

*For my parents*  
*Edward E. Tufte and Virginia James Tufte*

*To the memory of*  
*John W. Tukey (1915–2000)*

## *Introduction to the Second Edition*

This new edition provides high-resolution color reproductions of the many graphics of William Playfair, adds color to other images where appropriate, and includes all the changes and corrections accumulated during the 17 printings of the first edition.

This book began in 1975 when Dean Donald Stokes of Princeton's Woodrow Wilson School asked me to teach statistics to a dozen journalists who were visiting that year to learn some economics. I annotated a collection of readings, with a long section on statistical graphics. The literature here was thin, too often grimly devoted to explaining use of the ruling pen and to promulgating "graphic standards" indifferent to the nature of visual evidence and quantitative reasoning. Soon I wrote up some ideas. Then John Tukey, the phenomenal Princeton statistician, suggested that we give a series of joint seminars. Since the mid-1960s, Tukey had opened up the field, as his brilliant technical contributions made it clear that the study of statistical graphics was intellectually respectable and not just about pie charts and ruling pens.

After moving to Yale University, I finished the manuscript in 1982. A publisher was interested but planned to print only 2,000 copies and to charge a very high price, contrary to my hopes for a wide readership. I also sought to design the book so as to make it *self-exemplifying*—that is, the physical object itself would reflect the intellectual principles advanced in the book. Publishers seemed appalled at the prospect that an author might govern design.

Consequently I investigated self-publishing. This required a first-rate book designer, a lot of money (at least for a young professor), and a large garage. I found Howard Gralla who had designed many museum catalogs with great care and craft. He was willing to work closely with this difficult author who was filled with all sorts of opinions about design and typography. We spent the summer in

An early and most worthy use of a map to chart patterns of disease was the famous dot map of Dr. John Snow, who plotted the location of deaths from cholera in central London for September 1854. Deaths were marked by dots and, in addition, the area's eleven water pumps were located by crosses. Examining the scatter over the surface of the map, Snow observed that cholera occurred almost entirely among those who lived near (and drank from) the Broad Street water pump. He had the handle of the contaminated pump removed, ending the neighborhood epidemic which had taken more than 500 lives.<sup>6</sup> The pump is located at the center of the map, just to the right of the D in BROAD STREET. Of course the link between the pump and the disease might have been revealed by computation and analysis without graphics, with some good luck and hard work. But, here at least, graphical analysis testifies about the data far more efficiently than calculation.

<sup>6</sup> E. W. Gilbert, "Pioneer Maps of Health and Disease in England," *Geographical Journal*, 124 (1958), 172–183. Shown here is a redrawing of John Snow's map. For a reproduction and detailed analysis of the original map, see Edward Tufte, *Visual Explanations: Images and Quantities, Evidence and Narrative* (Cheshire, Connecticut, 1997), Chapter 2. Ideally, see John Snow, *On the Mode of Communication of Cholera* (London, 1855).

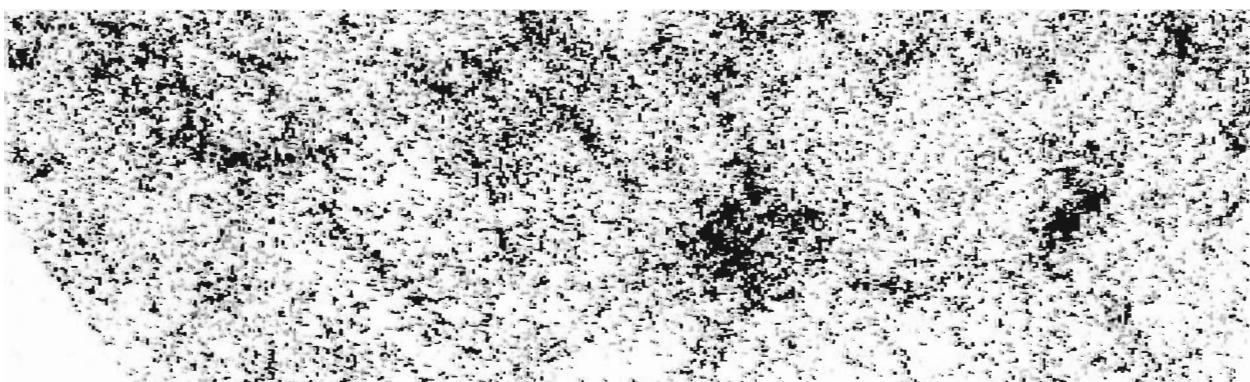


Charles Joseph Minard gave quantity as well as direction to the data measures located on the world map in his portrayal of the 1864 exports of French wine:

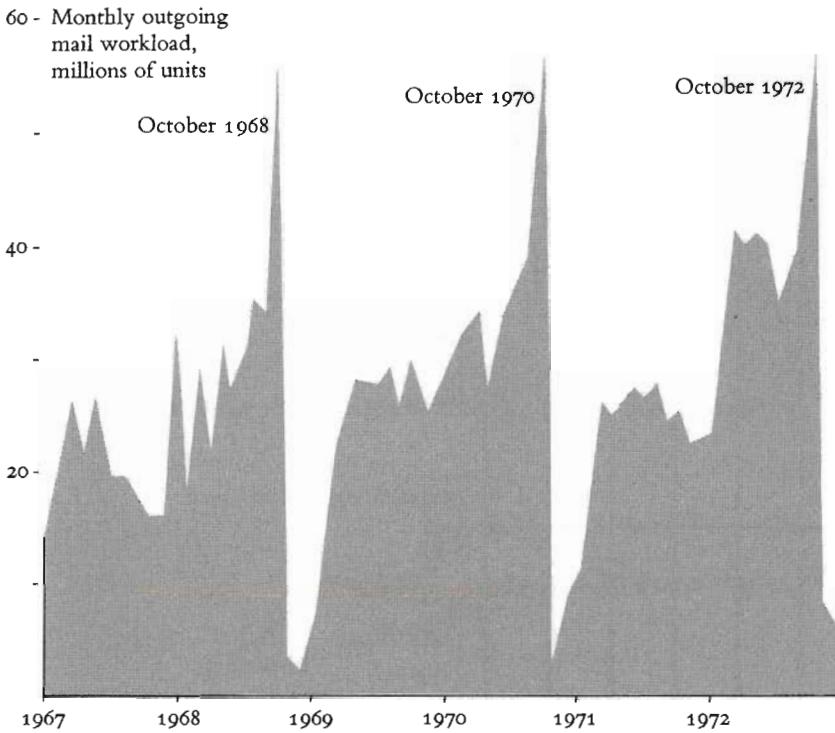
Computerized cartography and modern photographic techniques have increased the density of information some 5,000-fold in the best of current data maps compared to Halley's pioneering effort. This map shows the distribution of 1.3 million galaxies (including some overlaps) in the northern galactic hemisphere. The map divides the sky into  $1,024 \times 2,222$  rectangles. The number of galaxies counted in each of the 2,275,328 rectangles is represented by ten gray tones; the darker the tone, the greater the number of galaxies counted. The north galactic pole is at the center. The sharp edge on the left results from the earth blocking the view from the observatory. In the area near the perimeter of the map, the view is obscured by the interstellar dust of the galaxy in which we live (the Milky Way) as the line of sight passes through the flattened disk of our galaxy. The curious texture of local clusters of galaxies seen in this truly new view of the universe was not anticipated by students of galaxies, who had, of course, microscopically examined millions of photographs of galaxies before seeing this macroscopic view. Although the clusters are clearly evident (and accounted for by a theory of galactic origins), the seemingly random filaments may be happenstance. The producers of the map note the "strong temptation to conclude that the galaxies are arranged in a remarkable filamentary pattern on scales of approximately  $5^\circ$  to  $15^\circ$ , but we caution that this visual impression may be misleading because the eye tends to pick out linear patterns even in random noise. Indeed, roughly similar patterns are seen on maps constructed from simulated catalogs where no linear structure has been built in. . . ."<sup>7</sup>

The most extensive data maps, such as the cancer atlas and the count of the galaxies, place millions of bits of information on a single page before our eyes. No other method for the display of statistical information is so powerful.

<sup>7</sup> Michael Seldner, B. H. Siebers, Edward J. Groth and P. James E. Peebles, "New Reduction of the Lick Catalog of Galaxies," *Astronomical Journal*, 82 (April 1977), 249-314. See Gillian R. Knapp, "Mining the Heavens: The Sloan Digital Sky Survey," *Sky & Telescope* (August 1997), 40-48; Margaret J. Geller and John P. Huchra, "Mapping the Universe," *Sky & Telescope* (August 1991), 134-139.



The problem with time-series is that the simple passage of time is not a good explanatory variable: descriptive chronology is not causal explanation. There are occasional exceptions, especially when there is a clear mechanism that drives the Y-variable. This time-series does testify about causality: the outgoing mail of the U.S. House of Representatives peaks every two years, just before the election day:



The graphic is worth at least 700 words, the number used in a news report describing how incumbent representatives exploit their free mailing privileges to advance their re-election campaigns:

## FRANKED MAIL TIE TO VOTING SHOWN

## Testimony Finds the Volume Rises Before Elections

WASHINGTON, June 1 (AP)—New court testimony and documents show that much of the mail Congress sends at taxpayer expense is tied directly to the re-election campaigns of Senate and House members. According to material filed

in a lawsuit in Federal Court: Senate Republicans put two direct-mail experts on the public payroll to advise them on how to use their free mailing privileges to get voters.

An election manual prepared for Senate Democrats refers to newsletters as a "free forum," and sets up a timetable

for sending them as an integral part of a model re-election cam-

part of a model re-election campaign.

Senator John G. Tower, Republican of Texas, mailed more than 800,000 special-interest letters at taxpayer expense as part of his 1972 re-election effort and received campaign contributions from 1,000 individuals.

volunteer offers and donations in response.

Senator Jacob K. Javits, Republican of New York, gave written approval in 1973 for a tax-paid mail program intended to better his image and pay off at the polls. He focused his mail on areas where he needed voters.

The volume of "official" Congressional mail rises in election years and peaks just before the general election.

None of this activity necessarily violates any law or regulation, since Congress has wide discretion in the use of tax-paid mail. Congress gave itself the right to send official mail at Government expense at the

founding of the republic, and only Congress polices against

only Congress pounces against abuses of the free mailings. Complaints of political use of the free-mailing privilege, called the franking privilege, are heard every election year. Recently, however, the volume and cost of franked mail has multiplied. A new Federal law will limit what out-of-office challengers can spend to unseat incumbents.

In 1972, Congress passed a law prohibiting mass franked mailings within 28 days before an election. The sponsor of that legislation, Representative Morris K. Udall, Democrat of Arizona, said in an interview that further changes were needed to curtail political abuse of the frank.

Mr. Udall urged a 60-day pre-election cutoff for mass mailings and said he favored closing a loophole that recently allowed defeated Representative Frank M. Clark, Democrat of Pennsylvania, to send a

ranked newsletter to his old constituents after he had left office. Mr. Clark is seeking to regain his old post.

#### **Practice Documented**

Seldom has the political use of franked mail been so well documented as in recent testimony and documents filed in a Federal Court by Common Cause, the lobby group, which is suing for an end to tax-financed mass mailings by Con-

For example, Joyce P. Baker, political mail specialist, said in a 1973 job proposal that she wanted to set up direct-mail programs for Republican senators using franked mail. "The purpose of such a program is to help an incumbent senator get re-elected," she

She was put on the Senate payroll at \$18,810 a year in 1973 and 1974 and testified that during that time she aided Republican Senators Robert J. Dole and George H.W. Bush.

le of Kansas, Peter H. Domínguez of Colorado, Charles McC. Mathias Jr. of Maryland. Another political mail special...

Lee W. MacGregor, wrote proposal for the use of blank mail by his chief, Senator Javits, in 1973.

The over-all objective of the linked mail program can be to get the recipient of the mail to identify positively with a particular stand you have taken, a bill you have introduced; to kind of identification that can be translated into a vote.

the polls on election day" [fragments] MacGregor said.

Mr. Javits was out of the country and could not be reached. His administrative assistant, Donald Kellerman, recommended the use of franked mail.

"It is a standard device to meet voters, not voters but citizens, know what the Senator is doing here in Washington," he said.

Senator Tower's use of  
ranked mail in his 1972 cam-  
aign was documented by mem-  
andums.

Tom Loeffler, a high-ranking campaign aide, wrote in a memorandum dated Oct. 27, 72, that during the campaign Senator Tower had sent "31

Mr. Tower was not available for comment. His administrative assistant, Elwin Skiles, said the Senator's use of

Postal Service figures show that in the 12 months before November, 1973, Congress sent 2.9 million franked pieces of mail. But in the next 12 months, covering the election season of 1974, Congress sent 0.6 million, a jump of 57 per cent about what's happening," Mr. Skiles said.

## 2 Graphical Integrity

For many people the first word that comes to mind when they think about statistical charts is “lie.” No doubt some graphics do distort the underlying data, making it hard for the viewer to learn the truth. But data graphics are no different from words in this regard, for any means of communication can be used to deceive. There is no reason to believe that graphics are especially vulnerable to exploitation by liars; in fact, most of us have pretty good graphical lie detectors that help us see right through frauds.

Much of twentieth-century thinking about statistical graphics has been preoccupied with the question of how some amateurish chart might fool a naive viewer. Other important issues, such as the use of graphics for serious data analysis, were largely ignored.

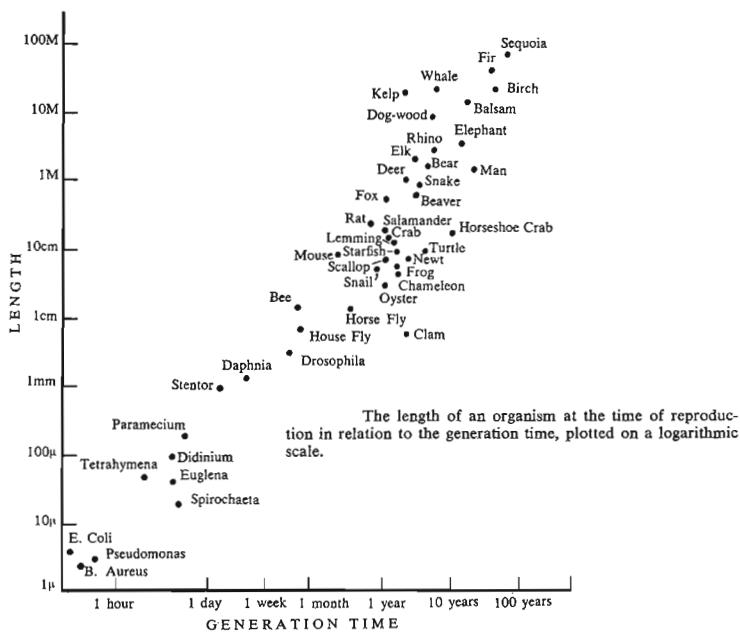
At the core of the preoccupation with deceptive graphics was the assumption that data graphics were mainly devices for showing the obvious to the ignorant. It is hard to imagine any doctrine more likely to stifle intellectual progress in a field. The assumption led down two fruitless paths in the graphically barren years from 1930 to 1970: First, that graphics had to be “alive,” “communicatively dynamic,” overdecorated and exaggerated (otherwise all the dullards in the audience would fall asleep in the face of those boring statistics). Second, that the main task of graphical analysis was to detect and denounce deception (the dullards could not protect themselves).

Then, in the late 1960s, John Tukey made statistical graphics respectable, putting an end to the view that graphics were only for decorating a few numbers. For here was a world-class data analyst spinning off half a dozen new designs and, more importantly, using them effectively to explore complex data.<sup>1</sup> Not a word about deception; no tortured attempts to construct more “graphical standards” in a hopeless effort to end all distortions. Instead, graphics were used as instruments for reasoning about quantitative information. With this good example, graphical work has come to flourish.

Of course false graphics are still with us. Deception must always be confronted and demolished, even if lie detection is no longer at the forefront of research. Graphical excellence begins with telling the truth about the data.

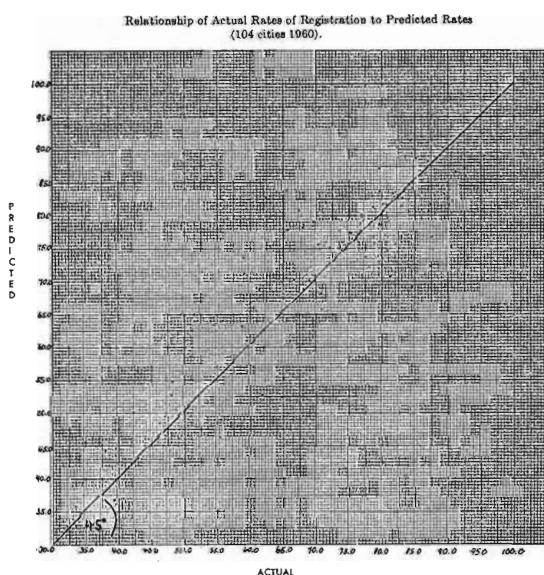
<sup>1</sup>John W. Tukey and Martin B. Wilk, “Data Analysis and Statistics: Techniques and Approaches,” in Edward R. Tufte, ed., *The Quantitative Analysis of Social Problems* (Reading, Mass., 1970), 370–390; and John W. Tukey, “Some Graphic and Semigraphic Displays,” in T. A. Bancroft, ed., *Statistical Papers in Honor of George W. Snedecor* (Ames, Iowa, 1972), 293–316.

Most of the ink in this graphic is data-ink (the dots and labels on the diagonal), with perhaps 10–20 percent non-data-ink (the grid ticks and the frame):



John Tyler Bonner, *Size and Cycle: An Essay on the Structure of Biology* (Princeton, 1965), p. 17.

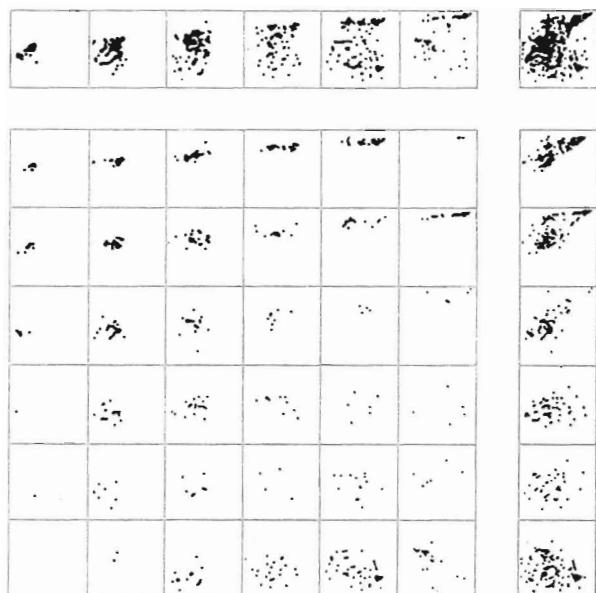
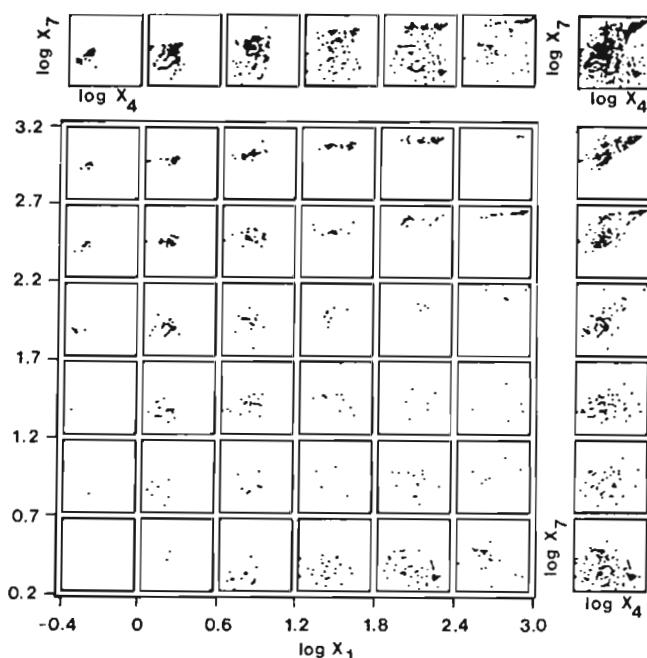
In this display with nearly all its ink devoted to matters other than data, the grid sea overwhelms the numbers (the faint points scattered about the diagonal):



The space occupied by the doubled grid lines consumes 18 percent of the area of this otherwise most ingenious design, a "multiwindow plot." Optical white dots appear at the intersections of the grid lines. (The plot shows the following: The large square contains  $X_4$ ,  $X_7$  scatterplots for the indicated levels of  $X_1$  and  $X_3$ . The marginal plots on the right are conditioned on  $X_3$  and the plots at the top on  $X_1$ . The upper right corner shows the unconditional  $X_4$ ,  $X_7$  scatter.) Redrawing eliminates the vibration:

Paul A. Tukey and John W. Tukey,  
"Data-Driven View Selection; Agglomeration and Sharpening," in Vic Barnett, ed., *Interpreting Multivariate Data* (Chichester, England, 1981), 231–232.

ULTIWINDOW PLOT OF PARTICLE PHYSICS MOMENTUM DATA

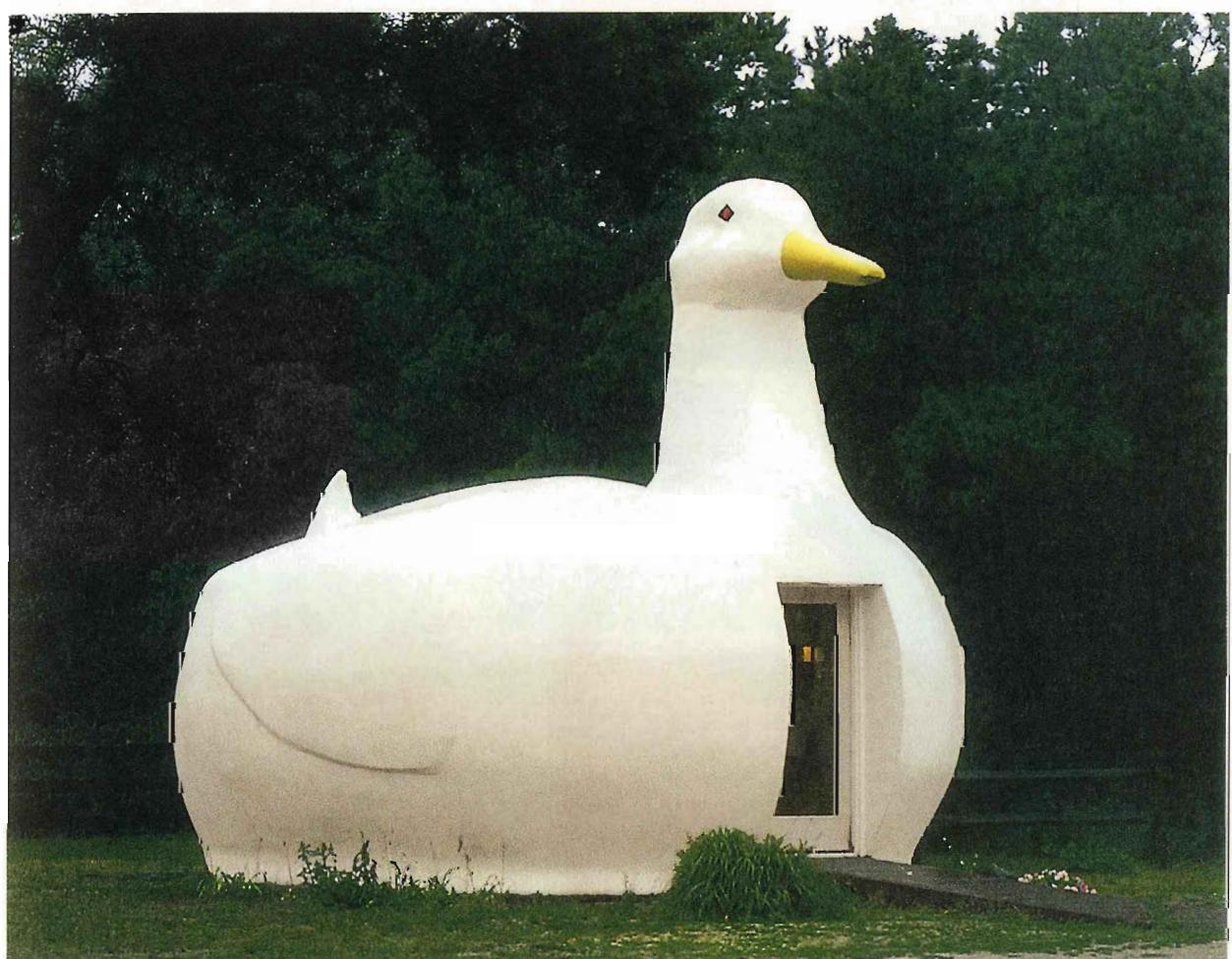


Brown, and Steven Izenour write about the ducks of modern architecture—and their thoughts are relevant to the design of data graphics as well:

When Modern architects righteously abandoned ornament on buildings, they unconsciously designed buildings that *were* ornament. In promoting Space and Articulation over symbolism and ornament, they distorted the whole building into a duck. They substituted for the innocent and inexpensive practice of applied decoration on a conventional shed the rather cynical and expensive distortion of program and structure to promote a duck. . . . It is now time to reevaluate the once-horrifying statement of John Ruskin that architecture is the decoration of construction, but we should append the warning of Pugin: It is all right to decorate construction but never construct decoration.<sup>2</sup>

<sup>2</sup> Robert Venturi, Denise Scott Brown, and Steven Izenour, *Learning from Las Vegas* (Cambridge, revised edition, 1977), p. 163. The initial statement of the duck concept is found on pp. 87–103.

*Big Duck*, Flanders, New York; photograph by Edward Tufte, July 2000.

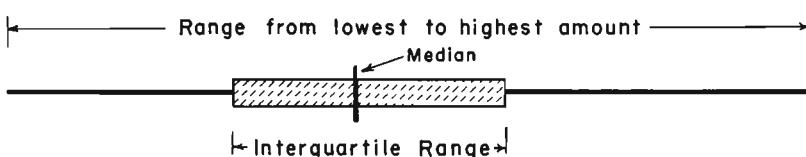


# 6 Data-Ink Maximization and Graphical Design

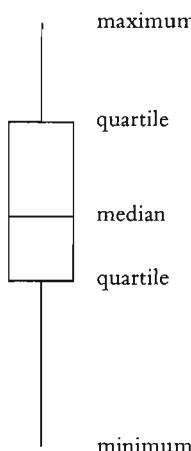
So far the principles of maximizing data-ink and erasing have helped to generate a series of choices in the process of graphical revision. This is an important result, but can the ideas reach beyond the details and particularities of editing? Is it possible to do what a theory of graphics is supposed to do, that is, to derive new graphical forms? In this chapter the principles are applied to many graphical designs, basic and advanced, including box plots, bar charts, histograms, and scatterplots. New designs result.

## Redesign of the Box Plot

Mary Eleanor Spear's "range bar"

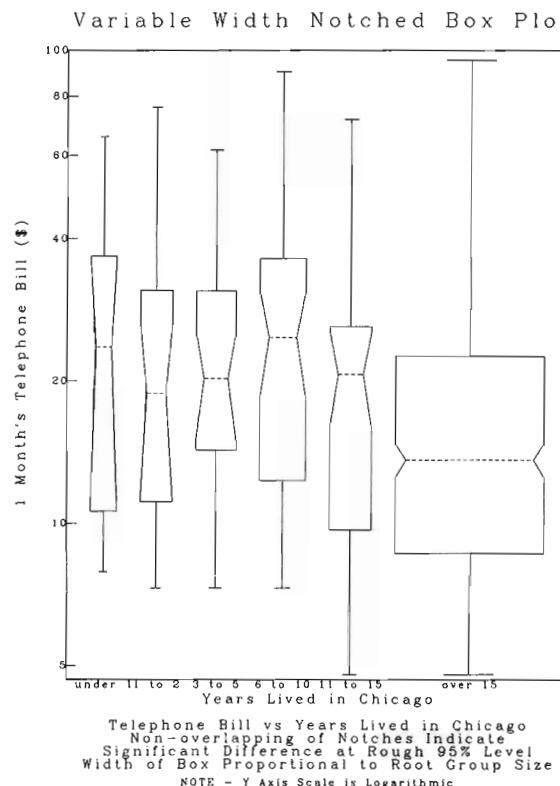


and John Tukey's "box plot"

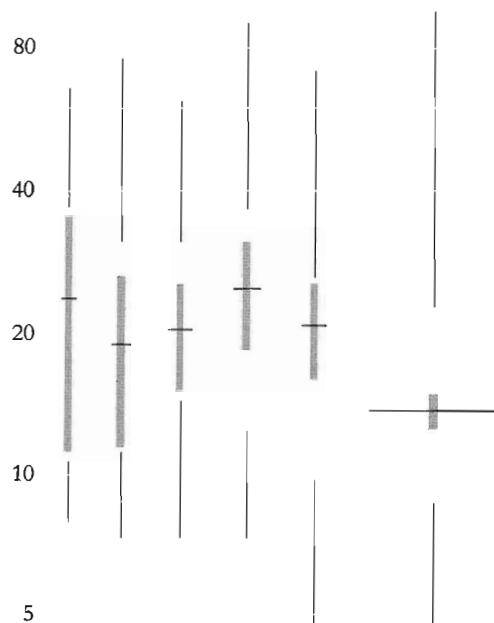


Mary Eleanor Spear, *Charting Statistics* (New York, 1952), p. 166; and John W. Tukey, *Exploratory Data Analysis* (Reading, Massachusetts, 1977).

Erasing and data-ink maximizing have induced changes in the plain old bar chart. The techniques—no frame, no vertical axis, no ticks, and the white grid—apply to other designs:



Robert McGill, John W. Tukey, and Wayne A. Larsen, "Variations of Box Plots," *American Statistician*, 32 (1978), 12–16.



Building data measures out of the data increases the quantitative detail and dimensionality of a graphic. The stem-and-leaf plot constructs the distribution of a variable with numbers themselves:

$0 | 9 = 900 \text{ feet}$

0   9	98766562
1	97719630
2	69987766544422211009850
3	876655412099551426
4	999844331929433361107
5	9766666554422210097731
6	89865441077761065
7	98855431100652108073
8	653322122937
9	377655421000493
10	0984433165212
11	4963201631
12	45421164
13	47830
14	00
15	676
16	52
17	92
18	5
19	39730

Stem-and-leaf displays:  
heights of 218 volcanoes, unit 100 feet.

$19 | 3 = 19,300 \text{ feet}$

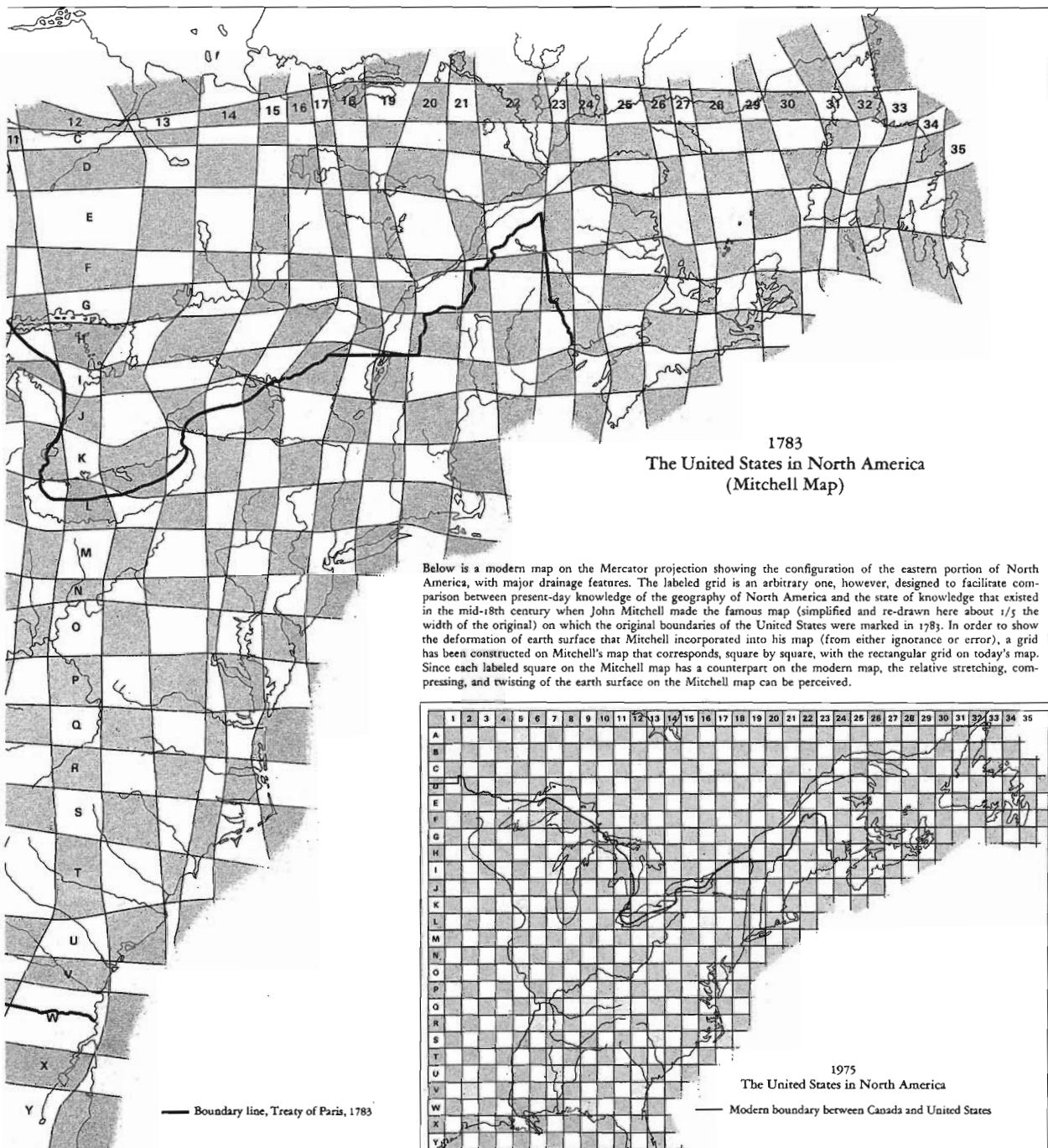
The idea of making every graphical element effective was behind the design of the stem-and-leaf plot. In presenting his invention, John Tukey wrote: "If we are going to make a mark, it may as well be a meaningful one. The simplest—and most useful—meaningful mark is a digit."<sup>2</sup>

Here, too, the data form the data measure. Note the bimodal distribution in the histogram of college students arranged by height.

<sup>2</sup> "Some Graphic and Semigraphic Displays," in T. A. Bancroft, ed., *Statistical Papers in Honor of George W. Snedecor* (Ames, Iowa, 1972), p. 296.



Brian L. Joiner, "Living Histograms," *International Statistical Review*, 43 (1975), 339–340. But, for further developments, see Mark F. Schilling, Ann E. Watkins, and William Watkins, "Is Human Height Bimodal?" *The American Statistician*, 56 (August 2002), 223–229.



Here the grid is the element of interest, rather than the map.

Lester J. Cappon, Barbara Bartz Petchenik, and John Hamilton Long, *Atlas of Early American History* (Princeton, 1976), p. 58.

Maps routinely present even finer detail. A cartographer writes that "the resolving power of the eye enables it to differentiate to 0.1 mm where provoked to do so. Clearly, therefore, conciseness is of the essence and high resolution graphics are a common denominator of cartography."<sup>2</sup> Distinctions at 0.1 mm mean 254 per inch.

How many statistical graphics take advantage of the ability of the eye to detect large amounts of information in small spaces? And how much information should graphics show? Let us begin by considering an empirical measure of graphical performance, the data density.

### **Data Density in Graphical Practice**

The numbers that go into a graphic can be organized into a data matrix of observations by variables. Taking into account the size of the graphic in relation to the amount of data displayed yields the *data density*:

$$\text{data density of a graphic} = \frac{\text{number of entries in data matrix}}{\text{area of data graphic}}$$

Data matrices and data densities vary enormously in practice. At one extreme, this overwrought display (originally printed in five colors) presents a data matrix of four entries, the names and the numbers for the two bars on the right. The left bar is merely the total of the other two. The graph covers 26.5 square inches (171 square centimeters), resulting in a data density of .15 numbers per square inch (.02 numbers per square centimeter), which is thin indeed.

<sup>2</sup>D. P. Bickmore, "The Relevance of Cartography," in John C. Davis and Michael J. McCullagh, eds., *Display and Analysis of Spatial Data* (London, 1975), p. 331.

based on about 50 numbers shown at the rate of 10 per square inch. Among the world's newspapers, the *Wall Street Journal*, *The Times* (London), and *Asahi* publish data-rich graphics, with data densities equal to those of the *Journal of the American Statistical Association*. Most of the American papers and magazines, along with *Pravda*, publish less data per graphic than the major papers of other industrialized countries.

Very few statistical graphics achieve the information display rates found in maps. Highly detailed maps portray 100,000 to 150,000 bits per square inch. For example, the average U.S. Geological Survey topographic quadrangle (measuring 17 by 23 inches) is estimated to contain over 100 million bits of information, or about 250,000 per square inch (40,000 per square centimeter).<sup>3</sup> Perhaps some day statistical graphics will perform as successfully as maps in carrying information.

<sup>3</sup> Morris M. Thompson, *Maps for America* (Washington, D.C., 1979), p. 187.

### High-Information Graphics

Data graphics should often be based on large rather than small data matrices and have a high rather than low data density. More information is better than less information, especially when the marginal costs of handling and interpreting additional information are low, as they are for most graphics. The simple things belong in tables or in the text; graphics can give a sense of large and complex data sets that cannot be managed in any other way. If the graphic becomes overcrowded (although several thousand numbers represented may be just fine), a variety of data-reduction techniques—averaging, clustering, smoothing—can thin the numbers out before plotting.<sup>4</sup> Summary graphics can emerge from high-information displays, but there is nowhere to go if we begin with a low-information design.

Data-rich designs give a context and credibility to statistical evidence. Low-information designs are suspect: what is left out, what is hidden, why are we shown so little? High-density graphics help us to compare parts of the data by displaying much information within the view of the eye: we look at one page at a time and the more on the page, the more effective and comparative our eye can be.<sup>5</sup> The principle, then, is:

Maximize data density and the size of the data matrix, within reason.

High-information graphics must be designed with special care. As the volume of data increases, data measures must shrink (smaller dots for scatters, thinner lines for busy time-series). The clutter of

<sup>4</sup> Paul A. Tukey and John W. Tukey, "Summarization: Smoothing; Supplemented Views," in Vic Barnett, ed., *Interpreting Multivariate Data* (Chichester, England, 1982), ch. 12; and William S. Cleveland, "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association*, 74 (1979), 829–836.

<sup>5</sup> It is suggested in the analysis of x-ray films to "search a reduced image so that the whole display can be perceived on at least one occasion without large eye movement." Edward Llewellyn Thomas, "Advice to the Searcher or What Do We Tell Them?" in Richard A. Monty and John W. Senders, eds., *Eye Movements and Psychological Processes* (Hillsdale, N.J., 1976), p. 349.

Page after page of Leonardo's manuscripts have a gentle but thorough integration of text and figure, a quality rarely seen in modern work:

Leonardo da Vinci, *Treatise on Painting* [Codex Urbinas Latinus 1270], vol. 2, facsimile (Princeton, 1956), p. 234, paragraph 827.

234.

chevai le cose vedute essere tanto minime che no che  
le membra ma il tutto quasi ripara impossibile a po-  
tere figurare come s'ell'occhio fosse, o, c'la bussa d'un  
giunto di braccio eguale alla tua tangua dipinta  
sia, a, b, distosta . m. dal occhio mezo  
braccio allora tu uedrai per esso spa-  
cio tutte le cose che . b a. ueder si porti dim-  
tro alla lungheza . o. d'no orizonte di  
cento miglia in tanta confusa diminuzione che no  
che figurav di quelle alcuna parte c'habbia figurav  
ma apena potrai porre si piccolo punto di penello che  
non sia maggiore c'hogni gran'casamento posto in  
dieci miglia di distanza.

*Perche li monti in longha distanza  
si dimostrano più scuri nella cima*

*che nella base -*

L'avia c'acquista gradi di grossezza in ogni grado dc  
la sua bassezza e della sua distanza e causa ch'esse  
cime de monti che più s'inalzano più mostrano la  
sua natu- rale oscu-  
rità per che máco  
sono impe- dite dalla  
grossezza de l'aria  
nella cima che nella  
loro base o nella vicinità che nella remozione, Pro-  
nuasi, o, p., d., s., c., v., a. & sono gradi dell'aria che sem-  
pre s'asottiglian' quanto più s'inalzano, a, f., f., h., h., k.  
sono li altri gradi transuersali dove l'aria agista'

Finally, a caveat: the use of words and pictures together requires a special sensitivity to the purpose of the design—in particular, whether the graphic is primarily for communication and illustration of a settled finding or, in contrast, for the exploration of a data set. Words on and around graphics are highly effective—sometimes all too effective—in telling viewers how to allocate their attention to the various parts of the data display.<sup>5</sup> Thus, for graphics in exploratory data analysis, words should tell the viewer *how* to read the design (if it is a technically complex arrangement) and not *what* to read in terms of content.

<sup>5</sup> Experiments in visual perception indicate that word instructions substantially determine eye movements in viewing pictures. See John D. Gould, "Looking at Pictures," in Richard A. Monty and John W. Senders, eds., *Eye Movements and Psychological Processes* (Hillsdale, N.J., 1976), 323–343.

aesthetics, graphical 177–191  
air pollution 42, 170  
*Akahata* ("Red Flag") 83  
Albers, Josef 183  
Alder, Berni 134, 145  
*American Education* 118  
*American Political Science Review* 167  
*Annuaire Statistique de la France* 167  
Anscombe, F. J. 14  
Anscombe's quartet 13–14  
area and quantity 69–73  
Arkin, Herbert 112  
*Asahi* 83, 167, 168  
Asch, S. E. 56  
astronomical graphics 26–29, 154, 166  
Ayres, Leonard P. 141, 155  
Ayres, Richard E. 95

Bamberger, Clara Francis 82  
Bancroft, T. A. 53, 140  
bar chart 96–97, 126–128, 129  
Barabba, Vincent P. 153  
Barnett, Vic 114, 168  
before-after time-series 39  
Beniger, James R. 20, 86  
Bertin, Jacques 112, 166, 169, 178  
Bickmore, D. P. 162  
Biderman, Albert D. 181  
'Big Duck' 116, 117  
bilateral symmetry 97  
*Biochemistry* 110  
*Biochimica et Biophysica Acta* 110  
Birkhoff, George D. 189  
blot maps 20, 139  
Blot, William J. 16  
Bonner, John Tyler 94  
boring statistics 79–80, 87  
Borissavlièvitch, Miloutine 189

Boutique Data Graphics 118  
Bowen, William G. 95  
box plot 97, 123, 129  
Brier, Stephen S. 14  
Brinton, Willard C. 112  
Brown, Denise Scott 116–117  
Brunés, Tons 189  
Bryan, Kirk 99  
*Business Week* 63, 83, 167

calculus 9  
calculus, graphical 46  
*California Water Atlas* 119  
Campbell, Angus 147  
Campbell, Donald T. 75  
cancer 16–20, 47, 171  
cancer maps 16–20, 121  
Cappon, Lester J. 146  
cars, frequency of repair 174  
Cartesian coordinates 9  
cause and effect 37, 47, 82, 187  
CEM 56  
chartjunk 107–121  
Chavannes, E. 21  
chemical elements 102–105  
Chernoff faces 97, 142  
Chernoff, Herman 136, 142  
Cheysson, É. 154  
chicken, 4,340-pound 73  
cholera map 24  
chromosomes 172–173  
Cleveland, William S. 38, 168  
color 153–154, 183  
color-deficient viewers 183  
Colton, Raymond R. 112  
computer graphics 26–27, 29, 42, 112,  
116, 120–121, 136, 153, 170  
Connecticut speeding 74–75

constant dollars 65–68  
*Consumer Reports* 80, 174  
context 74–75  
Converse, Philip E. 147  
copper, conductivity of 49  
Cornford, Eain M. 109  
*Cosmographia* 22  
county maps 16–20, 153  
Cox, Michael D. 99  
Crotty, William J. 95  
Curti, Merle 85

Dahl, Robert A. 85  
Dakin, Edwin F. 15  
Daniel, Cuthbeit 133  
data-based grid 148  
data-based labels 149–152  
data-built data-measures 139–144  
data density 161–169  
data-ink 91–105  
data-ink maximization 96–105,  
123–137, 175, 186, 187  
data-ink ratio 93–96  
data maps 16–27  
data matrix 167–169  
data measures 139–144  
data/text integration 180–182  
data variation 60–61  
Davis, John C. 162  
Day Mines, Inc. 54  
decoder, visual 153–154  
deflating money 65–68  
*Der Spiegel* 83  
design variation 60–63  
Dewey, Edward R. 15  
*Die Zeit* 83  
Diller, Burgoine 185  
distortion, graphical 55–59  
dogs 50

- Doll, R. 47, 82  
 dollars, constant 65–68  
 dot-dash-plot 133, 136, 139, 152  
 double-functioning labels 149–152  
 Duchamp, Marcel 36  
 ducks 116–121
- 'Easter Wings' 143  
 economic data 15, 32–34, 38, 54–55, 61–68, 70, 91–92, 108, 126, 148, 152, 158, 161, 180  
*The Economist* 63, 83, 167  
 educational tests, graphics 86  
 Ehrenberg, A. S. C. 178  
 election graphics 147, 179  
 electroencephalogram 93  
 elegance, graphical 177  
 Eliot, T. S. 100  
 encodings 178, 183  
 erasing principles 96–100, 136  
 Erbring, Lutz 120  
 excellence, graphical 13–51
- Fienberg, Stephen E. 14  
 'Fig.' 181  
 Finkner, Alva L. 153  
 Fiorina, Morris 66  
 Fisher, R. A. 133  
 Flanigan, William H. 85  
 Flury, Bernhard 97  
*Fortune* 167  
 Francolini, C. M. 153  
*Frankfurter Allgemeine* 83  
 Fraumeni, Joseph F. 16  
 French communes 166  
 French wine map 25–26  
 friendly data graphics 183  
 fuel economy graphic 57–59  
 Funkhouser, H. Gray 20, 28, 154
- Gabaglio, Antonio 72  
 galaxies map 26–27, 154, 166  
 gecko 36  
 Geological Survey maps 168  
 Gilbert, E. W. 24  
 Goldenberg, Edie N. 120  
 Goodin, W. R. 170  
 Gould, John D. 182  
 government spending 64–68, 158  
 grade school graphics 86  
 Graph of Magical Parallelepipeds 68  
 Graphical Hack 59  
 gray grid 116
- gray shades 154  
 grids 112–116  
 grids, data-based 145–148  
 grids, gray 116  
 grids, white 127–129, 136  
 Groth, Edward J. 26  
 Gurnett, Donald A. 29
- half-face 97, 136  
 Halley, Edmond 23  
 Hankins, Timothy H. 134  
 Hannibal's campaign 176–177  
 Hayward, Roger 102, 136  
 Herbert, George 143  
 Herdeg, Walter 80, 152  
 herring catches 172  
 hierarchy in graphics 153–159  
 high-information graphics 168–169  
 Hilgard, Ernest 85  
 histogram 126–128  
 histogram, living 140  
 Hjort, Johan 172  
 Ho, C. Y. 49, 150  
 Hoover, Robert 16  
 horizon, analogy to 186–187  
 horse paces 34–35  
 House of Representatives 37  
 Huot, Marie E. 109
- identification numbers 149–150  
 Israel, Atlas of 178  
 Italian post office 72  
 Izenour, Steven 117
- Japan, graphics in university entrance exams 86  
 Japanese beetle 43, 121  
 Japanese graphics 82–84  
 JASA style sheet 110, 150–151, 164  
 Joiner, Brian L. 140  
*Journal of the American Chemical Society* 110  
*Journal of American Medical Association* 167  
*Journal of the American Statistical Association* 110, 150–151, 164, 167–168  
*Journal of Biological Chemistry* 110  
*Journal of Chemical Physics* 110  
*Journal of the Royal Statistical Society* 167  
 Jupiter graphic 29
- Kahrl, William L. 119  
 Kelley, Stanley 95
- Kolers, Paul A. 79  
 Kooi, Kenneth 93  
 Kouchoukos, Nicholas 109  
 Kuznicki, James T. 100, 109
- labels on graphics 180–182, 183, 187  
 Lambert, Johann Heinrich 29, 32, 45–46  
*Lancet* 110  
 lane stripes 144  
 Larsen, Wayne A. 129  
 Lauer, A. R. 144  
 Law School Admission Test, graphics 86  
*Learning from Las Vegas* 116–117  
 Leonardo da Vinci 181–182  
 less is a bore 175  
 less is more 175  
 lettering 184  
 Lie Factor 57  
 lies, defense of 76–77  
 Liley, P. E. 49, 150  
 line weight 185  
 living histogram 140  
 Long, John Hamilton 146  
 Los Angeles smog 42, 170  
*Los Angeles Times* 42, 69  
 lung cancer 16–20, 47  
 lying graphics 76–77
- Macdonald-Ross, Michael 56  
 MacGregor, A. J. 112  
 magnetic monopole 39  
 mail, House of Representatives 37  
 Malcolm, Andrew H. 84  
 Manvel, Allen D. 81  
 map, blot 20, 139  
 map, cancer 16–20  
 map, cholera 24  
 map, data 16–27  
 map, French wine 25–26  
 map, galaxies 26–27, 154, 166  
 map, Israel 178  
 map, patch 20  
 map, thematic 16–27  
 map, Yü Chi Thu 20–21  
 Marey, E. J. 31, 34–36, 40, 115–116, 121  
 Mason, Thomas J. 16  
 Masterton, William 85  
 Matisse, Henri 160  
 maximizing data-density 168  
 maximizing data-ink 96–105, 123–137, 175, 186, 187  
 McClenaghan, William 85

- McCracken, Paul 48  
 McCullagh, Michael J. 162  
 McCutcheon, N. Bruce 100, 109  
 McGill, Robert 129  
 McKay, Frank W. 16  
 McRae, Gregory J. 42, 170  
 Meihoefer, H. J. 56  
 melanoma 171  
 Mies van der Rohe, Ludwig 175  
 Miller, Arthur H. 120  
 Miller, Warren E. 147  
 Minard, Charles Joseph 24–25, 39,  
     40–41, 51, 121, 176–177  
 Mitchell, H. L. 50  
 moiré vibration 107–112  
*Le Monde* 83, 167  
 Mondrian, Piet 185  
 Monkhouse, F. J. 112, 165  
 Monty, Richard A. 168, 182  
 Moore, Carol 152  
 Moscow, Napoleon's campaign 40–41  
     121, 176–177  
 mountain-to-the-sea 154  
 multifunctioning elements 139–159  
 multiple viewing angles 154–155  
 multiple viewing depths 154–155  
 multiwindow plots 114, 171  
 Murphy, Edmond A. 172  
 Museum of Modern Art 91  
 Napoleon's march 40–41, 51, 121,  
     176–177  
 narrative graphics 40–43  
 National Science Foundation 60  
*Nature* 110, 167  
 Needham, Joseph 20  
*New England Journal of Medicine* 110,  
     167  
 Newman, Barnett 177  
 Newman, L. Hugh 43  
*Newsweek* 167  
 Newton, Isaac 181  
 New York City weather 30, 121, 164  
*New York Times* 30, 54, 57, 61, 66, 71,  
     76, 80–83, 86, 121, 164, 167, 179, 180  
*Nihon Keizai* 83  
 non-data-ink 96, 107  
*Nude Descending a Staircase* 36  
 ocean currents 99  
 oil prices 61–63  
 optical art 107–108  
 optical dots 114  
 orthogonal lines 185–186  
 outliers 142  
 parallel schematic plot 125  
 patch maps 20  
 Pauling, Linus 85, 102  
 Pearson, Karl 145  
 Peebles, P. James E. 26  
 periodicity of elements 102–105  
 perpendicular lines 185–186  
 Petchenik, Barbara Bartz 146  
 Phillips curve 48  
 pie chart 178  
 Pittsburgh Civic Commission 55  
 Playfair, William 9, 32–34, 43–45, 52,  
     64–65, 73, 91–92, 126, 148, 188–189  
 population density map 155–157  
 Potter, D. E. 145  
 Powell, R. W. 49, 150  
 Prakash, Om 172  
*Pravda* 83, 167  
 Pravda School of Ordinal Graphics 76  
 presidential vote graphics 147, 179  
 press, graphical sophistication 82–84  
*Proceedings of the National Academy of  
     Sciences, U.S.A.* 110  
*Psychological Bulletin* 167  
 psychological experiments 55–56  
 Pugin 117  
 puzzle graphics 153  
 quartile plot 124, 136, 139  
 range-frame 130, 136, 139, 149, 152  
 redundant data-ink 96–100  
 Reinhardt, Ad 122  
 Reiss, Albert J. 14  
 relational graphics 43–50  
 Rhône bridge 39  
 Rickett, Barney J. 134  
 Riedwyl, Hans 97  
 Riley, Bridget 108  
 road stripes 144  
 Roberts, K. V. 145  
 Robinson, Arthur H. 20, 40  
 Robyn, Dorothy L. 20  
 Rogers, Anna C. 112  
 Ross, H. Laurence 75  
 royalty, genealogy 34–35  
 rugplot 135, 136  
 Ruskin, John 117  
 sampling error 172  
 Sampson, Roy J. 85  
 Samuelson, Paul 85  
 sans serif 183  
 SAS/GRAF 112  
 Satet, R. 69, 112  
 Sato, Isao 82, 85  
 Scarf, F. L. 29  
 scatterplot 130–135  
 Schlee, Susan 172  
 Schmid, Calvin F. 112  
 Schmid, Stanton E. 112  
*Science* 39, 83, 109, 110, 167, 173  
*Science Indicators*, 1974 60  
*Scientific American* 50, 73, 167  
 sea-horse 36  
 Seinfeld, J. H. 170  
 Seldner, Michael 26  
 semi-graphics 178–180  
 Senders, John W. 168, 182  
 Shahn, Ben 177  
 shape, graphical 186–190  
 Shinohara, Miyohei 82, 85  
 Shiskin, Julius 38  
 Shrink Principle 169  
 Shryock, Henry S. 113  
 Siebers, B. H. 26  
 Siegel, Jacob S. 113  
 Silverstein, Louis 80  
 Slowinski, Emil 85  
 small multiple 42, 48, 170–175  
 smog 42, 170  
 Snow, Dr. John 24  
*Social Indicators*, 1973 163  
 sophistication, graphical 82–86  
 Spear, Mary Eleanor 81, 112, 123  
 starfish 36  
*Statistical Abstract of the United States*  
     161, 167  
 stem-and-leaf plot 140  
 Stockhausen, Karlheinz 184  
 stock prices 15  
 Stokes, Donald E. 147  
 strangers 142  
 Strunk, William 81  
*Sunday Times* (London) 63  
 sunshine, British 165  
 supertable 179  
 Swift, Jonathan 106  
 Sypher, Wylie 139  
 table-graphic 145, 158–159  
 tables 56, 178–180  
 Tanur, Judith 85  
 television graphics 76, 81  
 Tell-A-Graf 112  
 Terpenning, Irma J. 38  
 textbooks 84–85  
 textbooks, statistical graphics 112

text-tables 178–180  
 Thissen, David 142  
 Thomas, Edward Llewellyn 168  
 Thomas, LaVerne 85  
 Thompson, Morris M. 168  
 Thrower, Norman J. W. 23  
 Thurber, James 52  
 tick marks 127–129, 149  
 Tilling, Laura 32, 45–46  
*Time* 62, 71, 79, 83  
*The Times* (London) 83, 167  
 time-series 28–39  
 time-series, before-after 39  
 Todd, Lewis Paul 85  
 train schedule 31, 98, 115–116, 121  
 Troy, Nancy J. 185  
 Tufte, Edward R. 29, 53, 75, 95, 184  
 Tukey, John W. 53, 114, 123, 125,  
     129, 136, 140, 168, 188

Tukey, Paul A. 114, 168  
 tungsten, conductivity of 150  
 typography 178–183  
 Valéry, Paul 153  
 Vasarely, Victor 108  
 Vauthier, L. L. 154  
 Venturi, Robert 117, 139, 175  
 viewing architecture 159  
 Wainer, Howard 79, 142, 153  
 Waldrop, M. Mitchell 39  
*Wall Street Journal* 83, 167–168  
*Washington Post* 62, 70, 83  
 White, E. B. 81  
 white grid 127–129, 136  
 White, Jan 79–80  
 white pines 50

Wilk, Martin B. 53  
 Wilkinson, H. R. 112, 165  
 Wilson, James Q. 85  
 Wiskenmann, Arthur 171  
 World War I graphic 141  
 Wurman, Richard Saul 90  
 Yajima, Yokichi 85  
 Youden, W. J. 143  
 Yü Chi Thu map 20–21  
 Yunis, Jorge J. 172  
 Zeeman, E. C. 50  
 Zingale, Nancy H. 85  
 Zusne, Leonard 97, 190

The following illustrations are reprinted by permission:  
**CHAPTER 1.** Anscombe's quartet © 1973 American Statistical Association. Map of the galaxies, P. James E. Peebles, Princeton University. Radio emissions of Jupiter, Donald A. Gurnett, © 1979 American Association for the Advancement of Science. New York City weather history © 1981 The New York Times Company. Text on franked mail © 1975 The New York Times Company. Magnetic monopole © 1982 American Association for the Advancement of Science. Bridge on Rhône, Bibliothèque de l'École Nationale des Ponts et Chaussées. Los Angeles smog, by Bob Allen © 1979 The Los Angeles Times, based on data from Gregory J. McCrae, California Institute of Technology. Japanese beetle © 1965 Aldus Books, London. Phillips curve, Organisation for Economic Cooperation and Development, Paris. Dogs © 1976 Scientific American, Inc. **CHAPTER 2.** Payments to travel agents © 1978 The New York Times Company. Drawing by CEM © 1961 The New Yorker Magazine, Inc. Fuel economy standards © 1978 The New York Times Company. OPEC oil prices © 1978 The New York Times Company. Oil prices © 1979 Time, Inc. Oil prices © 1979 The Washington Post. Real price of oil, *Business Week*, © 1979 McGraw-Hill, Inc. Real price of oil © 1979 The Sunday Times, London. Real price of oil © 1979 The Economist. Growth of government © 1977 Yale University Press. New York budget © 1976 The New York Times Company. Shrinking doctor, by Bob Allen and Pete Bentevoja, © 1979 The Los Angeles Times. Shrinking dollar © 1978 The Washington Post. Picture of oil derricks © 1981 The New York Times Company. **CHAPTER 3.** Views on economy © 1980 The New York Times Company. Pace of city life © 1976 The New York Times Company. **CHAPTER 4.** Electroencephalogram © 1971 Harper & Row. Generation time © 1965 Princeton University Press. Registration rates © 1967 American Political Science Association. Registration rates © 1970 Addison-

Wesley Publishing Company. Registration rates © 1970 Holt, Rinehart and Winston. Ocean currents, Kirk Bryan, Princeton University and the American Meteorological Society. Bar chart, taste experiment, James T. Kuznicki, Proctor & Gamble Company. Roger Hayward drawing of periodic properties of elements, Linus Pauling. **CHAPTER 5.** Vertical bars, glucose transfer, Eain M. Cornford, © 1981 American Association for the Advancement of Science. Multiwindow plot, Paul A. Tukey. Big Duck, Peter Blake. California crops, State of California, Office of Planning and Research. Newspaper content © 1979 American Political Science Association. **CHAPTER 6.** Parallel schematic plot © 1977 Addison-Wesley Publishing Company. Box plot variation © 1978 American Statistical Association. Pulsar signals © 1975 Academic Press. Empirical cumulative distribution © 1976 John Wiley. **CHAPTER 7.** Stem-and-leaf plot © 1972 The Iowa State University Press. Living histogram, Brian Joiner. Magnetohydrodynamics © 1970 Academic Press. 1783 map © 1976 Princeton University Press. 1956 and 1960 elections © 1966 John Wiley. **CHAPTER 8.** French communes © 1973 Mouton and Gauthier-Villars. Bertin small multiple © 1973 Mouton and Gauthier-Villars. Melanoma © 1972 Springer-Verlag. Normal distributions © 1964 Pergamon Press, Ltd. Chromosomes, Jorge J. Yunis, © 1982 American Association for the Advancement of Science. Automobile frequency of repair © 1982 Consumers Union of the United States, Inc. **CHAPTER 9.** How different groups voted for president © 1980 The New York Times Company. Forecasting © 1979 The New York Times Company. Leonardo da Vinci drawing, facsimile © 1956 Princeton University Press. Musical score, Karlheinz Stockhausen, © 1964 Verlag M. DuMont Schauberg. Seats-votes graph © 1973 American Political Science Association. Poster illustration based on Burgoine Diller, Geometrical Composition in Black, Red, and White, Yale University Art Gallery, Collection Société Anonyme.

## *Acknowledgments*

I am indebted to many for their advice and assistance with this book.

For leave and research support during several academic years,  
the Center for Advanced Study in the Behavioral Sciences, the  
John Simon Guggenheim Foundation, the Woodrow Wilson School  
of Princeton University, and Yale University.

For providing access to their superb collections, the Bibliothèque  
 Nationale and the Bibliothèque de l'École Nationale des Ponts et  
 Chaussées in Paris, and, at Yale University, the Historical Medical  
 Library and the Beinecke Rare Book and Manuscript Library.

For helping me appreciate the practicalities in the production  
 of statistical graphics, several members of the art department at the  
 *New York Times* and my students in Graphic Design and in the  
 Department of Statistics at Yale.

For assistance in establishing Graphics Press, Peter B. Cooper,  
 Earle E. Jacobs, Jr., and Trudy Putsche.

For design and artwork, Howard I. Gralla and Minoru Niijima.

For their help and hospitality in Paris during my work on the  
 Minard drawings, Michel Balinski, Jean Dubout, André Jammes,  
 and Claudine Kleb.

For providing examples and for suggesting improvements in the  
 manuscript, James Beniger, Inge Druckrey, Timothy Gregoire,  
 Joanna Hitchcock, Joseph LaPalombara, Kathryn Scholle, Stephen  
 Stigler, Howard Wainer, and Ellen Woodbury.

For their reviews of the manuscript and for their inspiration and  
 encouragement through all the years of this enterprise, Frederick  
 Mosteller and John W. Tukey.

*June 1982  
Cheshire, Connecticut*