Presentation Graphics

Leland Wilkinson

SPSS Inc., 233 South Wacker, Chicago, IL 60606

Department of Statistics, Northwestern University, Evanston, IL 60201

email: leland@spss.com
KEY WORDS: statistical graphics, charts

Abstract

This paper surveys briefly the history of presentation graphics, principles of usage, and applications.

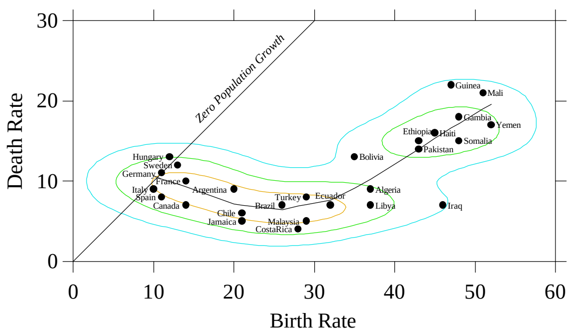
To appear in the *International Encyclopedia of the Social & Behavioral Sciences*.

1 Introduction

Figure 1 shows a graphic of death rates against birth rates per 100,000 population for 27 selected countries in a 1990 UN databank (Wilkinson, 1999). The contours in show two concentrations of countries. One, to the left, has relatively lower death rates and small-to-moderate birth rates. The second, toward the upper right, has high death rates and extraordinarily high birth rates. The curve in the middle of the contours shows that the overall relation between death and birth rates is curvilinear.

Figure 1

Plot of 1990 death rates against birth rates for selected countries.



The single plot in Figure 1 contains three graphs: (1) a point graph (collection of points) whose labels show country names, (2) a set of contours based on a kernel density estimate (Silverman, 1986) that represents the concentration of the countries, and (3) a LOESS smoother (Cleveland and Devlin, 1988) . The graphic also includes three guides for our understanding. The first is a general geometric object called a form that is, in this instance, a line delineating zero population growth. Countries to the left of this line tend to lose population, and countries to the right tend to gain. The other two are guides that demarcate axes for the represented space. Other examples of guides are legends and titles.

Each of the constituent graphical elements in Figure 1 has a long lineage in the history of statistical charts. Scatterplots of points and contours are at several centuries old, and regression smoothers are at least a century old. The use of a rectangular coordinate system to locate graphs within a frame demarcated by axes predates by centuries Decartes' eponymous Cartesian coordinate system. We will review briefly the **history** of statistical graphics and charts.

The lines, colors, typography, proportions, and other aspects of Figure 1 have been chosen to facilitate communication of the ideas and models underlying the chart. We need not resort to platitudes such as, "A picture is worth a thousand words" to understand that this graphic comprises a relatively substantial analysis of the data. We will review briefly the topic of graphics **principles**, which can help us devise effective ways to display analyzed data.

The world of charts is enormous. Principles can guide us through that world, but there are many alternatives for representing one or more variables in a graphic. In the final section, we will present some of these **choices** and use them to derive additional guidelines for developing effective presentations.

2 History

The contours, or level curves, in Figure 1 represent a map. This representation is based on a statistical function that maps birth rate and death rate onto a density. Any value on this density can be derived from a corresponding pair of birth rate and death rate values taken from the domain of values within the frame bounded by the two displayed scales. The geometric space in which this mapping is embedded is three-dimensional. The first two dimensions are birth rate and death rate, and the third dimension is density.

The technical mathematical terms in the previous paragraph (contours, level curves, map, function, onto, density, value, domain, frame, bounded, scales, geometric space, dimensions) are derived from ordinary language, often ancient, describing the physical world. This correspondence between the world of charts and the physical world is not coincidental. Charts are maps of abstract worlds. The word chart and the word cartography have the same root (Latin *charta*, a piece of paper or papyrus). The Greek word for geometry means land measurement. A scale (in Latin) is a ladder. We map in order to organize our world in our mind. Our world is larger than the domain of geography, however. As Pinker (1997) has argued, abstract reasoning is built on metaphors for reality. We manipulate abstractions by making them analogous to concrete objects.

Statistical charts are maps, but it does not suffice to describe them simply as abstract geography. The history of statistical charts contains remarkable leaps from a world of continuous rectangular (or spherical) space to other worlds that are categorical (worlds with gaps), multidimensional (worlds beyond three dimensions), and topological (worlds in abstract coordinate systems). Funkhouser (1937), Tilling (1975), Beniger and Robyn (1978), Fienberg (1979), Robinson (1982), Stigler (1983), Tufte (1983, 1990, 1997), Collins (1993), Wainer (1997), and Wainer and Velleman (2001) document these leaps. These references also contain numerous examples of graphics, ancient and modern, that represent significant landmarks in this history. The remainder of this section will highlight a few.

Physical coordinates were used by Egyptian surveyors as early as the third millenium B.C.E. for locating points on land (Beniger and Robyn, 1978). By the beginning of the second millenium C.E., coordinates were used to locate non-geographic entities. Funkhouser (1936) shows an example of a line graphic of planetary movements on a space-time grid produced by an unknown astronomer in the 10th or 11th century. Not surprisingly, one of the "planets" is the Sun.

Cartesian coordinates were devised by Descartes for locating points in abstract space, particularly for representing equations underlying his analytic geometry (Funkhouser, 1937). Relatively few 17th or 18th century embeddings of observed data points in Cartesian coordinates have been found, however (Tilling, 1975, Funkhouser 1937). Scientists of the Enlightenment sometimes regarded graphics as dilettantish, and if they used them at all, it was for geometrical diagrams or maps of algebraic functions rather than for plots of raw data (Beniger and Robyn, 1978). Social scientists of this period focused on collecting tables of state statistics that could be perused in detail without resort to graphical approximation (Beniger and Robyn, 1978, although see Stigler, 1983). Beniger and Robyn (1978) and Wainer and Velleman (2001) discuss notable exceptions to these observations (graphics by Wren, Halley, Huygens, Plot, Lister, Priestley) but conclude that the important varieties of what we now call the statistical chart did not arise until the work of Playfair in the late 18th century.

Playfair graphed the data of the state, the root of the word statistics. He invented the pie chart, the circle chart, the categorical bar chart, the area chart, and he adapted other geometric forms in novel ways to graph data over time and space. Wainer and Spence (1997) provide a brief biography of Playfair’s fascinating life and the charts he created.

Other widely used graphical forms were developed in the 19th and early 20th centuries. Beniger and Robyn (1978) and Stigler (1983) summarize a number of these. Density bars were used by Guerry in 1833 to display crime by age groupings. Pearson adopted this form for representing a frequency distribution and named it a histogram. Minard developed a number of schematic maps overlaying statistical data on geography. Florence Nightingale introduced polar area charts. The age-sex pyramid (a dual histogram) was introduced by F.A. Walker, Superintendent of the U.S. Census in 1874. The sample cumulative distribution plot was introduced by Galton in 1875.

If charts are maps of abstract worlds, we might conclude that guiding principles for graphics usage could be derived from the psychology of perception and cognition in the physical world. Indeed, recent psychological research has tended to support the simple statement that the more graphics adhere to the rules of the physical world, the more likely we are to extract and process the information in them reliably and accurately. The following section covers this area.

3 Principles

Over the centuries of their history, various charts have emphasized *communication* (e.g., Snow’s map of cholera deaths in London, reproduced in Tufte, 1983), *persuasion* (Fletcher’s map of the distribution of ignorance in 19th century England and Wales, reproduced in Wainer, 1997), and *graphic design* (e.g., Playfair’s chart of imports and exports of England to and from North America, reproduced in Tufte, 1983). These criteria are not exclusive, of course. Some charts incorporate all three (e.g., Minard’s compound chart of Napoleon’s troop losses in the Russian campaign, reproduced in Tufte, 1983).

Not surprisingly, modern critics of charts have tended to align themselves along these same dimensions. Communicators (e.g., Cleveland, Kosslyn) have conducted experiments to determine rules that improve the transfer of information from chart-maker to chart-viewer. Persuaders (e.g., Holmes, 1991) have evaluated charts on the basis of their psychological impact. Designers (e.g., Herdeg, 1981) have applied aesthetic criteria. Others (e.g., Tufte, 1983) have adopted the criteria of the communicator, and the methods of the designer. None of these approaches can claim exclusive validity. Since scientists are more interested in rules that foster accurate communication of replicable results, however, this section will focus on communication.

Evaluating graphical communication has been the province of psychology, in the subdisciplines of human-factors, ergonomics, and applied cognitive science. For almost a century, human factors psychologists have developed methods and criteria for evaluating graphics. Military psychologists, for example, have conducted randomized experiments to test the effectiveness of aircraft cockpit displays.

Psychologists gave much less attention to the perception of statistical graphics until a statistician, William Cleveland, published a series of experiments designed to identify aspects of charts that helped or hindered accurate decoding of quantitative data (Cleveland and McGill, 1984). (Beniger and Robyn, 1978, cite a number of statisticians’ informal studies of the effectiveness of popular charts in the 1920’s, but these did not induce many psychologists to study the topic more formally).

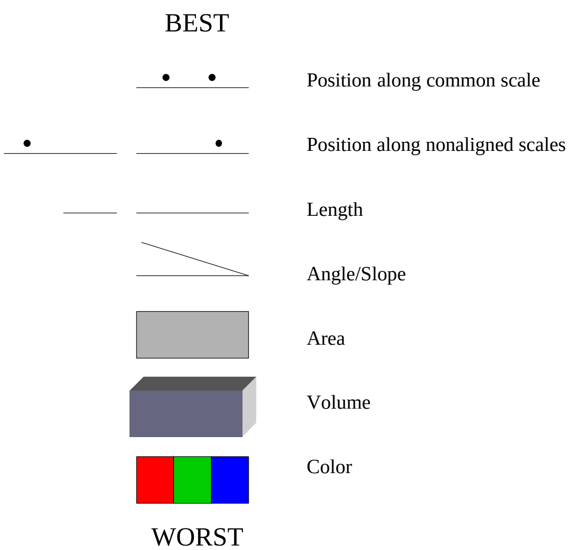
Cleveland used paper--and--pencil tests containing simple graphical elements -- points, lines, angles, areas, colors -- generated randomly by a computer. His subjects gave magnitude estimates of the values generating the instances of each element. From these estimates, he derived and analyzed error rates. Figure 2 summarizes the main result of these experiments. The top of the figure (BEST) represents elements that resulted in fewer errors and the bottom (WORST) represents those with more.

Cleveland’s hierarchy gives us a start in designing effective displays. All other things being equal, we might choose to use a bar chart instead of a pie chart because judgments of position on a common scale and judgments of length are both more accurate than judgments of angle. There are exceptions to these rules, however (see the discussion of Figure 3 below). A chart is more than the sum of its parts; elements interact in the context of a chart to suppress or enhance errors. Kosslyn (1994) summarizes relevant psychological studies in the extensive notes at the end of his book.

Others have provided numerous general rules for effective graphical communication. The following guidelines are derived from Bertin(1981), Tufte (1983), Cleveland(1994), and Kosslyn(1994).

- Avoid clutter. Especially, do not clutter the data region inside the plotting frame demarcated by axes. Tufte constructs a data/ink ratio and urges us to maximize it by removing irrelevant detail. This exhortation has value as long as we do not remove redundant features that can reinforce an accurate perception. Most writers agree that one should avoid jazzy textures (especially stripes), gratuitous colors, excessive tick marks and grid lines, ornate fonts, and unnecessary embellishment. Avoid visual illusions due to pseudo-3D and other gratuitous use of color, angles, area, or volume in a 2D plotting world.
- Seek simple geometric forms -- straight lines, circles, triangles, squares. When it helps to simplify a relationship, use transformations (logs, square roots) to convert curves to straight lines, for example. Use polar coordinates when variables are cyclical (astronomical, perceptual, directional).
- Sort and organize. This is especially important for complex graphics. Usually, one should sort the values of a dependent variable (the range of the graph), but other sorts can help viewers navigate and perceive structure. Bertin(1981) has numerous examples showing the value of simple and multivariate data sorts.Figure 12 below shows an example.
- Annotate extensively. Make legends comprehensive and informative. Figure captions should describe the source of the data and explain the relation of the data to the graphic. Note exceptions (see Figure 6).

Figure 2
Cleveland (1994) graphical features hierarchy.



4 Choices

Rules guide, but examples instruct. We need to consider specific choices for specific problems to understand exceptions to rules. The graphic representation problem is often expressed as one of choosing the proper chart type for a given set of data. This is an unhelpful way to express the problem, however. As Wilkinson(1999) shows, different chart types with their own popular names are often mirror images or simple transformations of each other. Instead, we will examine several basic problems confronting the designer of a scientific graphic: **geometry** (choosing the type of graphical element for a chart), **coordinates** (choosing a coordinate system for the chart), and **uncertainty** (representing random error).

4.1 Geometry

The following examples were selected to illustrate principles, not exhaust possibilities. It is organized by the configuration and number of variables, categorical and continuous, that we wish to graph.

4.1.1 One Variable

Figure 3 shows five different ways of representing values on one categorical variable. The data are from 1606 respondents to the 1993 General Social Survey (Davis, Smith, and Marsden, 1993). Respondents were asked, "How many sex partners have you had in the last 12 months?" Those reporting more than 4 partners (some reported up to 100) were consolidated into the last category. There were 1466 responses in the resulting 6 categories.

The upper left graphic is a bar chart of the frequencies of occurrence of each category. Bars are most suited for displaying either magnitudes referenced against zero (anchored bars, as in this figure) or a range of values on a continuous scale (range bars, as in Figure 7).

Anchoring bars at zero can sacrifice resolution when no data values exist near zero; this creates substantial white space in the plotting area (Cleveland, 1994). In these cases, dots are preferable to bars because we do not need to include zero on the scale. The upper right graphic shows a dot plot of the same variable. In this case, we would prefer the bar version because there are several values near zero and it is easier to anchor the bars visually on the partner categories. The locations of both the dots and the tops of the bars are easily decoded on the vertical scale.

The third graphic in the figure is a pie chart of the proportions of total category frequencies. The general antipathy statisticians have toward pie charts is misinformed. Pie charts are not always bad, although the issues on when to use them are complex. In general, bar charts are more suited to absolute judgments and pie charts are more suited to judging proportions (Simkin and Hastie, 1987). Pie sectors near a fourth or half of a pie in size are judged especially accurately. When there are many sectors, however, pies are not usually appropriate. In this example, the bar chart would be more suitable.

The fourth graphic is a divided bar chart, which is nothing more than a pie chart in rectangular coordinates. Divided bars do worse than pies in studies of graphical perception (e.g., Simkin and Hastie, 1987). Cleveland (1994) provides several examples to show why they are generally a bad choice.

The fifth graphic is a 3D pie chart. This is probably the worst of the popular ways to represent proportions. Three dimensional graphics are not in themselves harmful. What creates the problem with the 3D pie chart is the distortion of angles that the perspective projection produces. In the 3D pie, the single partner category seems to cover almost 75 percent; the actual coverage is 68 percent.

Some users and computer programs apply line or area elements (see Figure 7) to categorical variables. Popular examples are line graphics of means in analysis of variance, or area graphics of profiles across categories. These elements should generally be avoided with categorical data. Spacings between categories (even ordered categories) are not fixed, so slopes of line segments or profiles between categories are meaningless.

Figure 4 shows five ways of representing one continuous variable. The data are per-capita consumption of spirits for each of the 50 US states (Bureau of the Census, 1986).The top graphic is a classic histogram overlaid with a kernel density estimate (Silverman, 1986). Histogram bars look like ordinary bars, but they measure areas rather than intervals and they rest on a continuous scale rather than on a set of categories. Histograms are more useful than kernels when we are interested in displaying frequencies within intervals, especially when the intervals themselves are bounded by meaningful, round units. Kernels are more useful than histograms when we are looking for a smooth estimate of a continuous distribution.

The next lower graphic is a dot histogram (Chambers et al, 1983, Wilkinson, 1999). This display represents each observation with a dot located at its scale value. If several values coincide on the scale, the dots are stacked vertically. This display is most suited for small samples, when every value is to be displayed.

The next lower graphic is a stripe plot (Chambers et al, 1983). A vertical stripe occurs at each data value. This display can handle more cases than the dot plot, although it is not suitable for large datasets.

The next lower graphic is a jittered density (Chambers et al, 1983). Random error is added to each value so that the plotting symbols do not overlap. This display is designed to handle rougly the same conditions as the stripe plot.

The last graphic in Figure 4 is a box plot, or schematic plot (Tukey, 1977). As Tufte (1983) notes, Tukey’s plot is based on an earlier display that represents the quartiles of a distribution. In that earlier plot, the whiskers cover the range and the box covers the midrange. Tukey improved on this design by scaling the whiskers to allow for outliers. The display in Figure 4 shows how important this feature can be. Two states (Nevada and New Hampshire) have unusually large values (see Figure 6), possibly due to gambling and cheap liquor respectively. The box plot is a clean representation of the summary values we usually want to see – the median, the 25th and 75th centiles (hinges), the extreme values (usually bounding the range), and possible outliers (outside values).

4.1.2 Two variables

Figure 5 shows four ways of representing two categorical variables crossed. The data are from the 1993 General Social Survey used in Figure 3. The additional variable is general happiness, as measured by the response to the question, "Taken all together, how would you say things are these days – would you say that you are very happy, pretty happy, or not too happy?"

The upper left plot shows a 3D bar chart of happiness against number of sexual partners. The vertical scale is the proportion of respondents within each partner category. This graphic has the disadvantage that short bars tend to be hidden by long ones. It generally should be avoided.

The graphic to the right is a tiled bar chart. This graphic uses color or gray scale to indicate the relative frequency of each category combination. As Cleveland’s research has shown, however, the use of color or gray scale to encode continuous values is problematic (Cleveland, 1994).

The next graphic below shows a paneled categorical plot. Instead of using an extra dimension in 3D, we use an extra categorical dimension in 2D to panel the plot. This is probably the most effective method for handling the extra dimensionality required for representing the frequencies plus the two categories.The bottom panel shows a multiple divided bar chart. This layout has the same problems as the single divided bar chart (see Figure 3). Cleveland (1985) shows an example of this graphic and offers a more effective dot plot alternative.

Figure 6 shows the most popular way of representing two continuous variables crossed with each other – the scatterplot. The data are per-capita consumption of spirits for each of the 50 US states, used in Figure 4 (Bureau of the Census, 1986). The additional variable is number of deaths from chronic liver disease and cirrhosis per 100,000 people by state. This scatterplot has been enhanced in a number of ways. The bordering box plots help us to assess the marginal distributions of the variables and highlight the Nevada and New Hampshire outliers. The smoother shows the conditional mode of deaths (estimated mode of deaths given consumption). This was computed through kernel regression (Scott, 1992). This particular smoother limits itself to areas where there are relatively more data values, so it provides a conservative estimate of trend.

Figure 3

Reported number of sexual partners (General Social Survey, 1993).

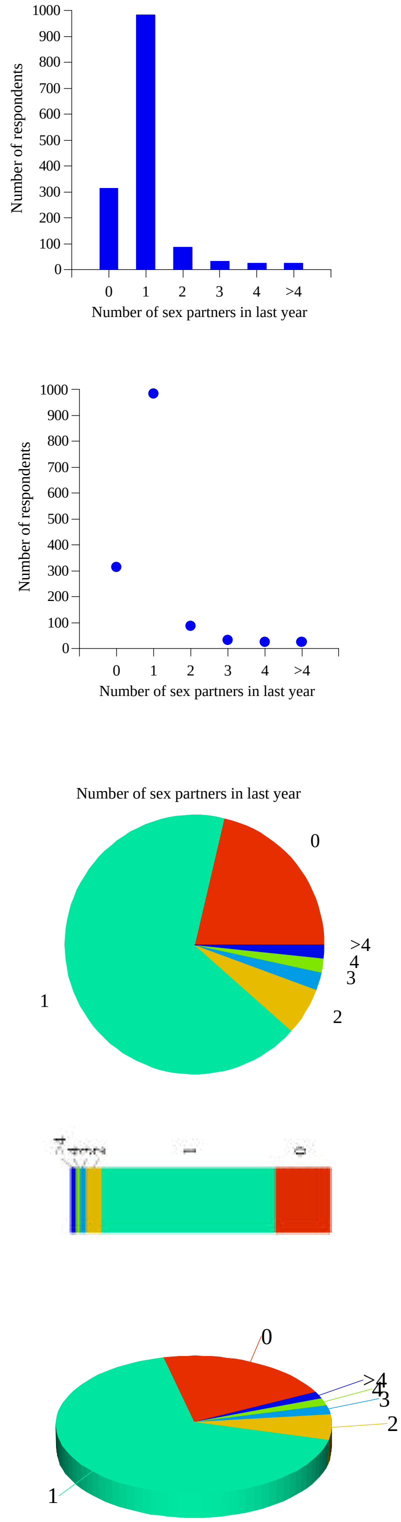


Figure 4

Consumption of spirits in gallons per capita by US state, 1986.

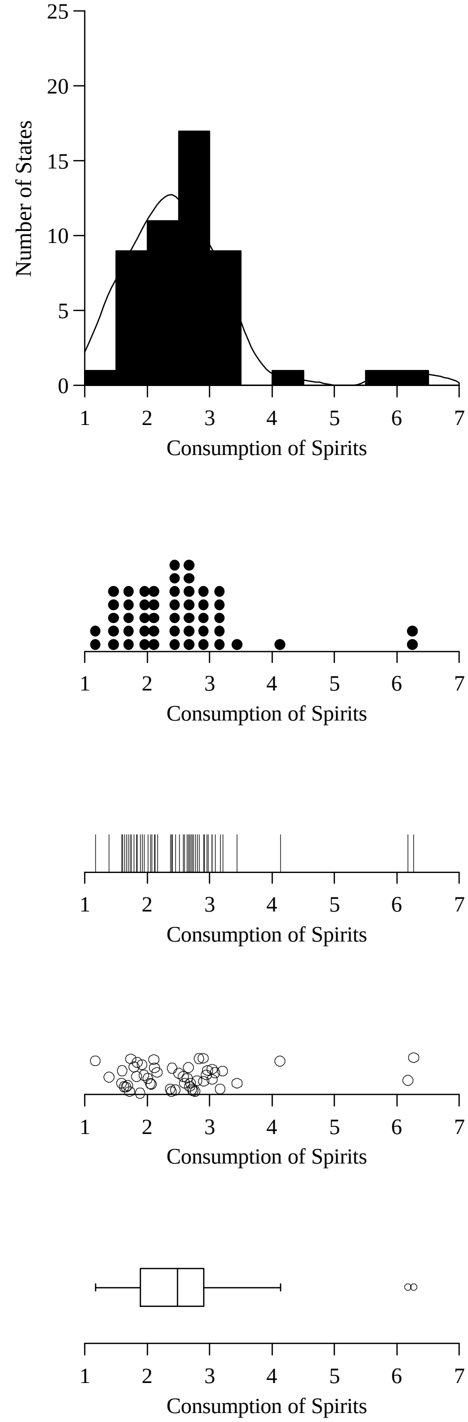


Figure 5
Number of reported sexual partners during previous year by reported level of happiness among 1462 respondents to 1993 General Social Survey. Third variable is proportion of respondents in each sexual partner category.

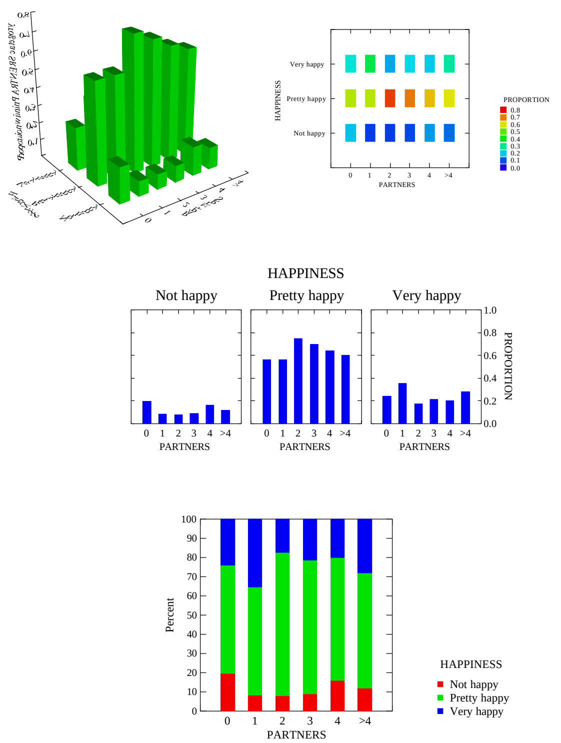
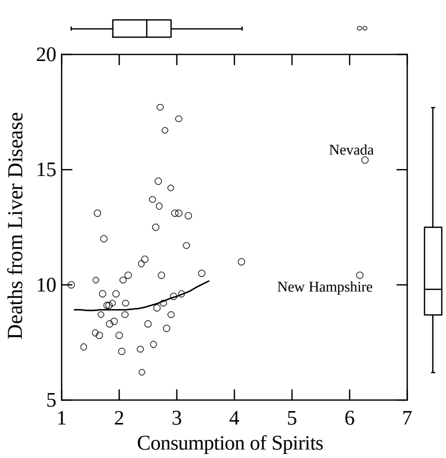


Figure 6
Consumption of spirits in gallons per capita by US state vs. Number of deaths from chronic liver disease and cirrhosis per 100,000 by US state, 1986



Time series data require special treatment. Figure 7 shows four examples. The data are number of US patents issued in the century from 1880 to 1980 (Wilkinson, Blank, and Gruber, 1996). The top panel shows a line element representing the series. Lines are most useful when a series is relatively smooth. The next panel shows a point element with a superimposed LOESS smoother. Points are best when we supplement them with a smoother; they tend to obscure time order if used alone, unless the series is especially smooth. The third panel shows spikes. These are vertical lines best used to reveal deviation from a constant level. This panel displays the residuals from the LOESS smooth. The lowest panel shows an area chart for the raw series. This highlights trend but prevents the use of confidence intervals and other enhancements to the plot. It should generally be avoided. Cleveland (1994) discusses other graphical representations for time series data.

Figure 7
US Patents issued from 1880 to 1980.

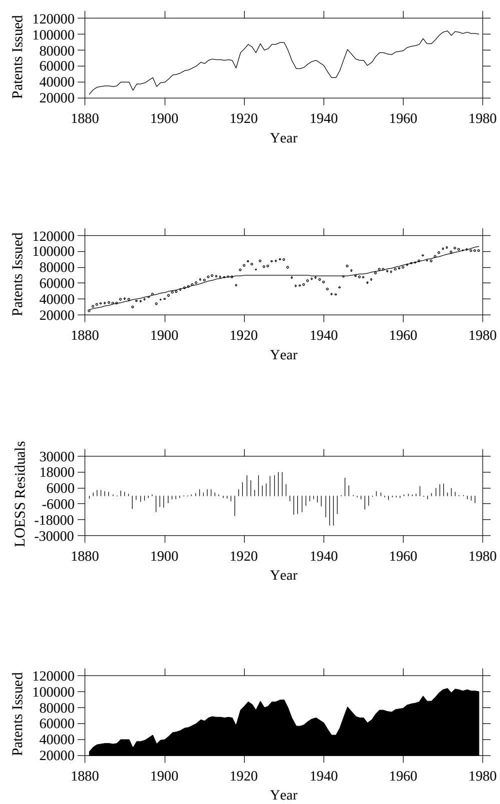
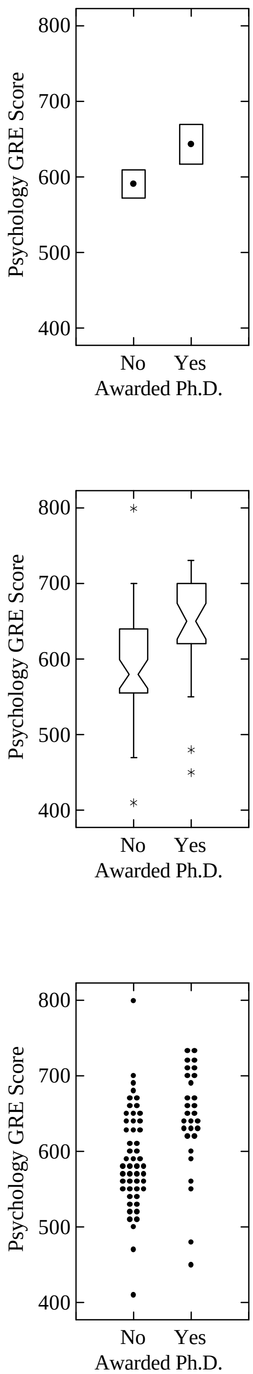


Figure 8 shows three ways of representing a categorical variable crossed with a continuous variable. The data are based on 80 graduate students over a ten year period in a US psychology department (Wilkinson, Blank, and Gruber, 1996). Graduate Record Examination Advanced Psychology Test scores are plotted against whether or not the students eventually received their Ph.D.

Figure 8
Advanced Psychology Graduate Record Examination (GRE) scores for 80 students taken from 10 years of a graduate psychology program. Students are grouped by whether or not they were awarded a Ph.D.



The first graphic employs a dot for the mean and a range bar to represent a 95 percent confidence interval on the mean. The bars do not overlap, which suggests that the advanced test can help identify those who achieve a Ph.D. (Interestingly, neither the Verbal nor Quantitative test scores predicted graduation in this sample.)

The second graphic is a notched box plot (McGill, Tukey, and Larsen, 1978). This is a variation on the box plot that not only conveys more of the important data landmarks than the classic confidence interval plot but also provides an approximate confidence interval of its own. If the data are independent samples from identically distributed populations that are lumpy in the middle (approximately normal in their midrange), then comparing the notches yields an approximate 95 percent test of the null hypothesis that the true medians in the population are equal. In this example, the notches do not overlap, reinforcing what we concluded from the confidence intervals on the means. Because it relies on the median instead of the mean, the notched box plot procedure is more robust against outliers.

The third graphic is a dot plot (Wilkinson, 1999). Used frequently in the medical literature, this plot shows every data point. In a small sample (e.g., clinical case study), a dot plot can be useful for readers who wish to consider every data value. It is of little use in making informal graphical inferences concerning group differences, however. It is best to think of grouped dot plots as unadorned one--dimensional scatterplots within categories.

4.1.3 Three variables

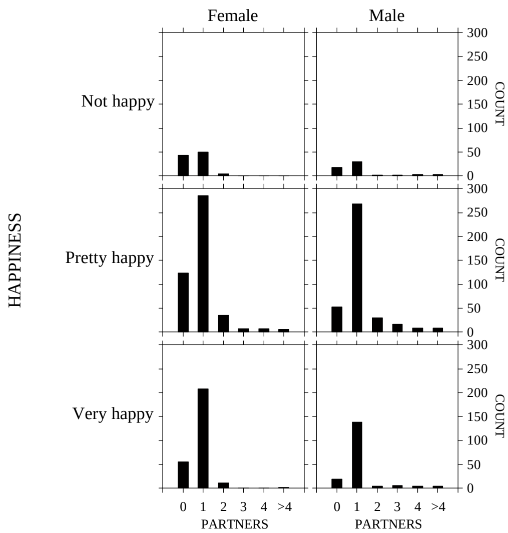
Figure 9 shows a triple crossing of categorical variables. The data are from the 1993 General Social Survey used in Figure 3 and Figure 5. The additional variable is gender (observed by the interviewer, not reported by the respondent). We ask whether the relationship between reported happiness and number of sexual partners differs by gender.

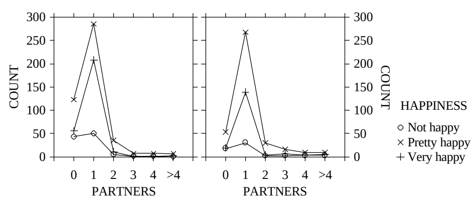
The top plot in the figure shows a paneled graphic. The format is similar to a Trellis display (Becker, Cleveland, and Shyu, 1996), but the labeling of the paneling variables (happiness and gender) is placed outside the plotting area to improve readability.

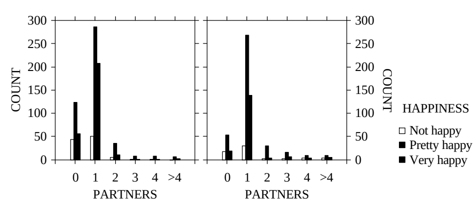
The middle plot collapses happiness into a legend in order to reduce the number of panels. This is a popular method for saving space, particularly when representing factorial layouts in ANOVA and other designs. It has several problems. First, the line segments used to highlight trends have slopes that depend on the particular spacing and arrangement of partner categories. Unless we employ an external scaling procedure to determine the spacing of the partners categories, we have scant justification for using these lines. Second, the collapsing introduces a symbol choice problem. It is difficult to find symbols that are easily distinguishable for more than a few categories. The symbols collide, as well, at the upper end of the horizontal scales.

Figure 9

Number of reported sexual partners during last year by reported level of happiness among 1462 respondents to 1993 General Social Survey. Third variable is proportion of respondents in each sexual partner category.



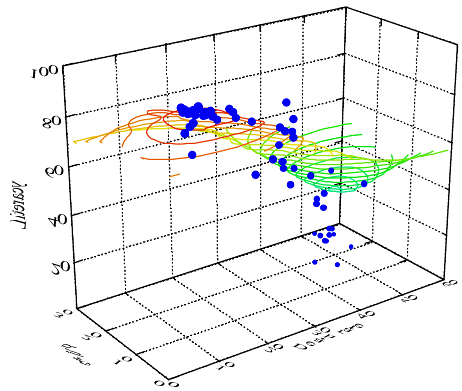


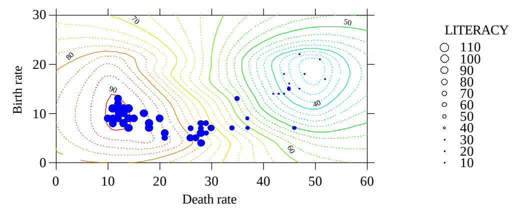


The bottom plot introduces an even less appealing alternative. Clustered bar charts are used widely, but they have several defects. First, it is difficult to discern separate patterns for the categories. One needs to focus on one set of bars to do this, but there is visual interference from the other bars in each cluster. Second, these bars can become quite thin with more than a few categories. Decoding this graphic is problematic.

Figure 10 shows a triple crossing of continuous variables. The data are from the UN databank used in Figure 1. The additional variable is literacy (percentage of the population that can read). We ask whether literacy is related to birth and death rates taken together.

Figure 10 Birth rate and death rate vs. literacy for selected countries.





The upper plot is a 3D frame that includes a contoured regression surface and a scatterplot. The contours (as opposed to a wire-frame or rendered surface) help viewers discern literacy levels in different elevations of the plot and keep the surface from hiding the data. Coloring the contours improves the coherence of the surface.

The lower plot is a 2D frame that includes the same contoured surface plus a scatterplot of symbols whose size is proportional to literacy. This plot makes it easier to decode specific *xyz* triplets.

Some critics tend to eschew 3D surface plots, but they have their uses. As Wilkinson (1999) argues, surfaces elicit a wholistic impression of a function. They are less useful for decoding individual values. Perhaps the best compromise is to present both displays simultaneously to provide different structural views. Another alternative is to categorize one or more of the variables and display the results in a paneled plot such as those in Figure 9.

4.2 Coordinates

We do not ordinarily think of charts being embedded in different coordinate systems, except for certain scientific graphics such as polar plots. Some chart types, such as radar charts, pie charts, and horizontal bar charts are nothing more than popular charts in unusual coordinate systems. Wilkinson (1999) provides numerous examples.

Figure 11
Difference between death month and birthday by occupational category.

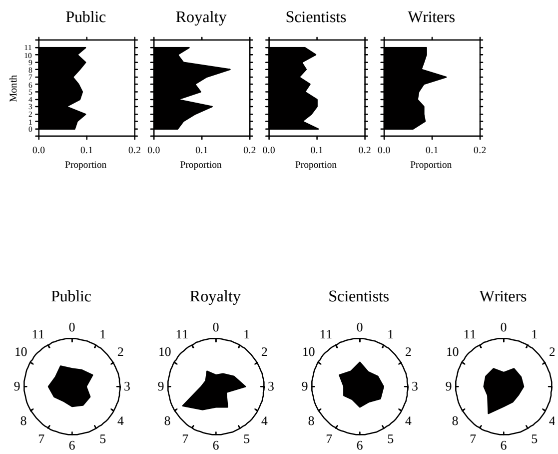


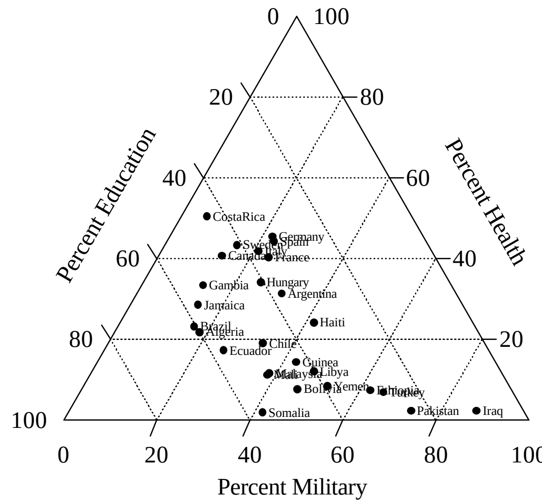
Figure 11 shows two different graphics in different coordinate systems, based on data from Andrews and Herzberg (1985), contributed by C. O'Brien. The data are months since their last birthday during which members of different social groups died. The upper plot shows a paneled horizontal area chart produced by rotating a vertical area chart 90 degrees (exchanging *x* and *y* coordinates in this case). This graphic reminds us that horizontal bar charts are a coordinate transformation of a vertical bar chart, not a new chart type.

If it makes sense to consider a death one month before a birthday as being more similar to a death one month after than to one six months after, then we might want to plot these variables in polar coordinates. The lower plot shows a profile plot of the same results.

Figure 12 contains a triangular coordinate plot. Mosteller and Tukey (1968), Bishop, Fienberg, and Holland (1975), and Wainer(1997) show examples of this type of plot. The data are from the UN dataset used in Figure 1. The variables represent expenditures in US dollars per person for health, education, and military.

The outside tick marks signify how to read the scales. Grid lines collinear with ticks demarcate values on the three scales. Iraq and Pakistan, for example, allocated most to military spending and least to education and health. Costa Rica spent more than 40 percent on education, almost 50 percent on health, and less than 10 percent on military. While triangular coordinates may be used most frequently in analysis of mixture models in industrial experiments, they also can be useful for social science models involving tradeoffs or compensatory mechanisms.

Figure 12
Percentage per-captia spending on Education, Military, and Health.



4.3 Uncertainty

There are two chief methods for representing error in a graphic. The first is sharp: error bounds are represented by clear edges, points, or lines. The second is fuzzy: error is represented by blurring an estimate.

Figure 13 shows two sharp ways of representing error in a graphic. The data are from the 1993 General Social Survey used in Figure 3. The additional variable is feelings about the Bible, as measured by the response to the question, "Which of these statements comes closest to describing your feelings about the Bible?" The responses coded are 1 (Word of God), 2 (Inspired word), and 3 (Book of fables). We assume that the dependent variable is continuous (biblical absolutism vs. relativism?) even though the responses are integers.

The upper plot shows error bars representing 95 percent confidence intervals on the means by category. Error bars are also used to represent one standard deviation or one standard error. It is important to make clear in accompanying titles or notes which type is used. If the spacing of values on the independent variable is meaningful, then we can fit a linear regression model to our data. We will further assume the number of sex partners to represent values of a continuous variable (promiscuity?). Assuming the assumptions for inference using a *t* distribution are appropriate, we can fit continuous confidence bounds on the regression line as in the lower panel of the figure. This plot uses jittering to reveal the concentration of data points where they are tied.

Figure 13
Sex partners vs. feelings about Bible (GSS survey responses)

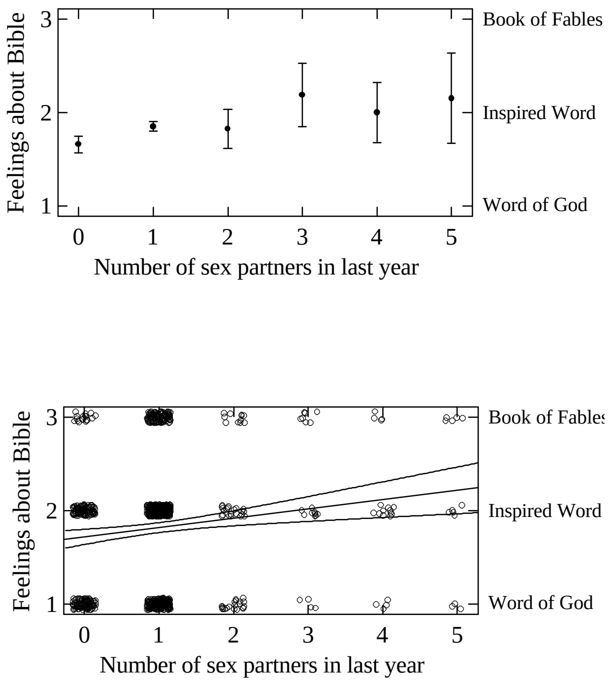


Figure 14
Bootstrapped data from Gonnelli et al. (1996)

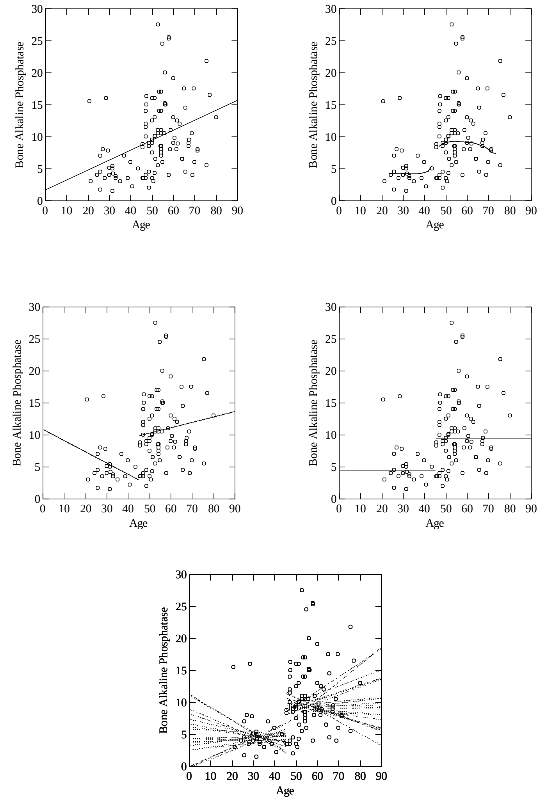


Figure 14 shows one way to represent error by fuzziness. The data are from Gonnelli et al. (1996). They represent concentration of bone alkaline phosphatase (BAP) in a sample of women of different ages. The authors fit a linear regression (shown in the upper left panel) to argue that BAP levels increase with age.

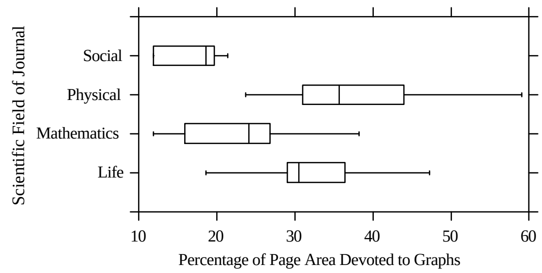
A modal regression in the upper right panel indicates that there is a discontinuity in this relationship at age 45 or so, corresponding most likely to the onset of menopause. Accordingly, we fit separate linear regressions to the two subgroups split at age 45 (third panel). These regressions appear to be sensitive to outliers, however, so we fit linear models using robust regression with *t* weighting (fourth panel). This fit indicates that a plausible model for predicting BAP from age involves level differences but no slope differences.

It is not easy to compute confidence intervals on the robust linear fits, so we resort to bootstrapping to provide an estimate of error. The bottom panel shows the result of 20 bootstrap robust fits displayed as faint dashed lines. The non-overlapping envelopes of the fits indicates that our level-change model is reasonable.

5 Conclusion

Figure 15 shows the principal result of a survey by Cleveland (1984) on the use of graphics in scientific articles. The horizontal axis represents the percentage of total page area devoted to graphs in 47 articles sampled from the 1980-81 volumes of selected journals. As Cleveland noted, the substantial differences among disciplines remain when one examines number of graphs per article instead of page area.

Figure 15
Use of graphics in journal articles (data from Cleveland, 1984).



In Cleveland's survey, social science journals ranked lowest in the use of graphics. We might speculate that this is a legacy of the aversion toward graphics among 18th century social scientists that Beniger and Robyn (1978) describe. Whatever the reason, there is little justification for this to continue. Software packages such as S-Plus, SAS, SPSS, and SYSTAT (used for the graphics in this article) now offer statistical graphics. And publishing standards such as Adobe PDF format now make it possible to display graphics as easily as text.

References

Andrews, D. F. and Herzberg, A. M. (1985). *Data: A Collection of Problems from Many Fields for the Student and Research Worker.* New York: Springer-Verlag.

Becker, R. A., Cleveland, W. S., and Shyu, M-J. (1996). The design and control of Trellis display. *Journal of Computational and Graphical Statistics, 5*, 123-155.

Beniger, J. R. and Robyn, D. L. (1978). Quantitative graphics in statistics: A brief history. *The American Statistician*, *32*, 1-11.

Bertin, J. (1967). *Sémiologie Graphique.* Editions Gauthier-Villars, Paris. English translation by W. J. Berg and H. Wainer (1983). *Semiology of Graphics.* Madison, WI: University of Wisconsin Press,

Bertin, J. (1981). *Graphics and Graphic Information-Processing*. English translation by W J Berg and P Scott. New York: Walter de Gruyter.

Bishop, Y. M., Fienberg, S. E., and Holland, P. W. (1975). *Categorical Data Analysis.* Cambridge, MA: MIT Press.

Bureau of the Census (1986). *State and Metropolitan Area Data Book.* Washington, DC: US Government Printing Office.

Chambers, J. M., Cleveland, W. S., Kleiner, B., and Tukey, P. A. (1983). *Graphical Methods for Data Analysis*. Monterey, CA: Wadsworth.

Cleveland, W. S. (1984). Graphs in scientific publications. *The American Statistician, 38,* 19-26

Cleveland, W. S. (1993). A model for studying display methods of statistical graphics (with discussion). *Journal of Computational and Graphical Statistics, 2,* 323-343.

Cleveland, W. S. (1994). *The Elements of Graphing Data* (Rev. Ed.). Summit, NJ: Hobart Press.

Cleveland, W. S. (1995). *Visualizing Data*. Summit, NJ: Hobart Press.

Cleveland W. S. and Devlin, S. (1988). Locally weighted regression analysis by local fitting. *Journal of the American Statistical Association, 83*, 596-640.

Cleveland, W. S. and McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association, 79*, 531-554.

Collins, B. M. (1993). Data visualization -– has it all been seen before? In R.A. Earnshaw and D. Watson (Eds.), *Animation and Scientific Visualization: Tools and Applications.* New York: Academic Press, 3-28.

Davis, J. A., Smith, T. W., and Marsden, P. V. (1993). *The General Social Survey.* Chicago: National Opinion Research Center.

Fienberg, S. (1979). Graphical methods in statistics. *The American Statistician, 33*, 165-178.

Funkhouser, H. G. (1936). Note on a tenth century graph. *Osiris, 1*, 260-262.

Funkhouser, H. G. (1937). Historical development of the graphical representation of statistical data. *Osiris, 3,* 269-404.

Gonnelli, S., Cepollaro, C., Montagnani, A., Monaci, G., Campagna, M. S., Franci, M. B., and Gennari, C. (1996). Bone alkaline phosphatase measured with a new immunoradiometric assay in patients with metabolic bone diseases. *European Journal of Clinical Investigation, 26,* 391-396.

Herdeg, W. (1981). *Graphis Diagrams: The Graphic Visualization of Abstract Data.* Zürich: Graphis Press.

Holmes, N . (1991). *Designer's Guide to Creating Charts and Diagrams.* New York: Watson-Guptill Publications.

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

McGill, R., Tukey, J. W., and Larsen, W. A. (1978). Variations of box plots. *The American Statistician, 32*, 12-16.

Mosteller, F. and Tukey, J. W. (1968). Data analysis, including statistics. In G Lindzey and E Aronson, Eds., *Handbook of Social Psychology, 2nd edition, Vol. 2.* , Reading, MA: Addison-Wesley, 80-203

Pinker, S. (1997). *How the Mind Works.* New York: W W Norton & Company.

Robinson, A. H. (1982). *Early Thematic Mapping in the History of Cartography.* Chicago: University of Chicago Press.

Scott, D. W. (1992). *Multivariate Density Estimation: Theory, Practice, and Visualization.* New York: John Wiley & Sons.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis.* London: Chapman & Hall.

Simkin, D., and Hastie, R. (1987). An information processing analysis of graph perception. *Journal of the American Statistical Association, 82*, 454-465.

Stigler, S. (1983). *The History of Statistics.* Cambridge, MA: Harvard University Press.

Tilling, L. (1975). Early experimental graphics. *The British Journal for the History of Science, 8*, 193-213.

Tufte, E .R. (1983). *The Visual Display of Quantitative*

Information. Cheshire, CT: Graphics Press.
Tufte, E .R. (1990). *Envisioning Data.* Cheshire, CT: Graphics Press.
Tufte, E .R. (1997). *Visual Explanations.* Cheshire, CT: Graphics Press.
Tukey, J. W. (1977). *Exploratory Data Analysis.* Reading, MA: Addison-Wesley.
Wainer, H. (1997). *Visual Revelations: Graphical Tales of Fate and Deception from Napoleon Bonaparte to Ross Perot.* New York: Springer-Verlag.
Wainer, H., and Spence, I. (1997). Who was Playfair? *Chance*, *10,* 35-37.
Wainer, H., Velleman, P. F. (2001). Statistical graphs: Mapping the pathways of science. *The Annual Review of Psychology*, *52.*
Wilkinson, L. (1999). Dot plots. *The American Statistician*, *53*, 276-281.
Wilkinson, L. (1999). *SYSTAT*. Chicago: SPSS Inc.
Wilkinson, L. (1999). *The Grammar of Graphics.* New York: Springer-Verlag.
Wilkinson, L., Blank, G., and Gruber, C. (1996). *Desktop Data Analysis with SYSTAT*. Upper Saddle River, NJ: Prentice-Hall.