

D.J. Treiman. (2009). *Quantitative Data Analysis: Doing Social Research to Test Ideas*. San Francisco, CA: Jossey-Bass.

*Reviewed by Michele Tarsilla*

As a doctoral student in evaluation, I have a great interest in gaining a better understanding of quantitative data analysis. Too often in the past, I have had to depend on professional statisticians in order to analyse data that I have collected. Therefore, I thought that reading Donald Treiman's book, *Quantitative Data Analysis: Doing Social Research to Test Ideas* would be beneficial. In particular, I started reading the book (a total of 417 pages divided into 16 chapters) with the anticipation that, by the end of it, I would feel more comfortable analysing complex datasets and reduce the need for external assistance. I also hoped that reading this book would help me increase my knowledge and understanding of such seemingly difficult statistical techniques as fixed-effects modeling and multinomial and ordinal logistic regression.

My expectations for this book were also quite high given the author's impressive academic and professional background. Unfortunately, the trepidation with which I started reading this book faded after the first 20 pages. By the end of the third chapter on cross-tabulation, I was already wondering how Treiman's book could have been used by students in statistics classes for almost three decades (as Treiman himself explains in the preface, his students at UCLA have been helping him over the years to fine tune the content of the book to fit students' needs). I actually found it quite difficult to categorize *Quantitative Data Analysis* as a textbook (i.e., as a book that a neophyte in quantitative data analysis could easily learn from) for three main reasons.

First, the book often lacks the simplicity of language that should characterize any reading specifically catering to students. Concepts were not always clearly explained and technical terms without adequate definitions were used so widely throughout the book that reading particularly difficult paragraphs more than once left me puzzled at times. Second, the book suffers tremendously from an unfortunate placement of statistics tables away from the accompanying text (i.e.,

a certain analysis was presented on one page and the corresponding table was found two or three pages later). As a result, I found it quite inconvenient, if not distracting, to go back and forth between the narrative of a certain analysis and the tables or graphs which the analysis referred to. Third, while Treiman sometimes provides the reader with practical tips on how to conduct data analysis with popular statistical software programs, most of the author's references included in the text referred to STATA and SAS rather than SPSS which, to my knowledge, is more widely used than the other two by those researchers with either a basic or intermediate statistical knowledge who represent the book's target audience. As a result, I would warn SPSS users (who do not have familiarity with either STATA or SAS) that most of the references made by the author on how to conduct quantitative data analysis would not necessarily apply to them.

In the end, despite Treiman's promise in the preface, I did not really feel that *Quantitative Data Analysis* empowered me "to produce publishable papers in my field or first-rate dissertation chapters." That said, despite the weaknesses highlighted earlier, there were several features of this book that I especially enjoyed.

First, the online availability of several datasets discussed in detail in some of the chapter provides motivated readers with the unique opportunity to put in practice some of the techniques learned from the chapters. In particular, the datasets touched on a wide variety of topics, including the following: Life Histories and Social Change in Contemporary China Survey (1996), Social Stratification in Eastern Europe after 1989 (1994), The Survey of Socioeconomic Opportunities and Achievement in South Africa (1994), and the US General Social Survey (2007).

Second, the inclusion of Technical Point tables throughout the text offers the reader a more in-depth explanation of concepts introduced earlier in the chapters. Interestingly enough, some of the boxes included in the text presented the biographies of illustrious statisticians often unknown to the general public (i.e., Paul Lazarsfeld, Hans Zeisel, R.A. Fisher, Philip Hauser) whose work has contributed immensely to the development of data analysis over the decades.

Third, I found Chapter 7 (Multiple Regression Tricks: Techniques for Handling Special Analytic Problems), Chapter 15 (Improving Causal Inferences: Fixed and Random Effect Modeling), and Chapter 16 (Final Thoughts and Future Directions: Research and Interpretation

Issues) particularly enlightening. Chapter 7 alone taught me how to construct and interpret linear splines as a way of representing abrupt changes in slope and how to decompose the difference between two means. Chapter 15 provided a concise presentation of fixed and random effects modeling in a more straightforward manner than any other statistics textbooks that I had previously consulted on this topic. Chapter 16, while reiterating some concepts I was already familiar with (i.e., sample selection bias, multilevel analysis, autocorrelation, probability sampling), provided me with quite useful tips on how to conduct better quantitative data analyses in the future including (a) using the Heckman correction to address a sample selection bias, (b) relying on ethnographic studies before invoking the concept of superpopulation, (c) modifying surveys to reflect relationships unanticipated by my theories but observed in the data of cross-sectional studies conducted in the past, and (d) conducting sensitivity analysis of my general linear model framework to both gain confidence and inspire confidence on the part of my readers.

In conclusion, despite the book's limitations, I would recommend this book to people who have had at least a year of intense graduate-level statistics, possess a moderate degree of familiarity with quantitative data analysis, are comfortable with algebra, and normally use (or are open to learning how to use) STATA to conduct data analysis. Before starting *Quantitative Data Analysis*, readers may want to remember, though, that this is not one of those books meant to be read on a train or before going to bed. In order to gain a better understanding of the very complex concepts presented and maximize the overall comprehension of the material presented in this book, readers should read it with a certain degree of concentration and at a relatively moderate pace. In doing so, the reader should also make sure to have an Internet connection available so as to access the databases and additional resources mentioned throughout the book and an adequate statistical analysis software (i.e., STATA or SAS) to put in practice the data analysis techniques presented in all 16 chapters.