

# 7

## Creating Effective Charts

### SOLUTIONS

1. Figure 7A is missing a legend; 7B is missing axis titles, axis labels, and units of measurement.
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, continuous), then state which type of chart would be most appropriate.
  - a. Three-way association between one continuous and one nominal predictor (date and gender, respectively), and a categorical outcome (number of people receiving degrees). Multiple line chart, to show trends by date on the (*x*-axis) in the number of people receiving college degrees (on the *y*-axis), with a separate line for each gender (shown in the legend).
  - b. Two-way (bivariate) association between transportation mode (nominal) and cost (continuous). Simple bar chart, with one bar for each transportation mode and cost on the *y*-axis.
  - c. Composition (univariate) of a nominal variable. Pie chart to illustrate the percentage (or dollar value) of market share for each cola producer.
  - d. Distribution (univariate) of many-valued continuous variable. Line chart with SAT scores on *x*-axis and number or percentage of cases on *y*-axis.
  - e. Distribution of one categorical variable (educational attainment) within another categorical variable (continent). Stacked bar chart, with one bar for the U.S. and for each continent of origin and one slice for each educational attainment level. Each bar totals 100% of that continent's immigrants (on the *y*-axis) to illustrate composition while controlling for different numbers of immigrants from each continent.
  - f. Association between one nominal variable (archeological artifact) and one continuous variable (date). High/low/close chart ("high" and "low" show either standard error or 95% CI) with artifacts on *x*-axis and date on *y*-axis. Alternatively, a bar chart with error bars added.
  - g. Association between two continuous variables. Scatter chart (perhaps with a fitted line), to allow for the possibility of several cases with the same value (e.g., several people with the same weight but different blood pressure).
  - h. Distribution of a nominal variable (contraceptive type) by an interval variable (age group). Clustered bar chart with one cluster for each contraceptive type and a different bar color for each age group (in the

- legend). Y-axis shows % of each age group that uses that method. A stacked bar chart cannot be used because contraceptive types are not mutually exclusive (each woman could use more than one type).
- i. Two bivariate associations between pairs of continuous variables: temperature by date and CO<sub>2</sub> concentration by date. Line chart with two different Y-scales to portray concurrent trends in two continuous variables, each of which is measured in different units.
5. Stacked bar chart to present the data shown in table 7A. Same title as table 7A.
- a. Counties arranged on the x-axis in descending order of total number of unhealthy ozone days.
  - b. A different color slice for each level of ozone warning, identified in the legend.
  - c. Number of unhealthy ozone days goes on the y-axis.