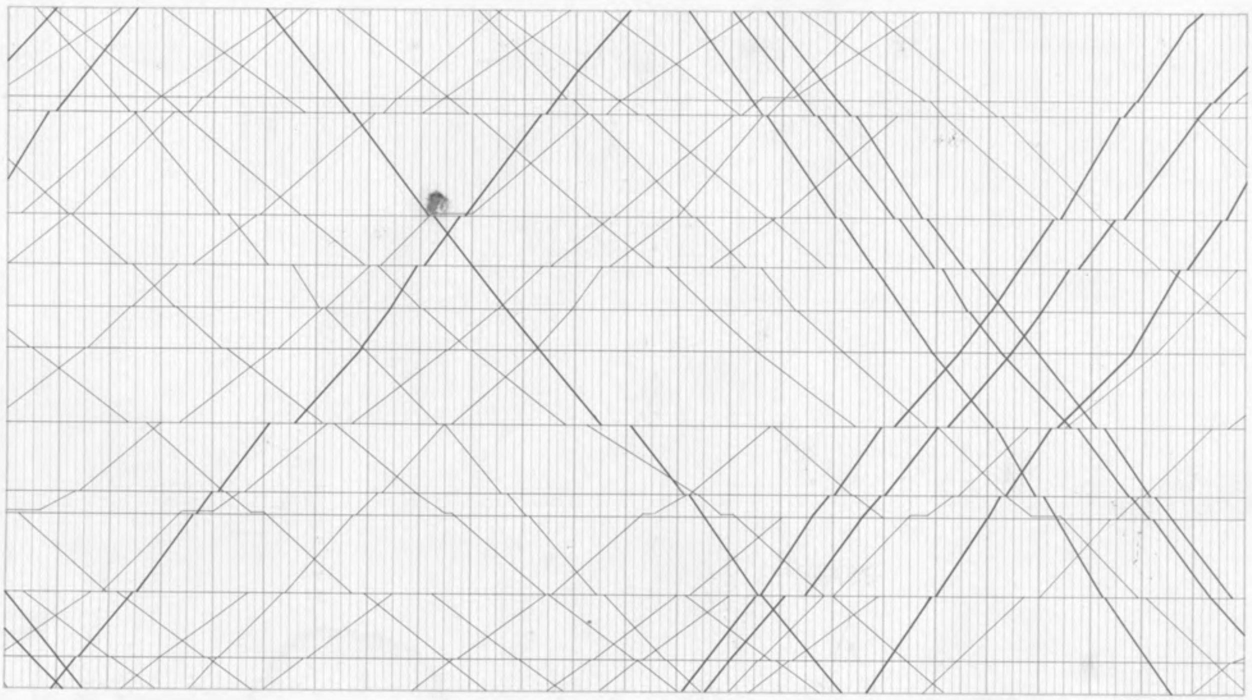


SECOND EDITION

The Visual Display of Quantitative Information

EDWARD R. TUFTE

7 *Multifunctioning Graphical Elements*

The same ink should often serve more than one graphical purpose. A graphical element may carry data information and also perform a design function usually left to non-data-ink. Or it might show several different pieces of data. Such *multifunctioning graphical elements*, if designed with care and subtlety, can effectively display complex, multivariate data.[1]

Consider, for example, the multifunctioning blot of the blot map. It simultaneously locates the geographic unit on a two-dimensional surface, describes the shape of the geographic unit, and indicates the level of the variable displayed by color or intensity of shading. That is a great deal of information for a small patch of ink—and the different pieces of information are not confounded and mixed together.

In contrast, the conventional graphical frame performs only a modest design function, the separation of the grid and data measures from the labels. And it is a place to hang the grid ticks. With all that ink doing so little, it is a prime candidate for mobilization as a double-functioning graphical element. Hence the range-frame, the quartile frame, and the dot-dash-plot.

The principle, then, is:

> Mobilize every graphical element, perhaps several times over, to show the data.

The danger of multifunctioning elements is that they tend to generate graphical puzzles, with encodings that can only be broken by their inventor. Thus design techniques for enhancing graphical clarity in the face of complexity must be developed along with multifunctioning elements.

[1] The idea of double-functioning elements appears in architectural criticism; see Robert Venturi, *Complexity and Contradiction in Architecture* (New York, second edition, 1977), ch. 5. Venturi in turn cites Wylie Sypher, *Four Stages of Renaissance Style* (Garden City, N.Y., 1955).

Data-Built Data Measures

The graphical element that actually locates or plots the data is the *data measure*. The bars of a bar chart, the dots of a scatterplot, the dots and dashes of a dot-dash-plot, the blots of a blot map are data measures. The ink of the data measure can itself carry data; for example, the dots of the scatterplot can take on different shadings in response to a third variable.

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 feet

0	98766562
1	97719630
2	69987766544422211009850
3	876655412099551426
4	9998844331929433361107
5	97666666554422210097731
6	898665441077761065
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

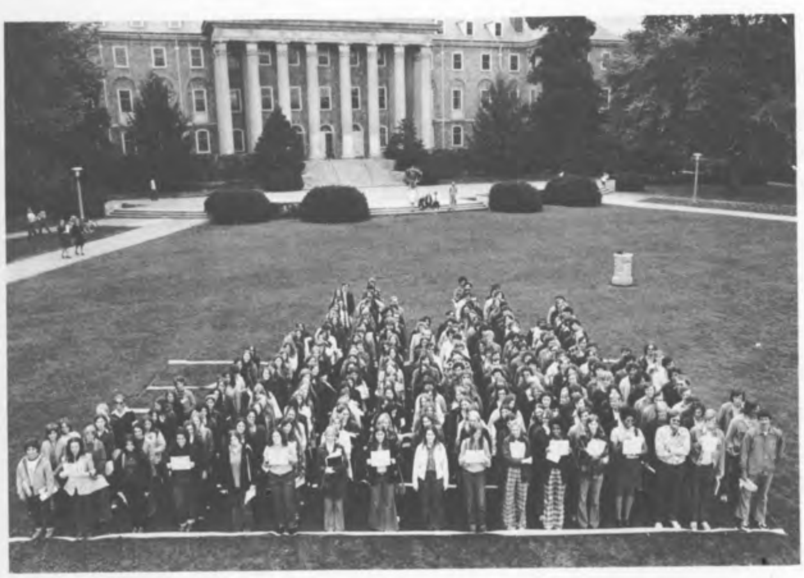
Stem-and-leaf displays:
heights of 218 volcanoes, unit 100 feet.

19|3 = 19,300 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."[2]

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

[2] "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.

A distinguished graphic that builds data measures out of data was designed by Colonel Leonard P. Ayres for his statistical history of World War I, a book with several notable graphics all done by typewriter and rule. Constructing the data measures out of each American division's name (a numerical designation) turns what might have been a routine time-series into an elegant display. (Note that the cumulative design depends on the fact that none of the divisions returned before October 1918.) The triple-functioning data measure shows: (1) the number of divisions in France for each month, June 1917 to October 1918; (2) what particular divisions were in France in each month; and (3) the duration of each division's presence in France.

Leonard P. Ayres, *The War with Germany* (Washington, D.C., 1919), p. 102.

																8
																38
																31
															34	34
															86	86
															84	84
															87	87
														40	40	40
														39	39	39
														88	88	88
														81	81	81
														7	7	7
														85	85	85
													36	36	36	36
													91	91	91	91
													79	79	79	79
													76	76	76	76
												29	29	29	29	29
												37	37	37	37	37
												90	90	90	90	90
												92	92	92	92	92
												89	89	89	89	89
												83	83	83	83	83
												78	78	78	78	78
											80	80	80	80	80	80
											30	30	30	30	30	30
											33	33	33	33	33	33
											6	6	6	6	6	6
											27	27	27	27	27	27
											4	4	4	4	4	4
											28	28	28	28	28	28
											35	35	35	35	35	35
											82	82	82	82	82	82
										77	77	77	77	77	77	77
									3	3	3	3	3	3	3	3
									5	5	5	5	5	5	5	5
								32	32	32	32	32	32	32	32	32
						41	41	41	41	41	41	41	41	41	41	41
					42	42	42	42	42	42	42	42	42	42	42	42
			26	26	26	26	26	26	26	26	26	26	26	26	26	26
		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct

1917 1918

Encoding of data measures can be far more elaborate. The plotted points here are Chernoff faces, which reduce well, maintaining legibility even with individual areas of .05 square inches as shown.[3] The analyst would observe the standard X-Y scatterplot and then turn to the within-scatter detail, seeking clusters of similar observations over the X-Y plane. Outlying faces and those inconsistent with others in the neighborhood—they are, of course, *strangers*—should be identified by observation number or name.

[3] Herman Chernoff, "The Use of Faces to Represent Points in k-Dimensional Space Graphically," *Journal of the American Statistical Association* 68 (June 1973), 361–368. For an application of faces located over two dimensions, see Howard Wainer and David Thissen, "Graphical Data Analysis," *Annual Review of Psychology*, 32 (1981), 191–241.

With cartoon faces and even numbers becoming data measures, we would appear to have reached the limit of graphical economy of presentation, imagination, and, let it be admitted, eccentricity.

But let us consider this shaped poem, "Easter Wings" by George Herbert (1593–1633), which uses space—the length of each line—to depict quantity, all done 150 years before Playfair. The lines double-function: the longer lines describe wealth, plenty, largesse, and rising to flight; shorter lines tell of poverty and becoming "most thinne"; and lines of intermediate length indicate transition and change (decaying, rising, combining, becoming):

Easter-wings.

LOrd, who createdst man in wealth and store,
Though foolishly he lost the same,
Decaying more and more,
Till he became
Most poore:
With thee
O let me rise
As larks, harmoniously,
And sing this day thy victories:
Then shall the fall further the flight in me.

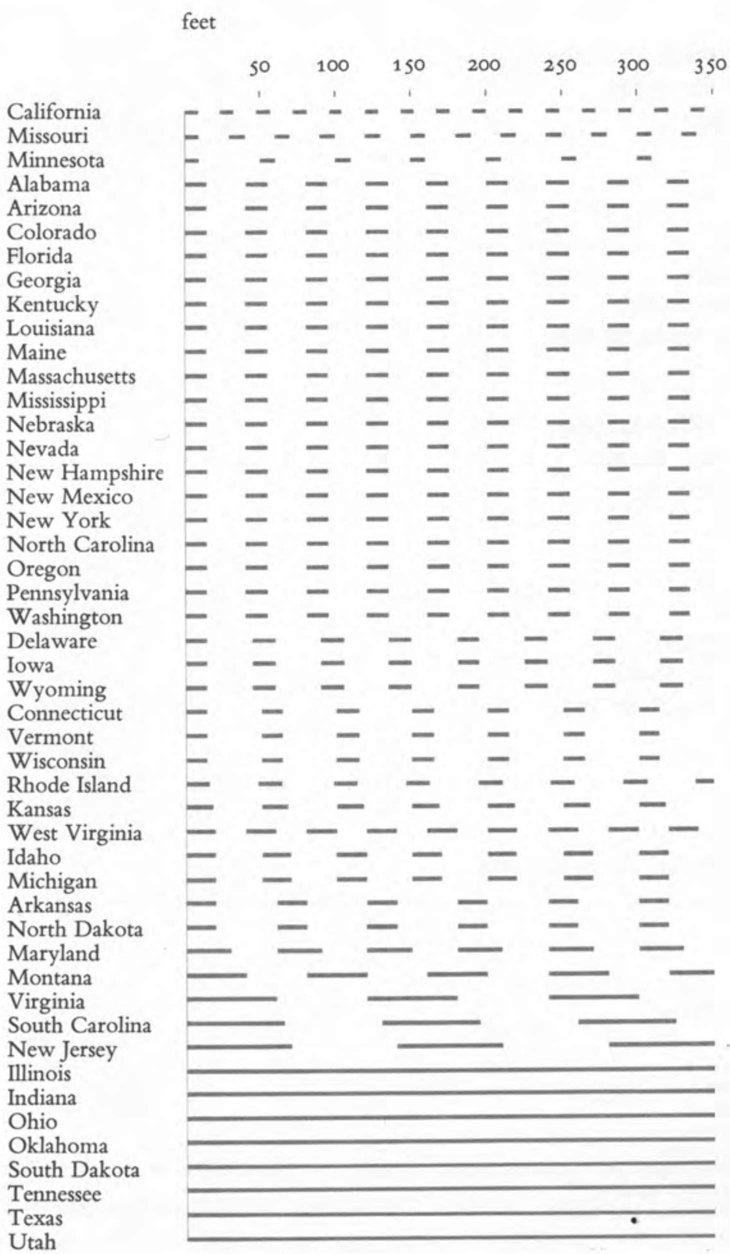
My tender age in sorrow did beginne:
And still with sicknesses and shame
Thou didst so punish sinne,
That I became
Most thinne.
With thee
Let me combine
And feel this day thy victorie:
For, if I imp my wing on thine,
Affliction shall advance the flight in me.

And the typographical delight of the statistician W. J. Youden:

THE
NORMAL
LAW OF ERROR
STANDS OUT IN THE
EXPERIENCE OF MANKIND
AS ONE OF THE BROADEST
GENERALIZATIONS OF NATURAL
PHILOSOPHY ◆ IT SERVES AS THE
GUIDING INSTRUMENT IN RESEARCHES
IN THE PHYSICAL AND SOCIAL SCIENCES AND
IN MEDICINE AGRICULTURE AND ENGINEERING ◆
IT IS AN INDISPENSABLE TOOL FOR THE ANALYSIS AND THE
INTERPRETATION OF THE BASIC DATA OBTAINED BY OBSERVATION AND EXPERIMENT

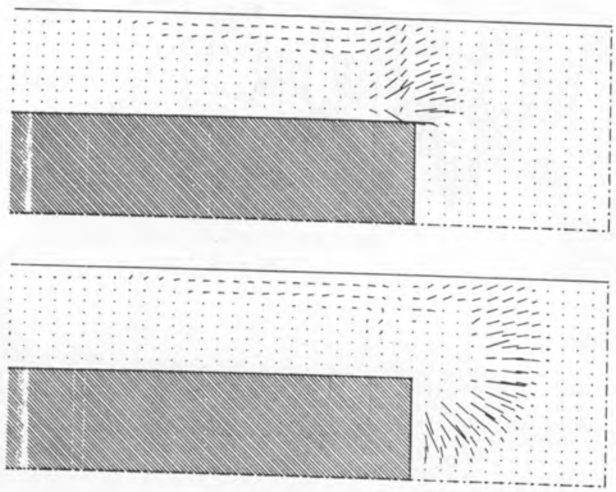
Finally, this graphical pun: the visual data as the data measure, as in the living histogram. The chart shows how states once differed in their engineering standards for painting lane stripes on road pavement. Some states marked the road lanes with short dashes and long gaps; others used only solid lines. Portrayed in the graphic is the actual physical pattern painted on the road, with 48 U.S. states ordered by the length of the painted mark:

Redrawn from A. R. Lauer, "Psychological Factors in Effective Traffic Control Devices," *Traffic Quarterly*, 5 (January 1951), 94.



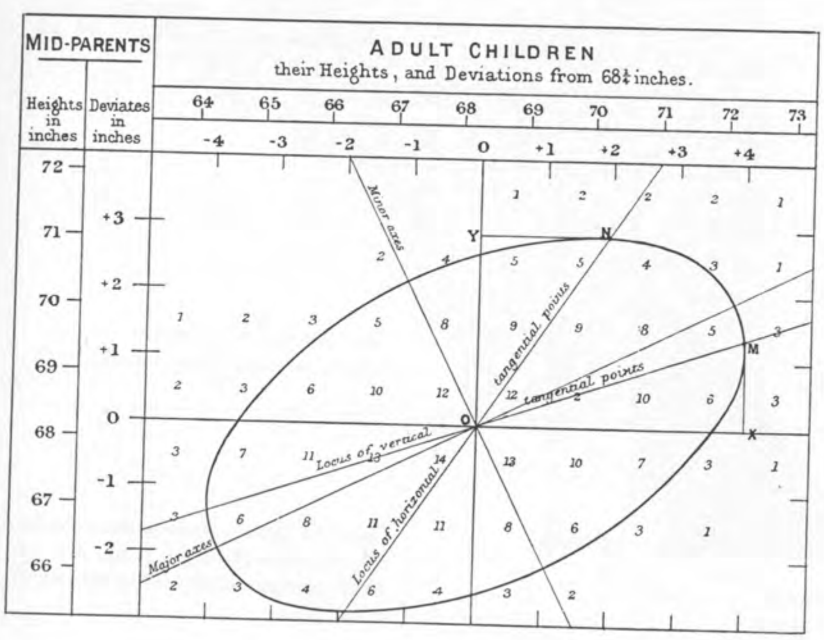
Data-Based Grids

Very occasionally the grid can report directly on the data. This grid is formed by the location of measurement instruments; the plain dots register a zero reading, in contrast with the white background where no readings were taken. Erasing the grid would erase measured data (rather uneventful, to be sure). Such is not the case for most grid dots, ticks, and lines.

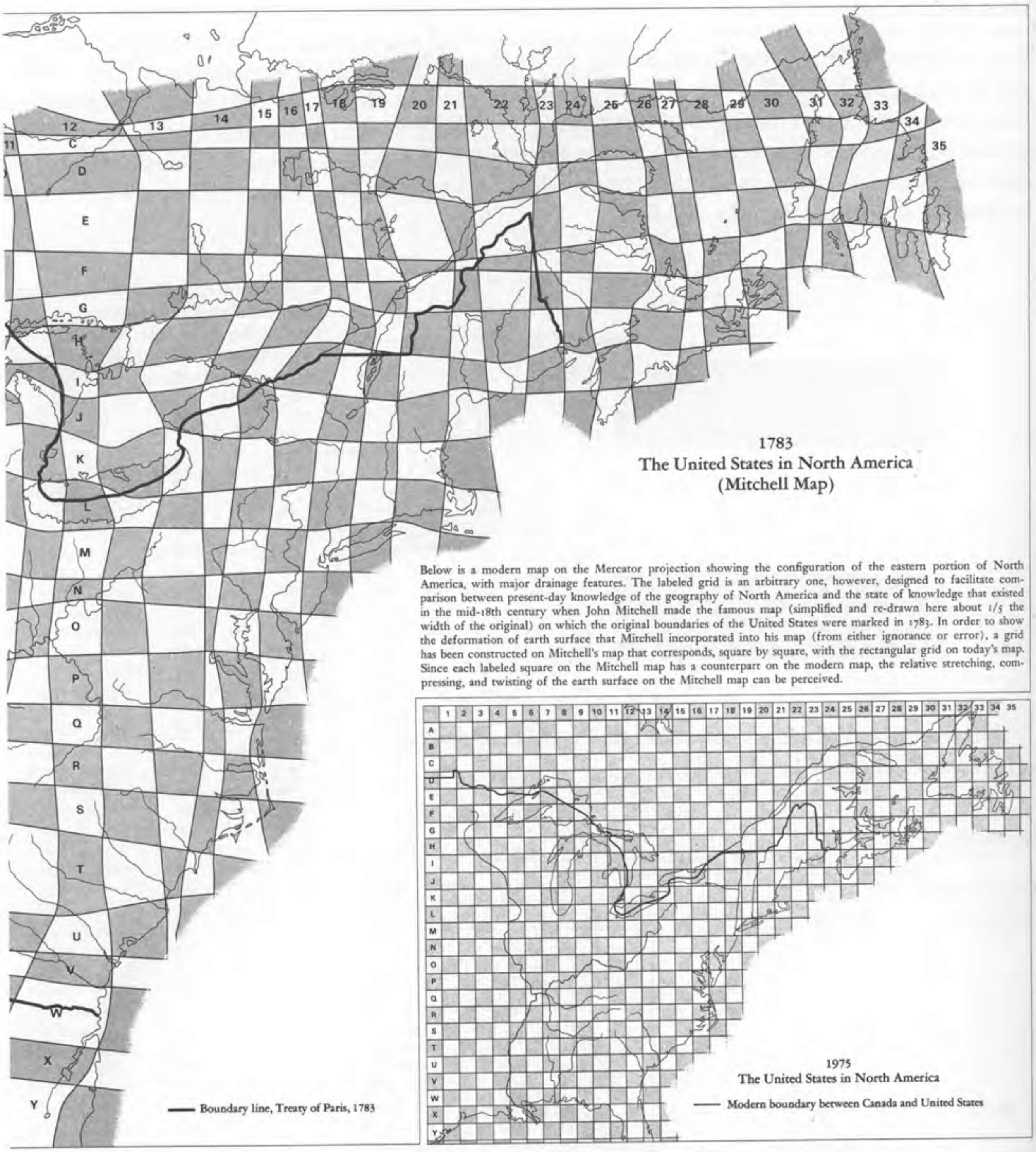


K. V. Roberts and D. E. Potter, "Magnetohydrodynamic Calculations," in Berni Alder, et al., eds., *Methods in Computational Physics: Volume 9, Plasma Physics* (New York, 1970), p. 402.

The arrangement of data in this table-graphic yields an internal grid, a rare example of data as grid:



Karl Pearson, *The Life, Letters and Labours of Francis Galton* (Cambridge, 1930), vol. III-A, 14.

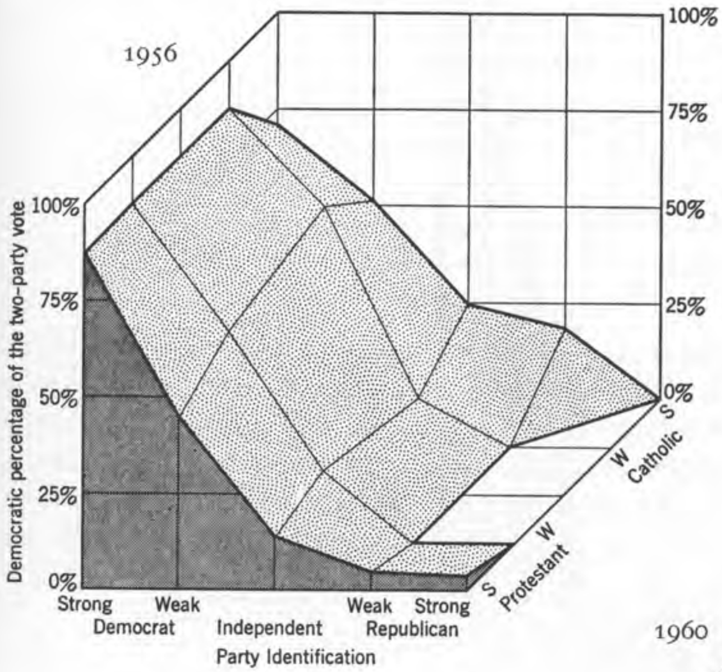


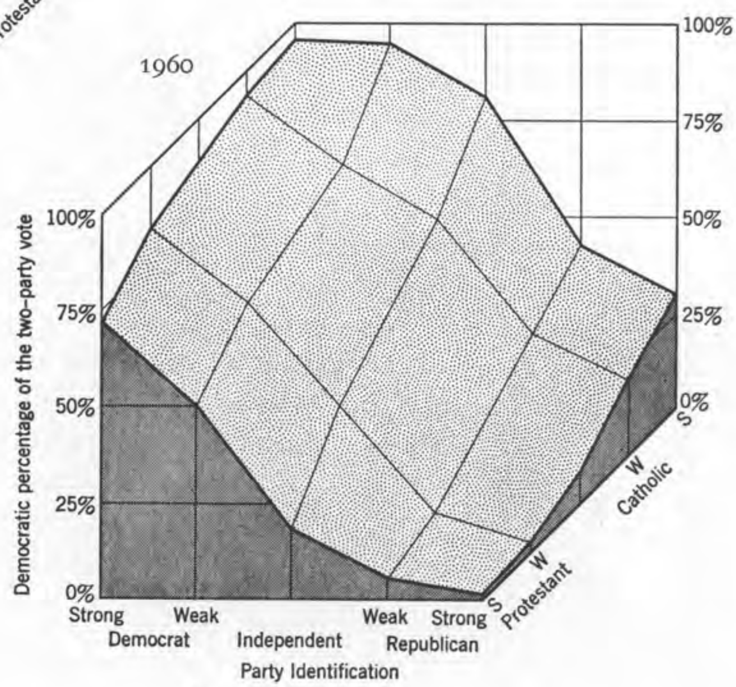
Here the grid is the element of interest, rather than the map.

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

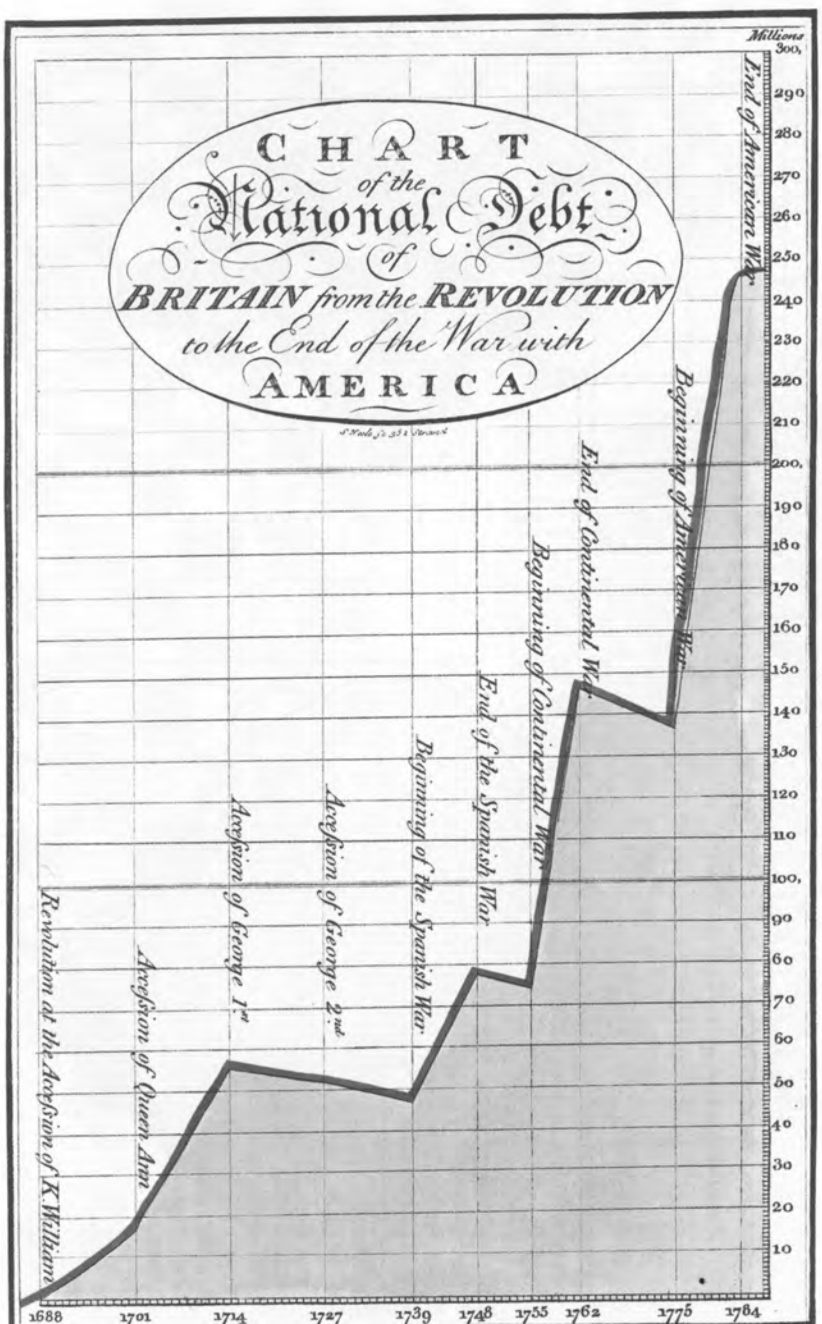
The grid that follows presents the data on the surface of the rock; on the sides, the grid is conventional. The two displays compare the effect of religion, taking into account party affiliation, on a person's vote for president in 1956 and in 1960 (when a Catholic ran for president). Note there is no reliable slope associated with religion in 1956, once party is controlled; in 1960, a systematic effect is found. Reading the slopes in the other direction shows the persistent effect of party in both elections:

Philip E. Converse, "Religion and Politics: The 1960 Election," in Angus Campbell, Philip E. Converse, Warren E. Miller, and Donald E. Stokes, *Elections and the Political Order* (New York, 1966), 102–103.



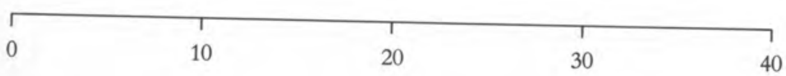


Playfair tied the grid to the data in his skyrocketing debt graphic. Although the implicit plotting coordinates are based on regular intervals, the vertical grid lines in the published version are irregularly spaced, keyed to significant events. The data-based grid is a shrewd graphical device, serving rather than fighting with the data. It is a technique underused in contemporary graphical work.

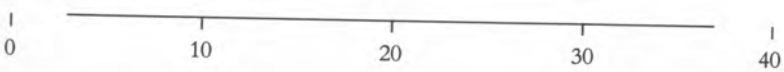


Double-Functioning Labels

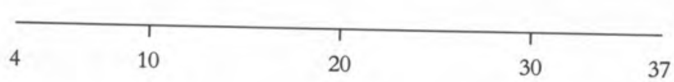
Data-based coordinate lines lead to *data-based labels,* as, for example, at the bottom of Playfair's debt graphic. Again, the issue is the same: why not use the ink to show data? Beginning with conventionally labeled frame



and erasing to the range-frame



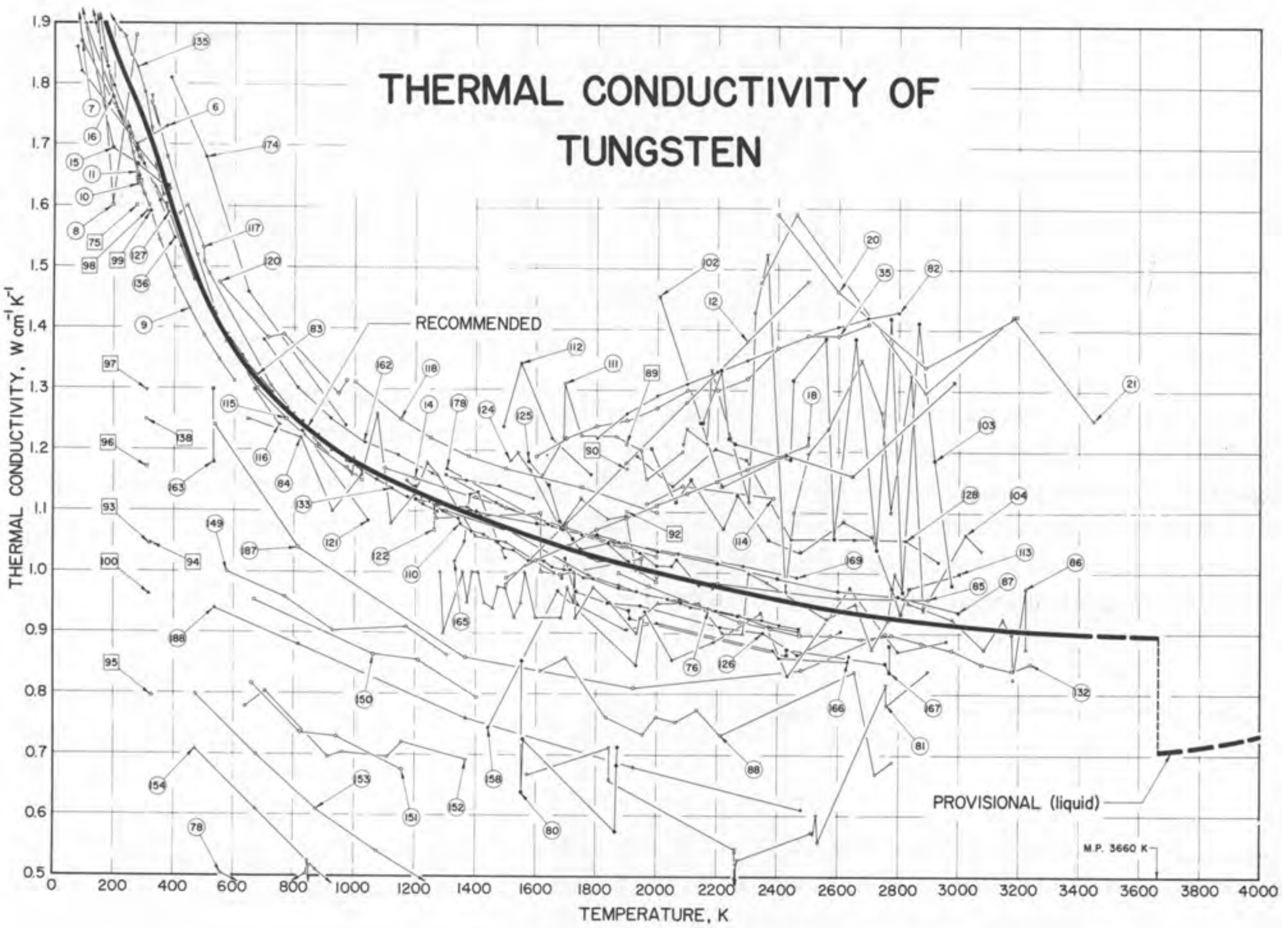
leaves those lonely ticks and numbers out on the tails, working to help the eye get a better reading on where the line of the range-frame ends. But that job can be done better by investing the same ink in data: rather than showing the minimum round number and the maximum round number at the ends of the frame, show the actual minimum and maximum realized in the data:



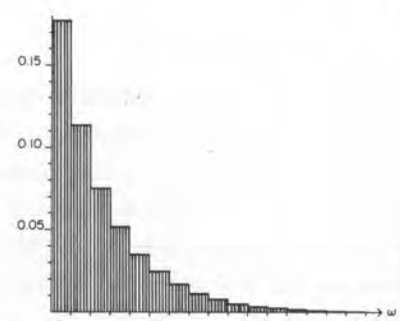
With its greater precision and two tick-marks less of non-data-ink, the range-frame with range-labels is superior to the range-frame with round number labels. Both improve on the standard, passive frame.

Numbers also double-function when used both to name things (like an identification number) and to reflect an ordering. In this graphic (in which the circled numbers fail to double-function), each number identifies a particular study of the thermal conductivity of tungsten, ordered alphabetically by the last name of the first author. If that list were ordered by date of publication instead, then the code would also indicate the time order in which the various conductivity determinations were made. Thus "1" would indicate the earliest study, and so on—or, alternatively, "61c" would be the third study published in 1961. Such information has interest, since we could see which of the early studies got the right answer. In addition, the movement of the studies toward the "correct" recommended values could be tracked. This extra information requires no additional ink.

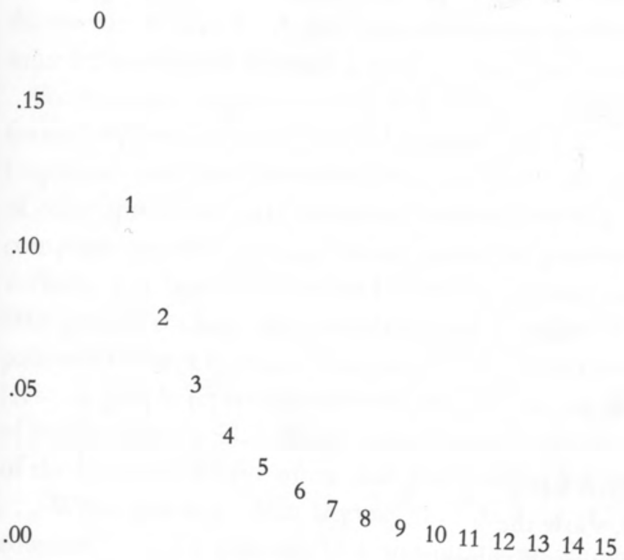
C. Y. Ho, R. W. Powell, and P. E. Liley, *Thermal Conductivity of the Elements: A Comprehensive Review*, supplement no. 1, *Journal of Physical and Chemical Reference Data*, 3 (1974), I–692.



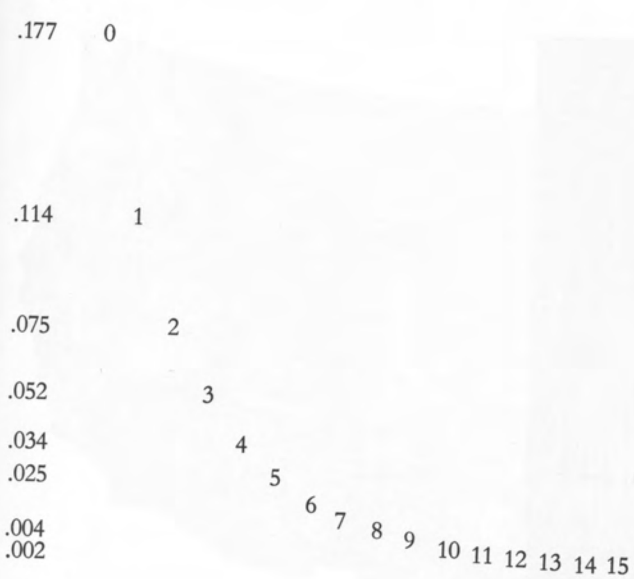
In most graphics, the coordinate labels are far from the data measures. Consequently the eye of the viewer must move back and forth between the path formed by the data and the coordinate positions arrayed along the margins of the graphic. Sometimes this eye-work can be eliminated entirely by turning the coordinate labels into data measures, another double-functioning maneuver. Take the example from the style sheet of the *Journal of the American Statistical Association*:



The grid increments of the X-axis are relocated upward to mark the path of the data:

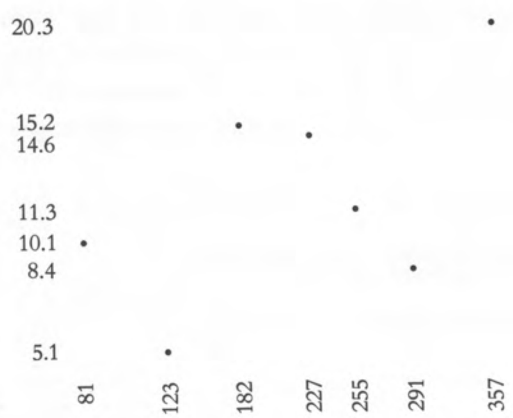


And since the issue in this display is the probability at each integer value, the round-number Y-scale is replaced by exact values:



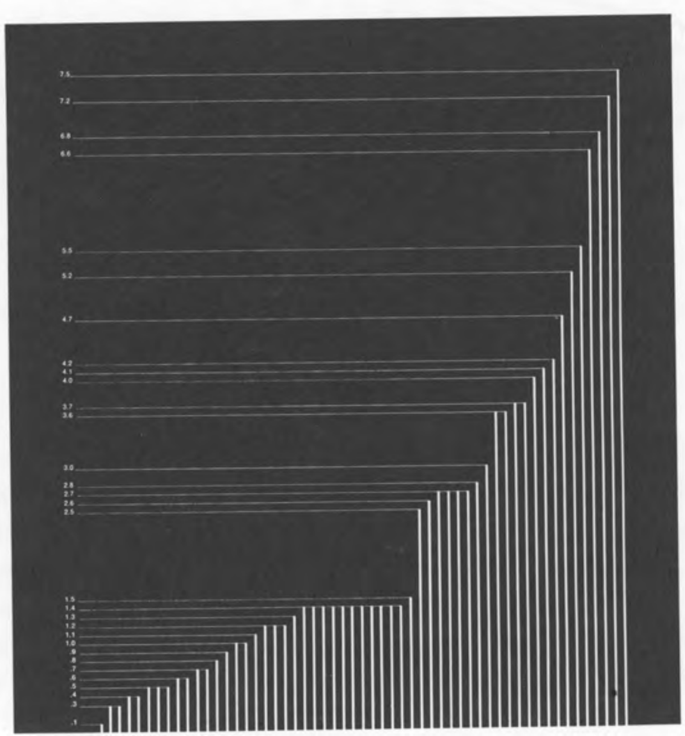
The Y-scale now resembles the dashes of the dot-dash-plot, with the vertical column of data-positioned numbers serving as the dashes to indicate the marginal distribution.

The method of data-based markers for the marginal distributions suggests a further enhancement of the dot-dash-plot:



Now the numbers in the margin eliminate the standard frame and even a range-frame, replace the coordinate ticks, show the marginal distribution of both variables, and record the exact values of the two measurements made on each unit of observation. This graphical arrangement performs better for smaller data sets (say 30 observations or less) and when a fine level of detail is required.

Finally, a striking design with data-based labels:



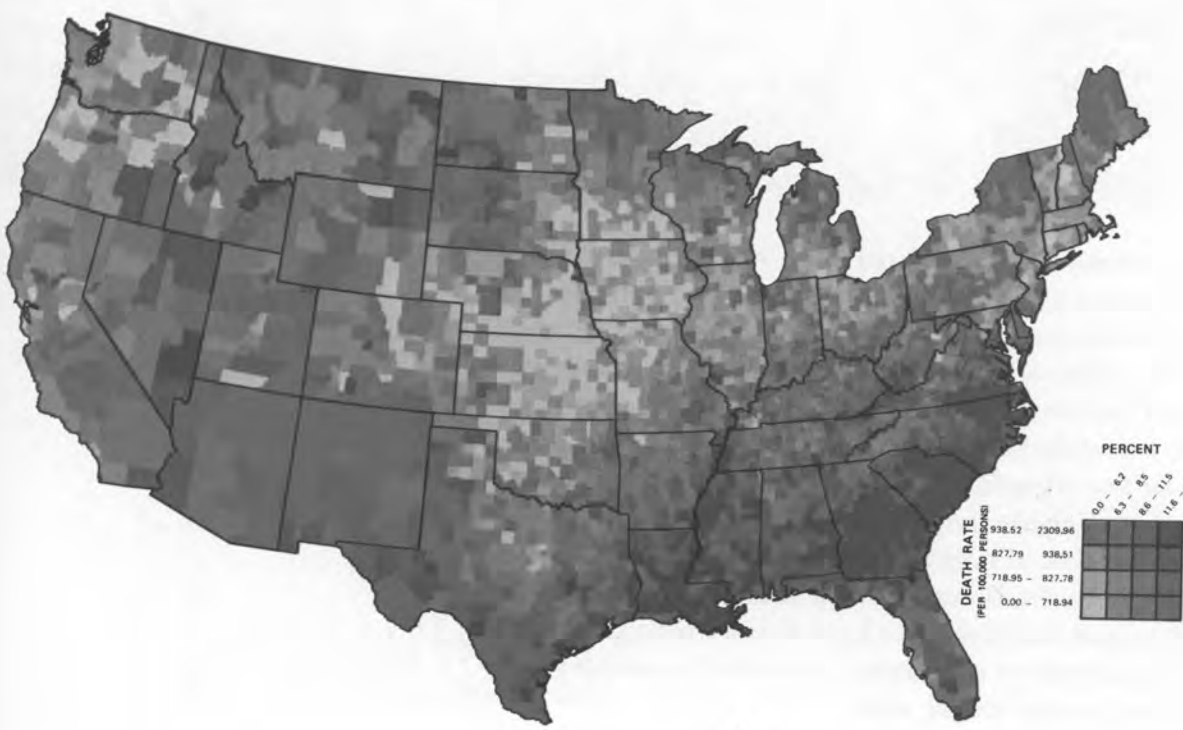
Designed by Carol Moore, Corporate Annual Reports, Inc., in Walter Herdeg, *Graphis/Diagrams* (Zurich, 1976), p. 23.

Puzzles and Hierarchy in Graphics

The complexity of multifunctioning elements can sometimes turn data graphics into visual puzzles, crypto-graphical mysteries for the viewer to decode. A sure sign of a puzzle is that the graphic must be interpreted through a verbal rather than a visual process.

For example, despite its clever and multifunctioning data measure, formed by crossing two four-color grids, this is a puzzle graphic. Deployed here, in a feat of technological virtuosity, are 16 shades of color spread on 3,056 counties, a monument to a sophisticated computer graphics system.[4] But it is surely a graphic experienced verbally, not visually. Over and over, the viewers must run little phrases through their minds, trying to maintain the right pattern of words to make sense out of the visual montage: "Now let's see, purple represents counties where there are both high levels of male cardiovascular disease mortality *and* 11.6 to 56.0 percent of the households have more than 1.01 persons per room. . . . What does that mean anyway? . . . And the yellow-green counties. . . ." By contrast, in a non-puzzle graphic, the translation of visual to verbal is quickly learned, automatic, and implicit —so that the visual image flows right *through* the verbal decoder initially necessary to understand the graphic. As Paul Valéry wrote, "Seeing is forgetting the name of the thing one sees."

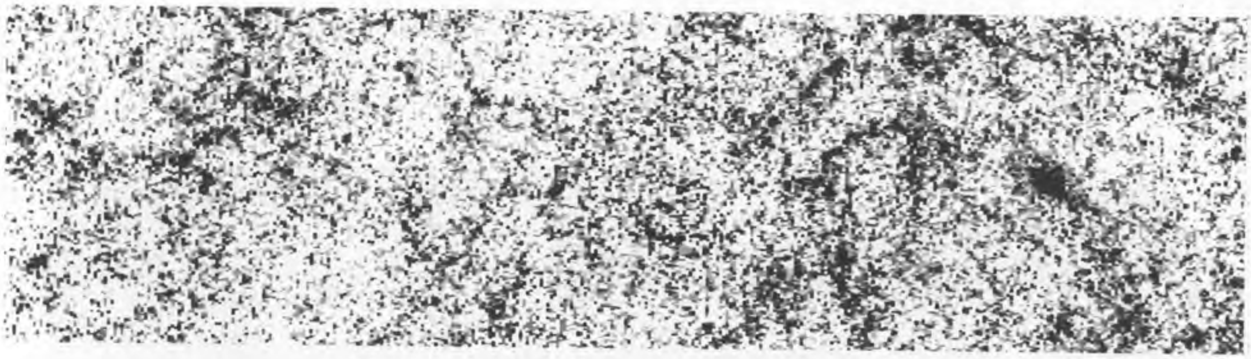
[4] The technique is described in Vincent P. Barabba and Alva L. Finkner, "The Utilization of Primary Printing Colors in Displaying More than One Variable," in Bureau of the Census, Technical Paper No. 43, *Graphical Presentation of Statistical Information* (Washington, D.C., 1978), 14–21. The maps are assessed in Howard Wainer and C. M. Francolini, "An Empirical Inquiry Concerning Human Understanding of Two-Variable Color Maps," *American Statistician*, 34 (1980), 81–93.



Color often generates graphical puzzles. Despite our experiences with the spectrum in science textbooks and rainbows, the mind's eye does not readily give a visual ordering to colors, except possibly for red to reflect higher levels than other colors, as in the hot spots of the cancer map. Attempts to give colors an order result in those verbal decoders and the mumbling of little mental phrases again—indeed, even mnemonic phrases *about* the phrases required for graphical decoding:

> A method of coloring ingenious in idea but not very satisfactory in practice was used by L. L. Vauthier. It was called the mountain-to-the-sea method. White was used for the representation of the greatest intensity of the fact because it indicated the summit of a mountain with its eternal snow, next came green representing the forests farther down the slopes, then yellow for the grain of the plains, and finally for the minimum the blue of the waters at sea level.[5]

Because they do have a natural visual hierarchy, varying shades of gray show varying quantities better than color. Ten gray shades worked effectively in the galaxies map:



The success of gray compared to the visually more spectacular color gives us a lead on how multifunctioning graphical elements can communicate complex information without turning into puzzles. The shades of gray provide an easily comprehended order to the data measures. This is the key. Central to maintaining clarity in the face of the complex are graphical methods that *organize and order the flow of graphical information* presented to the eye.

How can graphical architecture promote the ordered, sequenced, hierarchical flow of information from the graphic to the mind's eye? How can the data-information be arranged so that the viewer is able to peel back layer after layer of data from a graphic?

Multiple layers of information are created by *multiple viewing depths* and *multiple viewing angles*.

[5] H. Gray Funkhouser, "Historical Development of the Graphical Representation of Statistical Data," *Osiris*, 3 (1937), 326, who cites É. Cheysson, "Les méthodes de statistique graphique à l'Exposition universelle de 1878," *Journal de la Société de Statistique de Paris*, 19 (1878), 331.

Graphics can be designed to have at least three viewing depths: (1) what is seen from a distance, an overall structure usually aggregated from an underlying microstructure; (2) what is seen up close and in detail, the fine structure of the data; and (3) what is seen implicitly, underlying the graphic—that which is behind the graphic. Look at all the different levels of detail created by this population density map of the United States, a glory of modern cartography prepared by the Bureau of the Census. Each dot, except in urban centers, represents 500 people. Note the corridors connecting the major urban complexes; the effects of landforms on the population distribution (the central valley of California, the valleys and ridges of Appalachia, and the clusters along rivers); and the small towns along the highways, linked like a string of pearls. The map arrays, in effect, some 400,000 points on its implicit grid.

Different visual angles for different aspects of the data also organize graphical information. Each separate line of sight should remain unchanging (preferably horizontal or vertical) as the eye watches for data variation off the flat of the line of sight. For multivariate work, several clear lines can be created. Recall Ayres' display of American divisions in France. Even with its complex, interwoven data, the graphic is not a puzzle. Three separate visual angles make the flow of information coherent: the profile of the horizon for the upward-moving time-series, the vertical for the composition of the bar, and the horizontal for each division's stay. Thus while every drop of ink serves three different data display functions, each of the three comes to the eye with its own independence and integrity.

Population Distribution, Urban and Rural, in the United States: 1970


URBAN POPULATION
URBANIZED AREAS
Extent of areas
PLACES OUTSIDE URBANIZED AREAS
25,000 - 50,000
10,000 - 25,000
2,500 - 10,000
RURAL POPULATION
Places of 1,000-2,500
Each dot represents 500 of remaining population