**Proposal for a New Edition**

**Title:** Learn R: As a Language

**Author:** Pedro J. Aphalo

**Weak and strong points in current edition:**

1. **–** Although the book focuses on the R language itself, given the intended audience, basic programming concepts are not explained with enough clarity.
2. **–** The overall focus of the book is on the language and grammar. However, the chapter on plotting with ggplot2 is closer to a cookbook of plotting recipes.
3. – The chapter on model fitting and statistic summaries is brief and could have more depth and include a few additional procedures to demonstrate a broader range of methods. Too few of the recommended packages included in the R distribution are described. A few contributed packages should have been described.
4. **+** Building a solid understanding of R before briefly introducing the ‘tidyverse’. My views are similar to those of Matloff, although possibly less extreme (https://github.com/matloff/TidyverseSkeptic).
5. **+** The book fills a gap in the available books by emphasizing the use of the R language in scripts, while still presenting the whole scope of the language.
6. **+** “2021 CHOICE Outstanding Academic Title Award Winner”.

**List of major potential changes and additions:**

1. **(Improve clarity)** Chapters 2, 3 and 5. The main step towards better explaining the basic programming concepts and the constructs of the R grammar that implement them will be adding diagrams and flow charts integrated into improved textual descriptions.
2. **(Improve consistency of approach)** **Chapter 7 on plotting with ggplot2**. Remove part of the case examples on plotting with ‘ggplot2’ and expanding on the structure of ggplot objects, how the different components are connected and how plot rendering takes place. I will use package ‘gginnards’ to demonstrate how ggplot (“gg”) objects are structured and when computations and plot rendering takes place. I will also put more emphasis in the “packaging” of partial plots into lists, functions and use the recently released and very promising package ‘ggpackets’ (Doug Kelkhoff, 2022 in CRAN). (In other words, emphasize could be called “ggplot scripting”.)
3. **(Update, minor) Chapter 7 on plotting with ggplot2**. Track changes in package ggplot2 and extensions such as the linewidth aesthetic and parameter orientation allowing flipping of statistics.
4. **(Expand contents) Chapter 2 on statistics.** Expand the explanations and add a table listing the various accessors for fitted model objects. Add diagrams highlighting differences and similarities in formal parameters and accessors among the most common model fit functions.
5. **(Expand contents) Chapter 2 on statistics.** Add quantile regression with ‘quantreg’ and major axis regression with ‘smatr’ and/or ‘lmodel2”. Give more emphasis to numerical approximation and bootstrapping methods and some of their applications. Describe rlm and some other functions from package ‘MASS’ and describe ‘lme4’.
6. **(Update, minor) Chapter 6 on grammar of data.** Track changes to the tidyverse and other packages.
7. (**Minor**) **Introductory Chapter 1.** Mention uses of R in machine learning and add some references as further reading on this subject and on bioinformatics.
8. **(Update, minor) Chapter 8 on Data import and export.** Track changes to the tidyverse and other packages.

**List of minor potential changes and additions:**

1. In addition to correcting known typing mistakes and other small errors, update all code examples to not only work with current versions of R and packages but, when possible, edit them to make use of any new features that simplify the code or can make it easier to understand.
2. My own packages have evolved significantly since the first edition, and if needed for the new edition I will revise them. Most likely ‘gginnards’ will gain some new features while I revise the chapter on plotting.
3. Package ‘learnrbook’ supporting the book will be updated.

My aim is that the new edition will have at most 400 pages (compared to 350 pages in the first edition). To achieve this some text and several examples will be removed to make space for new content. Some of the “cookbook” examples for ggplots will be moved to on-line chapters, and the on-line chapter on connecting R to code written in other programming languages will remain as on-line only in revised form. Currently, there are 4 on-line only chapters and four appendixes with an additional 110 pages for the first edition (<https://www.learnr-book.info/learn-r-extra-web.pdf>). Almost all these on-line material will remain on-line updated for the new edition.

**Questionnaire:**

|  |
| --- |
| **1) What is your general opinion of the book? What are its main strengths and weaknesses? Please give details.** |
|  |
| **2) Please explain how you have used the book for teaching in the past. What was the course title, and what students were on it? Did you use the whole book or selected chapters? Did you supplement with any other materials?** |
|  |
| **3) Please suggest any specific improvements to the book – material that could be added, revised, restructured, or deleted.** |
|  |
| **4) Do you have any comments on the proposed revisions outlined by the authors above?** |
|  |
| **5) Do you believe there is scope for a new edition of the book? Would you, for example, use the new edition for your teaching?** |
|  |
| **6) What are the key related and/or competing titles? How do they compare with this book?** |
|  |
| **7) Do you have any additional comments to help improve the book?** |
|  |