BOOK REVIEW

EDUCATIONAL STATISTICS AND DATA SCIENCES R FOR TIDYING AND MANIPULATING DATA

Author: Srisuthiyakorn, S. (2018)

Faculty of Education, Chulalongkorn University: Bangkok. 288 pages. ISBN: 978-616-407-601-3

Supamas Chumkaew[1]

Data analysis is a statistical method used to process the collected data and present it in a descriptive form explicitly. Currently, there are many statistical data analysis methods and programs that are widely used, some in the form of software packages that have been developed for users to import or input data into the programs. Then, the selective menu or the statistics are used to analyze the data in answering questions according to the purposes of data analysis at that time. Some applications and programs can only be used to analyze data with certain statistical values. If the advanced statistics are needed to analyze the data, certain compact computer programs must be used to assemble the second, third or fourth programs. Using multiple programs to analyze data to answer questions for specific purposes of data analysis may be an inconvenient method for users who have to purchase multiple programs for data analysis at a time.

R program (R) is a new computer language that is widely used in data analysis because R language is a program distributed to the general public with the right to freely access, develop, enhance, modify or improve it without any charge in any case and the data analysis results are widely beautiful and colorful. Users can customize it according to the purpose of each task. The code used to analyze the data can be easily written and shared online, turning R into an interesting and selective computer program for data analysis. Therefore, the use of data analysis methods depends on the aptitudes and abilities of the data analyst to decide which program to analyze the data appropriately.

The book, "Educational Statistics and Data Sciences R for tidying and manipulating data", is an interesting alternative for those who want to use R programs for data analysis because it is introduced from the fundamentals of their basic use, whether the readers have ever used the R program before or not. Also, this book will educate the readers by laying out the fundamentals of the installation of the program, the programming environment, the nature of the data, as

[1] Lecturer, Ph.D., Registration and Evaluation Office Sukhothai Thammathirat Open University.

Email: supamas.tun@gmail.com

well as the writing of the data analysis instructions to obtain the results of the data analysis in various forms to meet the reader's needs.

Siwachoat Srisuthiyakorn, the author of this book, has collected and compiled knowledge from direct experience related to statistics and data sciences both as a user of compact computer programs in data analysis and as a teacher who teaches students to analyze data. This makes this book interesting and different from other R programming books on the market as it focuses on the fundamentals needed to use R programming for formal data analysis and data management prior to data analysis. Therefore, the essence of this book is packed with important concepts, methods, and clear examples to help readers learn and understand R language, as well as choose and write commands for data analysis correctly.

In this book, the details of data analysis with the R program are presented for readers to study in 7 chapters. Each chapter contains the following topics:

Chapter 1 Introduction: This chapter discusses the scope of content and the basic knowledge required for readers to download and install the R program, packages, and examples of commands and datasets used in the book.

Chapter 2 Fundamentals of R program: This chapter discusses R programming environment, basic mathematical calculation, functions, variables, vectors, matrices, list variables, data sets, and classification variables.

Chapter 3 Importing data into R program: This chapter discusses working directory size, importing and exporting flat files, importing and exporting .xlsx data files, importing data from the clipboard and importing data files with the .sav extension.

Chapter 4 Organizing datasets: This chapter discusses the exploration of the structure of datasets, data collection focused on row-based datasets and column-based datasets, column splitting dataset format modification, dealing with redundant datasets, and merging data.

Chapter 5 Data exploration: This chapter discusses the basic graphics functions, exploring the distribution of variables, exploring variable conditions with basic statistics and exploring the relationship between variables.

Chapter 6 Data processing: This chapter discusses variable selection and data unit screening, data transformation and summary with basic statistics.

Chapter 7 Missing value analysis: This chapter discusses the exploration and transcoding of missing values, missing value mechanisms, data missing mechanism diagnosis, solving missing values with traditional methods, and multi-missing-value replacement.

The book, "Educational Statistics and Data Sciences R for tidying and manipulating data", is available in both physical and electronic forms. Readers can shop at their convenience. This is a good and valuable book. The knowledge presented in this book will help readers or researchers who are interested in data analysis with the R program can study and use it to expand their own knowledge as well.