

REFERENCES

- Dunnington, G. W. (1955), *Carl Friedrich Gauss, Titan of Science: A Study of His Life and Work*, New York: Exposition Press.
- Gauss, C. F. (1966), *Disquisitiones Arithmeticae*, New Haven, CT: Yale University Press.

A Primer of Ecological Statistics.

Nicholas J. GOTELLI and Aaron M. ELLISON. Sunderland, MA: Sinauer, 2004, xviii + 510 pp., \$34.95 (P), ISBN: 0-87893-269-0.

In their preface the authors ask, “Why another book on statistics?” Their answer is that they have yet to find a single text that meets two specific needs of ecologists: a general introduction to probability theory including the assumptions of statistical inference and hypothesis testing and a detailed discussion of specific designs and analyses typical in ecology and environmental science. They suggest that their book should supplement more standard texts.

Do ecology and environmental science deserve such special treatment? Yes. Many important problems in ecology lead to investigations where assumptions in standard techniques are not met, but students rarely receive further guidance in such cases. Experimental design in ecology is frequently more complex than the customary agricultural or biological examples provided in biometry texts. The authors devote two chapters to experimental design problems in ecology and provide considerable discussion of potential difficulties. The valuable chapter on multivariate analysis should provide enough background for ecology students to gauge suitability of this approach for their research.

Chapter 5, “Three Frameworks for Statistical Analysis” is by far the most interesting and potentially the most revolutionary for teaching statistics to ecologists and environmental scientists. The authors pose the simple problem of estimating the number of ant colonies in a field and in a woodland. Hypothetical (unfortunately) data are presented of quadrat counts from both habitats, and then analyzed using a randomization procedure, a parametric analysis, and a Bayesian analysis. The concepts involved are well explained and the comparison in inferences is nicely illustrated.

My only regret is that this approach was not developed and extended as a basis for the whole book. There are many advances in dependent data analysis and statistical modeling, which the authors mention in passing, that if included, could provide the ecology student with better analytical techniques and greater possibilities in designing and analyzing studies. The authors note that the majority of students do not have a facility in programming and urge them to get one. Perhaps that kept the authors from extensively connecting ecological research with current developments in statistics. However, if more recent developments are given a thorough exposition, perhaps that will stimulate ecologists to acquire the techniques necessary to use them.

E. David FORD
University of Washington

PROC SQL: Beyond the Basics Using SAS.

Kirk Paul LAFLEER. Cary, NC: SAS Institute, Inc., 2004, xiv + 365 pp., \$48.95 (P), ISBN: 1-59047-534-8.

As a long-time SAS programmer, I read this text with great anticipation, eager to gain a different perspective on programming. Although assuming a modicum of familiarity with the SAS language and database constructs, Lafler’s prose is inviting and does not require a great deal of subject-specific expertise on the part of his audience. Indeed, the only prerequisites for enjoying this book are a rudimentary knowledge of SAS and a general understanding of relational tables and their operations (something with which all database users will be comfortable). The subject progression is, for the most part, natural (although reading the chapters in order is by no means required). Lafler’s development is example-driven, and each example is short and self-contained. After giving a brief description of the problem, the author provides SQL code and the resulting SAS output.

Overall, it is hard to find fault with the book. However, there are a few minor oversights in the introductory chapter that could potentially confuse the novice reader. For example, the author refers to certain concepts before he has formally defined them only to offer explanations in later sections. Two examples are the use of `primary key` (first used on page 6 and defined on page 9) and `order by` (first used on page 30 and defined on page 93).

Chapter 1 (“Designing Database Tables”) presents a quick introduction to database design. The author discusses tables and normalization. Despite its brevity, the summary is complete and provides an adequate overview on the

subject. Even though the title of the book (“Beyond the Basics”) hints at a level of familiarity, I found it interesting that the acronym SQL (Structured Query Language) is not formally defined.

Oddly lacking in Chapter 2 (“Working with data in PROC SQL”) is any mention of the general syntax of PROC SQL. (The author states as much in the preface—claiming that since other books address the topic of syntax, his will not.) Also missing is a discussion detailing how PROC SQL syntax differs from that of the other base SAS procedures. Some readers might appreciate a cursory overview of these topics encompassing no more than a few pages. Otherwise, the examples in this chapter are excellent and easy to follow. This is a first-rate overview of SQL operators, functions and predicates, as well as SAS numeric, character, and date values. A clever method of looking at dictionary tables through PROC SQL is mentioned at the end of the chapter, including an example with SAS system options and their current settings.

Chapter 3 (“Formatting Output”) is a summary of output customization methods, most of which will be familiar to the reader as they are shared with other SAS base procedures. Examples show how to exploit them in PROC SQL code. This section also covers the practical formatting capabilities of output delivery system. In particular, examples illustrate the mechanics of converting SAS output to either a SAS dataset, an RTF document, or an HTML file. Chapter 4 (“Coding PROC SQL Logic”) discusses conditional logic and its use within PROC SQL. Expressions such as `case`, `where`, and `coalesce` are covered in great detail and lay the groundwork for transitioning from basic SQL coding to more complicated programming such as embedding SQL in a host macro language.

Chapters 5 (“Creating, Populating, and Deleting Tables”) and 6 (“Modifying and Updating Tables and Indexes”) focus on the data management aspects of PROC SQL. Topics covered here include table design, the `create tables` statement, the `like` clause, and populating the tables via the `insert into` statement and the `select` query. There is an entire section devoted to integrity constraints and the consequences of violating them. Chapter 6 concludes with examples such as modifying and updating tables as well as creating and using indexes.

Chapter 7 (“Coding Complex Queries”) introduces the reader to perhaps the most exciting and commonly used aspect of PROC SQL—combining data from different tables. A wide variety of complex queries and joins are discussed in great detail. Throughout, Lafler provides numerous examples in an insightful, simple, and intuitive manner. In doing so, he subtly reminds the reader of the advantages PROC SQL provides over the more involved programming steps required to accomplish these tasks in the base language.

Chapter 8 (“Working with Views”) provides possibly all the information that the average SAS programmer might ever need regarding views. The author clearly and superbly outlines the definition of a view and its differences from a regular database table. In addition to highlighting their advantages, he is even-handed in his treatment and also alerts the reader to potential drawbacks (such as the longer processing times than tables). The author illustrates the uses of views, stepping through examples which include eliminating data redundancy, restricting data access, and hiding the complexities of a query.

Chapters 9 (“Troubleshooting and Debugging”) and 10 (“Tuning for Performance and Efficiency”) close the book with discussions of diagnostic techniques, remedial measures, and code enhancement. The author presents a thorough overview of common programming errors and discusses PROC SQL debugging options with clear explanations on when to use them. Perhaps the most enjoyable part of chapter 9 is a section with examples of errors that the author, himself, committed. Such candor is appreciated, as the reader can learn from the author’s experience. The examples follow a common structure: first presenting an error message issued by the SAS log, followed by the true description of the problem (which in some cases is quite different), the code that generated the error, and a suggestion of corrective action. These are a wonderful supplement for all levels of programmers, particularly novices and those still struggling to interpret the true meaning of SAS generated errors. Advanced users will most likely enjoy the final section with examples on constructing efficient queries.

In closing, all levels of SAS users will find this book useful. The valuable tips, examples, and suggestions found in this monograph make it a welcome addition to every programmer’s cache of code. Its clarity, simplicity, and logical layout make this book an indispensable guide for any PROC SQL programmer. The author states it best when he says that this book “complements the SAS SQL documentation very nicely.”

Izabela E. ANNIS
VNU Inc.