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 number of siblings.
    c. Siblings' heights.
    d. Current marital status of siblings.
    e. Temperature at 9 A.M. today.
    f. Form 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. Your stopwatch is accurate to the nearest tenth of a second. In nine trials, the average time for a mouse named Squeeky to solve a maze was 10.44444444 seconds. Write a sentence to report that average.
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 steps do you need to take before you can make the comparison?

Table 4A. Mass of four specimens

Specimen	Mass
1	1.2 lbs.
2	500 grams
3	0.7 kilograms
4	12 ounces

6. For each of the figures 4.3a through 4.3e (*Writing about Numbers*, 2nd Edition, 80–81), 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 and other information needed to answer 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. Can Leah go on the Ferris wheel at the amusement park?
   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 US 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 teenage girls. What is wrong with the graph?

Prevalence of smoking and teen pregnancy (%)

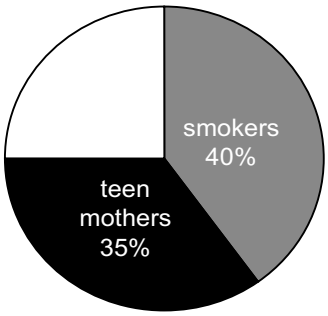

Figure 4A

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 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 Numbers*, 2nd Edition.
    b. Revise the question to correct the problems you identified in part a.

12. For each of figures 4.3a through 4.3e (*Writing about Numbers*, 2nd Edition, 80–81), come up with an example variable (topic) and context for which that distribution might be plausible.

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%.

14. What is wrong with the following fictitious set of instructions for authors from a scientific journal that frequently publishes results of inferential statistical tests?

"In the interest of saving space, round all numeric results to the nearest single decimal place."

15. Fill in table 4C to indicate whether the following values make sense for the concepts and units mentioned. If not, explain why not and suggest a plausible range of values for that variable.

Table 4c

Concept	Units	Value	Plausible (Y/N?)	If no, specify a plausible range of values
a. Height of an adult female in the US, 2014	Inches	65		
b. Height of an adult female in the US, 2014	Inches	120		
c. Height of a 6-year-old child in the US, 2014	Inches	65		
d. Height of a 6-year-old child in the US, 2014	Centimeters	120		
e. Observed daily low temperature, New York City, January	Degrees Fahrenheit	30		
f. Observed daily low temperature, New York City, January	Degrees Fahrenheit	−10		
g. Observed daily low temperature, New York City, January	Degrees Celsius	30		
h. Observed daily low temperature, New York City, January	Degrees Kelvin	−10		
i. Mean hourly wage, fast food employee, Los Angeles, CA, 2010	Dollars	8		
j. Mean annual wage, fast food employee, Los Angeles, CA, 2010	Dollars	8		

k. Share of population that is poor	Percentage	25		
l. Share of population that is poor	Percentage	–5		
m. Share of population that is poor	Proportion	25		
n. Share of population that is poor	Proportion	0.05		
o. Change in share of population that is poor	Percentage change	0.05		
p. Change in share of population that is poor	Percentage change	–5		

16. A weather scale based on daily observations is to be used to quantify weather in cities worldwide in all seasons. Dr. Blowhard suggests that the scale should be constructed for each observation (date, location) by adding together values of the following variables:
    - observed high temperature, °F
    - rainfall, inches
    - snowfall, inches
    - an indicator of whether the wind speed exceeded 20 MPH, coded 1 = yes, 2 = no
    - a variable coded 1 = sunny all day, 2 = sunny part of the day, 3 = cloudy all day

    a. Critique the scale construction method using the criteria discussed in chapter 4 of *Writing about Numbers*, 2nd Edition.
    b. If you think a valid scale could be constructed using the listed variables, provide instructions on how to do so.
    c. If you do not think a valid scale could be constructed using the listed variables, describe how you could use the variables (e.g., in a study of how weather affects immigration).

Table 4D. Coding of items about abortion attitudes, 2000 US General Social Survey

"Please tell me whether or not you think it should be possible for a pregnant woman to obtain a legal abortion . . ."	Yes	No	Don't know
If there is a strong chance of serious defect in the baby?	1	2	8
If she is married and does not want any more children?	1	2	8
If the woman's own health is seriously endangered by the pregnancy?	1	2	8
If she is not married and does not want to marry the man?	1	2	8
If she became pregnant as a result of rape?	1	2	8
If the woman wants it for any reason?	1	2	8

Source: National Opinion Research Center, 2000.

17. Answer the following questions about construction of an abortion attitudes scale using the six items shown in table 4D.
    a. How should "don't know" answers for each of the items be treated when computing the scale?
    b. If each respondent's scores on the individual items are *summed* to compute the value of the scale for them,
        i. what is the minimum possible valid value of the scale? The maximum?
        ii. what would be the substantive interpretation of a scale score of 12 for an individual respondent?
    c. If each respondent's scores on the individual items are *averaged* to compute the scale, what is the minimum possible valid value of the scale? The maximum?
    d. If the value of the scale for each respondent consists of a *tally* of the number of items with which that respondent agreed, what is the minimum possible valid value of the scale? The maximum?
    e. If only valid ("yes" or "no") answers to each abortion question were included in the calculation of the scale, what can you say about the set of respondents who would be included in the final analytic sample?