
Willem Mertens • Amedeo Pugliese •
Jan Recker

Quantitative Data Analysis

A Companion for Accounting and
Information Systems Research

 Springer

Preface

Quantitative Data Analysis for Accounting and Information Systems Research guides postgraduate research students and early career researchers in choosing and executing appropriate data analysis methods to answer their research questions. It supports researchers when they are planning to collect data and when they have data that they want to analyze. Using a variety of examples and lay language, this book provides hands-on guidelines on (1) when to use which data analysis method, (2) what kind of data and data structure are required for each of the methods, (3) what each method does and how the methods should be used, and (4) how to report results.

This book is not intended to provide an exhaustive overview of data-analysis methods, nor does it explain the methods in depth. Instead, it guides researchers in applying the right methods in the right way. More skilled researchers can also use the book to refresh their knowledge or as a checklist to avoid skipping important steps. It explains the most commonly used methods in an intuitive and hands-on way, pointing out more advanced resources along the way. As such, it does not aspire to compete with manuals like those of Stevens [1], Field [2], or Crawley [3]. We are not statisticians but researchers who apply statistics,¹ so the book covers the issues that commonly affect others like us, who are engaging in quantitative empirical research.

Quantitative Data Analysis for Accounting and Information Systems Research is the book we would have liked to have had as a support in our own research. Every chapter provides an unintimidating starting point for building your data-analysis skills, the information required to run the most common analyses and report them, and pointers to more extensive resources. At the risk of saying things that may not be *entirely* true in the purest statistical sense, we try to keep the language of this book as simple as possible. As such, the book is brief and written in a language that we hope everyone can understand—from students to researchers to people who wish to study the organizations in which they work. Our goal is to help you conduct academic research of high quality and do the right things right—not to make you a

¹If you are a statistician or simply more observant than we are, we invite you to tell us if you identify an error.

statistics expert—so this book is not about statistics but about *applying statistics* to the research questions that keep you awake at night. (We doubt these questions are about collinearity, but if they are, this may not be the book you are looking for.)

In brief, this book is a software-independent starting point for answering the question: What methods do I use to answer my research questions and how?

We hope you have fun!

Brisbane, QLD, Australia
Padova, Italy

Willem Mertens
Amedeo Pugliese
Jan Recker

References

1. Stevens JP (2009) Applied multivariate statistics for the social sciences. Taylor and Francis, LLC, London
2. Field AP (2013) Discovering statistics using IBM SPSS statistics, and sex and drugs and rock 'n' roll, 4th edn. Sage, London
3. Crawley MJ (2013) The R book, 2nd edn. Wiley, West Sussex

Acknowledgments

Many people have contributed to the development of this book and its content. We are grateful to everyone who helped us discover, explain, and write down our understanding of what matters in using statistics for data analysis. Although there are many of you out there, we would like to thank a few in particular.

First, we are grateful for the support of the QUT's School of Management, School of Accountancy and Information Systems School, for supporting us in the development and conduct of the Advanced Data Analysis workshop series; this book would not have been possible without it. Special thanks go to Professor Michael Rosemann and Professor Peter Green for their inspiring entrepreneurial spirit, flexibility, and support. Second, we are grateful that so many of our colleagues and students attended these workshops and discussed and challenged our understanding of data analysis methods and the way we taught them.

Finally, the ones that contributed—or perhaps suffered—most are our lovely wives, Laura, Claudia, and Laura. Thank you for your support, your patience, and for sharing some of our headaches. You make our lives 89% more enjoyable ($p < .001$, [75–100]).

September 2016

Willem Mertens
Amedeo Pugliese
Jan Recker

Contents

1	Introduction	1
1.1	Introduction to the Basics	2
1.2	Navigating the World of Statistics—And This Book	3
1.3	What This Book Does Not Cover	5
	References	6
2	Comparing Differences Across Groups	7
2.1	One or Two Groups	7
2.2	More than Two Groups: One-Way ANOVA	10
2.3	More than Two Grouping Variables: Factorial ANOVA	12
2.4	More than One Dependent Variable: Multivariate ANOVA	14
2.5	More Advanced Models: Covariance and Repeated Measures	16
2.6	When to Use Group Comparisons	17
	References	19
3	Assessing (Innocuous) Relationships	21
3.1	What Are Regression Models?	21
3.2	When Do We Use Regression Models?	24
3.3	How Do We Examine Regression Models?	26
3.4	How Do We Report Regression Analyses?	31
3.5	What If.	32
	References	36
4	Models with Latent Concepts and Multiple Relationships: Structural Equation Modeling	37
4.1	What Are Structural Equation Models?	37
4.2	When Do We Use Structural Equation Models?	41
4.3	How Do We Examine Structural Equation Models?	43
4.4	How Do We Report Structural Equation Model Analyses?	49
4.5	What If.	51
	References	57
5	Nested Data and Multilevel Models: Hierarchical Linear Modeling	61
5.1	What Are Hierarchical Linear Models?	61
5.2	When Do We Use HLMs?	63

5.3	How Do We Investigate HLMs?	64
	References	71
6	Analyzing Longitudinal and Panel Data	73
6.1	What Are Longitudinal and Panel Data?	73
6.2	Clustering as a Way to Deal with Nestedness	77
6.3	Which Models Can We Use to Analyze Longitudinal Data?	81
6.4	Estimating and Reporting Fixed-Effects and Random-Effects Models	84
6.5	When to Use OLS, Fixed-Effects, and Random-Effects Models . . .	92
6.6	Final Remarks, Suggestions and Your Best Tool: Thinking!	96
	References	97
7	Causality: Endogeneity Biases and Possible Remedies	99
7.1	Your Research Question Is Causal: What Does that Mean?	100
7.2	Self-Selection and Endogeneity	105
7.3	Specifying OLS Models to Minimize Endogeneity Concerns . . .	107
7.4	More Complex Designs to Support Causal Claims	111
7.5	Some Caveats and Limitations	132
	References	133
8	How to Start Analyzing, Test Assumptions and Deal with that Pesky p-Value	135
8.1	Structuring, Cleaning, and Summarizing Data	136
	8.1.1 Structuring Data	136
	8.1.2 Cleaning Data	137
	8.1.3 Exploring Data: Summary Statistics and Visualization . . .	142
8.2	Testing Assumptions	143
	8.2.1 Independence of Observations	143
	8.2.2 Normality	147
	8.2.3 Homogeneity of Variance and Homoscedasticity	148
	8.2.4 Linearity	149
	8.2.5 What if Assumptions Are Violated?	150
8.3	Mindfully Interpreting Statistics: The Case of the p -Value	152
	References	155
9	Keeping Track and Staying Sane	157
	Index	161