12

Writing about Data and Methods

1. For each of the following scenarios, list what information you would report in a data section for a scientific paper. Hint: what additional information would you want to know?
   a. A three-year study of a six-month drug rehabilitation program that recruited 200 subjects to examine cure and relapse rates
   b. A study of calcium intake among 50 pregnant women, based on their recall over a two-week period

2. Dr. Dollar is conducting a study of poverty patterns in the United States based on annual income data from the 2000 census. She defines a categorical measure of income group comparing family income (calculated from income of individual family members, alimony, and four types of social benefits) against the federal poverty thresholds. Classifications are defined in terms of multiples of the threshold: <.50, .50–.99, 1.00–1.84, 1.85–2.99, and 3.00 or greater. Search for "poverty" on the U.S. Census Web page (*http://www.census.gov*) for more detail. State how you would describe the poverty measure in
   a. a one-page summary of the study for a local newspaper;
   b. documentation of a new data set that has collected data on each of the income components as part of a written questionnaire;
   c. a journal article on poverty patterns, written for people who are familiar with poverty thresholds.

3. Making use of newly available data from a three-year panel study of a sample of 10,000 people drawn from the 2000 census, Dr. Dollar describes movement in and out of poverty and duration of poverty (in months) over the study period. Poverty was defined as family income below the threshold (<1.0). Data were collected annually, with retrospective recall of income in each of the previous 12 months. What information should be added to item 2.c to describe these data for this research question?

4. A researcher at the Panel Study of Income Dynamics accidentally erased a file containing information from two years' worth of data. Embarrassed, he went ahead and analyzed data for the other 30 years in the study. What assumptions did he implicitly make about the missing data?

5. Fauth et al. (2004) studied the effects of a residential mobility experiment, comparing outcomes of low-income adults who moved to public housing in low-poverty neighborhoods with outcomes for those who stayed in public housing in their original high-poverty neighborhoods. They studied the six neighborhood and housing quality measures shown in table 12A. What information about these variables should be included in a data section for a scientific paper about this study?

Table 12A. Means and standard deviations of neighborhood and housing characteristics, Yonkers Residential Mobility Program, 1994–1995

Measure	Mean	Standard deviation
Danger (3 items)	0.72	0.91
No. of victimizations in past year (1 item)	0.21	0.58
Disorder (5 items)	0.72	0.74
Cohesion (4 items)	0.52	0.32
Resources (5 items)	2.98	0.60
Housing quality (5 items)	0.35	0.43

Source: Adapted from Rebecca C. Fauth, Tama Leventhal, and Jeanne Brooks-Gunn, "Short-term effects of moving from public housing in poor to middle-class neighborhoods on low-income, minority adults' outcomes," *Social Science and Medicine* 59 (2004): 2271–84, table 1, *http://www.sciencedirect.com/science*.

6. For each of the following data, methods, and objectives, write a short discussion of strengths and limitations for the concluding section of a scientific article.
    a. Study: 20 subjects were interviewed at the Snooty Golf Club at noon on a Friday in early April regarding their preferred color and fit of jeans. Objective: a marketing study by the Gap clothing store.
    b. Study: two classes of second graders in the same school were given a math test in September. One class was then taught with a new math curriculum, the other with the standard curriculum. The classes were tested again in June. Objective: an evaluation of the new math curriculum.
    c. Study: data on hair color and age were collected for everyone aged 25 − 84 in a city of 200,000 people. Deaths over a two-year period were ascertained from death certificates. Two models were estimated: one with hair color as the independent variable and mortality as the dependent variable; the second with age as another independent variable. Objective: understand the potential benefit of hair dye in improving survival.

7. In her study "Gender, Preloss Marital Dependence, and Older Adults' Adjustment to Widowhood," Carr (2004) uses data on respondents who were widowed between waves 1 and 2 of the study, matched to control subjects who remained married at wave 2. (See table 9D for more on her study.) Carr's study used data from a longitudinal study over a seven-year period. In her methods section, she describes a model of attrition (nonparticipation at wave 2) from the sample between waves 1 and 2. She found that "age and anxiety increased the risk of nonparticipation, and home ownership decreased the risk of nonparticipation at wave 2." None of the other demographic, socioeconomic, or health characteristics were associated with attrition.
    a. Write an equation to convey her final specification for the model of attrition, including the dependent variable and type of model estimated.
    b. What questions is she trying to answer with that model?
    c. Write a short discussion of the implications of her attrition findings for interpretation of her results about psychological adjustment to widowhood.