Solution

a. The zip codes don't measure or count anything. The numbers are actually labels for geographic locations.

b. The ranks reflect an ordering, but they don't measure or count anything. The rank of 1 is from a book with sales substantially greater then the book with rank of 2, so the different numbers don't correspond to the magnitudes of the sales.

c. The numbers on the football jerseys don't measure or count anything. Those numbers are simply substitutes for names.

Example 6 involves data that do not justify the use of statistics such as the mean or median. Example 7 involves a more subtle issue.

Example 7 **Class Size**

It is well known that smaller classes are generally more effective. According to the National Center for Education Statistics, California has a mean of 20.9 students per teacher, and Alaska has a mean of 16.8 students per teacher. (These values are based on elementary and secondary schools only.) If we combine the two states, we might find the mean number of students per teacher to be 18.85, as in the calculation shown below, but is this result correct? Why or why not?

$$\bar{x} = \frac{20.9 + 16.8}{2} = 18.85$$

Solution

The combined states of California and Alaska do *not* have a mean of 18.85 students per teacher. The issue here is that California has substantially more students and teachers than Alaska, and those different numbers should be taken into account. Combining California and Alaska, we get a total of 6,539,429 students and 315,013 teachers, so the student/teacher ratio is $6{,}539{,}429/315{,}013 = 20.8$ (not the value of 18.85 from the above calculation). When using values from different sample sizes, consider whether those sample sizes should be taken into account.

As another illustration of the principle in Example 7, if you have a list of the 50 state per capita income amounts and you find the mean of those 50 values, the result is not the mean per capita income for the entire United States. The population sizes of the 50 different states must be taken into account. (See the *weighted mean* in the following subsection.)

Part 2: Beyond the Basics of Measures of Center

Calculating the Mean From a Frequency Distribution

The first two columns of Table 3-2 shown here are the same as the frequency distribution of Table 2-2 from Chapter 2. When working with data summarized in a frequency distribution, we don't know the exact values falling in a particular class, so we make calculations possible by pretending that all sample values in each class are equal to the class midpoint. For example, consider the first class interval of 50–69 with a