

4

Five More Technical Principles

PROBLEM SET

1. For each of the following topics, indicate whether the variable or variables used to measure it are continuous or categorical, and single or multiple response.
 - a. Respondent's current marital status
 - b. Respondent's current number of siblings
 - c. Siblings' current heights
 - d. Current marital status of siblings
 - e. Temperature at 9 A.M. today
 - f. The forms of today's precipitation
2. A new school is being considered in your hometown. Several possible grade configurations are being considered (Plan A: grades K–3, 4–5, 6–8, 9–12; Plan B: grades K, 1–4, 5–7, 8–12). The current configuration is K–5, 6–9, and 10–12. Design a question to collect information from school principals on the age distribution of students, making sure the data collection format provides the detail and flexibility needed to compare the different scenarios for the district now and in five years.
3. In a health examination survey, several hundred girls aged five to ten years were measured with a metric measuring tape marked in increments of millimeters. The estimated coefficient on age (years) from an OLS model of height was reported as 5.06666667 centimeters. Write a sentence to report that coefficient.
4. In a microbiology lab exercise, the size of viral cells being compared ranged from 0.000000018 meters (m) in diameter for Parovirus to 0.000001 m in length for Filoviridae (American Society for Microbiology 1999). What scale would you use to report those data in a table? In the text?

5. Write one or two sentences to compare the four specimens in table 4A. Which specimen is the heaviest? The lightest? By how much do they differ? What information do you need before you can make the comparison?

Table 4A. Mass of four specimens

| Specimen | Mass |
|----------|---------------|
| 1 | 1.2 pounds |
| 2 | 500 grams |
| 3 | 0.7 kilograms |
| 4 | 12 ounces |

6. For each of the figures 4.3a through 4.3e (*Writing about Multivariate Analysis*, 64–66), choose
 - a. a typical value;
 - b. an atypical value;
 - c. a plausible contrast (two values to compare).
 Explain your choices, with reference to range, central tendency, variation, and skewness.

7. Identify pertinent standards or cutoffs for each of the following questions.
 - a. Does Mr. Jones deserve a speeding ticket?
 - b. Is the new alloy strong enough to be used for the library renovations?
 - c. How tall is five-year-old Susie expected to be next year?
 - d. Does Vioxx increase the odds of a heart attack?
 - e. Is this year's projected tuition increase at Public U unexpected?
 - f. Should we issue an ozone warning today?

8. Indicate whether each of the following sentences correctly reflects table 4B. If not, rewrite the sentence so that it is correct. Check both correctness and completeness of the data.
 - a. Between 1964 and 1996, there was a steady decline in voter participation, from 95.8% in 1964 to 63.4% in 1996.
 - b. Voter turnout was better in 1996 (63.4%) than in 1964 (61.9%).
 - c. Almost all registered voters participated in the 1964 United States presidential election.
 - d. The best year for voter turnout was 1992, with 104,600 people voting.
 - e. Less than half of the voting age population voted in the 1996 presidential election.
 - f. A higher percentage of the voting-age population was registered to vote in 1996 than in 1964.

Table 4B. Voter turnout, United States presidential elections, 1964 through 1996

| Year | Total Vote (1000s) | Registered Voters (RV) (1000s) | Vote/RV (%) | Voting Age Pop. (VAP) (1000s) | Vote/VAP (%) |
|------|--------------------|--------------------------------|-------------|-------------------------------|--------------|
| 1964 | 70,645 | 73,716 | 95.8 | 114,090 | 61.9 |
| 1968 | 73,212 | 81,658 | 89.7 | 120,328 | 60.8 |
| 1972 | 77,719 | 97,329 | 79.9 | 140,776 | 55.2 |
| 1976 | 81,556 | 105,038 | 77.6 | 152,309 | 53.5 |
| 1980 | 86,515 | 113,044 | 76.5 | 164,597 | 52.6 |
| 1984 | 92,653 | 124,151 | 74.6 | 174,466 | 53.1 |
| 1988 | 91,595 | 126,380 | 72.5 | 182,778 | 50.1 |
| 1992 | 104,600 | 133,821 | 78.2 | 189,529 | 55.2 |
| 1996 | 92,713 | 146,212 | 63.4 | 196,511 | 47.2 |

Source: Institute for Democracy and Electoral Assistance 1999.

9. A billboard reads: “1 in 250 Americans is HIV positive. 1 in 500 of them knows it.”
 - a. According to the two statements above, what share of Americans are HIV positive and know it? Does that seem realistic?
 - b. Rewrite the second statement to clarify the intended meaning
 - i. as a fraction of HIV-positive Americans;
 - ii. as a fraction of all Americans.

10. An advertisement for a health education program included figure 4A to show the prevalence of two common health behavior problems among teenaged girls. What is wrong with the graph?

Prevalence of smoking and teen pregnancy (%)

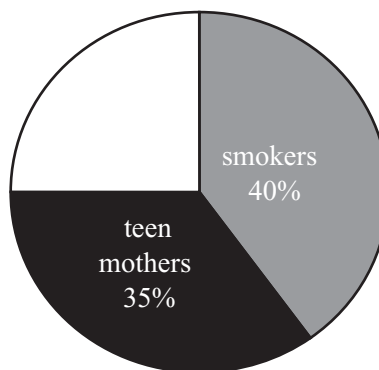


Figure 4A.

11. You are involved in a research team that is conducting a study of commuting. One of the team members submits the following question to be included on the questionnaire:
“How do you usually commute to work?
Car __ Public transportation __ Train __ Carpool __ Walk __”
 - a. Critique the wording of the question using the guidelines in chapter 4 of *Writing about Multivariate Analysis*.
 - b. Revise the question to correct the problems you identified in part a.
12. What is wrong with the following fictitious set of instructions for authors from a scientific journal? “In the interest of saving space, round all numeric results to the nearest single decimal place.”
13. Each of these statements contains an error. Identify the problem and rewrite the statement to correct the error. If additional information would be needed to make the correction, indicate what kind of information is needed.
 - a. The proportionate increase in income during the 1990s was 20%.
 - b. Male infants outnumbered females (sex ratio at birth = 0.95).
 - c. A majority of respondents (0.67) agreed that there should be a waiting period before buying a gun.
 - d. Cancer accounted for two out of every ten deaths, equivalent to a death rate of 20%.