STATISTICAL GRAPHICS FOR VISUALIZING DATA

William G. Jacoby
ICPSR and Michigan State University

ICPSR Summer Program
July 27-29, 2009

http://polisci.msu.edu/jacoby/icpsr/graphics

SELECTED REFERENCES

I. **Books and Monographs:**

Chambers, John M.; William S. Cleveland; Beat Kleiner; Paul A. Tukey. (1983) *Graphical Methods for Data Analysis*. Pacific Grove, CA: Wadsworth Brooks/Cole.

Chen, Chun-houh.; Wolfgang Härdle; Antony Unwin (Editors). (2008) *Handbook of Data Visualization*. Berlin, Germany: Springer.

Cleveland, William S. (1993) *Visualizing Data*. Summit, NJ: Hobart Press.

Cleveland, William S. (1994) *The Elements of Graphing Data (Revised Edition)*. Summit, NJ: Hobart Press.

Cook, R. Dennis and Sanford Weisberg. (1999) *Applied Regression, Including Computing and Graphics*. New York: Wiley-Interscience.

Cook, Dianne and Deborah F. Swayne. (2007) *Interactive and Dynamic Graphics for Data Analysis with R and Ggobi*. New York: Springer.

Fox, John. (2000) *Nonparametric Simple Regression: Smoothing Scatterplots*. Thousand Oaks, CA: Sage.

Fox, John. (2002) *An R and S-Plus Companion to Applied Regression*. Thousand Oaks, CA: Sage.

Friendly, Michael. (1991) *SAS System for Statistical Graphics*. Cary, NC: SAS Institute, Inc.

Friendly, Michael. (2000) *Visualizing Categorical Data*. Cary, NC: SAS Institute, Inc.

Gower, J.C. and David J. Hand. (1996) *Biplots*. London, UK: Chapman & Hall.

Jacoby, William G. (1997) *Statistical Graphics for Univariate and Bivariate Data*. Thousand Oaks, CA: Sage.

Jacoby, William G. (1998) *Statistical Graphics for Visualizing Multivariate Data*. Thousand Oaks, CA: Sage.

Kosslyn, Stephen M. (1994) *Elements of Graph Design*. New York: W. H. Freeman.

Kosslyn, Stephen M. (2006) *Graph Design for the Eye and Mind*. Oxford, UK: Oxford University Press.

Mitchell, Michael N. (2004) *A Visual Guide to STATA Graphics*. College Station, TX: StataCorp LP.

Murrell, Paul. (2006) *R Graphics*. Boca Rotan, FL: Taylor and Francis Group, LLC.

Playfair, William. (2005) *The Commercial and Political Atlas and Statistical Breviary*. Cambridge, UK: Cambridge University Press. (Originally published, 1786. This edition edited and introduced by Howard Wainer and Ian Spence).

Robbins, Naomi B. (2004) *Creating More Effective Graphs*. New York: Wiley.

Sarkar, Deepayan. (2008) *Lattice: Multivariate Data Visualization with R*. New York: Springer.

Swayne, Deborah F.; Dianne Cook; Andreas Buja; Duncan Temple Lang; Hadley Wickham; Michael Lawrence. (2006) *Ggobi Manual*. Manuscript available online at the Ggobi web site: `http://ggobi.org/docs/`.

Theus, Martin and Simon Urbanek. (2009) *Interactive Graphics for Data Analysis: Principles and Examples*. Boca Raton, FL: Chapman & Hall/CRC.

Tufte, Edward R. (2001) *The Visual Display of Quantitative Information (Second Edition)*. Cheshire, CT: Graphics Press.

Tukey, John W. (1977) *Exploratory Data Analysis*. Reading, MA: Addison-Wesley Publishing Company.

Unwin, Antony; Martin Theus; Heike Hofmann. (2006) *Graphics of Large Datasets: Visualizing a Million*. New York: Springer

Wainer, Howard. (1997) *Visual Revelations*. New York: Copernicus.

Wainer, Howard. (2005) *Graphic Discovery: A Trout in the Milk and Other Visual Adventures*. Princeton, NJ: Princeton University Press.

Wainer, Howard. (2009) *Picturing the Uncertain World*. Princeton, NJ: Princeton University Press.

Wilkinson, Leland. (2005) *The Grammar of Graphics (Second Edition)*. New York: Springer-Verlag.

Young, Forrest W.; Pedro Valero-Mora; Michael Friendly. (2006) *Visual Statistics: Seeing Data with Dynamic Interactive Graphics*. New York: Wiley.

II. **Articles, Chapters, and Manuscripts:**

Becker, R. A. and W. S. Cleveland. (1996) "Trellis Graphics User's Manual." Unpublished manuscript (although all of this material is included in the documentation for the S-Plus computing language). A PDF version of this document is available on the web site for this course: `http://polisci.msu.edu/jacoby/icpsr/graphics`.

Becker, R. A.; W. S. Cleveland; M. Shyu. (1996) "The Visual Design and Control of Trellis Display." *Journal of Computational and Graphical Statistics* 5: 123-155. A PDF version of an earlier draft of this manuscript is available on the web site for this course: `http://polisci.msu.edu/jacoby/icpsr/graphics`.

Becker, R. A.; W. S. Cleveland; M. Shyu; S. P. Kaluzny. (1995) "A Tour of Trellis Graphics." Unpublished manuscript. A PDF version of this document is available on the web site for this course: `http://polisci.msu.edu/jacoby/icpsr/graphics`.

Denby, Lorraine and Colin Mallows. (2009) "Variations on the Histogram." *Journal of Computational and Graphical Statistics*. 18: 21-31.

Fox, John. (1990) "Describing Univariate Distributions." In John Fox and J. Scott Long (Editors). *Modern Methods of Data Analysis*. Newbury Park, CA: Sage.

Gabriel, K.R. (1971) "The Biplot Graphical Display for Matrices with Application to Principal Components Analysis." *Biometrics* 58: 453-467.

Jacoby, William G. (2000) "Loess: A Nonparametric, Graphical Tool for Depicting Relationships Between Variables." *Electoral Studies* 19: 577-613.

Jacoby, William G. (2006) "The Dotplot: A Graphical Display for Labeled Quantitative Values." *The Political Methodologist* 14 (1): 6-14. Available online at `http://polmeth.wustl.edu`.

Spence, Ian and Stephan Lewandowsky. (1990) "Graphical Perception." In John Fox and J. Scott Long (Editors). *Modern Methods of Data Analysis*. Newbury Park, CA: Sage.

Wainer, Howard and Paul F. Velleman. (2001) "Statistical Graphics: Mapping the Pathways of Science." *Annual Review of Psychology* 52: 305-335.

Young, Forrest W.; Richard A. Faldowski; Mary M. McFarlane. (1993) "Multivariate Statistical Visualization." In C. R. Rao (Editor) *The Handbook of Statistics, Volume 9*. New York: Elsevier Science Publishers.

III. **Software:**

ARC. A small, but powerful, software system for regression analysis. It is intended to accompany the Cook and Weisberg (1999) text listed above. Available at `www.stat.umn.edu/arc/`.

Ggobi. A data visualization system for viewing high-dimensional data. Available at: `www.ggobi.org`.

Mondrian. A general purpose statistical data-visualization system, containing several tools that are useful for categorical data. Available at: `http://stats.math.uni-augsburg.de/Mondrian`.

R. An extremely powerful language and environment for statistical computing and graphics. R is virtually identical to the S language, and the commercially-available software environment, S-Plus. Available at: `www.r-project.org`.

ViSta. The Visual Statistics System. A software package based on the premise that a researcher can interact with data and statistical models through graphical, as well as numeric, representations. Available at: `www.visualstats.org`.