

Preface

The purpose of this volume is twofold: to help readers who are new to political research to learn the basics of how to use R and to provide details to intermediate R users about techniques they may not have used before. R has become prominent in political research because it is free, easily incorporates user-written packages, and offers user flexibility in creating unique solutions to complex problems. All of the examples in this book are drawn from various subfields in Political Science, with data drawn from American politics, comparative politics, international relations, and public policy. The datasets come from the types of sources common to political and social research, such as surveys, election returns, legislative roll call votes, nonprofit organizations' assessments of practices across countries, and field experiments. Of course, while the *examples* are drawn from Political Science, all of the *techniques* described are valid for any discipline. Therefore, this book is appropriate for anyone who wants to use R for social or political research.

All of the *example and homework data*, as well as copies of all of the example code in the chapters, are available through the Harvard Dataverse: <http://dx.doi.org/10.7910/DVN/ARKOTI>. As an overview of the examples, the following list itemizes the data used in this book and the chapters in which the data are referenced:

- 113th U.S. Senate roll call data (Poole et al. 2011). *Chapter 8*
- American National Election Study, 2004 subset used by Hanmer and Kalkan (2013). *Chapters 2 and 7*
- Comparative Study of Electoral Systems, 30-election subset analyzed in Singh (2014a), 60-election subset analyzed in Singh (2014b), and 77-election subset analyzed in Singh (2015). *Chapters 7 and 8*
- Democratization and international border settlements data, 200 countries from 1918–2007 (Owsiak 2013). *Chapter 6*
- Drug policy monthly TV news coverage (Peake and Eshbaugh-Soha 2008). *Chapters 3, 4, and 7*
- Energy policy monthly TV news coverage (Peake and Eshbaugh-Soha 2008). *Chapters 3, 7, 8, and 9*

- Health lobbying data from the U.S. states (Lowery et al. 2008). *Chapter 3*
- Japanese monthly electricity consumption by sector and policy action (Wakiyama et al. 2014). *Chapter 9*
- Kansas Event Data System, weekly actions from 1979–2003 in the Israeli-Palestinian conflict (Brandt and Freeman 2006). *Chapter 9*
- Monte Carlo analysis of strategic multinomial probit of international strategic deterrence model (Signorino 1999). *Chapter 11*
- National Supported Work Demonstration, as analyzed by LaLonde (1986). *Chapters 4, 5, and 8*
- National Survey of High School Biology Teachers, as analyzed by Berkman and Plutzer (2010). *Chapters 6 and 8*
- Nineteenth century militarized interstate disputes data, drawn from Bueno de Mesquita and Lalman (1992) and Jones et al. (1996). Example applies the method of Signorino (1999). *Chapter 11*
- Optimizing an insoluble party electoral competition game (Monogan 2013b). *Chapter 11*
- Political Terror Scale data on human rights, 1993–1995 waves (Poe and Tate 1994; Poe et al. 1999). *Chapter 2*
- Quarterly U.S. monetary policy and economic data from 1959–2001 (Enders 2009). *Chapter 9*
- Salta, Argentina field experiment on e-voting versus traditional voting (Alvarez et al. 2013). *Chapters 5 and 8*
- United Nations roll call data from 1946–1949 (Poole et al. 2011). *Chapter 8*
- U.S. House of Representatives elections in 2010 for Arizona and Tennessee (Monogan 2013a). *Chapter 10*

Like many other statistical software books, each chapter contains example code that the reader can use to practice using the commands with real data. The examples in each chapter are written as if the reader will work through all of the code in one chapter in a single session. Therefore, a line of code may depend on prior lines of code within the chapter. However, no chapter will assume that any code from previous chapters has been run during the current session. Additionally, to distinguish ideas clearly, the book uses fonts and colors to help distinguish input code, output printouts, variable names, concepts, and definitions. Please see Sect. 1.2 on p. 4 for a description of how these fonts are used.

To the reader, are you a beginning or intermediate user? To the course instructor, in what level of class are you assigning this text? This book offers information at a variety of levels. The first few chapters are intended for beginners, while the later chapters introduce progressively more advanced topics. The chapters can be approximately divided into three levels of difficulty, so various chapters can be introduced in different types of courses or read based on readers' needs:

- The book begins with basic information—in fact Chap. 1 assumes that the reader has never installed **R** or done any substantial data analysis. Chapter 2 continues by describing how to input, clean, and export data in **R**. Chapters 3–5 describe graphing techniques and basic inferences, offering a description of the techniques

as well as code for implementing them **R**. The content in the first five chapters should be accessible to undergraduate students taking a quantitative methods class, or could be used as a supplement when introducing these concepts in a first-semester graduate methods class.

- Afterward, the book turns to content that is more commonly taught at the graduate level: Chap. 6 focuses on linear regression and its diagnostics, though this material is sometimes taught to undergraduates. Chapter 7 offers code for generalized linear models—models like logit, ordered logit, and count models that are often taught in a course on maximum likelihood estimation. Chapter 8 introduces students to the concept of using *packages* in **R** to apply advanced methods, so this could be worthwhile in the final required course of a graduate methods sequence or in any upper-level course. Specific topics that are sampled in Chap. 8 are multilevel modeling, simple Bayesian statistics, matching methods, and measurement with roll call data. Chapter 9 introduces a variety of models for time series analysis, so it would be useful as a supplement in a course on that topic, or perhaps even an advanced regression course that wanted to introduce time series.
- The last two chapters, Chaps. 10 and 11, offer an introduction to **R** programming. Chapter 10 focuses specifically on matrix-based math in **R**. This chapter actually could be useful in a math for social science class, if students should learn how to conduct linear algebra using software. Chapter 11 introduces a variety of concepts important to writing programs in **R**: functions, loops, branching, simulation, and optimization.

As a final word, there are several people I wish to thank for help throughout the writing process. For encouragement and putting me into contact with Springer to produce this book, I thank Keith L. Dougherty. For helpful advice and assistance along the way, I thank Lorraine Klimowich, Jon Gurstelle, Eric D. Lawrence, Keith T. Poole, Jeff Gill, Patrick T. Brandt, David Armstrong, Ryan Bakker, Philip Durbin, Thomas Leeper, Kerem Ozan Kalkan, Kathleen M. Donovan, Richard G. Gardiner, Gregory N. Hawrelak, students in several of my graduate methods classes, and several anonymous reviewers. The content of this book draws from past short courses I have taught on **R**. These courses in turn drew from short courses taught by Luke J. Keele, Evan Parker-Stephen, William G. Jacoby, Xun Pang, and Jeff Gill. My thanks to Luke, Evan, Bill, Xun, and Jeff for sharing this information. For sharing data that were used in the examples in this book, I thank R. Michael Alvarez, Ryan Bakker, Michael B. Berkman, Linda Cornett, Matthew Eshbaugh-Soha, Brian J. Fogarty, Mark Gibney, Virginia H. Gray, Michael J. Hanmer, Peter Haschke, Kerem Ozan Kalkan, Luke J. Keele, Linda Camp Keith, Gary King, Marcelo Leiras, Ines Levin, Jeffrey B. Lewis, David Lowery, Andrew P. Owsiak, Jeffrey S. Peake, Eric Plutzer, Steven C. Poe, Julia Sofia Pomares, Keith T. Poole, Curtis S. Signorino, Shane P. Singh, C. Neal Tate, Takako Wakiyama, Reed M. Wood, and Eric Zusman.

Political Analysis Using R

Monogan III, J.E.

2015, XIII, 242 p. 38 illus., 19 illus. in color., Softcover

ISBN: 978-3-319-23445-8