



---

# *Journal of Statistical Software*

March 2005, Volume 12, Book Review 2.

<http://www.jstatsoft.org/>

---

Reviewer: Thomas Lumley  
University of Washington

---

## **A Visual Guide to Stata Graphics**

Michael N. Mitchell  
Stata Press, College Station, TX, 2004.  
ISBN 1-881228-85-1. 395 pp. \$49.75 (p).

---

**Stata** 8.0 brought a long-awaited update to the graphical capabilities of the package, which were becoming increasingly outdated. The new graphics provide more attractive default layouts, ways to customize these defaults, and much greater programming flexibility. As a result, the level of skill needed to produce **Stata** graphs has suddenly shot up. This book goes some way towards providing a solution.

Most of the book consists of six annotated examples per page, with a **Stata** command, the resulting graph, and a description. Each graph is developed incrementally over several examples. The description explains the changes and the relevant code additions are indicated by color. As with the old **Stata** graphics system a graph must be produced with a single command, but the incremental approach illustrated by Mitchell is likely to be the best strategy for putting together this command. For example, Chapter 2 starts with a scatterplot of home ownership by home cost across the US states, and shows how to annotate it, subdivide by region, add regression lines and smoothers, and then goes on to deal similarly with the residuals from the linear regression.

After handling scatterplots and scatterplot matrices, Mitchell goes on to cover boxplots, dot plots, pie charts, bar charts and density plots, and time-series plots in similar detail. The book begins with an overview of the available graphs and how they are created, and the last three chapters cover options that can be added to most or all graph commands. In addition to simple changes to the rendering style or layout, these options include the ability to add text, lines, or other markers to annotate a graph. An appendix briefly describes the other built-in statistical graphs, ranging from simple summaries such as survival and ROC curve to periodograms, autocorrelation plots, and leverage-vs-residual plots. As would be expected, there is an accompanying website with the data and everything else needed to reproduce the graphs.

The excellent coverage of standard graph types is at the expense of the other two innovations in **Stata** 8: graphical schemes and graphical programmability. The appendix has  $2\frac{1}{2}$  pages on creating schemes, which allow all the graphical defaults to be customized. Varied schemes are used throughout the book, showing what is possible, but little information is given on how to

harness this power. Similarly, the appendix provides a few more complicated examples (back-to-back histograms, scatterplots with two-way error bars) but no indication of how one might create new effects similar to the shaded confidence bands that appeared around regression lines in Chapter 2, or write new commands for, say, a stratified scatterplot smoother.

This book will be useful to anyone needing to produce fairly standard graphs in **Stata**. The format, with small graphs on the outside of each page, is ideally suited to flipping through. Simply copying the code from an example is likely to be faster than using the **Stata** menu system to design a graph, and users will then gradually learn the (fairly complex) syntax. However, there is no discussion of when one graph type or color scheme might be preferable to another, and little information on creating new graph types, limitations that are more likely to be important to readers of this journal than to most audiences.

**Reviewer:**

Thomas Lumley  
University of Washington  
Department of Biostatistics  
Box 357232  
Seattle, WA 98195-7232, United States of America  
E-mail: [tlumley@u.washington.edu](mailto:tlumley@u.washington.edu)