CHAPTER 5

Measures of Central Tendency

The most commonly used descriptive statistics are measures of central tendency. As you can guess from this title, measures of central tendency attempt to locate the mid-most or center point in a group of data. For example, what was the average starting salary of the students who graduated from the MPA program last year? On average over the past 5 years, how many MPA students accepted job offers from nonprofit organizations? On last week's midterm exam, what was the mid-most score (i.e., the score that divided the observations in half)? On average, how many employees of the Mechanicsburg city government report job-related accidents every month? What was the mid-most amount of monetary donations received by the Agency for Civic Renewal such that half the donations received were larger and half were smaller? The measures of central tendency give a shorthand indication of the main trend in the data.

A **measure of central tendency** is a number or score or data value that represents the average in a group of data. Three different types of averages are calculated and used most often in public and nonprofit management. The first is the *mean,* which is the arithmetic average of the observations; the second is the *median,* which is the observation that falls exactly in the middle of the group; and the third is the *mode,* or the data value that occurs with greatest frequency. This chapter shows how to calculate the three measures of central tendency for both ungrouped and grouped data (data that have been assembled into a frequency distribution) and discusses the use and interpretation of these measures. It concludes with a discussion of the relationship between the measures of central tendency and the different levels of measurement—interval, ordinal, and nominal—you learned about in Chapter 2.

The Mean

The first measure of central tendency is the mean. The **mean** is the arithmetic average of a set of numbers. To calculate the mean, add all the numbers (the data points or observations), and divide this new number (the sum) by the total number of observations in the set, which we labeled **N** in Chapter 4.

To illustrate, suppose that the head of the Bureau of Records wants to know the mean length of government service of the employees in the bureau's Office of

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Table 5.1 **Years of Government Service**

Employee	Years	Employee	Years
Bush	8	Obama	9
Clinton	15	Gore	11
Reagan	23	Cheney	18
Kerry	14	Carter	20

Computer Support. Table 5.1 displays the number of years that each member of the office has been employed in government.

To calculate the mean length of government service of the employees in the bureau's Office of Computer Support, add the years of service of each of the employees. You should come up with a total of 118 years. Divide this sum by the number of employees in the office (8). This procedure yields the mean number of years of government service of the employees in the office (14.75).

The procedure for calculating the mean can be presented as a formula:

$$\mu = \frac{\sum_{i=1}^{N} X_i}{\mathbf{N}}$$

This formula is not as formidable as it seems. The Greek letter μ (mu) on the left side of the equal sign is the statistician's symbol for the population mean; it is pronounced "mew." The mean μ is equal to the formula on the right side of the equal sign. Another statistician's symbol is Σ; it means add (or sum) all the values of X (in our example, these were years of government service). The subscripts below and above the Σ indicate to sum all the *X's* from the first X_1, all the way through X_N, the last observation (here, the years of government service of the eighth employee). Finally, the formula says to divide the sum of all the *X*'s by **N**, the number of items being summed.

The formula for calculating the mean of a sample or subset of data, rather than the entire population, is identical except that different symbols are used to distinguish sample and population. The sample mean is denoted $\overline{X}$ ("x-bar"), and the number of observations in the sample **n**. Chapter 11 reviews these distinctions and elaborates their importance. That chapter introduces the topic of how the analyst can use a sample of data to make an inference to the larger population from which it is drawn.

The mean has several important characteristics:

1. Every item in a group of data is used to calculate the mean.
2. Every group of data has one and only one mean; as mathematicians would say, the mean is rigidly determined.
3. The mean can take on a value that is not realistic. For example, the average U.S. family had exactly 1.7 children, 2.2 pets, 1.1 automobiles, and made financial contributions to 3.4 charitable organizations in the past year.

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

4. An extreme value, sometimes called an **outlier**, has a disproportionate influence on the mean and thus may affect how well the mean represents the data. For example, suppose that the head of the Bureau of Records decides to shake up the Office of Computer Support by creating a new position in the office with responsibility to expedite operations. To fill the position, the head appoints a newly graduated MPA with a fine background in computers but only 1 year of prior service in government. In the space provided, calculate the mean years of government service of the employees in the expanded Office of Computer Support.

If you performed the calculations correctly, you should get a mean length of government service of 13.22 years. (The sum of the observations is 119; dividing by 9, the number of observations, yields a mean of 13.22.) This number tends to understate the years of government service of the employees in the Office of Computer Support. Why?

The Median

The second measure of central tendency is the median. The **median** is the middle observation in a set of numbers when the observations are ranked in order of magnitude. For example, the Stermerville City Council requires that all city agencies include an average salary in their budget requests. The Stermerville City Planning Office has seven employees. The director is paid $62,500; the assistant director makes $59,500. Three planning clerks are paid $42,600, $42,500, and $42,400. The secretary (who does all the work) is paid $37,500, and a receptionist is paid $36,300.

The planning director calculates the mean salary and finds that it is $46,186. Check this result (you should get a total payroll of $323,300 for the seven employees). This result disturbs the director, because it makes the agency look fat and bloated. The secretary points out that the large salaries paid to the director

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

and the assistant director are distorting the mean. The secretary then calculates the median, following these steps:

Step 1: List the salaries in order of magnitude. You may start with the largest or the smallest; you will get the same answer. The secretary prepares the following list:

Director	$62,500
Assistant Director	59,500
Clerk 1	42,600
Clerk 2	42,500
Clerk 3	42,400
Secretary	37,500
Receptionist	36,300

Step 2: Locate the middle item. With seven persons, the middle observation is easy to find; it is the fourth item, or the salary paid to clerk 2 ($42,500). For larger data sets, the rule is to take the number of observations (7, in this case) and add 1 to it (7 + 1 = 8). Divide this number by 2, and that number (8 ÷ 2 = 4) tells you that the median is the fourth observation (once data values have been put in order). Note that it makes no difference whether you select the fourth item from the top or the fourth from the bottom. The median is $42,500.

The planning director reports the median salary to the Stermerville City Council. It is lower than the mean—why? Nevertheless, the Stermerville mayor tells the planning director that because of the local tax revolt, the planning office must fire one person. The planning office responds, as all bureaucracies do, by firing the receptionist. After this action, what is the median salary of the planning office? Calculate it in the space provided.

After arranging the salaries in order, you may have discovered that with six observations, no middle item exists. The formula (N + 1) ÷ 2 seems to offer no help because (6 + 1) ÷ 2 = $3\frac{1}{2}$. But the median is actually that, the $3\frac{1}{2}$th item from the top (or bottom). Because observations cannot be split in half, we define $3\frac{1}{2}$ as halfway between the third and fourth items, in this case halfway between the salaries of clerk 2 and clerk 1. Because clerk 2 makes $42,500 and clerk 1 makes $42,600, the median is $42,550 [(42,600 + 42,500) ÷ 2]. Whenever the number of items (**N**) is an even number, the median will be halfway between the middle observations. This is easy to remember if you think of the median as the measure of central tendency that divides a set of numbers so that exactly half are smaller than the median and exactly half are larger than the median.

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

The median has several important characteristics:

1. The median is not affected by extreme values.
2. Although every observation is used to determine the median, the actual value of every item is not used in the calculations. At most, only the two middle items are used to calculate the median.
3. If items do not cluster near the median, the median will not be a good measure of the group's central tendency.
4. The median usually does not take on an unrealistic value. The median number of children per family in the United States, for example, is 2.
5. The median is the 50th percentile in a distribution of data, because half the observations fall above it and half fall below it. (Percentiles are measures of relative ranking. You may have seen them reported on a standardized test, such as the Graduate Record Examination, or GRE. As with the median, they express what percentage of the scores fall below a given score. For example, 79% of the scores lie below the 80th percentile.) In percentage distributions, you can use this fact to locate the median quickly by observing the data value at which the distribution of scores or observations crosses the 50th percentile. That value is the median.

Although the median conveys less precise information than the mean (knowing where the middle of a set of observations falls is less exact than is the precise numerical average), as you will see later in the chapter, the median is sometimes used in preference to the mean. In certain distributions of data, the median is more descriptive of the central tendency. For instance, an outlier typically has a much greater distorting effect on the mean than on the median. In addition, when a phenomenon cannot be measured on an (equal) interval scale (for example, job attitudes, client satisfaction, quality of life, volunteer interest in career development, and so forth), the median is especially useful as a measure of central tendency.

The Mode

The final measure of central tendency is the mode. The **mode** is simply the data value that occurs most often (with greatest frequency) in any distribution. In the frequency distribution in Table 5.2, what is the mode number of tickets issued? The value that occurs most often is 3 tickets issued, so 3 is the mode.

The distribution in Table 5.2 has only one mode; thus, it is called unimodal. The distribution in Table 5.3 is bimodal; it has two modes. Because Kapaun had 9 arrests for 14 weeks and 11 arrests for 14 weeks, the modes are 9 and 11. Distributions also can be trimodal, tetramodal, and so on.

Statisticians generally relax the definition of the mode(s) to include the distinct peaks or clusters in the data that occur with high frequency. You should, too. Table 5.4 presents an example—the number of research methods and

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Table 5.2 **Tickets Issued by Woodward Police, Week of January 28, 2011**

Number of Tickets	Number of Police Officers
0	2
1	7
2	9
3	14
4	3
5	2
6	1
	38

Table 5.3 **Arrests per Week, Kapaun Air Station, 2011**

Number of Arrests	Number of Weeks
7	2
8	4
9	14
10	8
11	14
12	10
	52

statistics courses required for graduation by a sample of MPA-granting schools with concentrations in nonprofit administration. The distribution is bimodal, with modes at one and three courses. Even though the latter mode occurs slightly less often (19 vs. 23), it would be misleading to ignore it, so it, too, is considered a mode.

The mode has several important characteristics:

1. Because the most common value in a distribution of data can occur at any point, the mode need not be "central" or near the middle.

2. Unlike the mean and the median, the mode can take on more than one value. No other measure of central tendency has this characteristic. As a result, in some complex data distributions, such as a bimodal or trimodal distribution, the mode is the statistic of choice for summarizing the *central tendencies.* When a distribution has more than one distinct mode, it usually indicates something important in the data that cannot be captured so well by the mean or median.

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Table 5.4	Number of Required Courses in Research Methods and Statistics
Number of Courses	**Number of Schools**
0	3
1	23
2	5
3	19
	50

3. More often than not, when a variable is measured on a numerical or interval scale (number of arrests, feet of snow plowed, and so forth), the mode may be of little interest. However, with variables measured at less precise levels—nominal (classifications, such as race or religion) and ordinal (rank orderings, such as responses on opinion questions)—the mode is much more useful and in certain instances is the preferred measure of central tendency. We return to this issue later in the chapter.

The Mean versus the Median

In most situations, numerical data can be summarized well with the mean. However, because situations can arise in which the mean gives a misleading indication of central tendency, the median is preferred. When extreme values or outliers occur on a variable, the mean is distorted or pulled toward them. The statistical term is *skewed*. By contrast, because the median is the value of the middle case—once the data points have been arranged in order—it will remain in the middle of the distribution even if the variable has an extreme value. In this situation, the median is the preferred measure of central tendency. (Chapter 6 returns to this issue.)

For example, suppose that a nonprofit administrator needed to estimate the average price of houses in a city in order to apply for a federal grant. She takes a random sample of 10 homes sold recently and discovers 9 of them sold for between \$180,000 and \$220,000, and 1 home sold for well over \$1 million. The mean housing price will be grossly inflated by the one outlying case and will yield a value unrepresentative of the price of houses in the city. The median will not be affected by the deviant case, however, and will have a value near the middle of housing prices, between \$180,000 and \$220,000. Thus, the median will be more typical of housing prices and should be used in preference to the mean. Remember that the purpose of measures of central tendency is to identify accurately the central point(s) in the distribution. Select and use the measures of central tendency that most clearly describe or represent the distribution of data.

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Levels of Measurement and Measures of Central Tendency

Chapter 2, "Measurement," introduced the concept of levels of measurement. Students of public and nonprofit administration are most accustomed to interval measurement: variables that can be measured on a numerical scale, such as the budget of an agency or work group in dollars, the number of clients assisted, or the amount of overtime hours logged by the cafeteria staff last week. Public and nonprofit managers (and those employed in business) also use variables measured at other, less precise levels. In discussing the median and the mode, we referred to these levels earlier in the chapter: Ordinal variables allow rank ordering of information—for example, how satisfied a client is with a nonprofit organization's response to her inquiry (very satisfied, satisfied, neutral, dissatisfied, or very dissatisfied) or a citizen's overall assessment of the services provided by the public library (very good, good, neutral, poor, or very poor). Nominal variables are classifications that have no metric information—for example, the gender of employees, their religion, their marital status, and so forth.

Why devote so much attention to the levels of measurement of the variables encountered in public and nonprofit administration? The answer is both simple and important: The statistics that can be appropriately calculated to summarize the distribution of single variables (the subject of this chapter) and to describe the relationship between variables (the subject of later chapters) differ from level to level. It is easy to see the source of the differences. Each of the levels expresses a different amount of information about a variable, and this idea is reflected directly in the kind of statistics that may be calculated and used.

For example, if a variable is measured at the interval level, we usually know everything about it that we may wish to know. It is possible to locate precisely all the observations along a scale: $32,749 yearly income; 4.57 prostitution arrests per week; 38 years of age; 247 cubic feet of sewage; 10,106 hours volunteered to United Way agencies last month. Because for these variables an equal distance separates each whole number on the measurement scale (dollars, arrests, years, cubic feet, and hours), all mathematical operations can be performed. Thus, as we saw earlier in the chapter, the scores of a group of cases or observations can be added and the sum divided by the number of observations to obtain the mean income, number of arrests, age, cubic feet of sewage, and hours volunteered. It is also possible to find the *median* or the middle score of the group for each of these variables. And, of course, the *mode*—the value occurring most frequently in a distribution—presents no problem.

Table 5.5 displays the number of pilots at selected air bases. At what level of measurement are these data? Be sure that you can calculate the mean (a total of 11,886 pilots ÷ 7 air bases = 1,698), median (896), and mode (0) for this distribution. If you have any problems, review the earlier portions of the chapter where these statistics are discussed. With interval-level data, the manager can calculate and use all three measures of central tendency.

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Table 5.5 **Pilots at Selected Air Bases**

Air Base	Number of Pilots
Minot	0
Torrejon	2,974
Kapaun	896
Osan	0
Andrews	6,531
Yokota	57
Guam	1,428

Now consider *ordinal* data. At this level of measurement we are able to rank objects or observations, but it is not possible to locate them precisely along a scale. A citizen may "strongly disapprove" of the Springhill mass transit system's performance, but no number is available that expresses her exact level of disapproval or how much less she approves than if she had said "disapprove." Because there are no numerical scores or numbers attached to the responses—which, in the case of interval variables, are added and divided to compute the mean—*it is not possible to calculate the mean for a variable measured at the ordinal level.*

How about the median? Can it be calculated for ordinal data? Suppose an interviewer obtains from 11 citizens their responses to the question, "Do you approve or disapprove of the Springhill mass transit system's performance? Strongly approve, approve, neutral, disapprove, strongly disapprove?"

Ordinal data are commonly displayed in a frequency distribution. Table 5.6 presents the frequency distribution for the Springhill mass transit system. Consider the "cumulative total" column in the table. Note that the mid-most case is the sixth citizen, who has the response "disapprove," which is, therefore, the median. As presented earlier in the chapter, the rule to find the mid-most case is to add 1 to the total number of cases, or **N,** and divide this result by 2. In this example, 11 + 1 = 12, and 12 divided by 2 = 6, so that the sixth case is in the middle of the distribution and has the median value, "disapprove."

It is important to note that the median is *not* "neutral." Although "neutral" is the middle response category, it does not tell us anything about the middle response given by the 11 citizens. "Neutral" falls in the middle of the scale but not in the middle of the 11 citizens. Note also that the median is the middle score of the 11 citizens, or the response "disapprove."

We have now shown that the median can be calculated for ordinal variables. So can the mode. In Table 5.6, the most frequently mentioned response is "strongly disapprove," given by four citizens. Therefore, it is the mode or modal response.

Finally, at the *nominal* level of measurement, it is not possible to assign numerical scores to cases (interval level). A score of 1.7 on religion or 458 on

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Table 5.6 **Frequency Distribution of Citizens' Responses Concerning Springhill's Mass Transit System**

Response	Number of Citizens	Cumulative Total
Strongly approve	1	1
Approve	2	3
Neutral	2	5
Disapprove	2	7
Strongly disapprove	4	11

nationality would be arbitrary and would make no sense. *Thus, it is not possible to calculate the mean for nominal data.*

Furthermore, the values of a group of cases on a nominal variable cannot be ranked in any kind of meaningful ordering of least to most, or vice versa (ordinal level). There is no meaningful or correct way to order the categories of race, religion, sex, or any other nominal variable. (Usually, we place them in alphabetical order for convenience, but that ordering is not a numerical or measurement scale.) Because the median is predicated on the ability to rank cases or observations of a variable so that the middle or median value may be found, *it is not possible to calculate the median for nominal data*.

However, the mode can be found for nominal data. For the data in Table 5.7, which is the modal occupation of the employees of the Civil Service Commission?

The mode is "lawyer," because it is the occupation of the largest number of people (192) in the distribution. Usually, the percentage of observations in the modal category is given, here 61%. Make sure that you can calculate the percentage distribution. If you have any difficulty, see Chapter 4.

Table 5.7 **Civil Service Commission Employees by Occupation**

Occupation	Number of People	Percentage
Lawyer	192	61
Butcher	53	17
Doctor	41	13
Baker	20	6
Candlestick maker	7	2
Indian chief	3	1
	N = 316	100

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Hierarchy of Measurement

We can summarize this discussion of levels of measurement and measures of central tendency in a convenient table. In Table 5.8, place an X in the column of a row if the designated measure of central tendency (mean, median, mode) can be calculated for a given level of measurement (nominal, ordinal, interval).

If you have completed the table correctly, the X's will form a triangle pattern that slopes upward from the bottom left entry (Nominal-Mode) to the top right entry (Interval-Mean). If you did not find this pattern, you should review earlier parts of the chapter.

The lesson of Table 5.8 is that any statistic that can be calculated for a variable at a lower (less precise) level of measurement can also be calculated at all higher (more precise) levels of measurement. If the statistic can be calculated at a level that places fewer numerical restrictions on the data, it can be calculated at higher levels that meet these numerical requirements—and more. Thus, the mode is available at all three levels, the median at the ordinal and the interval levels, and the mean only at the interval level. This rule is usually stated as the "hierarchy of measurement" to indicate the ascending power of the higher levels of measurement. To the degree possible, then, it is always to your advantage to construct, collect, and use variables measured at higher levels.

With this knowledge, you are now in a position to describe phenomena of interest in quantitative terms. For example, consider your work organization. The mean age of employees may be 37.1 years, the median 35, and the mode 39. The median opinion of employees with respect to contracting out for information technology services may be "disapprove"; perhaps the modal opinion is "strongly disapprove." Most of the employees may be white and male; and so on.

Table 5.8 Hierarchy of Measurement

	Level of Measurement		
Measure of Central Tendency	Nominal	Ordinal	Interval
Mean			
Median			
Mode			

Some Cautions

Two cautions regarding this discussion of levels of measurement should be kept in mind. First, most of the time you will not calculate statistics yourself; instead, a computer program will compute them for you. In order to store

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

information compactly in a computer, the substantive category labels or names for ordinal variables—such as strongly agree, agree, neutral, disagree, and strongly disagree—as well as for nominal variables—such as white, African American, and Hispanic—are entered and stored in the computer as numbers. These numbers are usually called *codes.* The computer may also store the labels or names for the codes, but it performs all calculations based on the codes, not the value labels assigned to them. For example, the coding schemes in Table 5.9 may apply.

Because the computer calculates all statistics based on the numerical codes entered for the variables (not the value labels), strange things can happen to the unwary analyst. For example, if instructed to do so, the computer can and will calculate a mean or a median for nominal variables or a mean for ordinal variables based on the codes—even though these statistics have no meaning at these levels of measurement. It is up to you as the analyst to recognize such statistics as a mean attitude of 2.7 or a median race of 1 for what they are: garbage. Note that for interval variables the codes are the actual data values (7.123, 5.6, 10075.9,14, and so forth), so this problem does not arise.

The second caution is in part a consequence of the first. Because ordinal variables frequently are coded for computer utilization in the manner shown earlier— 1 = strongly agree, 2 = agree, 3 = neutral, and so on—some students (and practicing managers) have jumped to the incorrect conclusion that these codes are actually meaningful numbers on a scale that expresses the precise level of an individual's agreement or disagreement with an interviewer's question. In other words, they have assumed that the coding categories are actual numbers and can be treated as such for statistical calculations—just as if they were interval data. This practice, which is rather common in all the social sciences, including political science and public and nonprofit administration, has led to the ordinal-interval debate. (Don't feel bad if you have missed it; it is not exactly a household term.) As the title suggests, the debate has focused on the justification for—or the lack of it—and the advantages of treating ordinal data as interval.

Table 5.9 Examples of Two Coding Schemes

Coding Scheme 1		Coding Scheme 2	
Code	Response	Code	Response
1	Strongly agree	1	White
2	Agree	2	African American
3	Neutral	3	Hispanic
4	Disagree		
5	Strongly disagree		

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Both sides have produced some persuasive evidence, and the debate has yet to be resolved definitively.

Our recommendation for students just starting out in quantitative work in public or nonprofit management is that you adopt the stance of the statistical purist—that you calculate and use for ordinal data only those statistics that are clearly appropriate for that level of measurement. For now, when you need to summarize or describe the distribution of an ordinal variable, rely on the median and mode. In the future, if you decide to continue your work in statistics, you can read some of the literature in the ordinal-interval debate and come to your own conclusions.

You may be wondering why we are placing such emphasis on the ordinal-interval debate. When you see the steps involved in examining and interpreting *relationships* among ordinal variables—as compared with those involved in the analysis of interval data (see the chapters on regression)—the significance of this debate will become clear. But that is the purpose of Chapters 15 through 17.

Chapter Summary

Measures of central tendency are values used to summarize a body of data by indicating middle (or central) points in the distribution. Each measure of central tendency has a distinct meaning and method of calculation. The mean is the arithmetic average of all data points. The median is the data value that is greater than 50% of all the data points and less than 50% of all the data points. The mode is the data value that occurs most frequently in the distribution. This chapter illustrates the main features of all three measures of central tendency.

The chapter shows how the level of measurement of a variable (interval, ordinal, nominal) determines the measures of central tendency that can appropriately be calculated for it. The hierarchy of measurement illustrates this concept. More specifically, the hierarchy of measurement indicates that the more precise the measurement, the more statistics that can be calculated and used. Thus, the mean, the median, and the mode can be calculated for internal-level data. The median and the mode can be calculated for ordinal-level data. Only the mode can be calculated for nominal-level data.

In the examples in this chapter and in the problems that follow, we use small numbers of cases to ease the burden of calculation of measures of central tendency while still illustrating the crucial concepts and points. In actual situations in public and nonprofit administration, you will typically deal with much larger numbers of cases, and a computer will perform the necessary calculations. The concepts and pointers for proper use and interpretation remain the same, however.

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Problems

5.1 During a recent crackdown on speeding, the Luckenbach, Texas, police department issued the following number of citations on seven consecutive days: 59, 61, 68, 57, 63, 50, and 55. Calculate the mean and median numbers of speeding citations.

5.2 The dean of Southwestern State University is concerned that many faculty members at SSU are too old to be effective teachers. She asks each department to send her information on the average age of its faculty. The head of the sociology department does not wish to "lie" with statistics, but, knowing the preference of the dean, he would like to make the department appear youthful. The names and ages of the sociology department's members are listed in the accompanying table. Calculate both the mean age and the median age. Should the department send the dean the mean age or the median age?

Member	Age
Durkheim	64
Campbell	31
Weber	65
Likert	27
Stanley	35
Katz	40
Lazarsfeld	33

5.3 The average number of sick leave days used per employee per year in Normal, Oklahoma, is 6.7. The Normal city manager feels the Public Works Department is abusing its sick leave privileges. The frequency distribution for the department follows. Calculate the percentage distribution. In what categories do the mean and median of sick leave days fall? Is the Public Works Department abusing its sick leave?

Number of Days of Sick Leave Taken	Number of Employees
0–2	4
3–5	7
6–8	7
9–11	14
12–14	6

5.4 The U.S. Army is allowed only five test firings of the Lance missile. The following figures represent the number of feet by which the missiles missed the target: 26, 147, 35, 63, and 51. Calculate the mean and the median. Which should the army report?

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

5.5 The Department of Welfare wants to know the average outside income for all welfare recipients in the state. Use the data in the accompanying table to present the percentage distribution. In what category does the median fall? In what category does the mode fall?

Income	Number of Families
0–300	25
300–600	163
600–900	354
900–1,200	278
1,200–1,500	421
1,500–1,800	603
1,800–2,100	211
2,100–2,400	84
2,400–2,700	32
2,700–3,000	5

5.6 When should the median be used in preference to the mean?

5.7 The legislature has limited the Bureau of the Audit to a monthly average of 34 employees. For the first 9 months of the year, the employment figures were 31, 36, 34, 35, 37, 32, 36, 37, and 34. Does it appear that the bureau will make the target? How many employees can the bureau have over the next 3 months and still meet the target?

5.8 The collective bargaining agreement between Family Services Agency and the Federation of Social Workers specifies that the average case load for caseworkers cannot exceed 45. Using the accompanying data, the agency claims compliance, yet the union argues that the agency has violated the agreement. Who is correct?

Caseworker	Case Load
A	43
B	57
C	35
D	87
E	36
F	93
G	45
H	48
I	41
J	40

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

5.9 Refer to Problem 5.1. The Luckenbach police chief is concerned that after the crackdown on speeding, the police department has not continued to enforce speed limits in town. She obtains data on speeding citations issued in the past 7 days. The numbers of citations issued over the past 7 days are: 41, 48, 51, 47, 44, 45, and 49. Calculate the mean and median for these data. Based on your analysis, does the police chief appear to be correct?

5.10 The director of the doctor of public administration (DPA) program at Federal University is developing a report to the faculty on the entering class of Ph.D. students. The director wants to present a statistical profile of the new students, including their grade point average (GPA) earned in master's degree programs. The GPAs for the eight entering students are 3.1, 3.7, 3.6, 3.2, 3.8, 3.5, 2.9, and 4.0. Calculate the mean GPA and median GPA earned by these students in their master's degree studies.

5.11 Some faculty at Federal University have complained to the director of the DPA program that Ph.D. students typically have strong verbal skills but lack mathematical preparation. (Fortunately, a statistics book is available in public administration to meet the needs of these students.) In response to the complaint, the Ph.D. director assembles the scores on the verbal and quantitative sections of the Graduate Record Examination (GRE) for the class of eight entering Ph.D. students. Each section is scored on a scale of 200 to 800. The GRE scores of each student are listed below.

GRE Verbal	GRE Quantitative
590	620
680	510
630	550
700	600
610	540
650	570
620	590
670	580

Based on your analysis of these data, evaluate the complaint lodged by the faculty members at Federal University.

5.12 The head of the Data Processing Department in Springhill wants to estimate the amount of waste and inefficiency in her department. She conducts a survey of employees in the department. One question asks, "To what extent does this department have waste and inefficiency in its operations?" The responses to the item are given in the accompanying table.

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Response	Number of Employees
To a very great extent	42
To a great extent	31
To a moderate extent	19
To some extent	12
Not at all	7

(a) At what level of measurement are these data?

(b) Calculate the percentage distribution and the appropriate measures of central tendency.

(c) According to these data, does the department appear to have a problem with waste and inefficiency? Explain your answer.

5.13 The civic center of Kulture City is badly in need of refurbishment. However, before the city council allocates funds for this purpose, its members want to get a better idea of how frequently the residents of Kulture City actually use the center. To find out, they hire a public opinion polling firm to survey citizens. The pollsters ask a random sample of residents the following question: "In the last year, how many times have you attended performances or activities at Kulture City civic center?" The responses of the sample are listed in the accompanying table.

Number of Performances or Activities Attended	Number of Citizens
0	587
1	494
2	260
3	135
4	97

(a) At what level of measurement are these data?

(b) Calculate the appropriate measures of central tendency and the percentage distribution.

(c) Should the city council allocate money to refurbish the civic center?

(d) Write a short memorandum in which you use these results to make a recommendation to the city council.

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

5.14 The head of a city's Recreation Department feels that the employees of the department are the best in the city. Each year, all city employees receive a standardized performance evaluation that rates them on a scale of performance: "far above expectations," "above expectations," "meets expectations," "below expectations," and "far below expectations." (Each rating has an operational definition.) The department head feels that his assessment of employees will be justified if at least 90% of them fall into the top two categories. The ratings received by department employees are shown in the accompanying table.

Rating	Number of Employees
Far above expectations	15
Above expectations	22
Meets expectations	77
Below expectations	9
Far below expectations	8

(a) At what level of measurement are these data?

(b) Calculate the percentage distribution and the appropriate measures of central tendency.

(c) What can you tell the head of the Recreation Department?

5.15 The director of a city's Personnel Office is concerned about racial diversity in the city's workforce. One variable she uses to measure diversity is race. According to records kept by the Personnel Office, the city employs 59 African Americans, 73 whites, 41 Hispanics, 38 Asians, and 17 from other ethnic groups.

(a) At what level of measurement are these data?

(b) Prepare the percentage distribution for these data, and calculate appropriate measures of central tendency.

(c) What can you tell the director of the Personnel Office?

Copyright 2010 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.