.2	Graphing Categorical Data . . . . .	94
5.4	Graphing Relationships . . . . .	96
5.4.1	Simple Scatterplots . . . . .	96
5.4.2	Enhanced Scatterplots* . . . . .	97
5.4.3	Scatterplot Matrices . . . . .	100
5.4.4	Point Identification in Scatterplots and Scatterplot Matrices . . . . .	100
5.4.5	3D Scatterplots* . . . . .	104
5.4.6	Plotting Means . . . . .	104
<b>6</b>	<b>Simple Statistical Tests</b>	<b>109</b>
6.1	Tests on Means . . . . .	109
6.1.1	Independent Samples Difference-of-Means $t$ Test . . . . .	109
6.1.2	One-Way Analysis of Variance . . . . .	111
6.1.3	Two-Way and Higher-Way Analysis of Variance . . . . .	113
6.2	Tests on Proportions . . . . .	120
6.3	Tests on Variances . . . . .	123
6.4	Nonparametric Tests . . . . .	125
6.5	Other Simple Tests* . . . . .	127
<b>7</b>	<b>Fitting Linear and Generalized Linear Models</b>	<b>129</b>
7.1	Linear Regression Models . . . . .	129
7.2	Linear Models with Factors* . . . . .	132
7.2.1	Linear-Model Formulas . . . . .	132
7.2.2	The Principle of Marginality . . . . .	133
7.2.3	Examples Using the Canadian Occupational Prestige Data . . . . .	133
7.2.4	Dummy Variables and Other Contrasts for Factors . . . . .	138
7.3	Fitting Regression Splines and Polynomials* . . . . .	141
7.3.1	Polynomial Terms . . . . .	141

7.3.2	Regression Splines . . . . .	141
7.4	Generalized Linear Models* . . . . .	145
7.5	Other Regression Models* . . . . .	149
7.6	Visualizing Linear and Generalized Linear Models* . . . . .	149
7.6.1	Partial Residuals in Effect Plots . . . . .	152
7.7	Confidence Intervals and Hypothesis Tests for Coefficients . . . . .	156
7.7.1	Confidence Intervals . . . . .	156
7.7.2	Analysis of Variance and Analysis of Deviance Tables* . . . . .	156
7.7.3	Tests Comparing Models* . . . . .	159
7.7.4	Testing Linear Hypotheses* . . . . .	160
7.8	Regression Model Diagnostics* . . . . .	163
7.9	Model Selection* . . . . .	167
<b>8</b>	<b>Probability Distributions and Simulation</b>	<b>171</b>
8.1	Working with Probability Distributions . . . . .	171
8.1.1	Continuous Distributions . . . . .	171
8.1.2	Discrete Distributions . . . . .	174
8.2	Plotting Probability Distributions . . . . .	175
8.3	Simple Random Simulations . . . . .	177
8.3.1	Setting the Seed of the R Pseudo-Random-Number Generator . . . .	178
8.3.2	A Simple Simulation Illustrating the Central Limit Theorem . . . .	178
<b>9</b>	<b>Using R Commander Plug-in Packages</b>	<b>183</b>
9.1	Acquiring and Loading Plug-Ins . . . . .	183
9.2	Using the <b>RcmdrPlugin.TeachingDemos</b> Package . . . . .	184
9.3	Survival Analysis with the <b>RcmdrPlugin.survival</b> Package* . . . . .	186
9.3.1	Survival Analysis Examples Based on Criminal Recidivism Data . .	189
	<b>Appendix A Guide to the R Commander Menus</b>	<b>199</b>
	<b>References</b>	<b>207</b>
	<b>Author Index</b>	<b>211</b>
	<b>Subject Index</b>	<b>213</b>



**Taylor & Francis**

Taylor & Francis Group

<http://taylorandfrancis.com>

---

# Preface

---

The R Commander is a point-and-click *graphical user interface* (*GUI*) for R, providing access to R statistical software through familiar menus and dialog boxes instead of by typing potentially arcane commands. I expect that this book, which explains how to use the R Commander, will be of interest to students and instructors in introductory and intermediate-level statistics courses, to researchers who want to use R without having to contend with writing commands, and to R users who will eventually transition to the command-line interface but who prefer to start more simply.

In particular, in a basic statistics course, the central goal (in my opinion) should be to teach fundamental statistical ideas—distribution, statistical relationship, estimation, sampling variation, observational vs. experimental data, randomization, and so on. One doesn't want a basic statistics course to devolve into an exercise in learning how to write commands for statistical software, letting the software tail wag the statistical dog. I initially wrote the R Commander for this reason: to provide a transparent, intuitive, point-and-click graphical user interface to R, implemented using familiar menus and dialog boxes, running on all commonly used operating systems (Windows, Mac OS X, and Linux/Unix systems), and distributed and installed as a standard R package—called the **Rcmdr** package.

Although it was originally intended for use in basic statistics classes, the current capabilities of the R Commander, described in the various chapters of this book, extend well beyond basic statistics. The current version of the **Rcmdr** package contains nearly 15,000 lines of R code (exclusive of comments, blank lines, documentation, etc.). The R Commander is, moreover (like R itself), designed to be extensible by *plug-in packages*, standard R packages that augment or modify the R Commander's menus and dialogs (see [Chapter 9](#)).

I caught wind of R in the late 1990s and incorporated it shortly thereafter in my teaching—in a graduate social statistics course in applied regression analysis and generalized linear models. In 2002, I published a book on using R (and S-PLUS) for applied regression analysis (Fox, 2002).<sup>1</sup>

I also wanted to use R to teach basic statistics to social science graduate students and undergrads, but I felt that R's command line interface was an obstacle. I expected that eventually someone would introduce a graphical user interface to R, but none materialized, and so, around 2002, I resolved to tackle this task myself. After some experimentation, I decided to use the Tcl/Tk GUI builder because the basic R distribution comes with the **tcltk** package, which provides an R interface to Tcl/Tk. This choice permitted me to write an R GUI—the R Commander—that runs on all operating systems supported by R and that is coded entirely in R, minimizing the necessity to install additional software.

This book provides background information on R and the R Commander. I explain how you can obtain and install R and the R Commander on your computer. Finally, I show you how to use the R Commander to perform a variety of common statistical tasks.

I hope that the book proves useful to you, and I invite you to contact me at [jfox@mcmaster.ca](mailto:jfox@mcmaster.ca) with corrections, comments, and suggestions.

---

<sup>1</sup>The second edition of this book (Fox and Weisberg, 2011) focuses solely on R and is coauthored with Sanford Weisberg.

---

## Acknowledgments

Michael Friendly, Allison Leanage, and several anonymous reviewers contributed helpful suggestions on a draft of this book.

I'm grateful to the numerous individuals whose contributions to the **Rcmdr** package are acknowledged in the package documentation (see the *Help > About Rcmdr* menu). In particular, Richard Heiberger made a number of contributions to the early development of the **R Commander**, not least of which was the original code for rendering **R Commander** plug-in packages self-starting. Miroslav Ristic substantially improved the code for the probability distribution dialogs. Milan Bouchet-Valat joined me as a developer of the **Rcmdr** package in 2013 and helped to modernize the **R Commander** interface and to adapt it better to various computing platforms.

Peter Dalgaard, a member of the R Core team, made a critical indirect contribution to the **R Commander**—in effect, making it possible—by incorporating the **tcltk** package in the standard R distribution. Similarly, Milan and I used Philippe Grosjean's **tcltk2** package to enhance the **R Commander** interface, and Philippe has been a valuable source of information on Tcl/Tk. Brian Ripley, another member of R Core, has been very generous with his time, helping me to solve a variety of problems in the development of the **R Commander**.

I'd also like to thank John Kimmel, my editor at Chapman & Hall/CRC Press, Shashi Kumar (for his L<sup>A</sup>T<sub>E</sub>X expertise), and the production staff at Chapman & Hall/CRC Press, for all their help and encouragement.

Finally, my work on the **R Commander** was partly supported by grants from the Social Sciences and Humanities Research Council of Canada, and from McMaster University through the Senator William McMaster Chair in Social Statistics.

## Introducing R and the R Commander

---

This chapter introduces R and the R Commander, explaining what they are and where they came from. The chapter also outlines the contents of the book, shows how to access the web site for the book, and describes the typographical conventions used in the text.

---

### 1.1 What Are R and the R Commander?

The R Commander—the subject of this book—is a point-and-click *graphical user interface* (*GUI*) for R, allowing you to use R statistical software through familiar menus and dialog boxes instead of by typing commands. Throughout the book, I assume that the statistical methods covered are familiar to you—or that you’re concurrently learning them in a statistics class or by independently reading a complementary statistics text. The object of the book is to show you how to perform data analysis with the R Commander employing common statistical methods, not to teach the statistical methods themselves.

An implication of this approach is that you should feel free to skip those parts of the book that take up statistical methods with which you’re unfamiliar. For example, most of the material in [Chapter 7](#), on working with statistical models in the R Commander, is beyond the level of a typical basic statistics course. Sections that deal with relatively advanced or difficult material are marked with asterisks.

R is highly capable, free, open-source statistical software. Although it is hard to know with any certainty how many people use R, it is—for example, judging by Internet traffic—possibly the most popular statistical software in the world. R is, in any event, very widely used, and its use is growing rapidly!

R incorporates a programming language that is finely adapted to the development of statistical applications. R descends from the S programming language, originally developed in the 1980s by statisticians and computer scientists at Bell Labs, led by John Chambers (see, e.g., Becker et al., 1988). Indeed, R can be regarded as a dialect of S. Eventually incorporated in a commercial product called S-PLUS, S was popular among statisticians prior to the development of R. At present, the free, open-source R has entirely eclipsed its commercial cousin S-PLUS.

R is *free* software in Richard Stallman’s famously dual sense of the term (Stallman, 2002):<sup>1</sup> It is free in the obvious sense of being costless, but also in the deeper sense that users may freely modify and distribute R. Moreover, R is licensed under the Free Software Foundation’s General Public License (GPL)—a “viral copy-left” that prevents individuals or companies from restricting users’ freedom to modify further and redistribute R. Freedom in the second sense essentially presupposes that R is *open source*: that is, that R is distributed not only as an executable program but also that the source code for R (written in a variety of programming languages, including in R itself) is available to interested users. For more

---

<sup>1</sup>Richard Stallman is the founder of the Free Software Foundation, with which the R Project for Statistical Computing is associated.



information about R, visit the R web site at <https://www.r-project.org/>. The R Commander is also free, open-source software distributed under the GPL.

Analyzing data with R doesn't necessarily entail writing programs in the R language, because the basic R distribution comes with impressive built-in statistical functionality. The capabilities of the standard R distribution, however, are greatly extended by (as I write this) nearly 8000 user-contributed R add-on *packages*, freely available on the Internet through the Comprehensive R Archive Network (abbreviated CRAN, and alternatively pronounced as “kran” or “see-ran”; see <https://cran.r-project.org/>). Moreover, roughly 1000 additional R packages are available through the closely associated Bioconductor Project (<http://bioconductor.org/>), which develops software primarily for bioinformatics (genonomics).

Whether you write your own R programs or use pre-packaged programs, standard data analysis in R consists of typing commands in the R language. As a simple example, to compute the mean of the variable `income`, you might type the command `mean(income)`, invoking the standard R *mean function* (program). Similarly, to perform a linear least-squares regression of `income` on `years of education` and `years of labor-force experience`, you might issue the command `lm(income ~ education + experience)`, employing the `lm` (linear-model) function. Learning to write R commands like these is an important skill and ultimately is the most efficient way to use R (see [Section 1.4](#)), but it can present a formidable obstacle to new, occasional, or casual users of R.

---

## 1.2 A Brief History of R and the R Commander

R began around 1990 as the personal project of Robert Gentleman and Ross Ihaka, two statisticians then at Auckland University in New Zealand (see Ihaka and Gentleman, 1996). Gentleman and Ihaka in effect grafted the syntax of the pre-existing statistical programming language S onto the Scheme dialect of Lisp, a programming language usually associated with work in artificial intelligence. This turned out to be a propitious choice, because, as mentioned, the S language was already widely used by statisticians.

Eventually Ihaka and Gentleman advertised their work on the Internet, attracting several other developers to the project, including John Chambers, the principal developer of S. Then, in 1997, the *R Project for Statistical Computing* was formalized, with a Core team of nine developers, a number that has since expanded to 20, many of whom are significant figures in the field of statistical computing. The R Core team is responsible for the continued development and maintenance of the basic R distribution.

As I explained, R is distributed under the free-software General Public License. The copyright to R is held by the R Foundation, which comprises the members of the R Core team along with about a dozen other individuals; I'm an elected member of the R Foundation.

The growth of R has been nothing short of amazing. [Figure 1.1](#) shows, for example, the expansion of the CRAN R package archive over the 14-year period for which I was able to obtain data.<sup>2</sup> The horizontal axes of the graph record R versions and corresponding dates, while the vertical axes show the number of CRAN packages on a logarithmic scale, so that a linear trend represents exponential growth.<sup>3</sup> The line on the graph was fit by least-squares

---

<sup>2</sup>This graph is updated from Fox (2009), where I discuss the social organization and trajectory of the R Project.

<sup>3</sup>If you're unfamiliar with logs, don't be concerned: The essential point is that the scale gets more compressed as the number of packages grows, so that, for example, the distance between 100 and 200 packages on the log scale is the *same* as the distance between 200 and 400, and the same as the distance between 400 and 800—all of these equal distances represent *doubling* the number of packages.

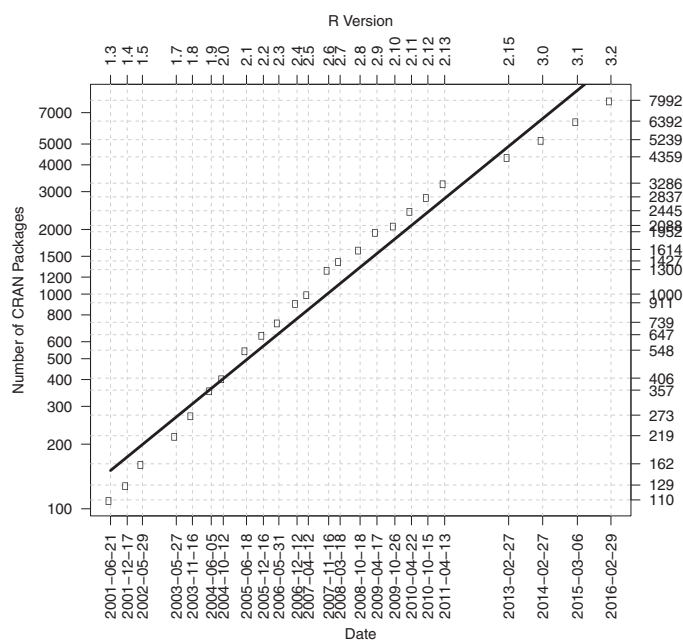


FIGURE 1.1: The growth of CRAN. The vertical axes show the number of CRAN packages on a log scale, while dates and corresponding minor R versions are shown at the bottom and top of the graph, respectively. The line on the graph was fit to the points by least-squares regression. Two R versions, 1.6 and 2.14, are omitted because their recorded dates were very close to the dates of the previous versions.

Source: Updated from Fox, “Aspects of the social organization and trajectory of the R Project,” *The R Journal*, 1(2): 5–13, 2009.

Source of Data: Downloaded from <https://svn.r-project.org/R/branches/> on 2016-03-03.

regression.<sup>4</sup> You can see from Figure 1.1 that, while the growth of CRAN was originally approximately exponential, its rate of growth has more recently slowed down. The slope of the least-squares line suggests that, on average over this period, CRAN expanded at a rate of about 35 percent a year.

As I mentioned in the Preface to this book, I began to work on the **R Commander** around 2002, and I contributed version 0.8-2 of the **Rcmdr** package to CRAN in May 2003. The first “non-beta” version, 1.0-0, appeared two years later, and was described in a paper in the *Journal of Statistical Software* (Fox, 2005), an on-line journal of the American Statistical Association. In March 2016, when I wrote the chapter you’re reading, that paper had been downloaded nearly 140,000 times—despite the fact that it was more than 10 years out of date!

I have continued to develop the **R Commander** in the intervening period: Version 1.1-1, which also appeared in 2005, introduced the capability to translate the **R Commander** interface into other languages, a feature supported by R itself, and there are now 18 such translations (counting Chinese and simplified Chinese as separate translations). In 2007,

<sup>4</sup>Again, if you’re unfamiliar with the method of least squares, don’t worry: You’ll almost surely study the topic in your basic statistics course. The essential idea is that the line comes as close (in a sense) to the points on average as possible.

Version 1.3-0 first made provision for R Commander plug-in packages, and there are currently about 40 such plug-ins on CRAN. In 2013, Milan Bouchet-Valat joined me as a developer of the R Commander, and version 2.0-0, released in that year, featured a revamped, more consistent interface—for example, featuring tabbed dialogs.

---

## 1.3 Chapter Synopses

[Chapter 2](#) describes how to download R from the Internet and install it and the R Commander on Windows, Mac OS X, and Linux/Unix systems. If you have already successfully installed R and the R Commander, then feel free to skip the chapter. There is, however, some troubleshooting information, to which you can make reference if you experience a problem.

[Chapter 3](#) introduces the R Commander graphical user interface (GUI) by demonstrating its use for a simple problem: constructing a contingency table to examine the relationship between two categorical variables. In developing the example, I explain how to start the R Commander, describe the structure of the R Commander interface, show how to read data into the R Commander, how to modify data to prepare them for analysis, how to draw a graph, how to compute numerical summaries of data, how to create a printed report of your work, how to edit and re-execute commands generated by the R Commander, and how to terminate your R and R Commander session—in short, the typical work flow of data analysis using the R Commander. I also explain how to customize the R Commander interface.

[Chapter 4](#) shows how to get data into the R Commander from a variety of sources, including entering data directly at the keyboard, reading data from a plain-text file, accessing data stored in an R package, and importing data from an Excel or other spreadsheet, or from other statistical software. I also explain how to save and export R data sets from the R Commander, and how to modify data—for example, how to create new variables and how to subset the active data set.

[Chapter 5](#) explains how to use the R Commander to compute simple numerical summaries of data, to construct and analyze contingency tables, and to draw common statistical graphs. Most of the statistical content of the chapter is covered in a typical basic statistics course, although a few topics, such as quantile-comparison plots and smoothing scatterplots, are somewhat more advanced.

[Chapter 6](#) shows how to compute simple statistical hypothesis tests and confidence intervals for means, for proportions, and for variances, along with simple nonparametric tests, a test of normality, and correlation tests. Many of these tests are typically taken up in a basic statistics class, and, in particular, tests and confidence intervals for means and proportions are often employed to introduce statistical inference.

[Chapter 7](#) explains how to fit linear and generalized linear regression models in the R Commander, and how to perform additional computations on regression models once they have been fit to data.

[Chapter 8](#) explains how to use the R Commander to perform computations on probability distributions, to graph probability distributions, and to conduct simple random simulations.

[Chapter 9](#) explains how to use R Commander *plug-in packages*. The capabilities of the R Commander are substantially augmented by the many plug-in packages for it that are available on CRAN. Plug-ins are R packages that add menus, menu items, and dialog boxes to the R Commander. I show you how to install plug-in packages, and illustrate the application of R Commander plug-ins by using the **RcmdrPlugin.TeachingDemos** package and the **RcmdrPlugin.survival** package as examples.

An appendix to the book displays the complete set of R Commander menus, along with cross-references to the text.

---

## 1.4 What's Next?

If you become a frequent user of R, you'll likely graduate from the R Commander to writing your own R commands and possibly your own R programs. There are several reasons to employ the command-line interface to R in preference to a GUI like the R Commander:

- The R Commander GUI provides access to only a small fraction of the capabilities of R and the many R packages available on CRAN. To take full advantage of R, therefore, you'll have to learn to write commands.
- Even if you limit yourself to the capabilities in the R Commander and its various plug-ins, frequent users of R find the command-line interface more efficient. Once you remember the various commands and their arguments, you'll learn to work more quickly at the command line than in a GUI.
- You'll find that a little bit of programming goes a long way. Writing simple scripts and programs is often the quickest and most straightforward way to perform data management tasks, for example.

If you do decide to learn to use R via the command-line interface, there are many books and other resources to help you. For example, I and Sanford Weisberg have written a text (Fox and Weisberg, 2011) that introduces R, including R programming, in the context of applied regression analysis. See the *Documentation* links on the R home page at <https://www.r-project.org/> for many alternative sources, including free resources.

The R Commander is designed to facilitate the transition to command-line use of R: The commands produced by the R Commander are visible in the *R Script* tab. The contents of the *R Script* tab may be saved to a file and reused, either in the R Commander or in an R programming editor. Similarly, the dynamic document produced in the R Commander *R Markdown* tab may be saved, edited, and executed independently of the R Commander. These features are briefly discussed in [Chapter 3](#).

Both the Windows and Mac OS X implementations of R come with simple programming editors, but I strongly recommend the RStudio *interactive development environment (IDE)* for command-line use of R. RStudio incorporates a powerful programming editor and is ideal both for routine data analysis in R and for R programming, including the development of R packages—and RStudio supports R Markdown documents. Like R and the R Commander, RStudio is free, open-source software: Visit the RStudio web site at <https://www.rstudio.com/products/rstudio/> for details, including extensive documentation.

## 1.5 Web Site

I have created a web site to support this book with a variety of resources, including:

- all the data files used in examples that appear in the text
- detailed (and potentially updated) installation instructions, including trouble-shooting information, beyond the instructions in [Chapter 2](#)
- information about significant updates to the R Commander following the publication of this book, along with errata correcting errors in the book (as they, almost inevitably, reveal themselves)
- a manual for authors of R Commander plug-in packages

A note about software versions: Although some of the “screenshots” and output in this book were produced with earlier versions, the book is current as of version 3.2.3 of R and version 2.1-4 of the **Rcmdr** package. Significant changes to the R Commander or changes to R that affect the R Commander will be addressed on the web site for the book.

Chapman and Hall maintains a link to the web site for the book at <https://www.crcpress.com/Using-the-R-Commander-A-Point-and-Click-Interface-for-R/Fox/p/book/9781498741903>, which can also be accessed at <http://tinyurl.com/RcmdrBook>.

---

## 1.6 Typographical Conventions

Different typefaces and fonts are used to distinguish the following elements:

- Computer software, such as operating systems (Windows, Mac OS X, Linux) and statistical software (R, the R Commander) are shown in a **sans serif typeface**.
- Graphical user interface components, such as menus (the *Edit* menu) and windows (the *R Console* window, the *Two-Way Table* dialog box), are shown in an *italic typeface*.
- Submenu and menu item selection is indicated by *>* (a greater than sign). Thus, for example, *Statistics > Summaries > Numerical summaries...* means “left-click on the *Statistics* menu, then on the *Summaries* submenu, and finally on the *Numerical summaries...* menu item.” Three dots following a menu item (as in *Numerical summaries...*) indicates that selecting the item leads to a dialog box, rather than performs a direct action. I will usually omit the three dots from menu items, however.
- Keys (e.g., *Tab*) and key combinations (*Ctrl-c*) are also shown in an *italic typeface*. The key combination *Ctrl-c*, for example, means “hold down the *Ctrl* key and press the *c* key.”
- R packages (such as the **Rcmdr** and **car** packages) are shown in **boldface**.
- Text meant to be typed directly (such as the R command `library(Rcmdr)`, or text to be typed into an R Commander dialog box) is shown in a **typewriter font**, as is R output, and as are the names of R objects, such as functions (`mean`), data sets (`States`), and variables in data sets (`pay`). Generic text (e.g., *variable-name*) meant to be replaced with specific text (e.g., `income`) is shown in an *italic typewriter font*.

- Files (`GSS.csv`) and file paths (`C:\Program Files\R\R-x.y.z\`) are also shown in `typewriter` font, with generic text again in *typewriter italics*.
- Internet URLs (addresses) are shown in a sans serif typeface (e.g., `https://cran.r-project.org`).
- When important, possibly unfamiliar, terms are introduced, they are set in *italics* (e.g., *rectangular data set*).



**Taylor & Francis**

Taylor & Francis Group

<http://taylorandfrancis.com>

---

## Installing R and the R Commander

---

This chapter describes how to download R from the Internet and install it and the R Commander on Windows, Mac OS X, and Linux/Unix systems. If you have already successfully installed R and the R Commander, then feel free to skip the chapter. There is, however, some troubleshooting information, to which you can make reference if you experience a problem.

---

### 2.1 Acquiring and Installing R and the R Commander

More detailed (and potentially more up-to-date) information on installing R and the **Rcmdr** package appears on the web site for this book. Please consult the web site if the information provided here proves insufficient, or if you encounter difficulties not discussed here.

R and R packages, like the **Rcmdr** package, are available on the Internet from CRAN (the Comprehensive R Archive Network—see [Chapter 1](#)) at <https://cran.r-project.org>. It is best not to download R and R packages directly from the main CRAN web site, however, but rather to use a CRAN *mirror site*. A link to a list of CRAN mirrors appears at the upper left of the CRAN home page (the top of which is shown in [Figure 2.1](#)). I suggest that you use the first “0-Cloud” mirror, which is generally both reliable and fast.

Regardless of whether you are a Windows user, a Mac OS X user, or a Linux/Unix user, I recommend that you install the current version of R, say R version  $x.y.z$ . In this generic version number, “ $x$ ” represents the *major version*, “ $y$ ” the *minor version*, and “ $z$ ” the *patch version* of R. A new minor version of R,  $x.y.0$ , is released by the R Core team each spring, and patch versions are released as needed, typically to fix bugs. Major versions appear infrequently, and only when substantial modifications are made to the base R software. As I’m writing this book, the current version of R is 3.2.3.<sup>1</sup>

There are (at most) five steps to installing R and the R Commander:

1. Download and install R.
2. On Mac OS X only, download and install the XQuartz windowing software.
3. Start R and install the **Rcmdr** package.
4. Load the **Rcmdr** package and, when asked, allow it to install additional packages.
5. If desired, optionally download and install Pandoc and L<sup>A</sup>T<sub>E</sub>X software for producing enhanced reports (as described in [Section 3.6](#)).

Specific instructions for Windows, Mac OS X, and Linux/Unix systems follow.

---

<sup>1</sup>The book, of course, was written over a period of time; R 3.2.3 was current when I was finalizing the text.