Learning Base R

Hao Zhao

To cite this article: Hao Zhao (2022): Learning Base R, Journal of Quality Technology, DOI: 10.1080/00224065.2022.2097967

To link to this article: https://doi.org/10.1080/00224065.2022.2097967

Published online: 20 Jul 2022.
R is one of the most common programming languages used for various statistical modeling and data analysis tasks. Learning Base R (2nd edition) by Lawrence M. Leemis provides an accessible approach to R language for beginners with little to no programming exposure. Unlike other introductory R language books, Learning Base R is more in-depth from a statistical perspective, giving a fundamental overview of R language.

In this book, Chapter 1 provides a comprehensive introduction to R language, as well as some tricks to make the R session more efficient. The following two chapters introduce basic arithmetic operations. Then, three elementary data structures, vector, matrices, and arrays, are introduced in Chapters 4 to 6. Chapters 7 and 8 describe built-in and user-written functions, and Chapter 9 introduces some useful utilities. Notably, some new functions have been added to these chapters in this new edition, such as assign, append, and attributes.

The next three chapters introduce three other types of elements that can be stored in data structures, complex numbers, character strings, and logical elements. Chapters 13 and 14 introduce the methods for comparing elements with relational operators and coercing elements to specific data types. In addition, a new table summarizing the concept of “is family” of functions is provided in this chapter. Two more advanced data structures, lists and data frames, are introduced in the next two chapters. Chapter 17 shows some built-in data sets in R. Chapter 18 concerns input/output; a more sophisticated application of scan is also introduced here.

After introducing these essential topics, some advanced topics are illustrated in the following chapters. Chapter 19 focuses on some suitable functions associated with the probability distributions of random variables. Chapters 20 and 21 give, in very fine detail, how to generate high-level graphics and custom graphics, and Chapters 22 to 24 introduce many of R’s programming capabilities. Chapter 25 explains the topic of the Monto Carlo simulation. Furthermore, Chapters 26 to 28 have the most modification comparing to the previous version. Some brief introductions to statistical inference methods are given in Chapter 26, which includes univariate data analysis, analysis of variance, regression, and time series analysis. Chapter 27 introduces linear algebra functions. Chapter 28 covers some popular packages for data visualization and data analysis, such as ggplot2, lubridate, Ipolve, and other packages in the exercises section.

There are over 400 exercises in total, an increase of 265 new exercises (an average of 9–10 new exercises per chapter) from the previous edition, to enhance the reader’s knowledge of R. The book also includes plenty of instructional videos and code for readers to explore, which are available on the author’s website.

In conclusion, this book covers the R programming language and all its details in a practical way. For those who are just starting to learn R, this book, along with the relevant examples and exercises, will give you enough knowledge to start using R for various data analysis tasks. There are also some extra tips and tricks to help you write clean and optimized code. If you are keen to learn R from scratch or just want a refresher, I would highly recommend this book.

Hao Zhao
Arizona State University,
Tempe, AZ
hzhao83@asu.edu

© 2022 American Society for Quality
https://doi.org/10.1080/00224065.2022.2097967