15 Communicating Evidence Visually

Most readers grasp quantitative evidence more easily in tables, charts, and graphs than they do in words. But some visual forms suit particular data and messages better than others. In this chapter, we show you how to choose the graphic form that best helps readers both grasp your data and understand how they support your argument.

15.1 CHOOSING VISUAL OR VERBAL REPRESENTATIONS

When the data are few and simple, readers can grasp them as easily in a sentence as in a table:

> In 2013, on average, men earned $50,033 a year and women $39,157, a difference of $10,876.

TABLE 15.1. Male-female salaries ($), 2013

Men	50,033
Women	39,157
Difference	10,876

But if you present more than a few numbers, readers will struggle to keep them straight:

> Between 1970 and 2010, the structure of families changed in two ways. In 1970, 85 percent of families had two parents, but in 1980 that number declined to 77 percent, then to 73 percent in 1990, to 68 percent in 2000, and to 64 percent in 2010. The number of one-parent families rose, particularly families headed by a mother. In 1970, 11 percent of families were headed by a single mother. In 1980, that number rose to 18 percent, in 1990 to 22 percent, and to 23 percent in 2000. There were some marginal changes among single fathers (headed 1 percent of the families in 1970, 2 percent in 1980, 3 percent in 1990, and 4 percent in 2000). Families headed by no adult remained stable at 3–4 percent.

A note on terminology: We use the term *graphics* for all visual representations of data. Traditionally, graphics are divided into *tables* and *figures.* A table is a grid with columns and rows. Figures are all other graphic forms, including graphs, charts, photographs, drawings, and diagrams. Figures that present quantitative data are divided into *charts* and *graphs.* Charts typically consist of bars, circles, points, or other shapes; graphs consist of continuous lines.

15.2 CHOOSING THE MOST EFFECTIVE GRAPHIC

When you graphically present data as complex as in that paragraph, the most common choices are tables, bar charts, and line graphs, each of which has a distinctive rhetorical effect.

A table seems precise and objective. It emphasizes discrete numbers and requires readers to infer relationships or trends on their own (unless you state them in an introductory sentence).

TABLE 15.2. Changes in U.S. family structure, 1970–2010

Percentage of total families					
Family type	1970	1980	1990	2000	2010
2 parents	85	77	73	68	64
Mother	11	18	22	23	27
Father	1	2	3	4	4
No adult	3	4	3	4	4

Charts and line graphs present a visual image that communicates values less precisely than do the exact numbers of a table but with more impact. But charts and graphs also differ. A bar chart emphasizes contrasts among discrete items:

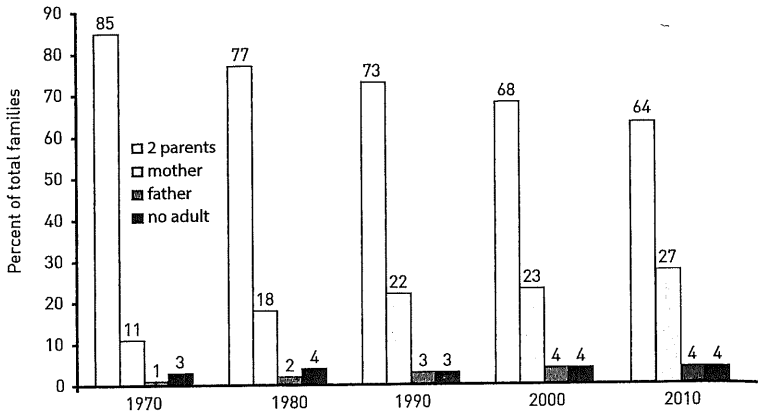

FIGURE 15.1. Changes in U.S. family structure, 1970–2000

A line graph suggests continuous change over time:

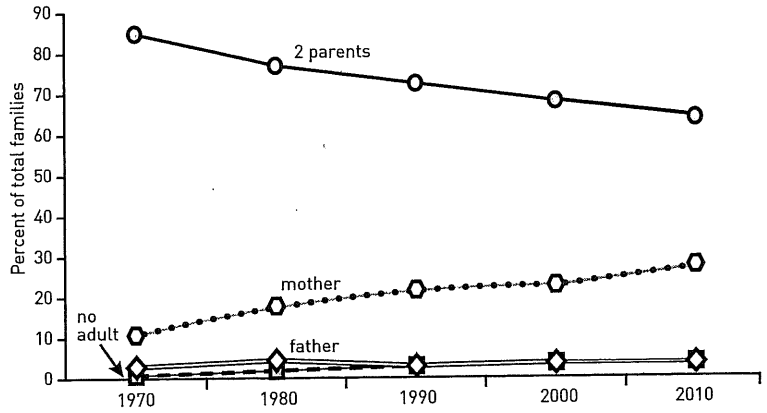

FIGURE 15.2. Changes in U.S. family structure, 1970–2000

Choose the form that achieves the effect you want, not the one that comes to mind first.

How many choices you should consider depends on your experience. If you're new to quantitative research, limit your choices to basic tables, bar charts, and line graphs. Your computer software offers more choices, but ignore those that you aren't familiar with. If you're doing advanced research, readers will expect you to draw

from a larger range of graphics favored in your field. In that case, consult table 15.7, which describes the rhetorical uses of other common forms. You may have to consider even more creative ways of representing data if you are writing a dissertation or article in a field that routinely displays complex relationships in large data sets. (See the bibliography for additional resources.)

What follows is a guide to the basics of tables, charts, and graphs.

15.3 DESIGNING TABLES, CHARTS, AND GRAPHS

Computer programs create graphics so dazzling that many writers let their software determine their design. That's a mistake. Readers don't care how fancy a graphic looks if it doesn't communicate your point clearly. Here are some principles for designing effective graphics. To follow them, you may have to change default settings in your graphics software.

15.3.1 Frame Each Graphic to Help Readers Understand It

A graphic representing complex numbers rarely speaks for itself. You must frame it to show readers what to see in it and how to understand its relevance to your argument:

1. Label every graphic in a way that describes its data. For a table, the label is called a *title* and is set flush left above the table; for a figure, the label is called a *legend* and is set flush left below the figure. Keep titles and legends short but descriptive enough to distinguish every graphic from every other one.

   - Avoid making the title or legend a general topic.

     NOT: Heads of households

     BUT: Changes in one- and two-parent heads of households, 1970–2010

   - Do not give background information or characterize what the data imply.

     NOT: Weaker effects of counseling on depressed children before professionalization of staff, 1995–2004

     BUT: Effect of counseling on depressed children, 1995–2004

- Be sure labels distinguish graphics presenting similar data.

  NOT: Risk factors for high blood pressure

  BUT: Risk factors for high blood pressure among men in Cairo, Illinois

  OR: Risk factors for high blood pressure among men in St. Louis, Missouri

2. Insert into the table or figure information that helps readers see how the data support your point. For example, if numbers in a table show a trend and the size of the trend matters, indicate the change in a final column. If a line on a graph changes in response to an influence not mentioned on the graph, add text to the image to explain it.

   Although reading and math scores declined by almost 100 points following redistricting, that trend reversed when supplemental math and reading programs were introduced.

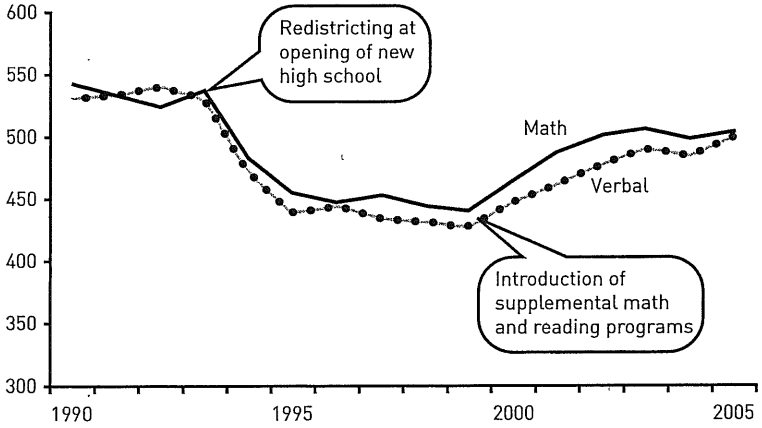

FIGURE 15.3. SAT scores for Mid-City High, 1990–2005

3. Introduce the table or figure with a sentence that explains how to interpret it. Then highlight what it is in the table or figure that you want readers to focus on, particularly any number or relationship mentioned in that introductory sentence. For example, we have to study table 15.3 to understand how it supports the sentence before it:

Most predictions about increased gasoline consumption have proved wrong.

TABLE 15.3. Gasoline consumption

	1970	1980	1990	2000
Annual miles (000)	9.5	10.3	10.5	11.7
Annual consumption (gal.)	760	760	520	533

We need a sentence to explain how the numbers support or explain the claim, a more informative title, and visual help that highlights what we should see in the table:

Gasoline consumption has not grown as predicted. Though Americans drove 23 percent more miles in 2000 than in 1970, they used 32 percent less fuel.

TABLE 15.4. Per capita mileage and gasoline consumption, 1970–2000

	1970	1980	1990	2000
Annual miles (000)	9.5	10.3	10.5	11.7
(% change vs. 1970)		8.4%	10.5%	23.1%
Annual consumption (gal.)	760	760	520	533
(% change vs. 1970)		0%	(31.5%)	(31.6%)

The added sentence tells us how to interpret the key data in table 15.4, and the shading tells us where to find them.

15.3.2 Keep All Graphics as Simple as Their Content Allows

Some guides encourage you to cram as much data as you can into a graphic. But readers want to see only the data relevant to your point, free of distractions. For all graphics:

1. Include only relevant data. If you include data only for the record, label it accordingly and put it in an appendix.
2. Keep the visual impact simple.
   - Box a graphic only if you group two or more figures.
   - Do not color or shade the background.

FOR TABLES

- Never use both horizontal and vertical dark lines to divide columns and rows. Use light gray lines only if the table is complex or you want to direct your reader's eyes in one direction to compare data.
- For tables with many rows, lightly shade every fifth row.

FOR CHARTS AND GRAPHS

- Use background grid lines only if the graphic is complex or readers need to see precise numbers. Make them light gray.
- Color or shade lines or bars only to show a contrast. Use color only if the text will be printed in color and not photocopied later. (Black-and-white photocopies make many colors look alike.)
- Never use iconic bars (for example, images of cars to represent automobile production) or add a third dimension merely for effect. Both look amateurish and can distort how readers judge values.
- Plot data on three dimensions only when your readers are familiar with such graphs and you cannot display the data in any other way.

3. Use clear labels.
   - Label all rows and columns in tables and both axes in charts and graphs.
   - Use tick marks and labels to indicate intervals on the vertical axis of a graph.
   - If possible, label lines, bar segments, and the like on the image rather than in a legend set to the side. Use a legend only if labels would make the image too complex to read.
   - When specific numbers matter, add them to bars or segments in charts or to dots on lines in graphs.

15.4 SPECIFIC GUIDELINES FOR TABLES, BAR CHARTS, AND LINE GRAPHS

15.4.1 Tables

Tables with lots of data can seem dense, so organize them to help readers.

- Order the rows and columns by a principle that lets readers quickly find what you want them to see. Do not automatically choose alphabetic order.
- Round numbers to a relevant value. If differences of less than 1,000 don't matter, then 2,123,499 is irrelevantly precise.
- Sum totals at the bottom of a column or at the end of a row, not at the top or left.

Compare tables 15.5 and 15.6.

TABLE 15.5. Unemployment in major industrial nations, 1990–2000

	1990	2001	Change
Australia	6.7	6.5	(0.2)
Canada	7.7	5.9	(1.8)
France	9.1	8.8	(0.3)
Germany	5.0	8.1	3.1
Italy	7.0	9.9	2.9
Japan	2.1	4.8	2.7
Sweden	1.8	5.1	3.3
UK	6.9	5.1	(1.8)
USA	5.6	4.2	(1.6)

Table 15.5 looks cluttered and its items aren't helpfully organized. In contrast, table 15.6 is clearer because it has an informative title, less visual clutter, and items organized to let us see the pattern more easily.

TABLE 15.6. Changes in unemployment rates of industrial nations, 1990–2000

English-speaking vs. non-English-speaking nations			
	1990	2001	Change
Canada	7.7	5.9	(1.8)
UK	6.9	5.1	(1.8)
USA	5.6	4.2	(1.6)
Australia	6.7	6.5	(0.2)
France	9.1	8.8	(0.3)
Japan	2.1	4.8	2.7
Italy	7.0	9.9	2.9
Germany	5.0	8.1	3.1
Sweden	1.8	5.1	3.3

15.4.2 Bar Charts

Bar charts communicate as much by visual impact as by specific numbers. But bars arranged in no pattern imply no point. If possible, group and arrange bars to create an image that matches your message. For example, look at figure 15.4 in the context of the explanatory sentence before it. The items are listed alphabetically, an order that doesn't help readers see the point.

> Most of the world's deserts are concentrated in North Africa and the Middle East.

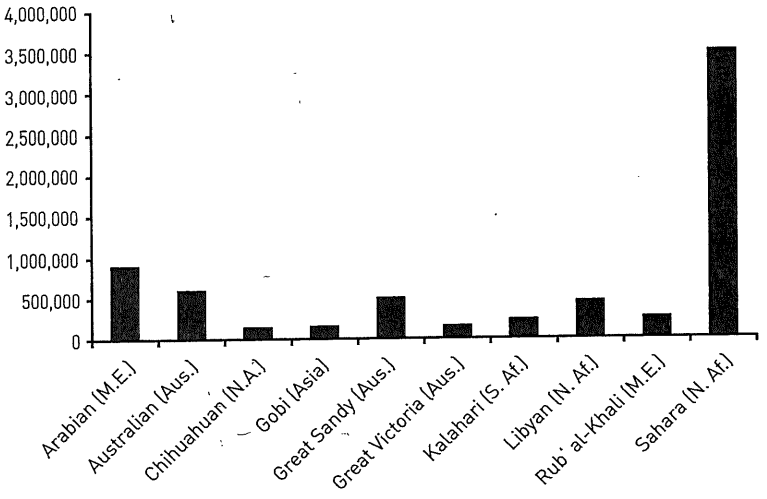

FIGURE 15.4. World's ten largest deserts

In contrast, figure 15.5 supports the claim with a coherent image.

> Most of the world's deserts are concentrated in North Africa and the Middle East.

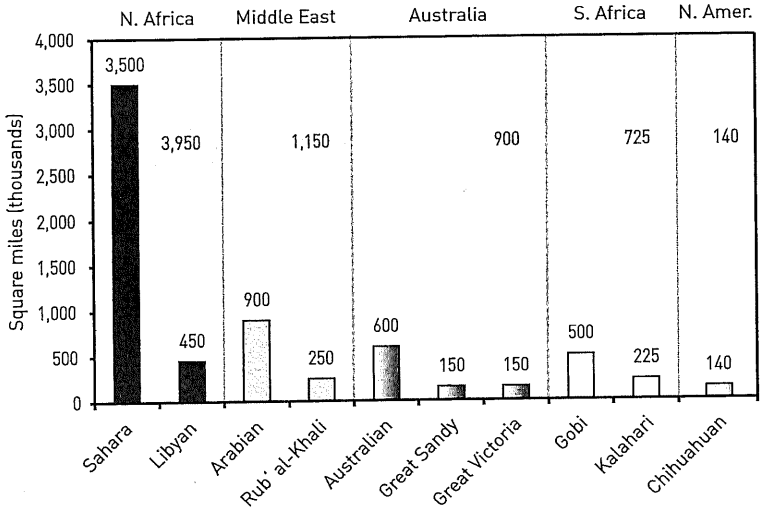

FIGURE 15.5. World distribution of large deserts

In standard bar charts, each bar represents 100 percent of a whole. But sometimes readers need to see specific numbers for parts of the whole. You can do that in two ways:

- Divide the bars into proportional parts, creating a "stacked" bar.
- Give each part of the whole its own bar, then group the parts into clusters.

Use stacked bars only when you want readers to compare whole values for different bars rather than their divided segments, because readers can't easily compare the proportions of segments by eye alone. If you do use stacked bars, do this:

- Arrange segments in a logical order. If possible, put the largest segment at the bottom in the darkest shade.
- Label segments with specific numbers and to assist comparisons; connect corresponding segments with gray lines.

Compare figures 15.6 and 15.7:

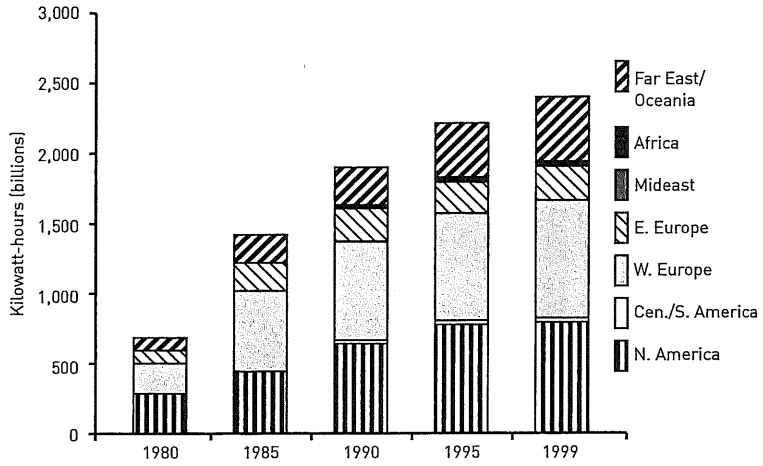

FIGURE 15.6. World generation of nuclear energy, 1980–1999

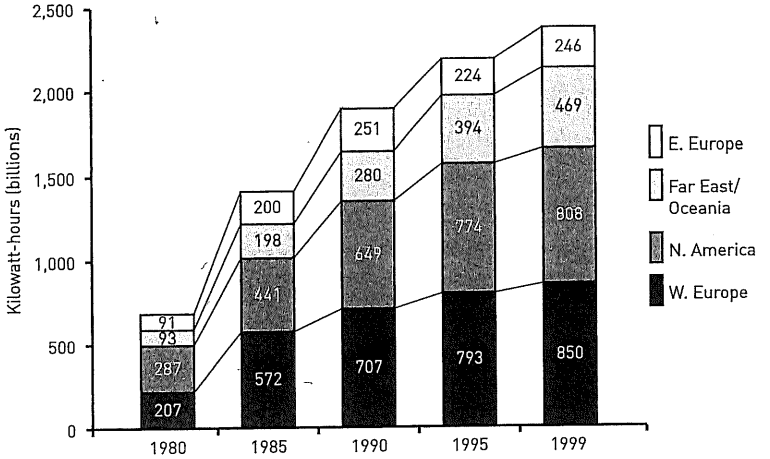

FIGURE 15.7. Largest generators of nuclear energy, 1980–1999

If you group bars because segments are as important as the wholes, do this:

- Arrange groups in a logical order; if possible put bars of similar size next to one another (order bars in the same way through all the groups).
- Label groups with the number for the whole, either above each group or below the labels on the bottom.

Most data that fit a bar chart can also be shown in a pie chart. Pie charts are popular in magazines, tabloids, and annual reports. While splashy, they are harder to read than bar charts. Readers must compare proportions of segments whose sizes are often hard to judge. But pie charts have their place, especially to communicate qualitative impressions about the comparative size of data, either to show that one segment is disproportionately larger than the rest or that the data is divided into many small segments. Avoid using pie charts to convey quantitative data, however. Use bar charts instead.

15.4.3 Line Graphs

Because a line graph emphasizes trends, readers must see a clear image to interpret it correctly. Do the following:

- Choose the variable that makes the line go in the direction, up or down, that supports your point. If the good news is a reduction (down) in high school dropouts, you can more effectively represent the same data as a rising line indicating increase in retention (up). If you want to emphasize bad news, find a way to represent your data as a falling line.
- Plot more than six lines on one graph only if you cannot make your point in any other way.
- If you have fewer than ten or so data points, indicate them with dots. If only a few are relevant, insert numbers to show their exact value.
- Do not depend on different shades of gray to distinguish lines, as in figure 15.8.

Compare figure 15.8 and figure 15.9:

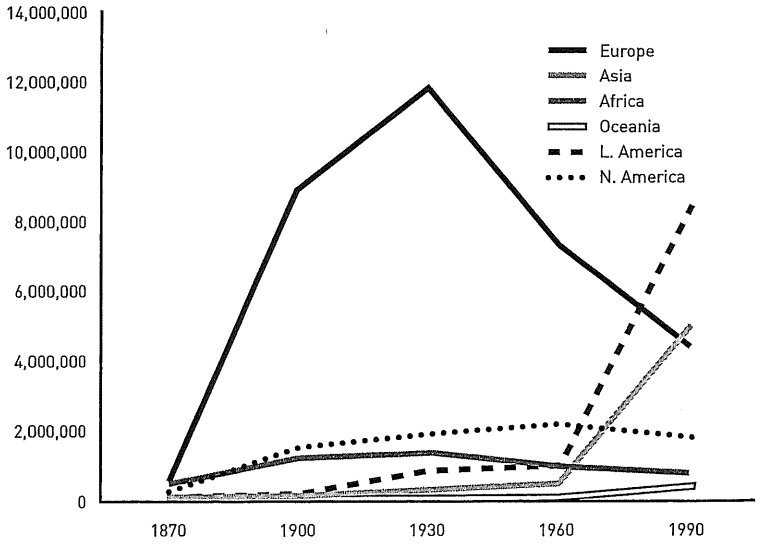

FIGURE 15.8. Foreign-born residents in the United States, 1870–1990

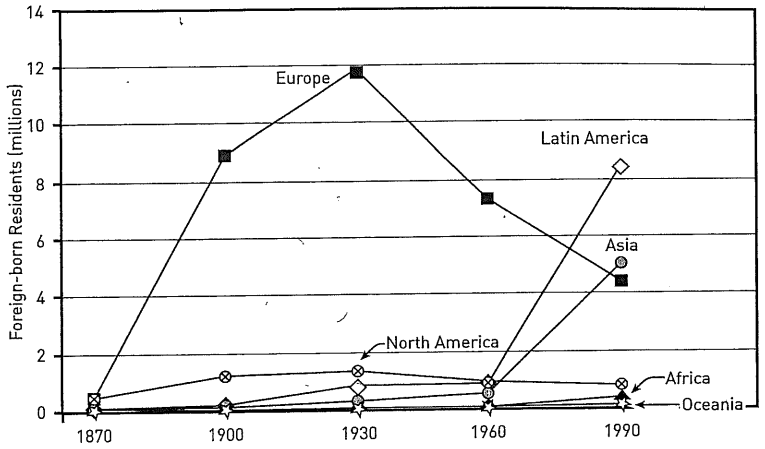

FIGURE 15.9. Foreign-born residents in the United States, 1870–1990

Figure 15.8 is harder to read because the shades of gray do not distinguish the lines well against the background and because our eyes have to flick back and forth to connect the lines to the legend. Figure 15.9 makes those connections clearer.

These different ways of showing the same data can be confusing. To cut through that confusion, test different ways of representing the same data. Construct alternative graphics; then ask someone unfamiliar with the data to judge them for impact and clarity. Be sure to introduce the figures with a sentence that states the claim you want the figure to support.

15.5 COMMUNICATING DATA ETHICALLY

Your graphic must be not only clear and accurate, but honest. Do not distort the image of the data to make your point. For example, the two bar charts below display identical data, yet imply different messages:

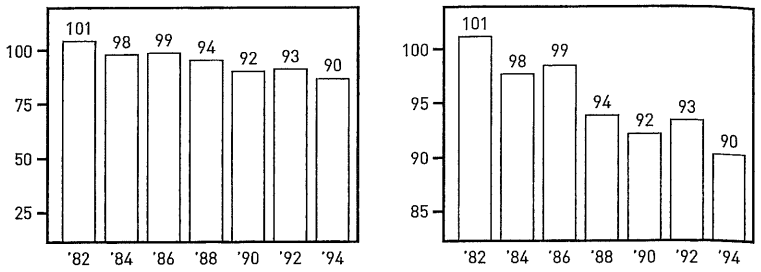

FIGURE 15.10. Capitol City pollution index, 1982–1994

The 0–100 scale in the figure on the left creates a fairly flat slope, which makes the drop in pollution seem small. The vertical scale in the figure on the right, however, begins not at 0 but at 80. When a scale is so truncated, it creates a sharper slope that exaggerates small contrasts.

Graphs can also mislead by implying false correlations. Someone might claim that unemployment goes down as union membership goes down and offer figure 15.11 as evidence. And indeed, in that graph, union membership and the unemployment rate do seem to move together so closely that a reader might infer one causes the other:

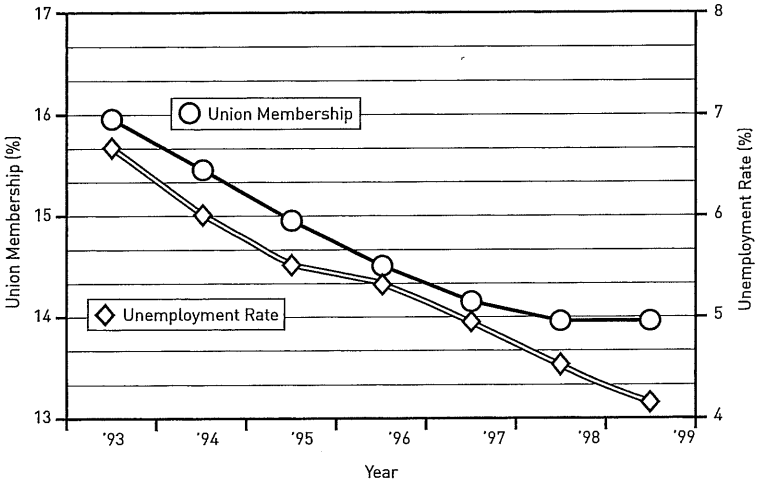

FIGURE 15.11. Union membership and unemployment rate, 1993–1999

But the scale for the left axis (union membership) differs from the scale for the right axis (the unemployment rate), making it seem that the two trends could be causally related. They may be, but that distorted image doesn't prove it.

Graphs can also mislead when the image encourages readers to misjudge values. The two charts in figure 15.12 represent exactly the same data but seem to communicate different messages:

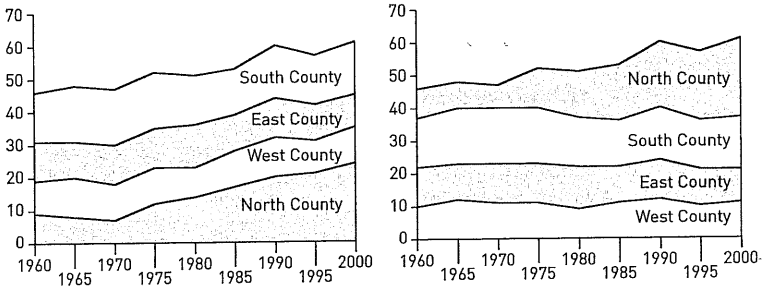

FIGURE 15.12. Representation of suburban counties in state university undergraduates (percent of total)

The charts in figure 15.12 are both stacked area charts. Despite their visual differences, they represent the same data. Area charts such as these represent changes in values not by the *angles* of the lines, but by the areas *between* them. In both charts, the bands for south, east, and west are roughly the same width throughout, indicating little change in the values they represent. The band for the north, however, widens sharply, representing a sharp increase in the numbers it represents. In the chart on the left, readers could easily misjudge the top three bands, because they are on top of the rising north band, making those bands seem to rise as well. In the chart on the right, on the other hand, those three bands do not rise because they are on the bottom. Now only the band for the north rises.

Here are four guidelines for avoiding visual misrepresentation:

- Do not manipulate a scale to magnify or reduce a contrast.
- Do not use a figure whose image distorts values.
- Do not make a table or figure unnecessarily complex or misleadingly simple.
- If the table or figure supports a point, state it.

TABLE 15.7. Common graphic forms and their uses

	Data	Rhetorical Uses
Bar Chart	Compares the value of one variable across a series of items called cases (e.g., average salaries for service workers$_{variable}$ in six companies$_{cases}$).	Creates strong visual contrasts among individual cases, emphasizing individual cases. For specific values, add numbers to bars. Can show ranks or trends. Vertical bars (called *columns*) are most common but can be horizontal if cases are numerous or have complex labels. See 15.4.2.
Bar Chart, Grouped or Split	Compares the value of one variable, divided into subsets, across a series of cases (e.g., average salaries$_{variable}$ for men and women service workers$_{subsets}$ in six companies$_{cases}$).	Contrasts subsets within and across individual cases; not useful for comparing total values for cases. For specific values, add numbers to bars. Grouped bars show ranking or trends poorly; useful for time series only if trends are unimportant. See 15.4.2.
Bar Chart, Stacked	Compares the value of one variable, divided into two or more subsets, across a series of cases (e.g., harassment complaints$_{variable}$ segmented by region$_{subsets}$ in six industries$_{cases}$).	Best for comparing totals across cases and subsets *within* cases; difficult to compare subsets across cases (use grouped bars). For specific values, add numbers to bars and segments. Useful for time series. Can show ranks or trends for total values only. See 15.4.2.
Histogram	Compares two variables, with one segmented into ranges that function like the cases in a bar graph (e.g., service workers$_{continuous\ variable}$ whose salary is \$0–5,000, \$5,001–10,000, \$10,001–15,000, etc.$_{segmented\ variable}$).	Best for comparing segments within continuous data sets. Shows trends but emphasizes segments (e.g., a sudden spike at \$5,000–10,000 representing part-time workers). For specific values, add numbers to bars.
Image Chart	Shows value of one or more variables for cases displayed on a map, diagram, or other image (e.g., states$_{cases}$ colored red or blue to show voting patterns$_{variable}$).	Shows the distribution of the data in relation to preexisting categories; de-emphasizes specific values. Best when the image is familiar, as in a map or diagram of a process.
Pie Chart	Shows the proportion of a single variable for a series of cases (e.g., the budget share$_{variable}$ of U.S. cabinet departments$_{cases}$).	Best for comparing one segment to the whole. Useful only with few segments or segments that are very different in size; otherwise comparisons among segments are difficult. For specific values, add numbers to segments. Common in popular venues, frowned on by professionals. See 15.4.2.

TABLE 15.7. *(continued)*

	Data	Rhetorical Uses
Line Graph	Compares continuous variables for one or more cases (e.g., temperature$_{variable}$ and viscosity$_{variable}$ in two fluids$_{cases}$).	Best for showing trends; deemphasizes specific values. Useful for time series. To show specific values, add numbers to data points. To show the significance of a trend, segment the grid (e.g., below- or above-average performance). See 15.4.3.
Area Chart	Compares two continuous variables for one or more cases (e.g., reading test scores$_{variable}$ over time$_{variable}$ in a school district$_{case}$).	Shows trends; deemphasizes specific values. Can be used for time series. To show specific values, add numbers to data points. Areas below the lines add no information but will lead some readers to misjudge values. Confusing with multiple lines/areas.
Area Chart, Stacked	Compares two continuous variables for two or more cases (e.g., profit$_{variable}$ over time$_{variable}$ for several products$_{cases}$).	Shows the trend for the total of all cases, plus how much each case contributes to that total. Likely to mislead readers on the value or the trend for any individual case, as explained in 15.5.
Scatterplot	Compares two variables at multiple data points for a single case (e.g., housing sales$_{variable}$ and distance from downtown$_{variable}$ in one city$_{case}$) or at one data point for multiple cases (e.g., brand loyalty$_{variable}$ and repair frequency$_{variable}$ for ten manufacturers$_{cases}$).	Best for showing the distribution of data, especially when there is no clear trend or when the focus is on outlying data points. If only a few data points are plotted, it allows a focus on individual values.
Bubble Chart	Compares three variables at multiple data points for a single case (e.g., housing sales,$_{variable}$ and distance from downtown,$_{variable}$ prices$_{variable}$ in one city$_{case}$) or at one data point for multiple cases (e.g., image advertising,$_{variable}$ repair frequency,$_{variable}$ and brand loyalty$_{variable}$ for ten manufacturers$_{cases}$).	Emphasizes the relationship between the third variable (bubbles) and the first two; most useful when the question is whether the third variable is a product of the others. Readers easily misjudge relative values shown by bubbles; adding numbers mitigates that problem.