

Smoking, Stephen Jay Gould and Convertibles:

Statistical Graphics for Data Presentation and Analysis

Some Web Sites of Interest

<http://hesweb1.med.virginia.edu/biostat/teaching/statcomp/>

Frank Harrell's (U Va) page on statistical computing and graphics – includes links to many of the sites referred to in this talk, and lots of great free stuff (from Frank and others) on interesting related topics.

<http://www.math.yorku.ca/SCS/Gallery/>

Michael Friendly's Gallery of Data Visualization, housing examples of some of the best and worst of statistical graphics. Includes the Ithaca Times and Stephen Jay Gould examples I cited.

<http://cm.bell-labs.com/cm/ms/departments/sia/wsc/index.html>

Web site of William Cleveland, including S-plus scripts and data sets for duplicating the plots in his book *Visualizing Data*. Also lots of great stuff on trellis display for multivariate graphics.

<http://lib.stat.cmu.edu/>

StatLib – THE system for distributing statistical software, data and info.

<http://www.stat.umn.edu/arc/>

Main page for *Applied Regression Including Computing and Graphics* by Cook and Weisberg. Includes the **ARC** software (which can be downloaded for free) and information on Xlisp-Stat and regression graphics.

<http://madison.hss.cmu.edu/>

A note on the research still going on into voting irregularities in Palm Beach, FL in the 2000 presidential election, with links galore.

<http://www.r-project.org/>

Home page for the R project. R is similar to S, but is free and available to all. While lacking some of the functionality and support of S-plus, the graphics can be just as terrific, and there are a lot of people working in R.

<http://www.stat.duke.edu/sites/java.html>

JAVA for statistical distributions, inference, and (some) visualization.

<http://www.stat.sc.edu/rsrch/gasp/>

Globally Accessible Statistical Procedures for doing lots of data analysis and educational things. Lots of nice interactive graphics here.

<http://www.bell-labs.com/topic/societies/asagraphics/index.html>

American Statistical Association Statistical Graphics Section - great links!

Thomas E. Love, PhD

E-mail: tel3@po.cwru.edu Web: www.chrp.org (slides as PDF)
Methods Seminar Series (IV), Center for Health Care Research & Policy
January 25, 2002

Smoking, Stephen Jay Gould and Convertibles:

Statistical Graphics for Data Presentation and Analysis

Some Books You Might Want to Know About

William S. Cleveland *The Elements of Graphing Data* (1994) Hobart Press.

This is a terrific look at the visual communication of data, and principles of graph construction. Some people (like Frank Harrell) consider this an essential discussion of “how” to do things well. The principles and methods are supported by a rigorous, scientific discussion of graphical perception, the visual decoding of information from data displays.

William S. Cleveland *Visualizing Data* (1993) Hobart Press.

This book takes *The Elements* as a starting point, and requires some familiarity with basic statistics and regression analysis. Some tools in the book are new and some are old, but all have a proven record of success in the analysis of common types of statistical data that arise in science and technology.

R. Dennis Cook and Sanford Weisberg *Applied Regression including Computing and Graphics* (1999) Wiley. (each man has also written other, similar texts)

See the web site. This is a great book that explains graphical regression and demonstrates it with the ARC software. Very software-specific, though.

Edward Tufte *The Visual Display of Quantitative Information* (1983) Graphics Press.

– *Envisioning Information* (1990) Graphics Press.

– *Visual Explanations: Images and Quantities, Evidence and Narrative* (1997) Graphics Press.

All three of Tufte’s books are intellectually stimulating and gorgeous to behold. Tufte, a self-publisher, takes extraordinary pains with design and production. *The Visual Display of Quantitative Information* focuses on charts and graphs that display numeric information. *Envisioning Information* explores similar territory emphasizing maps and cartography. *Visual Explanations* centers on dynamic data – information changing over time. Try www.edwardtufte.com

John Tukey *Exploratory Data Analysis* (1977) Addison-Wesley.

The orange book that serves as a bible for much of modern data analysis (actually, I believe Tukey’s writing is generally better with co-authors.)

Howard Wainer *Visual Revelations: Graphical Tales of Fate and Deception from Napoleon Bonaparte to Ross Perot* (2000) Lawrence Erlbaum.

Howard is a terrific speaker and a prolific and effective writer. He writes a quarterly column called *Visual Revelations* for *Chance* magazine, and this collection of his columns is a superb collection of great stories, including a couple I’ll tell today. Another collection is coming soon, I hear. Howard is also a translator of Jacques Bertin’s seminal *Semiologie Graphique* and author of *How to display data badly*. (1984) *The American Statistician*, 38, 137-147, and *Depicting error*. (1996) *The American Statistician*, 50, 101-111.

Leland Wilkinson *The Grammar of Graphics* (1999) Springer-Verlag.

Thomas E. Love, PhD

E-mail: tel3@po.cwru.edu Web: www.chrp.org (slides as PDF)

Methods Seminar Series (IV), Center for Health Care Research & Policy

January 25, 2002