

## BOOK REVIEW

### **Health Care Data and the SAS System**

By Marge Scerbo, Craig Dickstein, and Alan Wilson

ISBN 1-58025-865-4

Publication # 57638

Review by Jack Shoemaker, Accordant Health Services

There are hundreds, perhaps thousands, of SAS(r) consultants and contractors scattered about the United States who owe their livelihood, in large part, to being able to unravel the tangled web of health-care data. In a single, easy-to-read book full of useful and relevant examples, Scerbo, Dickstein, and Wilson, have placed the future employment prospects of these folks in serious jeopardy.

The first three chapters cover the data proper - what to expect, how they are organized, some of the pitfalls and 'gotcha's' to avoid. These chapters should become required reading for any user of health-care data regardless of analytic tool. A large part of the challenge for anyone new to health-care data is getting the lingo straight. Terms like "claims" and "encounters" are carefully explained and placed in the context of the health-care-delivery business.

Chapter 3 contains samples of the primary data collection forms as well as pictures of how those data might manifest themselves in record descriptions and data files. Chapter 4 is a tutorial on how to use SAS to read complex data structures. The narrative is peppered with code samples to amplify what they have written. The chapter builds up to a rather elaborate DATA step that includes pointer controls, conditional output, arrays, and iterative input.

Although the example at hand is a UB92 record, the techniques explained could just as easily apply to any complex record with repeating sets of fields. Garbage in; garbage out. If the data are noisy, the slickest program in the world won't help. In Chapter 5, the authors describe how to use the tools available in SAS to test the validity of the data. As with chapter 4, the narrative contains many code examples showing how to use `FREQ`, `MEANS`, `SORT`, `UNIVARIATE`, and `PRINT` to explore the data.

Chapter 6 is another tutorial on using Base SAS. Strip away the health-care theme, and this chapter could stand on its own merits as a sweeping survey of the many functions that SAS has at the ready to manipulate character strings, numbers, and dates. There's even a short section on using the SAS macro facility to accomplish simple, repetitive tasks.

Continuing with the tutorial theme of the middle chapters, chapter 7 provides sanguine advice about the most powerful, yet most error-prone, part of SAS - the tools used to combine data. The `MERGE` statement is explained in wonderful detail complete with all the caveats you would come to expect at this point. All sorts of merges are discussed - one-to-one, one-to-many, and the notorious many-to-many. The `MERGE` statement is contrasted with interleaving `SET` statements, `PROC APPEND`, and `PROC SQL`. Each tool has its place and the authors do an excellent job of laying out the reasons why you would use one technique over another. Once you understand the data content and you have sufficiently beaten them into submission; it's time for analysis.

Chapter 8 provides a review of the many tools SAS supplies for reporting and analysis. Careful consideration is given to data-reduction tools like `PROC SUMMARY` as well as data display tools like `PROC REPORT` and `PROC GPLOT`. In keeping with the prescription-for-use thread that weaves its way through the whole book, the chapter concludes with a no-nonsense table suggesting what `PROC` to use and when.

Chapter 9 brings all these concepts together as a case-study of emergency-room utilization in Maryland. The book then concludes with a perspicacious view of the future. There are new coding initiatives under foot; there are emerging technologies and treatment modalities; there are complex regulatory requirements and changing business climate. Perhaps the legion of SAS health-care consultants needn't search for new work just yet.