Reviewer: Paul Hofmarcher
Vienna University of Economics

An Introduction to **R** for Quantitative Economics

This book has been published in the "Springer Briefs in Economics" series of Springer, which has quite a broad scope. In its preface the author states that the book "...gives an introduction to R to build up graphing, simulating and computing skills to enable one to see theoretical and statistical models in economics in a unified way. [...] The book addresses the specific needs of economists, and helps them move up the R learning curve. Students of economics or economists keen to learn how to use R would find this book very useful."

On 109 pages, the book covers 14 chapters: It starts with an introduction to R on 6 pages in which the author explains, e.g., how to create vectors in R or how to install and load packages. Next to a very short introduction to **RStudio**, matrices, data frames and lists are introduced in Chapter 2. Importing data into R is discussed in Chapter 3, whereby solely the `read.delim` and `read.csv` functions are explained to the reader. The next two chapters use the **mosaic** package to discuss supply and demand functions and functional forms in general, while Chapter 6 is devoted to present Cobb-Douglas production functions. In doing so, the author presents various plots of these widely used production functions. Chapter 7 introduces matrices again, as well as the linear regression model (5 pages) while Chapter 8 gives an introduction to statistics including the central limit theorem and the t test (7 pages). The next two chapters solely present Anscombe's quartet data to discuss linear regression and carbon storage data to perform multiple regression. The last four chapters cover the comparison of treatment with control groups (8 pages), the Solow growth model (7 pages), the simulation of random walks (7 pages) and basic time series calculations (8 pages). Each of these 14 chapters concludes with a section "Exploring Further" in which the author proposes further reading material.

I like the idea of an R textbook for economists, but as relevant and fascinating as the subjects in this book are, it is flawed in many ways. First, for me the overall structure of the book is not convincing and to some extent even confusing. For example, data structures like matrices, data frames and lists are mentioned in Section 2 and are followed by a discussion of supply-demand issues and functions, while Section 7 is then again devoted to matrices. In both sections the author misses to explain differences between the data structures in R.

Even further, he states that "We use matrices to store data", which might be confusing and misleading for R beginners, as there are well known differences between data frames and matrices in R and usually R data frames correspond to what we call a "data set".

One of the great strengths of R is its ability to add user defined functions through the `function()` call. Dayal proposes an inconvenient way of writing functions via the **mosaic** package and the `makeFun` call, which provides a mechanism for creating functions via the formula interface but he neglects to explain or mention the concept of the formula interface in R which is indispensable for R beginners.

Obviously, Dayal is excited about the **mosaic** package – which is in fact a great R package. It is used as the one and only method for creating functions, for plotting (`plotFun`) and even functions which do not belong to the **mosaic** package, like `xyplot` from the **lattice** package (Sarkar 2008), are attributed to the **mosaic** package by Dayal (p. 23).

In most chapters, the author uses external data sets and I am convinced that the used data are very interesting. But unfortunately Dayal only provides external links to the data and does not provide them to the community, e.g., through an R package or as data files maintained by the author/publisher which would ease the replication of the examples.

Finally, the book covers a broad range of topics, on its 109 pages, at the cost that most issues can only be touched on the surface and explanations are very scarce. To give an example, Section 8 is – despite its title "Statistical Simulation" - simply concerned with the most basic functions for generating random numbers in R, i.e., introducing `rnorm, rbin` and `runif`. The same is true for Chapter 9, which presents Anscombe's data. Although Anscombe's data is a very interesting example to demonstrate the importance of graphing the additional informational value of this section is limited, e.g., compared to the associated Wikipedia (2015) entry.

On the other hand, I really liked that the book makes a good choice in graphing Cobb-Douglas functions (Section 6). Especially the isoquant and the 3-D plot of that widely-used production functions might give valuable insights to readers unfamiliar with that concepts. However, I think the reader would benefit if those concepts were presented with more verbal exposition and I missed a link to the *micEcon* project, which provides several tools for microeconomic analysis and microeconomic modeling using R (see `http://micEcon.org/`).

On balance, the book is not a treasure trove of insights, neither as an introduction into R nor appealing with regard to its discussion of empirical economic examples. The information value is relatively low, and there are a number of annoying omissions, particularly regarding data structures, functional programming and data availability.

References

Sarkar D (2008). ***lattice: Multivariate Data Visualization with R.*** Springer-Verlag, New York.

Wikipedia (2015). "Anscombe's Quartet — Wikipedia, The Free Encyclopedia." URL `https://en.wikipedia.org/wiki/Anscombe's_quartet`, accessed 2015-10-06.

Reviewer:

Paul Hofmarcher
Vienna University of Economics and Business
Department of Economics
1020 Vienna, Austria
E-mail: Paul.Hofmarcher@wu.ac.at
URL: http://www.wu.ac.at/vw1/m/hofmarcher/