

Preface to the Second Edition

It was nearly 10 years ago that I wrote the preface for the first edition of this book. The goals and scope of this book are still the same, but in this second edition you will find new data management features that have been added over the last 10 years. Such features include the ability to read and write a wide variety of file formats, the ability to write highly customized Excel files, the ability to have multiple Stata datasets open at once, and the ability to store and manipulate string variables stored as Unicode.

As mentioned above, Stata now reads many file formats. Stata can now read Excel files (see section 2.3), SAS files (see section 2.4), SPSS files (see section 2.5), and even dBase files (see section 2.6). Further, Stata has added the `import delimited` command, which reads a wide variety of delimited files and supports many options for customizing the importing of such data (see section 2.7.1).

Stata can now export files into many file formats. Stata can now export Excel files (see section 3.3), SAS XPORT 8 and SAS XPORT 5 files (see sections 3.4 and 3.5), and dBase files (see section 3.6). Additionally, the `export delimited` command exports delimited files and supports many options for customizing the export of such data (see section 3.7). Also, section 3.9 will illustrate some of the enhanced capabilities Stata now has for exporting Excel files, showing how you can generate custom formatted reports.

The biggest change you will find in this new edition is the addition of chapter 11, titled “Programming for data management: Part II”. Chapter 11 builds upon chapter 10, illustrating how Stata programs can be used to solve common data management tasks. I describe four strategies that I commonly use when creating a program to solve a data management task and illustrate how to solve 10 different data management problems, drawing upon these strategies as part of solving each problem. The concluding discussions of each example talk about the strategies and programming tools involved in solving the example. I chose the 10 examples in this chapter not only because the problems are common and easy to grasp but also because these programs illustrate frequently used tools for writing Stata programs. After you explore these examples and see these programming tools applied to data management problems, I hope you will have insight into how you can apply these tools to build programs for your own data management tasks.

Writing this book has been both a challenge and a pleasure. I hope that you like it!

Ventura, CA
May 2020

Michael N. Mitchell