Preface

With the increasing availability of data, there is an equally increasing need to train practitioners in data analytics methods aimed at delivering scientific insights. Like our first edition, this edition is written for the sole purpose of providing the foundations and the basic statistical knowledge with applications necessary for understanding the world around us. The goal of this book continues that of our first edition, that is, to recognize what questions can be answered relative to the data at hand, to identify the statistical methods needed to answer those questions, have a conceptual understanding of the basic mathematical equations behind the statistical methods used, to understand the limitations of the methods when assumptions are violated along with various remedies, to construct the SAS programming steps for generating the statistical output, and the ability to interpret the output and substantiate the tenability your hypothesized relationships.

Because the authors have over 40 years of combined experience in teaching statistics, we understand that not all people are created equal in mathematical ability or confidence. Given this reality, we have written this book with high performance anxiety in mind. The value of this SAS textbook is in its focus on conveying both statistical knowledge and application through the use of SAS software. The focus is interpreting statistics for applied reports and not professional journals. There is less focus on formulas, statistical theories, and literature from scholarly research journals, and more focus on statistical approaches commonly used in applied research. Our use of hands-on examples illustrates directly how methods can be applied to similar questions in the field.

By designing this textbook with demonstrations, it enables practicing social scientists to learn SAS. Important to learning SAS and statistics is the understanding of how other students have survived the process. The textbook includes prior students' responses to and interpretations of various SAS commands as well as highlights of how prior students came to understand confusing statistical or SAS material. Our hope is that prior students' perspectives and experiences will lower students' statistics and programming anxiety by understanding that many people experience and manage the same type of stress successfully.

There are several updates in this edition. First, we have updated, where necessary, all SAS programming code and output to reflect the latest version of SAS, namely SAS 9.4. In Chapter 2, where we had previously discussed how social scientists and

practitioners use statistics in their research, we incorporate additional information describing the statistical techniques most used by social scientists. In Chapter 3, where we describe the levels of measurement for variables, we now emphasize the role of levels of measurement in determining the appropriate statistical tests. When discussing measures of central tendency and exploring distributional properties, we emphasize the need to assess normality of a numeric continuous variable and describe how to generate the necessary SAS code to test the reasonableness of the normality assumption. In this edition, we moved the Chapter on getting to know your data with Proc Means to Chapter 5, followed by Chapter 6 which describes the importance of missing variables. We also spend additional space in Chapter 5 describing the structure of a code book (sometimes referred to as a data dictionary) and how it is directly linked to the survey used for collecting the data we analyze. In Chapter 7, we emphasize the need to apply data visualization techniques for understanding the shapes and trends in your data and how data visualization is a critical part of and the foundation of inferential statistics. We further emphasize the fact that understanding the variable types is critical in understanding what data visualization techniques should be employed. We describe and illustrate the SAS procedures that can be used to produce the data visualizations. Because the goal of our text is to aid in making sense of the world around us through inferential statistics, in Chapter 8, we are very methodical in providing the details of normal distributions, how to answer the probability that a numeric continuous variable has values within certain bounds by converting that normal distribution to a standard normal distribution, and extending that to sample means and how sampling distributions can be used to answer questions about means in the population. Specifically, the goal is to illustrate how conclusions can be made about the population by using a random sample of observations from that population. The remaining chapters emphasize how various research questions warrant specific statistical approaches and how the SAS software can be used to generate the output necessary for answering those questions.

Throughout this book we

- Move at a slower rate and cover basic programming and basic applied statistics material by providing many demonstrations addressing very basic statistics and highlight prior students' views of the learning process.

- Provide hands-on demonstrations and highlight SAS output in clear and concise ways to enable readers to rehearse techniques covered in the textbook and to distinguish when it is appropriate to use each statistical test.
- Highlight interpretation of statistical results to enable one to use statistics appropriately and prepare students and practicing professionals for additional self-guided learning of SAS.
- Incorporate space for student/professional annotation to reinforce skills, knowledge, and applications.

Once you complete this book, one should be able to determine if a particular method is appropriate to answer critical questions, why each method or is not appropriate. Caution should be taken after one finishes this book. It is an introduction to the most basic statistical methods. There are many more statistical techniques to be learned. Mastering all material covered in this book is only the beginning step to becoming proficient at statistical analysis and using SAS software.

Finally, it should be noted that a working knowledge of a programming language is an essential tool in being an effective and successful data analysts and that SAS is considered the best in the industry. The SAS Institute is a global leader is analytics, business intelligence and data management, serving customers in over 140 countries with software installations at more than 75,000 business, government, and university sites.

Acknowledgments

Many people have contributed directly or indirectly to the completion of this book. Applied practitioners and academicians reviewed drafts. We appreciate the time and effort that reviewers spent helping improve the book. We also thank students and friends who either field tested the book or served as a “box” contributor: Taras Fossett, André Patterson, Doris Palmer, Sara Kennedy, Heidi Wheeler, Elizabeth Eicher and others who did not want special mention.

Most of all, we want to thank Jeff (Joni’s husband) and Greg (Donna’s husband), and Kristen, Ashlee, and Jacob (Joni’s children) and Sophia and Hunter (Donna’s two youngest children who are still at home), and family members, as all endured some degree of frustration with us as we relentlessly pursued completion of the book.