Data Analysis Using *SAS Enterprise Guide*

This book presents the basic procedures for utilizing *SAS Enterprise Guide* to analyze statistical data. *SAS Enterprise Guide* is a graphical user (point-and-click) interface to the main SAS application. Each chapter contains a brief conceptual overview and then guides the reader through concrete step-by-step examples to complete the analyses.

The 11 sections of the book cover a wide range of statistical procedures, including descriptive statistics, correlation and simple regression, *t* tests, one-way chi-squares, data transformations, multiple regression, analysis of variance, analysis of covariance, multivariate analysis of variance, factor analysis, and canonical correlation analysis.

Designed to be used as either a stand-alone resource or an accompaniment to a statistics course, the book offers a detailed path to statistical analysis with *SAS Enterprise Guide* for advanced undergraduate and beginning graduate students, as well as professionals in psychology, education, business, health, social work, sociology, and many other fields.

Lawrence S. Meyers is Professor of Psychology at California State University, Sacramento. He teaches undergraduate and graduate courses in research design, data analysis, data interpretation, testing and measurement, and the history and systems of psychology. He was the coauthor of a textbook on research methods in the 1970s, has recently coauthored books on multivariate research design and analysis of variance, and has more than three dozen publications; some of his relatively recent work has been in areas such as measurement and testing and positive psychology. He received his doctorate from Adelphi University and worked on a National Science Foundation Postdoctoral Fellowship at the University of Texas, Austin and Purdue University.

Glenn Gamst is Professor and Chair of the Psychology Department at the University of La Verne, where he teaches the doctoral advanced statistics sequence. He received his doctorate from the University of Arkansas in experimental psychology. His research interests include the effects of multicultural variables, such as client–therapist ethnic match, client acculturation status and ethnic identity, and therapist cultural competence, on clinical outcomes. Additional research interests focus on conversation memory and discourse processing.

A. J. Guarino is on the faculty at Alabama State University, where he teaches graduate statistics courses in the Psychology Department. He received his bachelor's degree from the University of California, Berkeley, and he earned a doctorate in statistics and research methodologies from the University of Southern California through the Department of Educational Psychology.

Data Analysis Using *SAS Enterprise Guide*

Lawrence S. Meyers
California State University, Sacramento

Glenn Gamst
University of La Verne

A. J. Guarino
Alabama State University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo

Cambridge University Press
32 Avenue of the Americas, New York, NY 10013-2473, USA

www.cambridge.org
Information on this title: www.cambridge.org/9780521130073

First published 2009

Printed in the United States of America

A catalog record for this publication is available from the British Library

Library of Congress Cataloging in Publication data

Meyers, Lawrence S.
Data analysis using *SAS Enterprise Guide* / Lawrence S. Meyers, Glenn
Gamst, A. J. Guarino.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-521-11268-0 (hbk.) – ISBN 978-0-521-13007-3 (pbk.)
1. SAS (Computer file) 2. Enterprise guide. 3. Social sciences – Statistical methods – Data processing.
4. Mathematical statistics – Data processing. I. Gamst, Glenn. II. Guarino, A. J. III. Title.
HA32.M499 2009
330.0285′555–dc22 2009012822

ISBN 978-0-521-11268-0 Hardback
ISBN 978-0-521-13007-3 Paperback

Contents

Preface

The present book, *Data Analysis Using SAS Enterprise Guide*, provides readers with an overview of *Enterprise Guide*, the newest point-and-click interface from SAS. *SAS Enterprise Guide* is a graphical user (point-and-click) interface to the main SAS application, having relatively recently replaced the Analyst interface, which itself had replaced the original Assist interface. *Enterprise Guide* makes it easier than ever to access many SAS statistical analyses without learning to write the SAS code underlying its procedures.

We have written this book for readers who have little or no knowledge of *SAS Enterprise Guide* but who may wish to employ it for statistical analysis. Some of these readers will be students in an introductory statistics or data-analysis course; other readers will have taken an introductory statistics course and possibly a research methods course at some time in their past; and still other readers may have had several statistics and research design courses as a part of their background. We have therefore included in this book a relatively wide range of statistical procedures to meet the needs of various readers. There are chapters devoted to the more basic procedures such as descriptive statistics, correlation and simple linear regression, *t* tests, and one-way chi-square analysis. In addition, we have also included statistical procedures at a somewhat higher level; these include data transformations and other types of computations, multiple linear regression, logistic regression, and some analysis of variance designs. Finally, we have incorporated topics that are more advanced for those readers who might have the need to use such techniques as analysis of covariance, multivariate analysis of variance, factor analysis, and canonical correlation analysis.

Given the wide range and level of topics that we cover, it may not be surprising that the present book is intended to be neither a stand-alone statistics text nor a SAS "cookbook." Rather, our intent is to instruct readers on how to use *SAS*

Enterprise Guide to perform the statistical data analyses covered in the book as well as to understand the concepts underlying those procedures. That is, it is our belief that an exclusive and isolated "select this, then select that" robotic or cookbook synopsis of the steps involved in a given statistical analysis does not serve the needs of most readers. For this reason, we supply for each chapter some analytic and methodological context for the particular statistical procedure that we are describing, enabling readers to gain a sense of the research and statistical framework within which the particular procedure can be used. We also provide interpretations of the statistical results rather than just discussing how to read the output tables that were obtained from *SAS Enterprise Guide*.

There are 33 chapters in this book. They are organized into the following 11 sections.

Section I, "Introducing *SAS Enterprise Guide*," consists of two chapters presenting the basics of *SAS Enterprise Guide*. The software is designed to work on "projects." Chapter 1 describes what projects are and focuses on creating projects and navigating within them. Chapter 2 describes how to import data into projects, how to enter data directly into projects, and how to save projects.

Section II, "Performing Analyses and Viewing Output," consists of two chapters describing how to use *SAS Enterprise Guide*. Chapter 3 informs readers about how to select the statistical procedure they intend to use and how to interact with the dialog screens presented by *SAS Enterprise Guide* in the process of structuring the analysis. Chapter 4 addresses the management and viewing of output.

Section III, "Manipulating Data," contains three chapters focusing on some ways to organize existing data and generate new variables. Chapter 5 deals with sorting data and selecting a subset of the cases in the data set. Chapter 6 discusses how to recode variables into new or existing variables. Chapter 7 shows how to compute new variables.

Section IV, "Describing Data," consists of four chapters focused on descriptive statistical and graphical summary procedures. Chapter 8 focuses on computing measures of central tendency and variability. Chapter 9 shows how to graph data in different ways. Chapters 10 and 11 demonstrate how to generate standardized scores based on the sample mean and standard deviation (Chapter 10) and based on existing norms (Chapter 11).

Section V, "Score Distribution Assumptions," contains three chapters concerning some of the assumptions underlying most of the statistical procedures covered in this book. Chapter 12 explains what statistical outliers are and how to detect them. Chapter 13 focuses on the assessment of normality. Chapter 14 demonstrates how to perform data transformations in order to drive skewed distributions toward normality.

Section VI, "Correlation and Prediction," contains five chapters dealing with correlation as well as linear and nonlinear regression. Chapter 15 demonstrates how to perform a bivariate correlation analysis by using the Pearson product–moment correlation (*r*) and Spearman rho. Chapters 16 and 17 cover simple and multiple linear (ordinary least squares) regression, respectively. Chapters 18 and 19 describe the procedures involved in performing simple and multiple logistic regression, respectively.

Section VII, "Comparing Means: The *t* Test," contains three chapters encompassing different types of *t* tests. Chapters 20, 21, and 22 demonstrate how to conduct independent-groups *t* tests, correlated-samples *t* tests, and single-sample *t* tests, respectively.

Section VIII, "Comparing Means: ANOVA," contains four chapters. Chapters 23, 24, 25, and 26 describe the steps involved in computing analysis of variance (ANOVA) designs for a one-way between-subjects design, a two-way between-subjects design, a one-way within-subjects design, and a two-way mixed design ANOVA, respectively.

Section IX, "Nonparametric Procedures," consists of three chapters presenting some ways of analyzing frequency and rank-ordered data. Chapters 27 and 28 cover one-way and two-way contingency (chi-square) tables, respectively. Chapter 29 examines nonparametric one-way comparisons of means based on ranked data.

Section X, "Advanced ANOVA Techniques," is the first section focusing on advanced topics. It contains two chapters extending our treatment of ANOVA to more complex designs. Chapter 30 describes how to perform an analysis of covariance (ANCOVA). Chapter 31 demonstrates how to conduct a one-way multivariate analysis of variance (MANOVA).

Section XI, "Analysis of Structure," completes our book with two additional chapters on advanced topics, this time covering structural analysis. Chapter 32 describes how to perform and interpret an exploratory factor analysis. Chapter 33 focuses on canonical correlation analysis.

With the exception of those chapters in the first section in which we introduce the software and its interface, the chapters are generally structured in the following manner. We begin with an overview of the topic.

We then present some historical information on the statistical procedure where it is appropriate. We follow this by a numerical example – a data set that we subject to the statistical procedure that is the topic of the chapter. Most of the examples are based on data sets that we have created for this book, but a few draw on real data sets that we or our students have collected in the past; we make clear which is which when we present the data. We also very briefly describe the research design elements involved in the data collection to provide the context for the data sets. For

each numerical example, we also include a description of the *SAS Enterprise Guide* data set structure and a screen shot showing at least a portion of the data set.

We follow the numerical example by presenting step-by-step guidelines for setting up the analysis in *SAS Enterprise Guide*. Our presentation includes a narration of what has to be done and why it has to be done. This is accompanied by screen shots of the various dialog windows. Finally, we offer step-by-step guidelines for reading and interpreting the output (the printed results) of the analysis. These, too, are accompanied by screen shots of the output.

Acknowledgments

We wish to acknowledge and thank the following individuals for their efforts in maximizing the quality of this book. Lauren Cowles, our editor, has been most helpful and supportive to us during the entire writing process, and her assistant, David Jou, has been very responsive to our inquiries and requests. Peter Katsirubas, our Project Manager at Aptara, kept the production process moving and helped us through that stage of preparing the book. Finally, we are extremely grateful to Susan Zinninger for her marvellous copyediting skills; her time and effort have made the narrative smoother, more readable, and more consistent than what it was when we mailed it to Lauren.