

14. *Writing about Distributions and Associations*

SOLUTIONS

1. Descriptions of the specified tables and charts.
 - a. “Table 5.3 shows the demographic composition of the study sample. Just over half of the 2,058 respondents were female (51.4%). Persons aged 40 to 64 years were the largest single age group in the sample (41.4%), just edging out persons aged 18–39 (37.8%). Elderly persons (aged 65 and older) made up about one-fifth of the sample.
“The most common racial/ethnic group was non-Hispanic whites, with 2½ times as many respondents as the second largest racial/ethnic group, non-Hispanic blacks (55.6% and 22.1%, respectively). Hispanics comprised the third-largest group (15.9%), followed by Asians (4.2%) and persons of other racial/ethnic origin (2.2%).”
 - b. “In 2000 in the United States, human resources comprised by far the largest single category of federal outlays (61% of the \$1.8 trillion spent that year; figure 6.2b). The second largest category—national defense—accounted for only about one-quarter as much as human resources (16% of the total). Net interest, physical resources, and other functions together comprised the remaining 23% of all outlays.”
3. Use the results from table 11A (Zimmerman 2003) to answer the given questions.
 - a. “Among Williams College students with SAT scores in the middle 70%, women’s GPAs averaged 0.11 points higher than men’s ($p < 0.01$).”
 - b. “Among students in the bottom 15% of SAT scores, there was no significant association between roommate’s math SAT score and student’s college GPA. Although the estimated coefficient suggests a GPA 0.15 points lower if roommate’s math SAT were 400 instead of 800, the finding was not statistically significant.”
5. In statistical terminology, a situation where the association between one independent variable (mode of interview) and the dependent variable (odds of reporting a sensitive behavior) depends on a second

independent variable (type of behavior) is called an interaction or effect modification. In GEE lingo, it is called an exception.

- a. Description of the pattern for a group of first-year high school students: “A recent study in Kenya found that the chances of reporting specific sensitive behaviors such as having had sex with a stranger or being coerced into sex differed depending on how the data were collected (table 12A). For the three most sensitive topics studied, boys were more likely to report having experienced those behaviors if they were interviewed using a self-administered computer-aided interview than if interviewed in person. On the other hand, they were more likely to report ever having had a girlfriend if interviewed in person than if they completed a computer-aided self-interview.”
- b. Description of the pattern for a group of graduating statistics majors: “A study by Mensch and colleagues of teenagers in Kenya found that method of data collection and type of sensitive behavior interact in their effect on odds of reporting sensitive behaviors such as having had sex with a stranger or being coerced into sex (table 12A). For the three most sensitive topics studied, the odds of reporting those behaviors were highest among boys interviewed using a self-administered computer-aided interview and lowest among those interviewed in person. In contrast, for the topic ‘ever having had a girlfriend’ the odds were highest among boys interviewed in person and lowest among those completing a computer-aided self-interview.”