

10

Writing about Data and Methods

PROBLEM SET

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 (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 would you want to add to item 2.c to describe these data for this research question?
4. A researcher in a meteorology lab accidentally erased a file containing information from two years’ worth of climatic data. Embarrassed, he went ahead and analyzed data for the other 28 years in the study. What assumptions did he implicitly make about the missing data?

5. For each of the following data, methods and objectives, write a short discussion of strengths and limitations for the concluding section of a general interest newspaper article.
 - a. Study: twenty 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.
 - 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 25–85 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 predictor and mortality as the outcome; the second with age as another predictor. Objective: understand the potential benefit of hair dye in improving survival.
6. Revise the material in box 10.2 for a general interest newspaper article. Consider detail, vocabulary, and citation style.