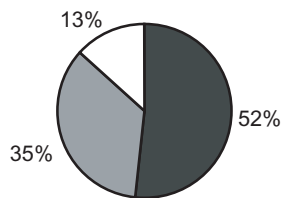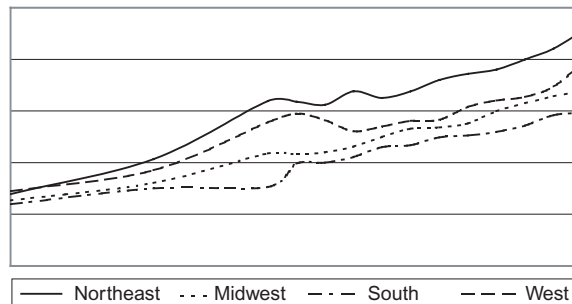# 7

## Creating Effective Charts

1. What is missing from the charts in figures 7A and 7B?

**Age distribution of the elderly population
United States, 2000**



**Figure 7A.**

**Median sales price of new one-family homes,
by region, United States, 1980–2000**



Northeast ···· Midwest —·— South ——— West

**Figure 7B.**

2. Answer the following questions for figures 7.3, 7.5a, and 7.9 in *Writing about Numbers*.
   a. Who is described by the data?
   b. To what date or dates do the data pertain?
   c. Where were the data collected?
   d. What criteria were used to organize the values of the variables on chart axes? (Hint: consider type of variable.)
   e. What are the units of measurement? Are they the same for all numbers shown in the chart?
   f. Are there footnotes to the chart? If so, why? If not, are any needed?

3. For each of the following topics, identify the type of task (e.g., univariate distribution, relationship between two variables, or relationship among three variables), and types of variables to be presented (e.g., nominal, ordinal, interval), then state which type of chart would be most appropriate.

   a. Annual number of people receiving college degrees, by gender, from 1980 to 2000.

   b. Average commuting costs per month, by mode of transportation (bicycle, bus, car, train, walk, other). (One number per type of transportation.)

   c. Current market share for Coca Cola, Pepsi, and other cola brands.

   d. Distribution of SAT mathematics scores in 2000. (Range = 200 to 800, in increments of 10.)

   e. Educational attainment distribution ($< HS$, $= HS$, $> HS$) for native-born United States residents and for immigrants from other North American countries, Africa, Asia, Australia & New Zealand, Europe, and Latin America, 2000.

   f. Estimates of dates for each of 15 archeological artifacts, with margin of error for each estimate.

   g. Relationship between systolic blood pressure (mm Hg) and percentage body fat for a sample of 150 elderly people.

   h. Type of contraceptive (condom, diaphragm, implant/injectable, oral contraceptive (the Pill), surgical sterilization, other, none) by 10-year age groups of women aged 15 to 44 in the United States in 1997. (Some women use more than one method.)

   i. Trends in mean annual global temperature (°F ) and carbon dioxide ($CO_2$) concentration (ppmv) from 1950 to 2000.

4. For each of the topics in question 3 that involve an *XY*-type chart, indicate which principle you would use to decide what order to display values on the *x*-axis; see chapter 6 of *Writing about Numbers* for a list of organizing principles.

5. Create a stacked bar chart to present the data shown in table 7A, allowing bar height to vary, reflecting total number of ozone days. To help you plan your chart, answer the following questions, then draw an approximate stacked bar chart, allowing the level to vary by county.
   a. Which variable goes on the *x*-axis, and what principle would you use to organize its values?
   b. Which variable goes in the slices (and legend)?
   c. Which variable goes on the *y*-axis, and in what units is it measured?

**Table 7A. Number of unhealthy ozone days by level of warning for selected counties in Indiana, 1996–1998**

| | Level of warning[a] | | |
|---|---|---|---|
| | Unhealthy for sensitive groups | Unhealthy | Very unhealthy |
| Allen | 25 | 0 | 0 |
| Clark | 29 | 3 | 1 |
| Elkhart | 15 | 0 | 0 |
| Floyd | 27 | 6 | 0 |
| Hamilton | 31 | 3 | 0 |
| Hancock | 28 | 2 | 0 |
| Lake | 29 | 2 | 0 |
| La Porte | 26 | 6 | 1 |
| Madison | 27 | 3 | 0 |
| Marion | 32 | 3 | 0 |
| Porter | 25 | 3 | 0 |
| Posey | 14 | 1 | 0 |
| St. Joseph | 21 | 1 | 0 |
| Vanderburgh | 32 | 2 | 0 |
| Vigo | 25 | 1 | 0 |
| Warrick | 40 | 3 | 0 |

[a]Unhealthy for sensitive groups = 0.085–0.104 parts per million (ppm); Unhealthy = 0.105–0.124 ppm; Very unhealthy = 0.125–0.374 ppm.

Source: American Lung Association.

6. Revise your chart from the previous question to illustrate the relative importance (share) of different levels of ozone warning in each county.
   a. What aspects of each chart remain the same as in the previous question? What aspects change?
   b. What are the advantages and disadvantages of the two versions of the chart? Be specific for this topic and data.