Author : Corey S. Sparks

Title of Work : Applied Demographic Data Analysis with R

Based on the materials provided by the author(s), we would like to have your advice on the following points:

1. Overall, do you feel that the textbook as outlined in the proposal will be a useful book for its intended audience?   
*Yes.*

2. Does it cover the topics mentioned adequately?  
*Yes.*

3. What is the relative strength of this proposal compared to textbooks that are currently used in similar courses?

*Bringing the macro and micro perspective of demography into the same framework with R being the main demonstration tool.*

4. Which parts of the concept would the author need to address before it is suitable for the market segment stated in the proposal? How could the author best address those concerns?

*The market segment is very broad. It presents the book as data analysis for social scientists in general. However, the topics in the proposal cover only a segment of the area of statistical demography. The book should be more specific about what it is, either in the form of changing the title, subtitle or the description.*

5. Which types of courses would benefit from having this textbook available? Are those self-study or classroom courses?

*Courses that cover survey analysis and statistical analysis of demographic data. It could be suited for both, self-study as well as classroom courses.*

6. If different from the readership stated in the proposal, who do you feel will be the primary and secondary audience for this textbook?

*Data scientists/analysts (non-demographers) who work with demographic data and want to learn more about statistical models to analyze it.*

*I’m not sure that I agree that “professionals/instructors who wish to learn how to use R” would be the primary users, as there is a huge amount of other R resources that cover learning R in much more depth than this book will.*

7. Are the planned pedagogical features of this textbook suitable and sufficient, and are they presented in a way that is engaging and attractive to the intended readership?

*I think so.*

8. Would this proposed textbook appeal to a global market? If there is a geographical bias that would make this more suitable for a limited audience, please indicate which audience this would be.

*The proposal reads like the book will not be geographically biased.*

9. Is the proposal written in an engaging style and is the English usage in the sample materials in accordance with the intended level and readership?

*Yes.*

10. Do you have other suggestions for improvement to the author(s)?

*I have a few general comments regarding the first Chapter:*

* *Code style:*
  + *The code would be more readable if spaces are used around all operators as suggested in many style guides, see e.g.* [*https://jef.works/R-style-guide*](https://jef.works/R-style-guide/)
  + *For a long time, it has been recommended to use TRUE and FALSE instead of T and F, see e.g.* [*https://www.r-bloggers.com/2012/12/r-tip-avoid-using-t-and-f-as-synonyms-for-true-and-false/*](https://www.r-bloggers.com/2012/12/r-tip-avoid-using-t-and-f-as-synonyms-for-true-and-false/)
* *I would recommend to carefully review the sequence of difficulty of the code. For example, details about functions could go after other R basics, or the use of pipes already on page 21 could be overwhelming.*
* *Demographic data often come in the “wide” format. Showing a conversion from “wide” to “long” in order to use ggplot would be useful.*
* *The book should be useful also for users who already know R and therefore can skip Chapter 1. As there will be users who might have different preferences regarding processing datasets (e.g. data.table users), my suggestion would be to make the code less dependent on dplyr. In fact, demographic data (especially micro data) often contain millions of rows. For such datasets, processing via data.table can increase the memory/time efficiency more than tenfold, in addition to shorter syntax. It would be a pity to make the book less valuable for users who do not want to pass on those benefits.*
* *The section on Variable Types is missing the (important) type “logical”.*

**\*\*\***