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.

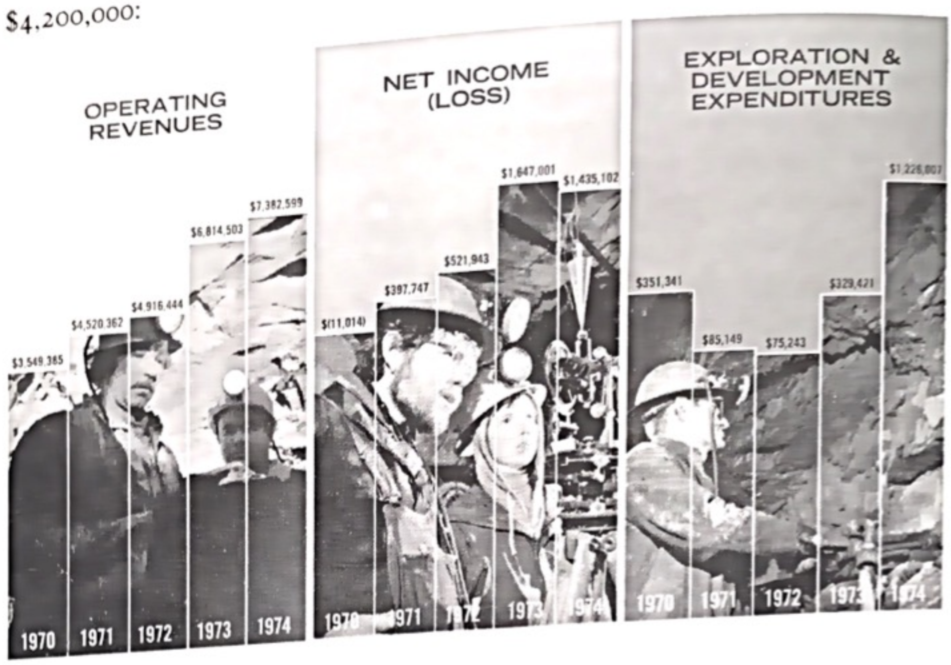
Much of twentieth-century thinking about statistical graphics has been preoccupied with the question of how some amateurish chart might fool a naive viewer. Other important issues, such as the use of graphics for serious data analysis, were largely ignored. At the core of the preoccupation with deceptive 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.[1] 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.

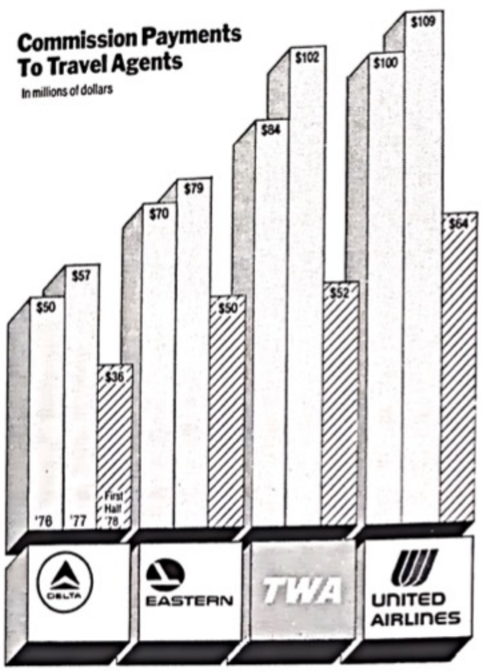
[1] John W. Tukey and Martin B. Wilk, "Data Analysis and Statistics: Techniques and Approaches," in Edward R. Tufte, ed., *The Quantitative Analysis of Social Problems* (Reading, Mass., 1970), 370–390; and John W. Tukey, "Some Graphic and Semigraphic Displays," in T. A. Bancroft, ed., *Statistical Papers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

Here are several graphics that fail to tell the truth. First, the case of the disappearing baseline in the annual report of a company that would just as soon forget about 1970. A careful look at the middle panel reveals a negative income in 1970, which is disguised by having the bars begin at the bottom at approximately minus $4,200,000:



Day Mines, Inc., 1974 Annual Report, 1.

This pseudo-decline was created by comparing six months' worth of payments in 1978 to a full year's worth in 1976 and 1977, with the lie repeated four times over.



New York Times, August 8, 1978, D-1.

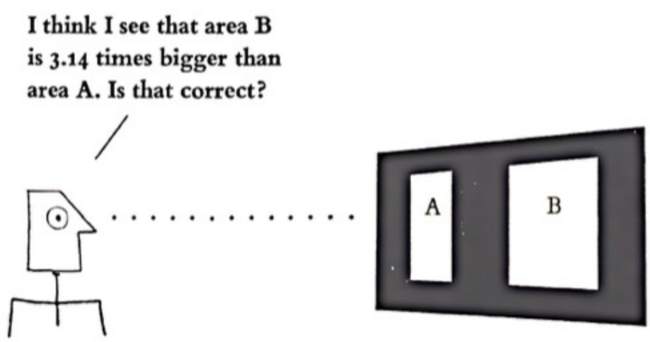
And sometimes the fact that numbers have a magnitude as well as an order is simply forgotten:

Pittsburgh Civic Commission, *Report on Expenditures of the Department of Charities* (Pittsburgh, 1911), 7.

What is Distortion in a Data Graphic?

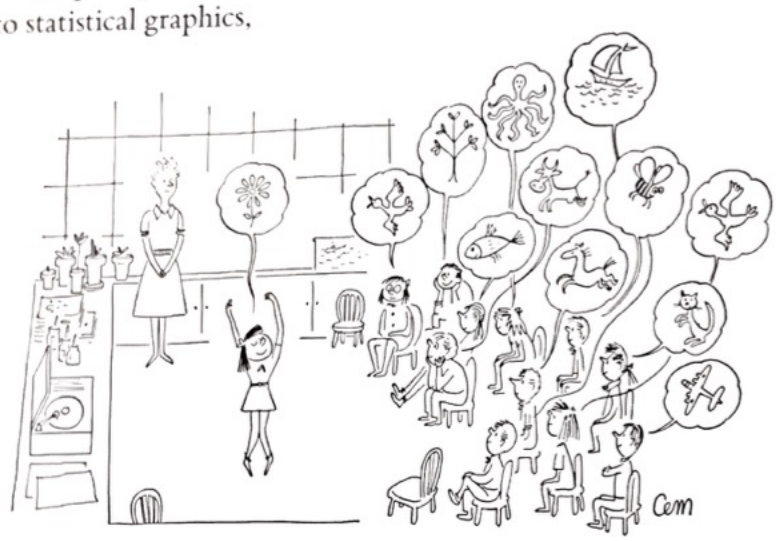
A graphic does not distort if the visual representation of the data is consistent with the numerical representation. What then is the "visual representation" of the data? As physically measured on the surface of the graphic? Or the *perceived* visual effect? How do we know that the visual image represents the underlying numbers?

One way to try to answer these questions is to conduct experiments on the visual perception of graphics—having people look at lines of varying length, circles of different areas, and then recording their assessments of the numerical quantities.



Such experiments have discovered very approximate power laws relating the numerical measure to the reported perceived measure. For example, the perceived area of a circle probably grows somewhat more slowly than the actual (physical, measured) area: the reported perceived area = $(\text{actual area})^x$, where $x = .8 \pm .3$, a discouraging result. Different people see the same areas somewhat differently; perceptions change with experience; and perceptions are context-dependent.[2] Particularly disheartening is the securely established finding that the reported perception of something as clear and simple as line length depends on the context and what other people have already said about the lines.[3]

Misperception and miscommunication are certainly not special to statistical graphics,



but what is a poor designer to do? A different graphic for each perceiver in each context? Or designs that correct for the visual transformations of the average perceiver participating in the average psychological experiment?

One satisfactory answer to these questions is to use a table to show the numbers. Tables usually outperform graphics in reporting on small data sets of 20 numbers or less. The special power of graphics comes in the display of large data sets.

At any rate, given the perceptual difficulties, the best we can hope for is some uniformity in graphics (if not in the perceivers) and some assurance that perceivers have a fair chance of getting the numbers right. Two principles lead toward these goals and, in consequence, enhance graphical integrity:

> The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities represented.
>
> Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.

[2] The extensive literature is summarized in Michael Macdonald-Ross, "How Numbers are Shown: A Review of Research on the Presentation of Quantitative Data in Texts," *Audio-Visual Communication Review*, 25 (1977), 359–409. In particular, H. J. Meihoefer finds great variability among perceivers; see "The Utility of the Circle as an Effective Cartographic Symbol," *Canadian Cartographer*, 6 (1969), 105–117; and "The Visual Perception of the Circle in Thematic Maps: Experimental Results," ibid., 10 (1973), 63–84.

[3] S. E. Asch, "Studies of Independence and Submission to Group Pressure. A Minority of One Against a Unanimous Majority," *Psychological Monographs* (1956), 70.

Drawing by CEM; copyright 1961, *The New Yorker*.

Violations of the first principle constitute one form of graphic misrepresentation, measured by the

$$\text{Lie Factor} = \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}}$$

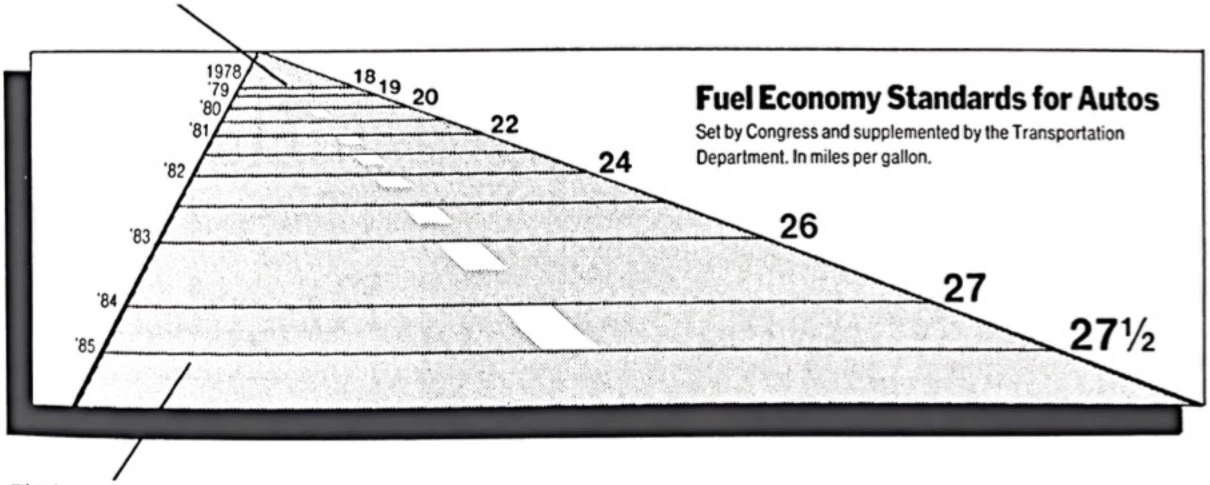
If the Lie Factor is equal to one, then the graphic might be doing a reasonable job of accurately representing the underlying numbers. Lie Factors greater than 1.05 or less than .95 indicate substantial distortion, far beyond minor inaccuracies in plotting. The logarithm of the Lie Factor can be taken in order to compare overstating (log LF > 0) with understating (log LF < 0) errors. In practice almost all distortions involve overstating, and Lie Factors of two to five are not uncommon.

Here is an extreme example. A newspaper reported that the U.S. Congress and the Department of Transportation had set a series of fuel economy standards to be met by automobile manufacturers, beginning with 18 miles per gallon in 1978 and moving in steps up to 27.5 by 1985, an increase of 53 percent:

$$\frac{27.5 - 18.0}{18.0} \times 100 = 53\%$$

These standards and the dates for their attainment were shown:

This line, representing 18 miles per gallon in 1978, is 0.6 inches long.



This line, representing 27.5 miles per gallon in 1985, is 5.3 inches long.

New York Times, August 9, 1978, D-2.

The magnitude of the change from 1978 to 1985 is shown in the graph by the relative lengths of the two lines:

$$\frac{5.3 - 0.6}{0.6} \times 100 = 783\%$$

Thus the numerical change of 53 percent is presented by some lines that changed 783 percent, yielding

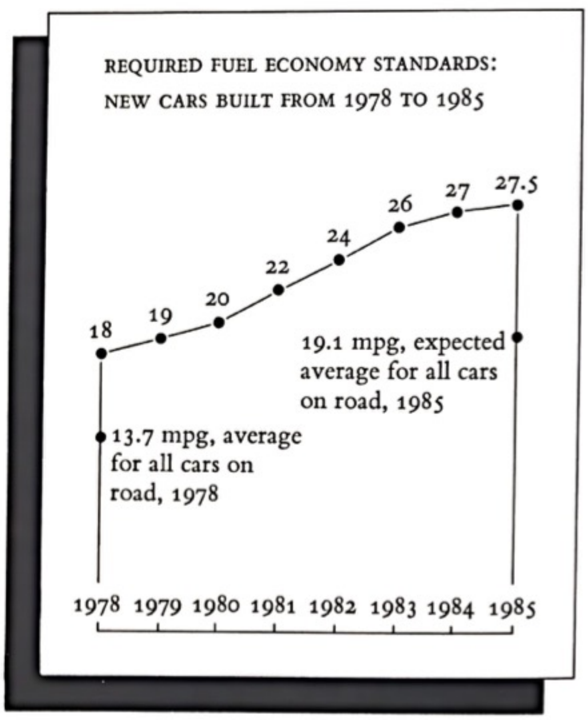
$$\text{Lie Factor} = \frac{783}{53} = 14.8$$

which is too big.

The display also has several peculiarities of perspective:

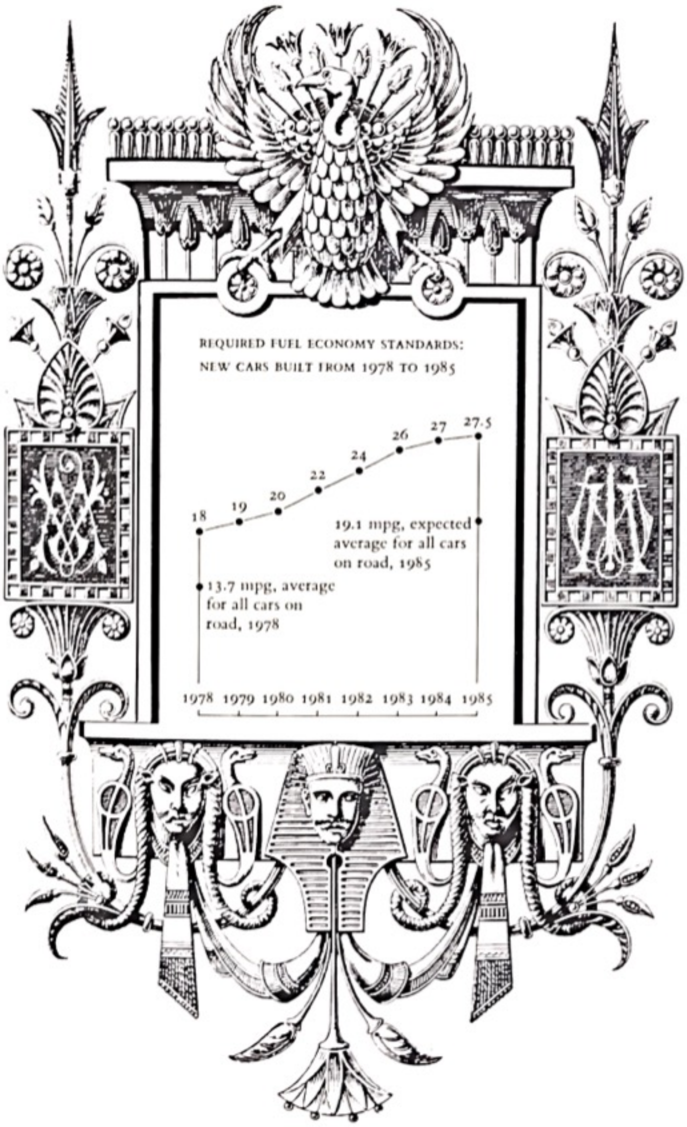
- On most roads the future is in front of us, toward the horizon, and the present is at our feet. This display reverses the convention so as to exaggerate the severity of the mileage standards.
- Oddly enough, the dates on the left remain a constant size on the page even as they move along with the road toward the horizon.
- The numbers on the right, as well as the width of the road itself, are shrinking because of two simultaneous effects: the change in the values portrayed and the change due to perspective. Viewers have no chance of separating the two.

It is easy enough to decorate these data without lying:



The non-lying version, in addition, puts the data in a context by comparing the new car standards with the mileage achieved by the mix of cars actually on the road. Also revealed is a side of the data disguised and mispresented in the original display: the fuel economy standards require gradual improvement at start-up, followed by a doubled rate from 1980 to 1983, and flattening out after that.

Sometimes decoration can help editorialize about the substance of the graphic. But it is wrong to distort the data measures—the ink locating values of numbers—in order to make an editorial comment or fit a decorative scheme. It is also a sure sign of the Graphical Hack at work. Here are many decorations but no lies:

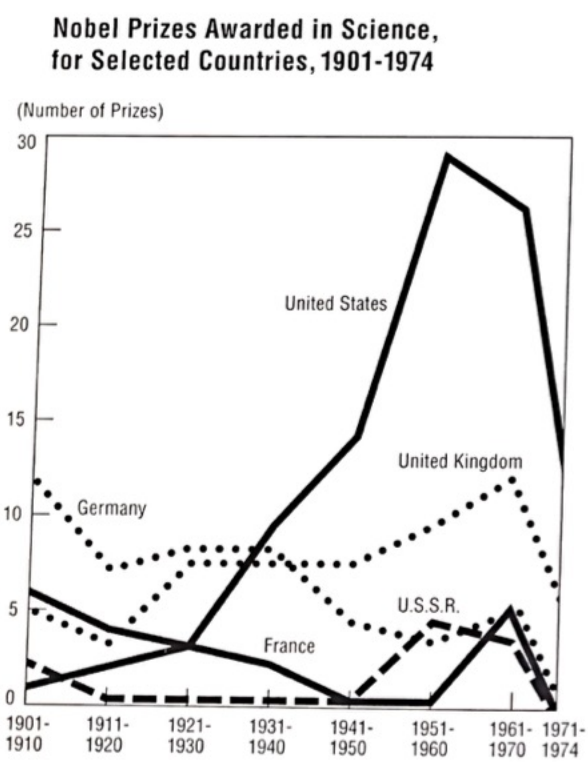


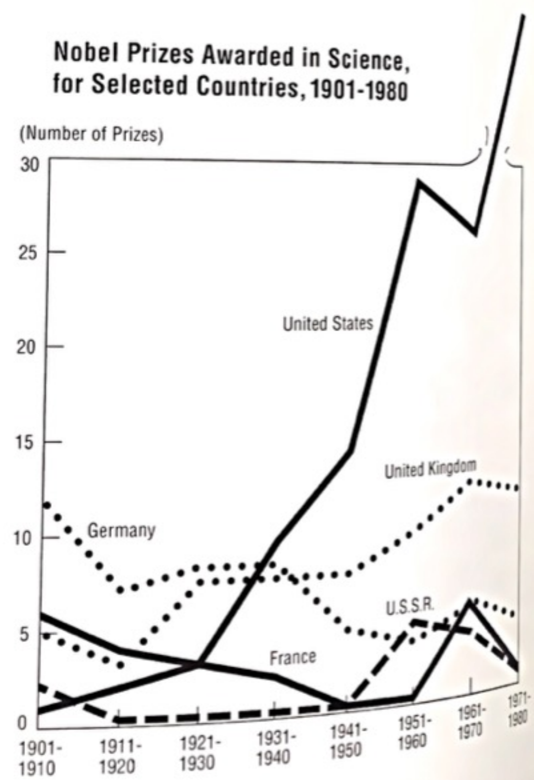
Design and Data Variation

Each part of a graphic generates visual expectations about its other parts and, in the economy of graphical perception, these expectations often determine what the eye sees. Deception results from the incorrect extrapolation of visual expectations generated at one place on the graphic to other places.

A scale moving in regular intervals, for example, is expected to continue its march to the very end in a consistent fashion, without the muddling or trickery of non-uniform changes. Here an irregular scale is used to concoct a pseudo-decline. The first seven increments on the horizontal scale are ten years long, masking the rightmost interval of four years. Consequently the conspicuous feature of the graphic is the apparent fall of curves at the right, particularly the decline in prizes won by people from the United States (the heavy, dark line) in the most recent period. This effect results solely from design variation. It is a big lie, since in reality (and even in extrapolation, scaling up each end-point by 2.5 to take the four years' worth of data up to a comparable decade), the U.S. curve turned sharply upward in the post-1970 interval. A correction, with the actual data for 1971–80, is at the right:

National Science Foundation, *Science Indicators, 1974* (Washington, DC, 1976), 15.

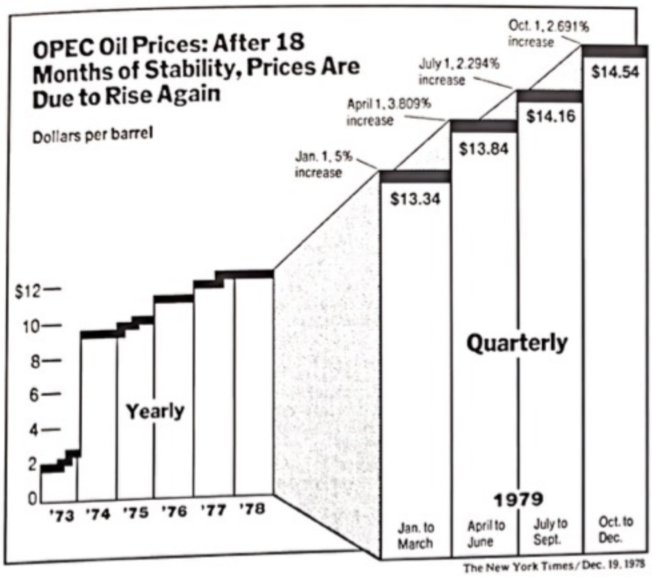




The confounding of *design variation* with *data variation* over the surface of a graphic leads to ambiguity and deception, for the eye may mix up changes in the design with changes in the data. A steady canvas makes for a clearer picture. The principle is, then:

Show data variation, not design variation.

Design variation corrupts this display:



New York Times, December 19, 1978, D-7.

Five different vertical scales show the price:

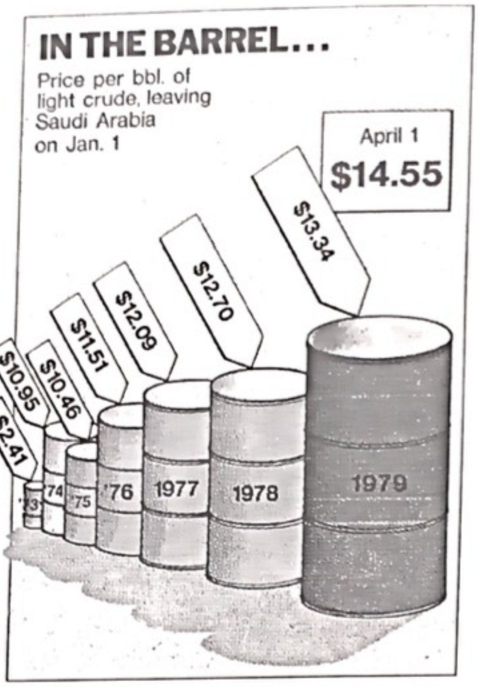
During this time	one vertical inch equals
1973–1978	$8.00
January–March 1979	$4.73
April–June 1979	$4.37
July–September 1979	$4.16
October–December 1979	$3.92

And two different horizontal scales show the passage of time:

During this time	one horizontal inch equals
1973–1978	3.8 years
1979	0.57 years

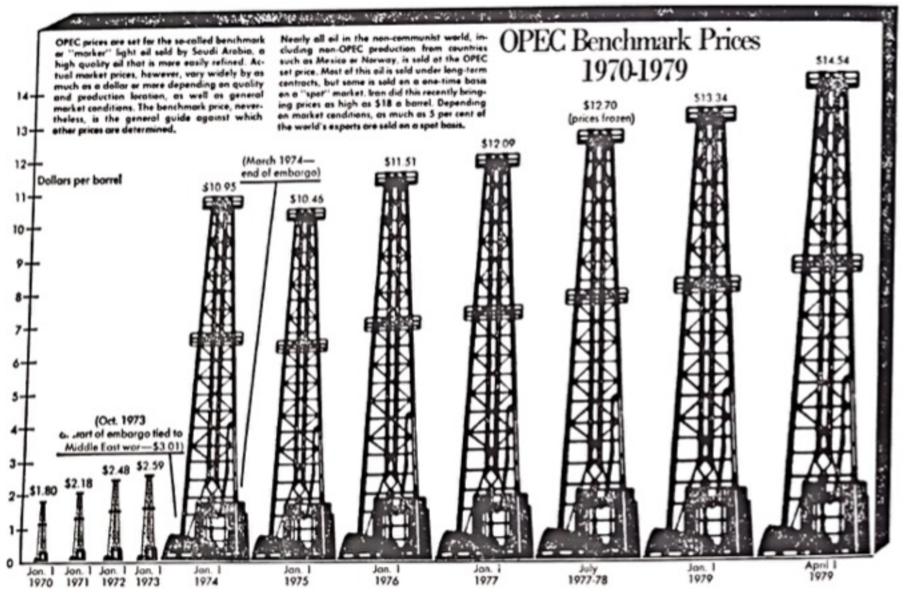
As the two scales shift simultaneously, the distortion takes on multiplicative force. On the left of the graph, a price of $10 for one year is represented by 0.31 square inches; on the right side, by 4.69 square inches. Thus exactly the same quantity is 4.69/0.31 = 15.1 times larger depending upon where it happens to fall on the surface of the graphic. *That* is design variation.

Design variation infected similar graphics in other publications. Here an increase of 454 percent is depicted as an increase of 4,280 percent, for a Lie Factor of 9.4:



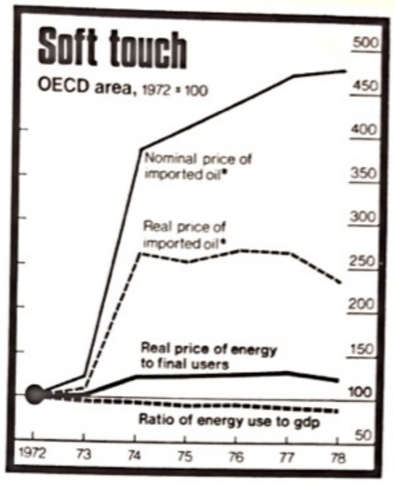
Time, April 9, 1979, 57.

And an increase of 708 percent is shown as 6,700 percent, for a Lie Factor of 9.5:

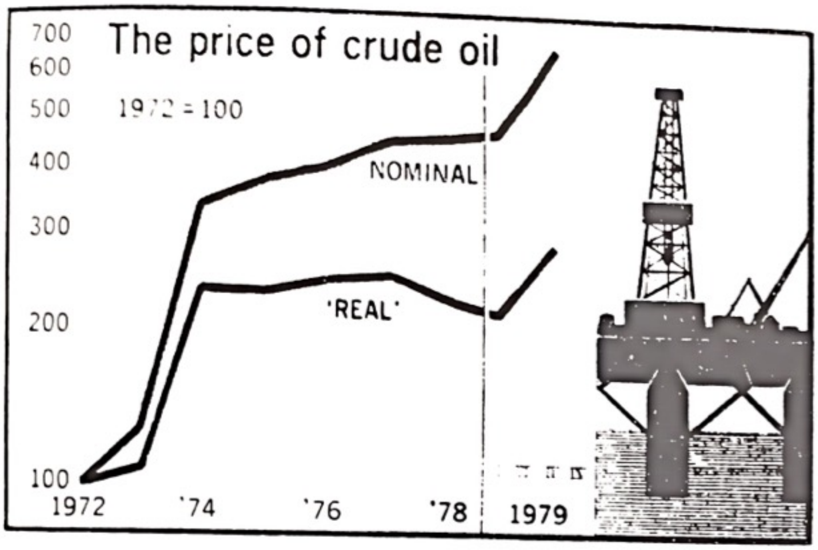


Washington Post, March 28, 1979, A-18.

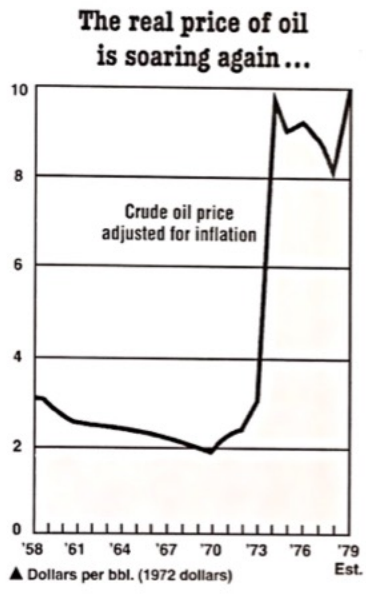
All these accounts of oil prices made a second error, by showing the price of oil in inflated (current) dollars. The 1972 dollar was worth much more than the 1979 dollar. Thus in sweeping from left to right over the surface of the graphic, the vertical scale in effect changes—design variation—because the value of money changes over the years shown. The only way to think clearly about money over time is to make comparisons using inflation-adjusted units of money. Several distinguished graphic designers did express the price in real dollars—and they also avoided other sources of design variation. The stars were *Business Week*, the *Sunday Times* (London), and *The Economist*.



The Economist, December 29, 1979, 41.

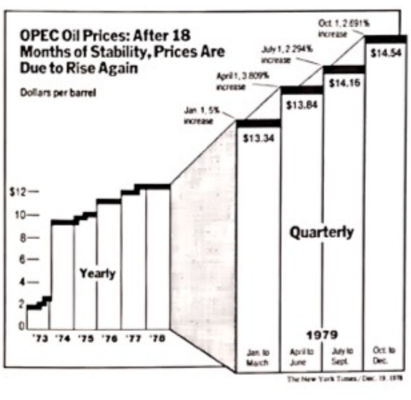


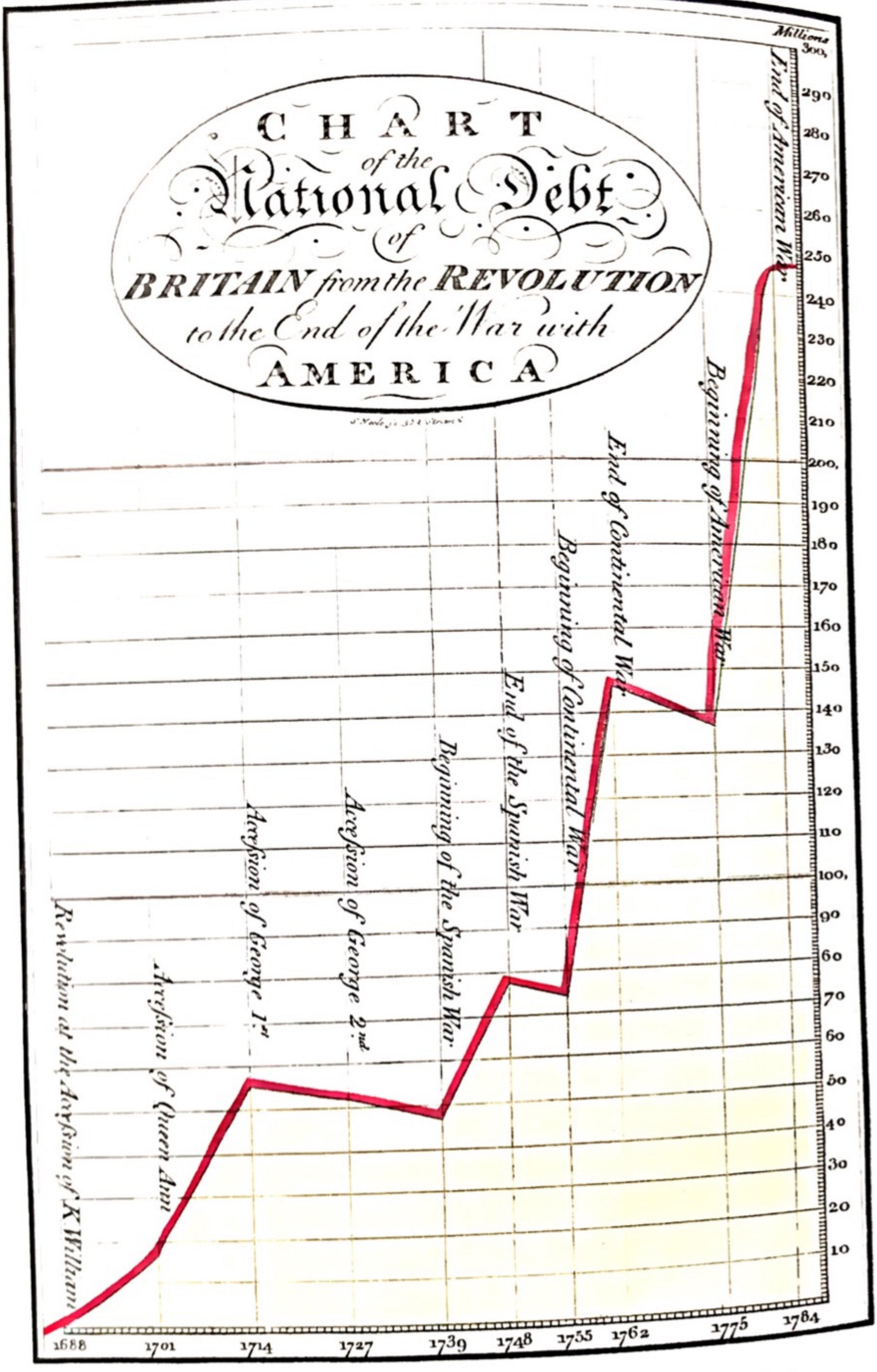
Sunday Times (London), December 16, 1979, 54.



Business Week, April 9, 1979, 99.

In the graphic we saw first, the two sources of design variation covered up an intriguing, non-obvious aspect of the data: in the four years prior to the 1979–1980 increases, the real price of oil had *declined*. Busy with decoration, the graphic had missed the news.





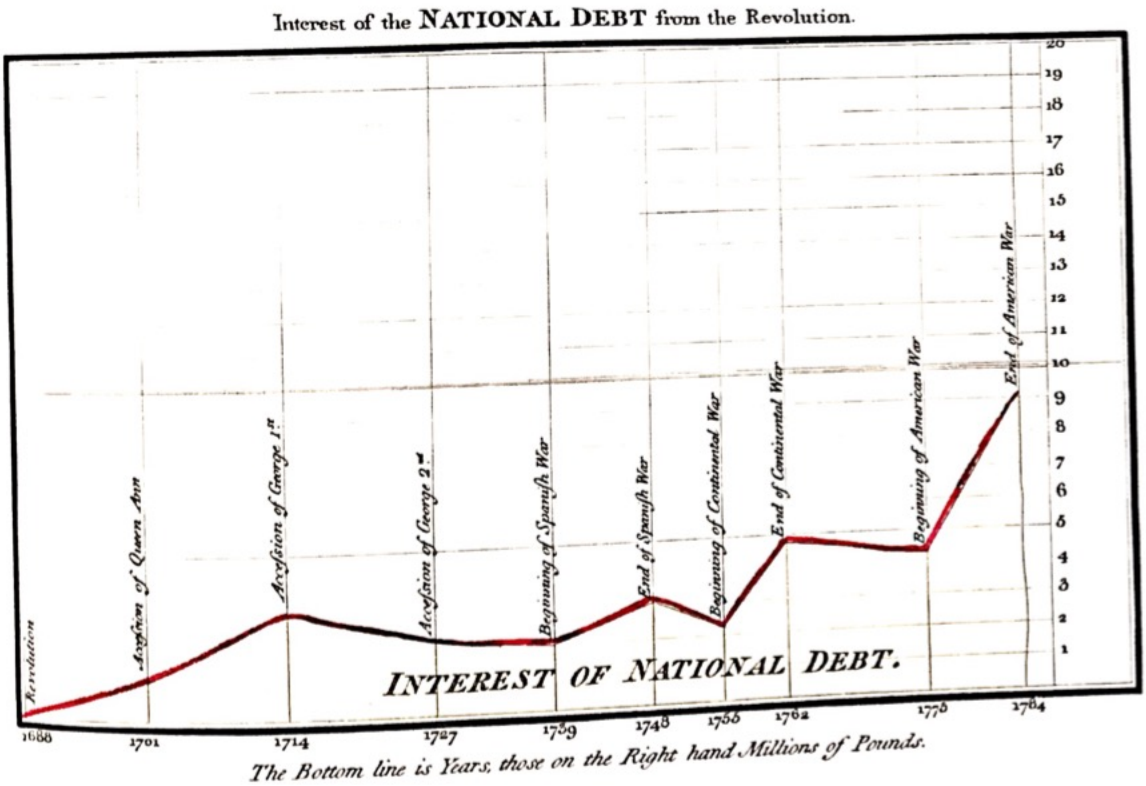
The Divisions at the Bottom are Years, & those on the Right hand Money.

The Case of Skyrocketing Government Spending

Probably the most frequently printed graphic, other than the daily weather map and stock-market trend line, is the display of government spending and debt over the years. These arrays nearly always create the impression that spending and debt are rapidly increasing.

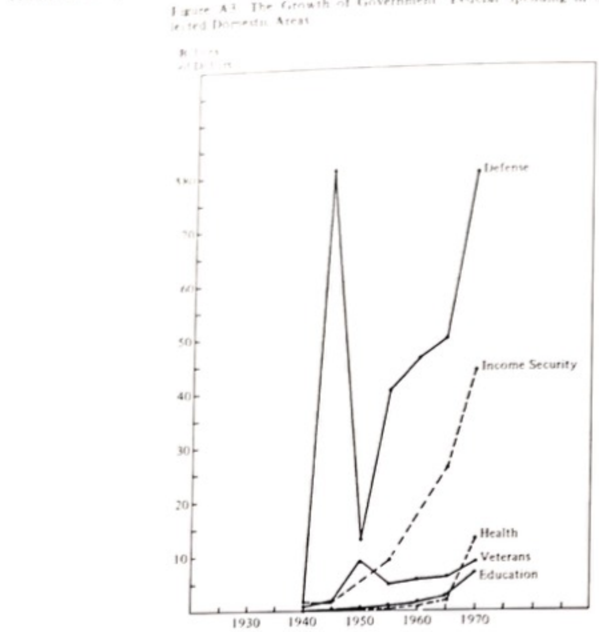
As usual, Playfair was the first, publishing this finely designed graphic in 1786. Accompanied by his polemic against the "ruinous folly" of the British government policy of financing its colonial wars through debt, it is surely the first skyrocketing government debt chart, beginning the now 200-year history of such displays. This is one of the few Playfairs that is taller than wide; less than one-tenth of all his graphics (about 90, drawn during 35 years of work) are longer on the vertical. The tall shape here serves to emphasize the picture of rapid growth. The money figures are not adjusted for inflation.

But Playfair had the integrity to show an alternative version a few pages later in *The Commercial and Political Atlas*. The interest on the national debt was plotted on a broad horizontal scale, diminishing the skyrocket effect. And, furthermore, "This is in real and not in nominal millions" (page 129):

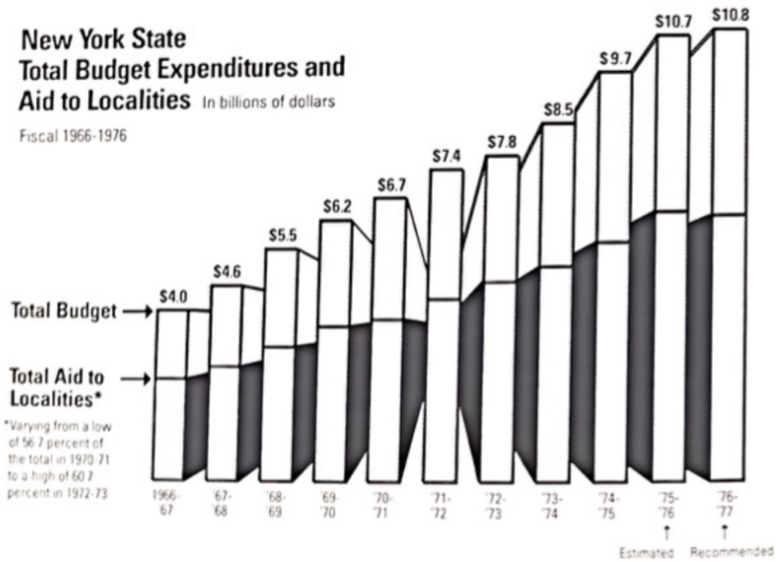


Although Playfair deflated money units over time in his work of 1786, the matter has proved difficult for many, eluding even modern scholars. This display helps its political point along by failing to discount for inflation and population growth and by using a tall and thin shape (the area covered by the data is 2.7 times taller than wide):

Morris Fiorina, *Congress: [Keystone of the] Washington Establishment* ([New Haven], 1977), 92.



Let us look, in detail, at another graphic on government spending:

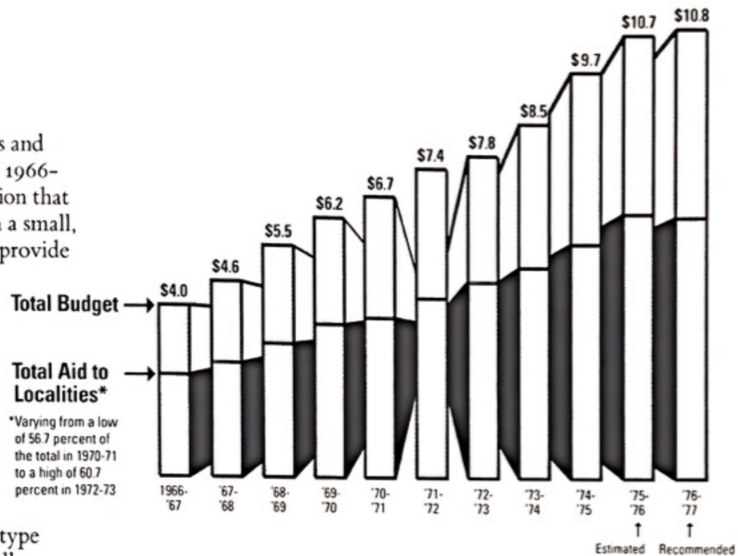


New York Times, February 1, 1976, IV-4

Despite the appearance created by the hyperactive design, the state budget actually did not increase during the last nine years shown. To generate the thoroughly false impression of a substantial and continuous increase in spending, the chart deploys several visual and statistical tricks—all working in the same direction, to exaggerate the growth in the budget. These graphical gimmicks:

These three parallelepipeds have been placed on an optical plane *in front* of the other eight, creating the image that the newer budgets tower over the older ones.

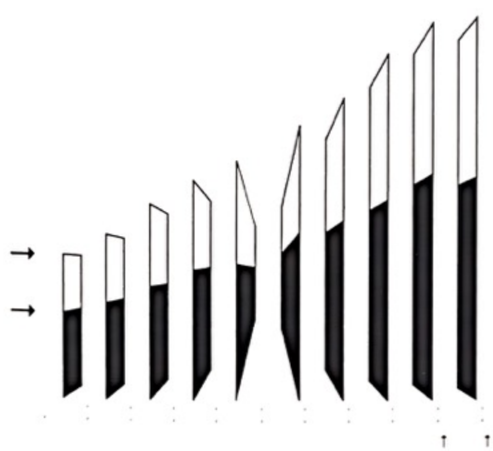
This cluster of type emphasizes and stretches out the low value for 1966–1967, encouraging the impression that recent years have shot up from a small, stable base. Horizontal arrows provide similar emphasis.

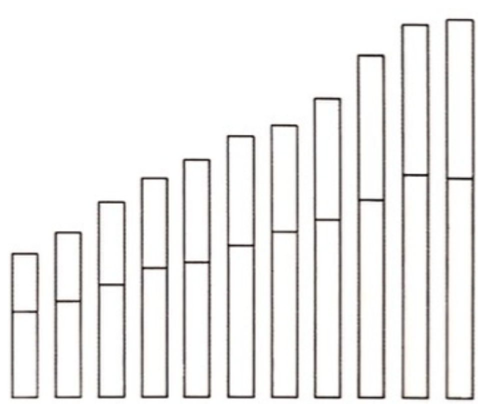


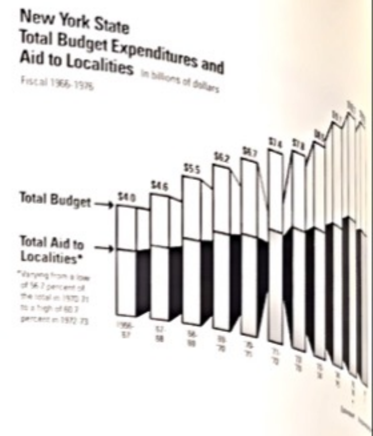
This squeezed-down block of type contributes to an image of small, squeezed-down budgets back in the good old days.

Arrows pointing straight up emphasize recent growth. Compare with horizontal arrows at left.

Leaving behind the distortion in the chartjunk heap at the left yields a calmer view:

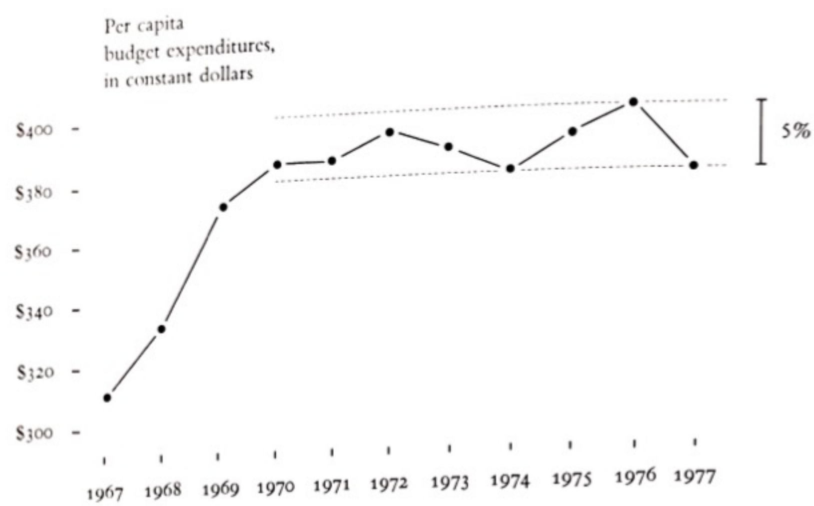






Two statistical lapses also bias the chart. First, during the years shown, the state's population increased by 1.7 million people, or 10 percent. Part of the budget growth simply paralleled population growth. Second, the period was a time of substantial inflation; those goods and services that cost state and local governments $1.00 to purchase in 1967 cost $2.03 in 1977. By not deflating, the graphic mixes up changes in the value of money with changes in the budget.

Application of arithmetic makes it possible to take population and inflation into account. Computing expenditures in *constant (real) dollars per capita* reveals a quite different—and far more accurate—picture:



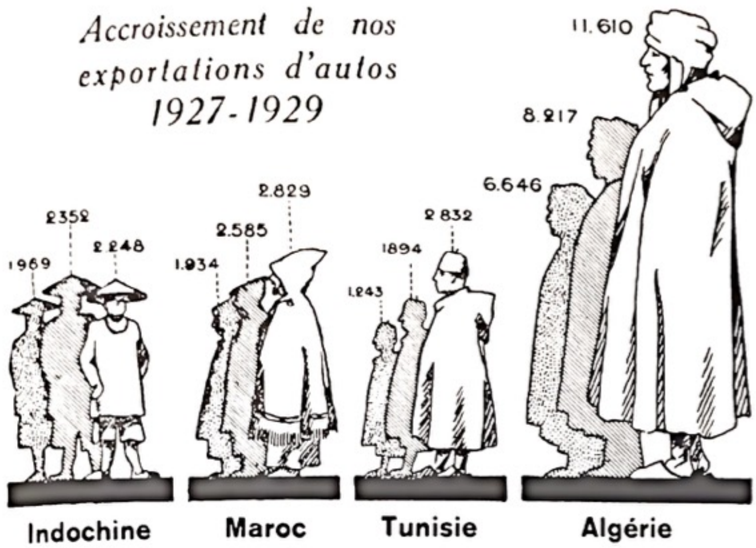
Thus, in terms of real spending per capita, the state budget increased by about 20 percent from 1967 to 1970 and remained relatively constant from 1970 through 1976. And the 1977 budget represents a substantial *decline* in expenditures. That is the real news story of these data, and it was completely missed by the Graph of the Magical Parallelepipeds. Of course no small set of numbers is going to capture the complexities of a large budget—but, at any rate, why tell lies?

The principle:

> In time-series displays of money, deflated and standardized units of monetary measurement are nearly always better than nominal units.

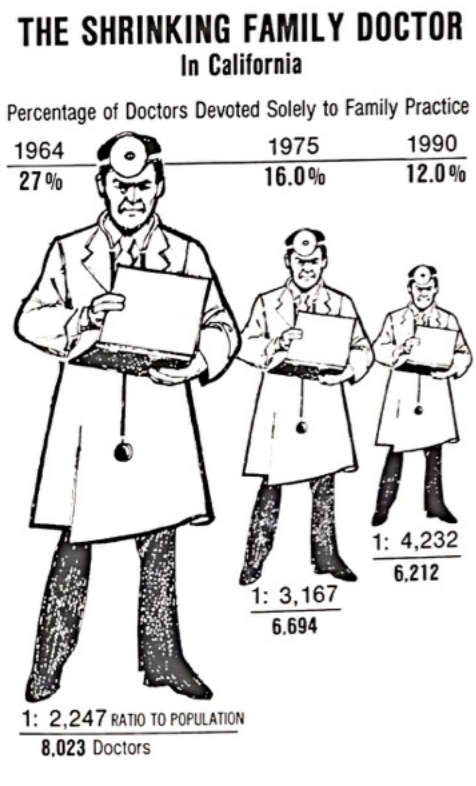
Visual Area and Numerical Measure

Another way to confuse data variation with design variation is to use areas to show one-dimensional data:



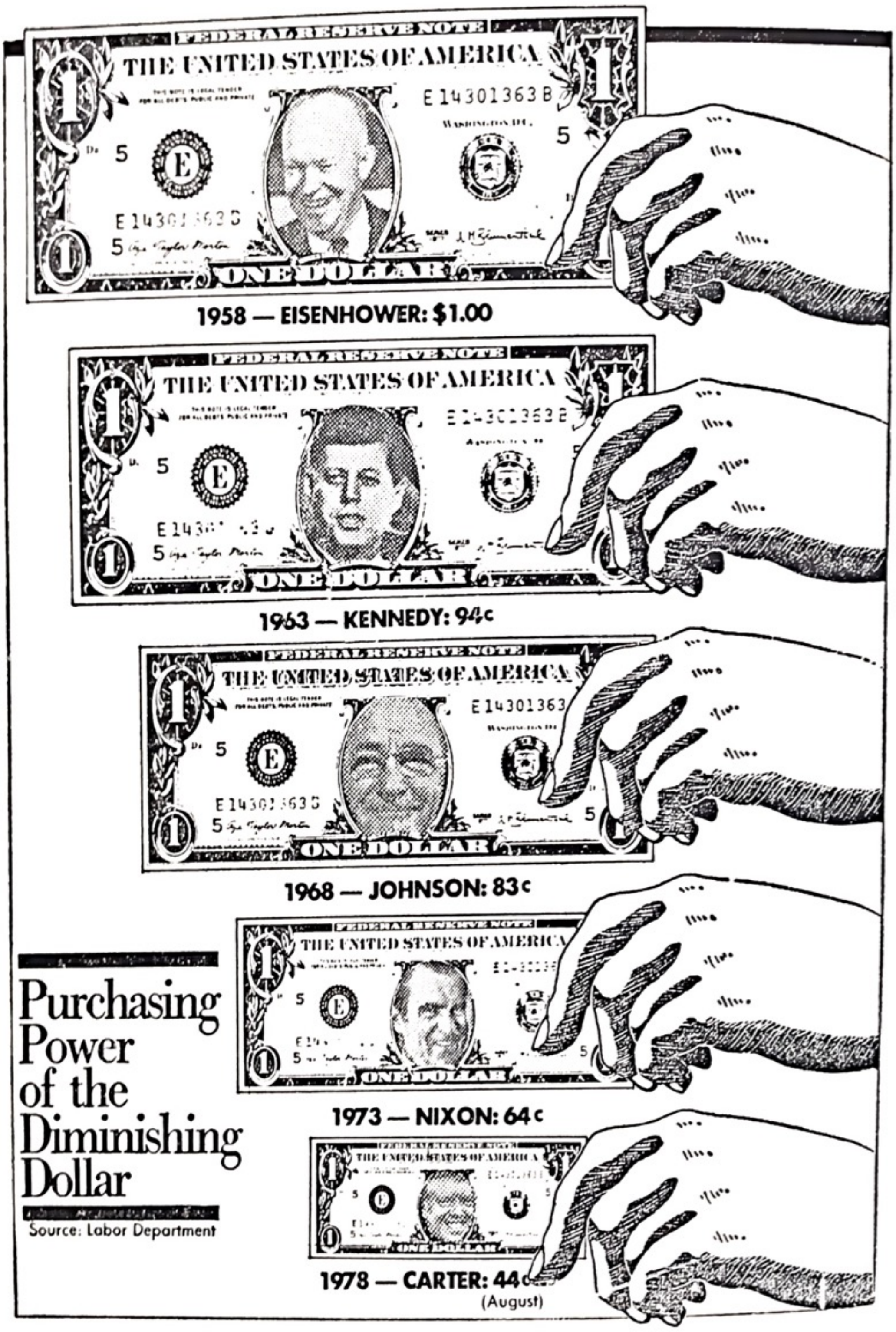
R. Satet, *Les Graphiques* (Paris, 1932), 12.

And here is the incredible shrinking doctor, with a Lie Factor of 2.8, not counting the additional exaggeration from the overlaid perspective and the incorrect horizontal spacing of the data:



Los Angeles Times, August 5, 1979, 3.

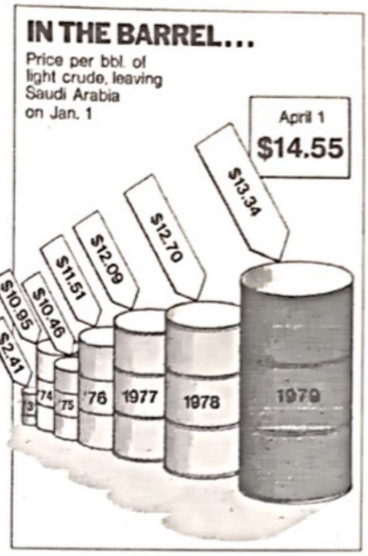
Many published efforts using areas to show magnitudes make the elementary mistake of varying both dimensions simultaneously in response to changes in one-dimensional data. Typical is the shrinking dollar fallacy. To depict the rate of inflation, graphs show currency shrinking on two dimensions, even though the value of money is one-dimensional. Here is one of hundreds of such charts:



Washington Post, October 25, 1978, 1.

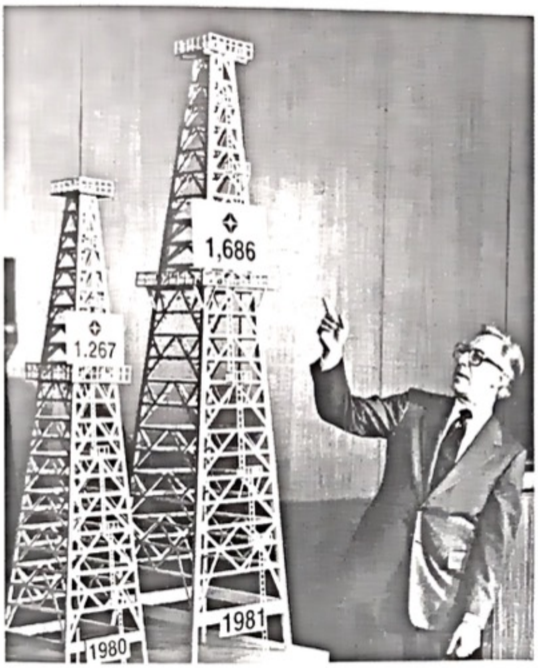
If the area of the dollar is accurately to reflect its purchasing power, then the 1978 dollar should be about twice as big as that shown.

There are considerable ambiguities in how people perceive a two-dimensional surface and then convert that perception into a one-dimensional number. Changes in physical area on the surface of a graphic do not reliably produce appropriately proportional changes in perceived areas. The problem is all the worse when the areas are tricked up into three dimensions:



By surface area, the Lie Factor for this graphic is 9.4. But, if one takes the barrel metaphor seriously and assumes that the *volume* of the barrels represents the price change, then the volume from 1973 to 1979 increases 27,000 percent compared to a data increase of 454 percent, for a Lie Factor of 59.4, which is a record.

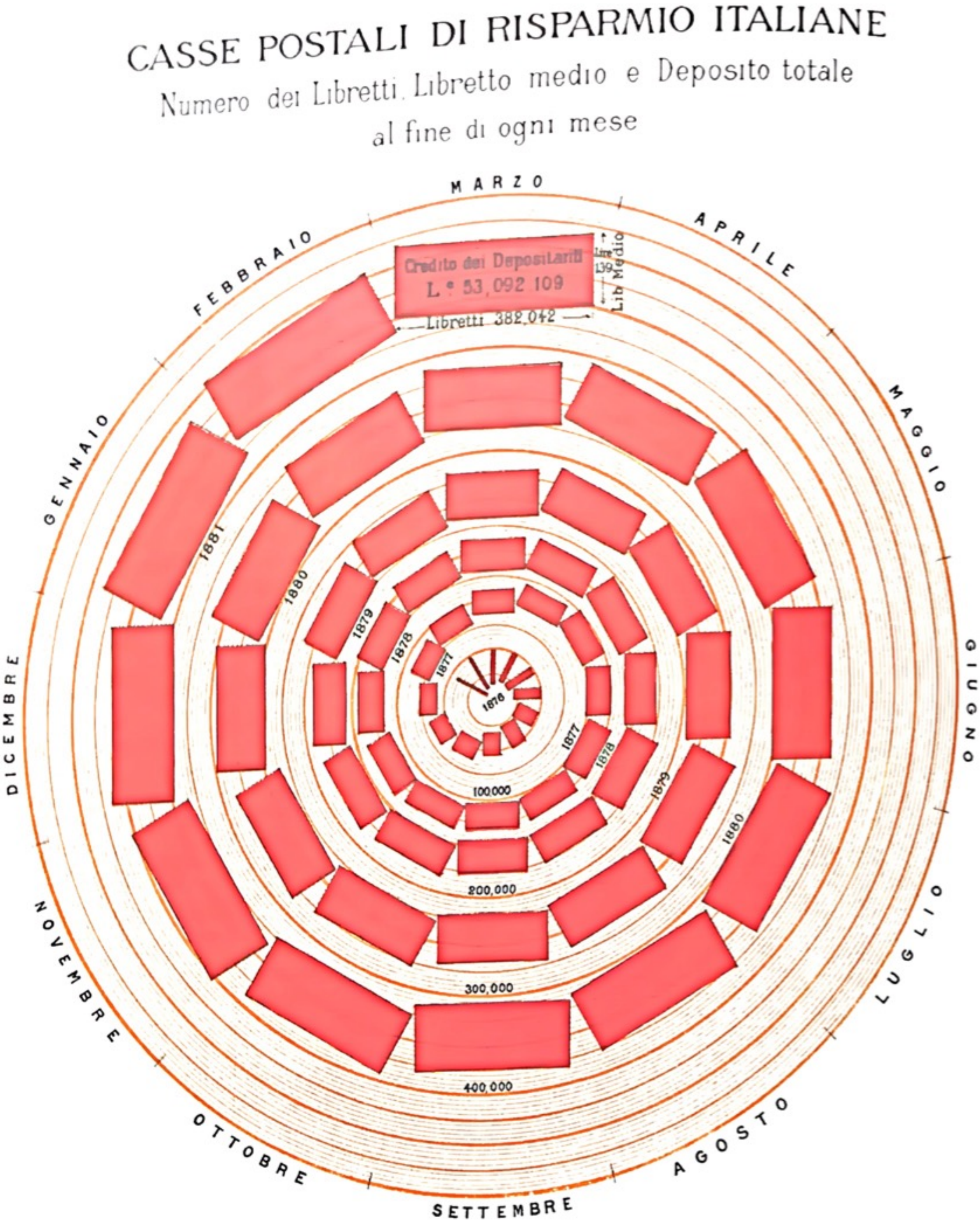
Similarly, a three-dimensional representation puffing up one-dimensional data:



New York Times, January 27, 1981, D-1.

Conclusion: The use of two (or three) varying dimensions to show one-dimensional data is a weak and inefficient technique, capable of handling only very small data sets, often with error in design and ambiguity in perception. These designs cause so many problems that they should be avoided:

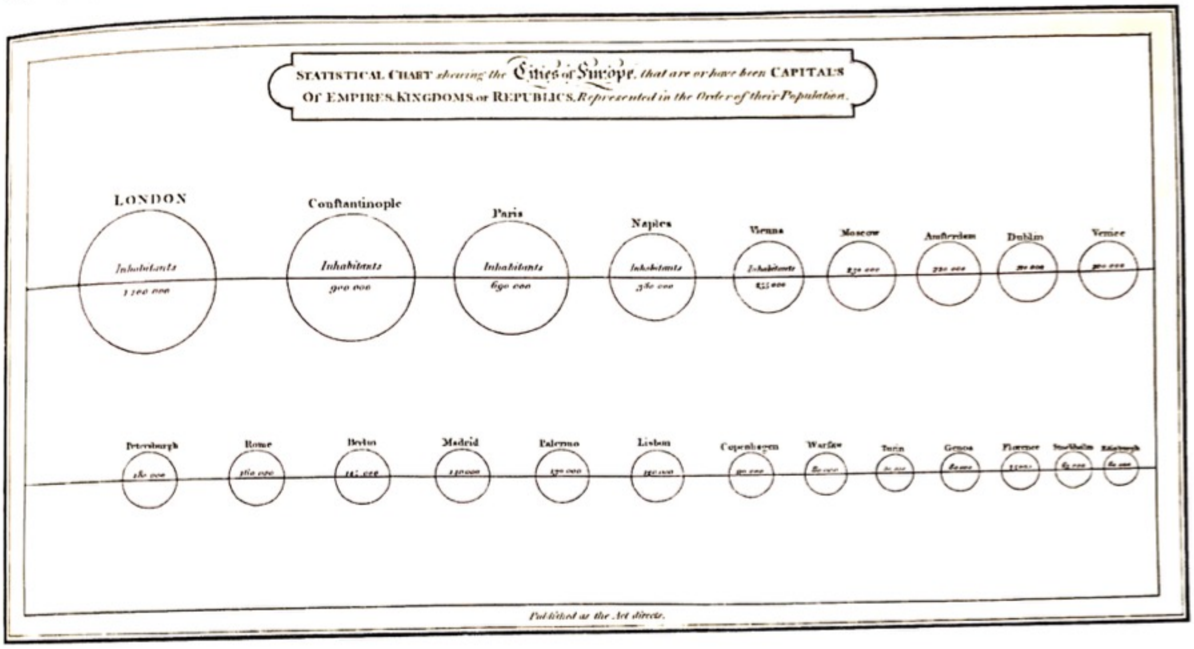
> The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.



This multivariate history of the Italian post office uses two dimensions in a way nearly consistent with this principle, with the number of postal savings books issued and the average size of deposits multiplying up to total deposits at the end of each month from 1876 to 1881.

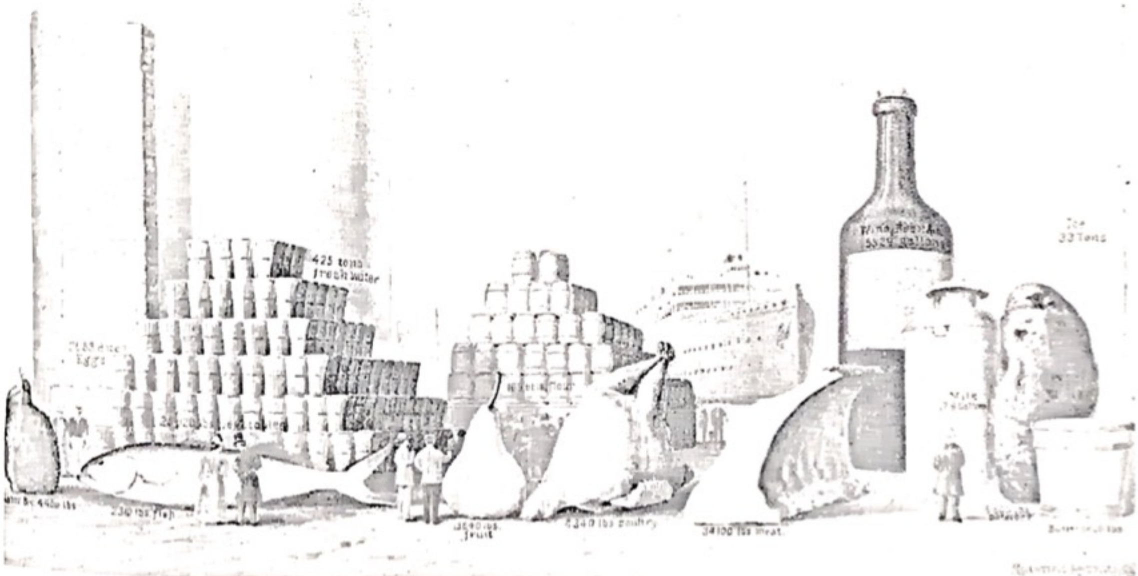
Antonio Gabaglio, *Teoria Generale della Statistica* (Milan, second edition, 1888).

But Playfair's circles, an early use of area to show magnitude, are not consistent with the principle, since the one-dimensional data (city populations) are represented by area:



Perhaps graphics that border on cartoons should be exempt from the principle. We certainly would not want to forgo the 4,340 pound chicken:

Scientific American Reference Book (New York, 1909), 280.

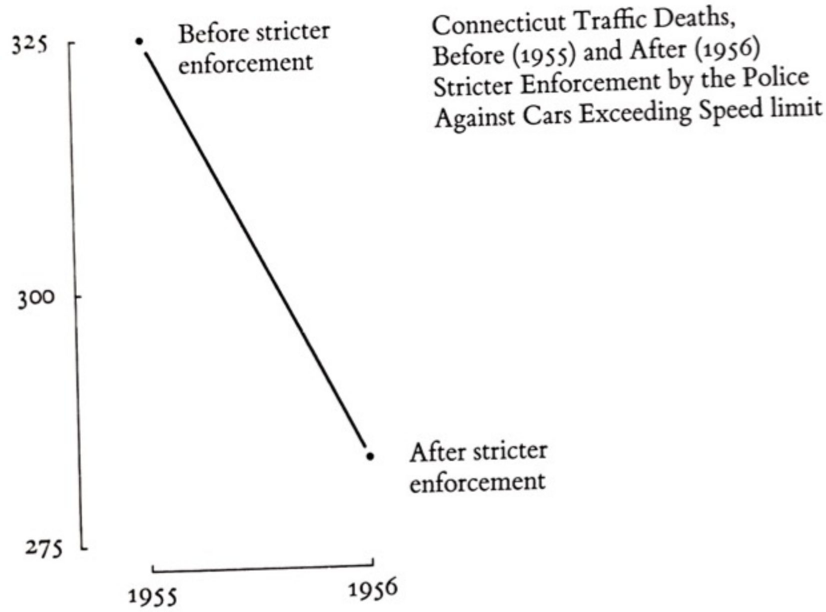


Context is Essential for Graphical Integrity

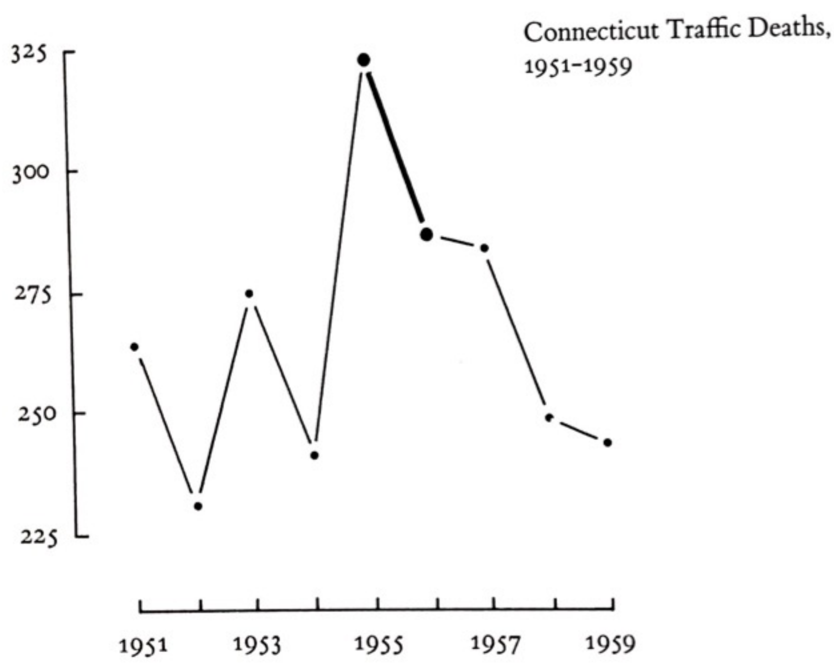
To be truthful and revealing, data graphics must bear on the question at the heart of quantitative thinking: "Compared to what?" The emaciated, data-thin design should always provoke suspicion, for graphics often lie by omission, leaving out data sufficient for comparisons. The principle:

Graphics must not quote data out of context.

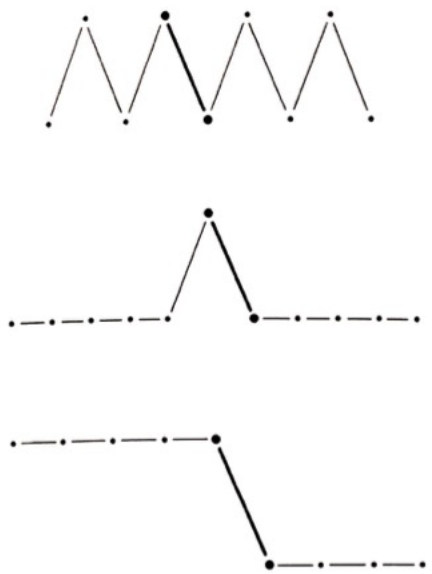
Nearly all the important questions are left unanswered by this display:



A few more data points add immensely to the account:

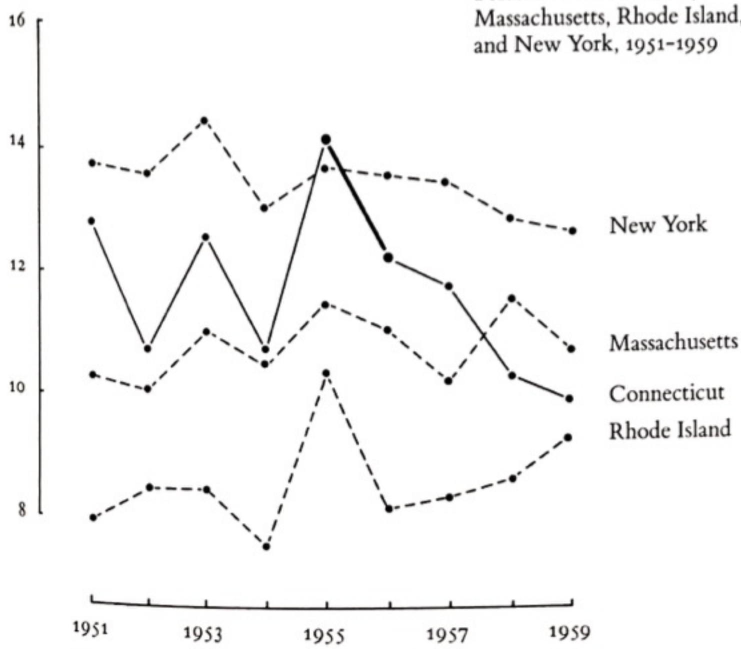


Imagine the very different interpretations other possible time-paths surrounding the 1955–1956 change would have:



Comparisons with adjacent states give a still better context, revealing it was not only Connecticut that enjoyed a decline in traffic fatalities in the year of the crackdown on speeding:

Traffic Deaths per 100,000 Persons in Connecticut, Massachusetts, Rhode Island, and New York, 1951–1959

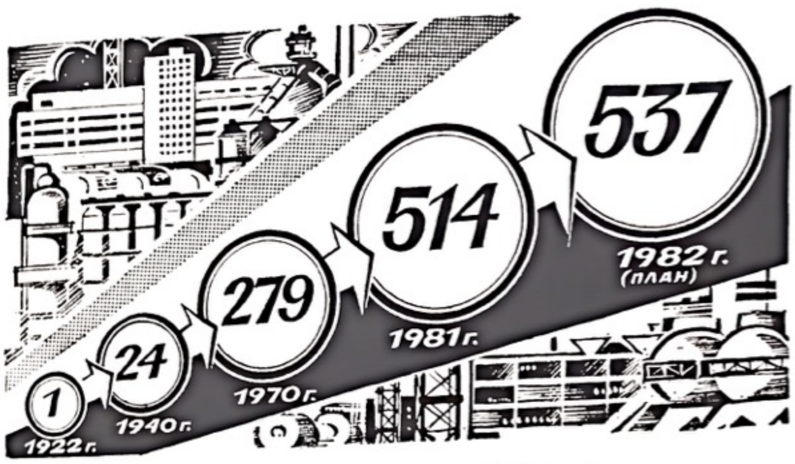


Donald T. Campbell and H. Laurence Ross, "The Connecticut Crackdown on Speeding: Time Series Data in Quasi-Experimental Analysis," in Edward R. Tufte, ed., *The Quantitative Analysis of Social Problems* (Reading, Massachusetts, 1970), 110–125.

Conclusion

Lying graphics cheapen the graphical art everywhere. Since the lies often show up in news reports, millions of images are printed. When a chart on television lies, it lies tens of millions of times over; when a *New York Times* chart lies, it lies 900,000 times over to a great many important and influential readers. The lies are told about the major issues of public policy—the government budget, medical care, prices, and fuel economy standards, for example. The lies are systematic and quite predictable, nearly always exaggerating the rate of recent change.

The main defense of the lying graphic is . . . "Well, at least it was approximately correct, we were just trying to show the general direction of change." But many of the deceptive displays we saw in this chapter involved fifteenfold lies, too large to be described as approximately correct. And in several cases the graphics were not even approximately correct by the most lax of standards, since they falsified the real news in the data. It is the special character of numbers that they have a magnitude as well as an order; numbers measure *quantity*. Graphics can display the quantitative size of changes as well as their direction. The standard of getting only the direction and not the magnitude right is the philosophy that informs the Pravda School of Ordinal Graphics. There, every chart has a crystal clear direction coupled with fantasy magnitudes.



Рост продукции промышленности [1922 г. = I].

Pravda, May 24, 1982, 2.

A second defense of the lying graphic is that, although the design itself lies, the actual numbers are printed on the graphic for those picky folks who want to know the correct size of the effects displayed. It is as if not lying in one place justified fifteenfold lies elsewhere. Few writers would work under such a modest standard of integrity, and graphic designers should not either.

Graphical integrity is more likely to result if these six principles are followed:

The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities represented.

Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.

Show data variation, not design variation.

In time-series displays of money, deflated and standardized units of monetary measurement are nearly always better than nominal units.

The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.

Graphics must not quote data out of context.