Journal of Official Statistics
Vol. 10. No. 4, 1994, pp. 467–468

Book Review

Books for review are to be sent to the Book Review Editor Gösta Forsman, Department of Mathematics, University of Linköping, S-581 83 Linköping, Sweden.

Cleveland, W.S., Visualizing Data. Hobart Press, Summitt, New Jersey, 1993. ISBN 0–9634884–0–6 CLOTH. 360 pp., $.

This text by William S. Cleveland is a superb addition to the understanding of methodology and technology for graphing and fitting equations to data. It is an excellent and timely follow-up to three earlier classics on graphical and exploratory data analysis. These three earlier classics were John Tukey's *Exploratory Data Analysis* in 1977; Edward Tufte's, *The Visual Display of Quantitative Data*, in 1983; and William S. Cleveland's *The Elements of Graphing Data* in 1985.

Visualizing Data provides a timely and thorough look at graphing and fitting equations to data. The impressive array of current computer graphics packages, such as S-Plus, are taken advantage of in this text and reference book.

One quite significant effect is that he addresses both modeling and graphics together. He examines a number of graphical methods including coplots, box plots, 3-D plots, M-D plots, residual dependence plots, sliced distribution plots and "banking" data plots. In addition, data fitting tools such as nonparametric local regression loess, other linear and nonlinear models, time series, data transformations and parametric bisquare methods are also addressed in this book. The author covers a substantial amount of material, but in a well organized form that flows smoothly from topic to topic.

In particular, the concept of enhancing visual perception by the use of graph "banking" to an optimum angle was quite enlightening. Cleveland explains that the data rectangle of a bivariate graph encloses the data on the graph. The aspect ratio of a graph is the height of that data rectangle divided by the width. Choosing the aspect ratio of a graph to enhance the visual perception of the orientation of line segments in the graph is called "banking". The term "banking" is used to describe this data display method and the name is analogous to the "banking" of a road to affect its desired slope.

The one surprise was the author's repeated use of one particular agricultural experiment data set from the 1930s. It definitely captured the reader's attention

which was the goal to illustrate the power of many of the techniques in the book. However, this is the one area where I thought the author may have overemphasized his conclusion, which was based on strong circumstantial evidence. In general, agricultural experiments that compare crop varieties can have difficulties in controlling the uniformity of the plots. In addition, there are numerous other potential recording or other nonsampling type errors that can creep into such an experiment. For example, say each variety was to receive the same rate of a chemical such as fertilizer or a pesticide or herbicide. But in actuality, the applicator device was partially clogged for a certain plot but then cleaned for other plots. I believe that the author did indeed provide very strong circumstantial evidence that one plot had the years mislabeled. However, I thought the author's statement that "the agronomists were less interested in the agricultural conclusions of the paper than brandishing the new method (ANOVA) of analysis" was not completely justified since some undetected errors can and often do creep into experiments.

Overall, this book is an excellent source and reference text for statisticians working with statistical graphics and data modeling. My compliments to the author and publisher for the excellent quality black and white printing of the graphics illustrations. This procedure kept the book more affordable but still highly informative and, as a result, will reach more end users.

If the market is ready, I would also encourage the author to consider a color based CD-ROM version of data visualization that could have rather astounding statistical graphics displayed.

References Cited in the Text

Cleveland, W.S. (1985). The Elements of Graphing Data. Belmont, CA: Wadsworth and Brooks-Cole.

Tufte, E.R. (1983). The Visual Display of Quantitative Data. Cheshire, CT: Graphics Press.

Tukey, J.W. (1977). Exploratory Data Analysis. Reading, MA: Addison-Wesley.

References Not Cited in the Text

Carr, D.B. (1994). Converting Tables to Row-Labeled Plots. Technical Report 101, Center for Computational Statistics, George Mason University, Fairfax, VA.

Cleveland, W.S. (1993). A Model for Studying Display Methods of Statistical Graphics. Journal of Computational and Graphical Statistics, 2, 323–343.

Kosslyn, S.M. (1994). Elements of Graphic Design. New York: Freeman and Co.

Tufte, E.R. (1990). Envisioning Information. Cheshire, CT: Graphics Press.

George A. Hanuschak
U.S. Department of Agriculture
National Agricultural Statistics Service
Fairfax, VA, U.S.A.