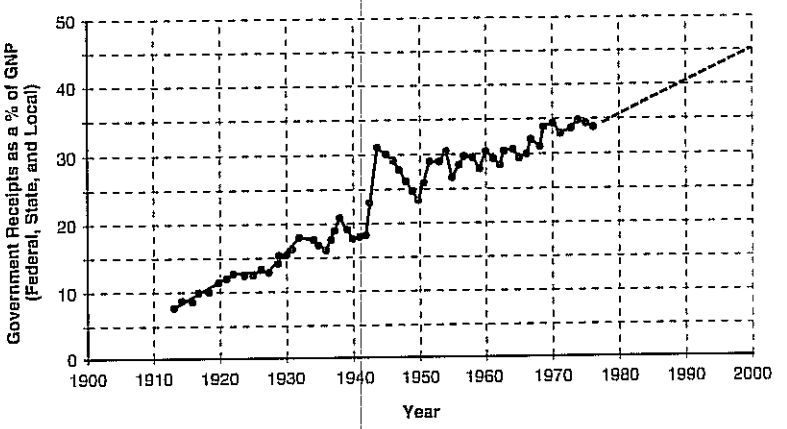



**FIGURE 3.11** An Illustration of a Time-Series Graph, "The Demise of Free Enterprise" from Augustine (1978).

## 3.11 TIME-SERIES GRAPHS

The *time-series graph*, a standard statistical figure in business and economics, is becoming common in some areas of education and psychology. It can be useful for identifying trends and changes in trends in ways that other representations of data cannot. A time-series graph is a line in which the $X$-axis, or baseline, is time and the vertical axis is a measure of the variable of interest. The time dimension can be measured in minutes, hours, days, weeks, months, or years, depending on the view that one wishes to take. Familiar examples of time-series graphs include the Dow-Jones stock price average plotted across days, the Consumer Price Index plotted across months, a patient's body temperature plotted across hours, and school enrollment plotted across years. Figure 3.11 is an illustrative time-series graph, "The Demise of Free Enterprise," provided by Augustine (1978). The extrapolated projections into the future are shown by the dashed line. Obviously, such projections into the distant future may have a large margin of error.

## 3.12 MISLEADING GRAPHS: HOW TO LIE WITH STATISTICS[15]

The ability to interpret properly, and not be misled by, information that is presented graphically is an important type of literacy for the both the layman and the professional. The general public is continually bombarded with data-based figures in newspapers and

[15]Finally, something practical that you can use!

magazines. Textbooks in all empirical disciplines are filled with graphs. Standardized achievement tests and university entrance tests are heavily weighted with graphic information that must be read critically (Tufte, 1983). Just as words can be misused to obscure the facts, so can pictures. At times, self-interest tempts one (including researchers) to use literal facts in such a way that the message is distorted. This may not be lying in a legal sense, but it accomplishes the same purpose. Graphs and charts can be organized so that they become propaganda rather than to illuminate the truth. Many, if not most, figures in the popular media are constructed to be as remarkable ("newsworthy") as possible; journalists are trained to tell an interesting story, regardless of whether words or pictures are used to tell the story. It behooves us to be on our toes so we are not credulous victims of misinformation.[16]

## Distorted Representation

A common, but not very subtle error, evident in many pictographs[17] found in the popular media is the linear-area fallacy. To get "more bang for the buck," graphic artists often represent the frequency in a category by the height of the figure (a linear distance), yet make only one figure per category. This lack of uniform representation of a frequency conveys a distorted picture of the data. Notice in Figure 3.12A how the amount of inflation is exaggerated across the five presidencies. The data are scaled by the *length* of the dollar bill, but it is the *area* of the bills that the reader perceives. The area of the Carter dollar is less than 20% that of the Eisenhower bill, whereas the proper comparison is 100 to 44, not 100 to 20. If the bills were of the same width as the Eisenhower dollar, but were fragments with different lengths, the representation would be fair and accurate. (Isn't the result dramatic enough without fudging?) The same flaw is seen in Figure 3.12B. Figure 3.12C gives a hopelessly distorted picture of the data.[19]

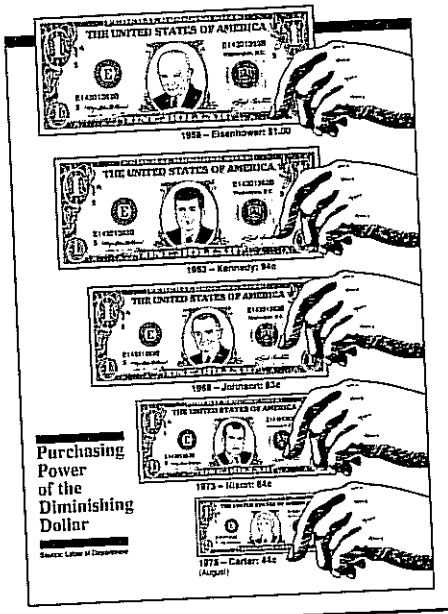



**FIGURE 3.12A** Misleading Graphs: Illustrations of Distorted Representation.[18]

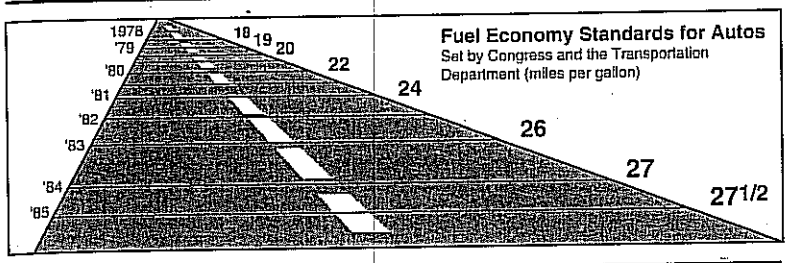



**FIGURE 3.12B** Misleading Graphs: Illustrations of Distorted Representation.

## Misleading Scaling and Calibration

There is no obvious calibration in Figure 3.12C. A more common shortcoming of graphs appears in Figure 3.13, where an arbitrary beginning scale value on the vertical axis is used. Variables that represent ratio scales should begin with zero to give a proper perspective for the visual interpretation. Figure 3.13A is a common method of perceptual exaggeration; the change over time is made to appear much larger than it is. Notice how different the magnitude of the change appears in a proper figure like Figure 3.13B. Many graphs (e.g., stock prices) typically ignore the zero point and thus perceptually exaggerate the magnitude of changes.

## Combination Graphs

Combination graphs can be one of the most devious ways of giving unwarranted credibility to graphic propaganda. All three graphs in Figure 3.14 use the same data, but note that the top and middle graphs lead to opposite conclusions! This is possible by an inappropriate scaling of both variables. Combination graphs need to be scrutinized (Wainer, 1992); the actual information is usually much less convincing than the graph. *Caveat emptor!*

[16]Several of the examples are from the excellent resource book by Tufte (1983).

[17]Histograms that use figures to represent frequencies.

[18]Sources for Figures. 3.12A: *Washington Post*, Oct. 25, 1978, p. 1; Figure 3.12B: *New York Times*, Aug. 9, 1978, p. D-2; Figure 3.12C: *Rocky Mountain News*, May 8, 1994, p. 81A.

[19]Why not construct a fair histogram to get the facts straight? Use "Year" as the baseline, and MPG as the vertical axis.

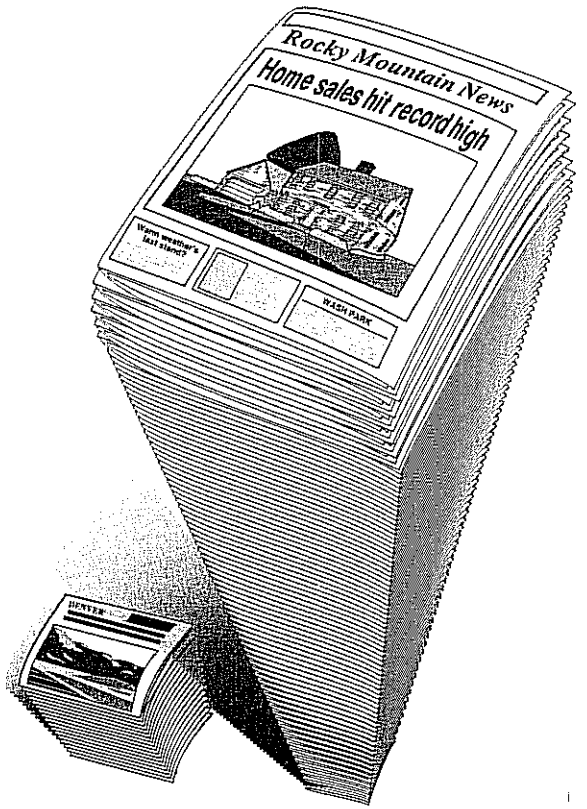
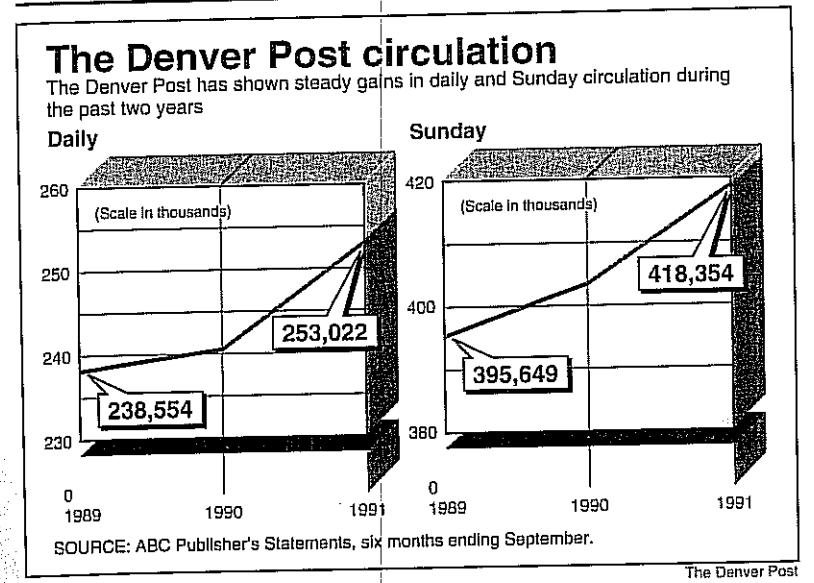



**FIGURE 3.12C** Misleading Graphs: Illustrations of Distorted Representation.

The lower graph in Figure 3.14 is fair, but still equivocal. How does one properly scale the SAT? Like all cognitive and affective measures, it has no meaningful zero point because it is not a ratio scale. Since 400 is the lowest possible score and 1600 the highest possible score, perhaps these should be used to anchor the scale.

One final point about graphic displays of information must be made. Just as in written



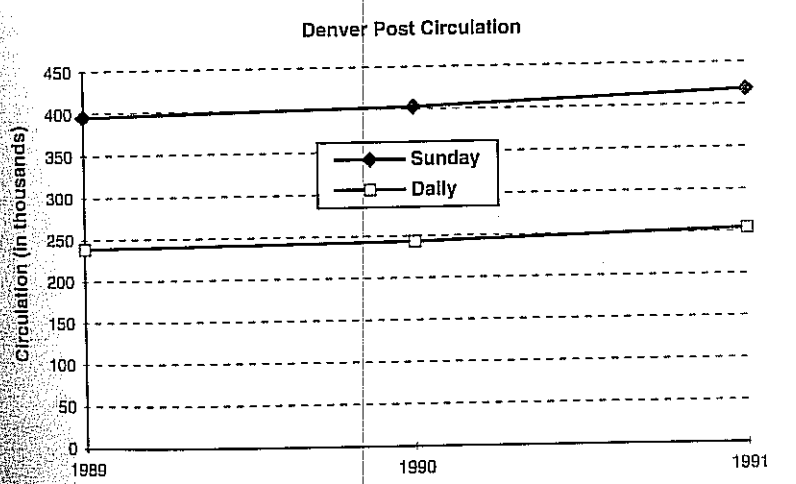




**FIGURE 3.13** An Example of Exaggerating a Trend by Ignoring the Zero Point of a Ratio Scale (top), Compared with a Fair Representation (bottom). *Denver Post*, November 3, 1991.

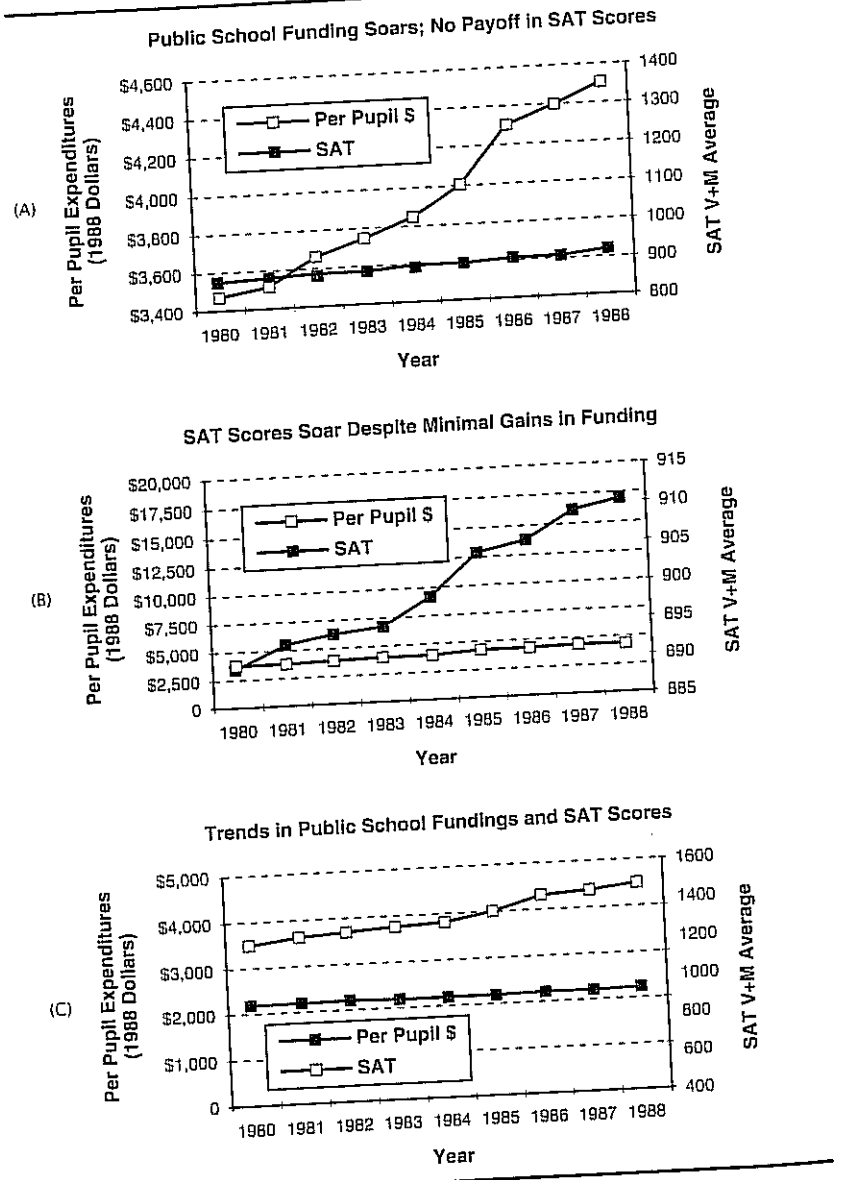



**FIGURE 3.14** Illustrations of Propagandizing with Data (graphs A and B); Only Graph C Does not Distort the Information.





communication, the intended audience has an important influence on what and how information is presented graphically. More details can be included when the graph is for an audience of professional peers than when the readers will be less interested and less sophisticated. Too much detail can be distracting and confusing to a lay audience. Figure 3.15 is an example of a graph that contains an abundance of information without becoming cluttered and confusing. The daily high and low temperatures are given with the normal temperatures as a backdrop. Note that the lines depicting the normal temperatures are regular, but not entirely smooth. These curves are based on daily averages over the past century or so; as the data base grows, these will become even less ragged. Observe that the range between the normal low and high is less during the winter than in the summer. The daily temperatures tend to increase until late July, then begin to decline until late January, although in any given year the pattern will vacillate considerably. The relative humidity did not vary greatly from month to month. On the average, New York City does not have any particularly wet or dry months.

## 3.13 CHAPTER SUMMARY

Statistical methods are tools for simplifying, organizing, and summarizing a set of observations. Statistical tables, figures, graphs, and charts can organize otherwise unwieldy quantities of data. The shape or configuration of a distribution becomes evident if the observations are grouped into intervals and displayed graphically.

Many variables in the behavioral and social sciences are normally distributed, but other kinds of distribution are also common. Skewed distributions result when observations pile up at the low or high end and drop off gradually toward the other end. A distribution is positively skewed when the tail points to the high scores; the opposite is true with negative skewness.

Frequency distributions can be expressed graphically using frequency and percentage histograms (bar graphs), polygons, box plots, and stem-and-leaf displays, among others. Histograms are appropriate when the baseline variable is categorical or nominal; frequency and percentage polygons require a quantitative baseline variable. Ogive (cumulative percentage) curves are useful for determining percentiles (the percentage of a distribution that falls below a given point), such as $Q_1$ ($P_{25}$), the median ($Q_2$ or $P_{50}$), and $Q_3$ ($P_{75}$). A box-and-whisker plot is a simple graph for conveying salient features of a distribution. A stem-and-leaf display conveys the general shape of a frequency distribution while retaining access to the values of all the original scores.

Graphs can be misused to distort data. When pictures are used to represent figures, all should be the same size. Scales having a true zero point should begin at that point to avoid exaggerating effects. Combination graphs can be particularly effective for propaganda, leading the reader to false conclusions.

## 3.14 CASE STUDY

Frequency distributions for each variable in the Chapman cholesterol study are shown in Figure 3.16. What can we learn from them? The more we know about the sample, the better we can define the population to which the findings can be generalized. We should note, for