



COMPUTING AND UNDERSTANDING AVERAGES

Means to an End

Difficulty Scale 😊😊😊😊
(moderately easy)

WHAT YOU WILL LEARN IN THIS CHAPTER

- Understanding measures of central tendency
- Computing the mean for a sample of scores
- Computing the median for a sample of scores
- Computing the mode for a sample of scores
- Understanding and applying scales or levels of measurement
- Selecting a measure of central tendency

You've been very patient, and now it's finally time to get started working with some real, live data. That's exactly what you'll do in this chapter. Once data are collected, a usual first step is to begin to understand all those bits of information using single numbers to describe the data. The easiest way to do this is through computing an average, of which there are several different types.

An **average** is the one value that best represents an entire group of scores. It doesn't matter whether the group of scores represents the number correct on a spelling test for 30 fifth graders or the typical batting percentage for all the baseball players on the New York Yankees or how voters feel about a congressional candidate. In all of these examples, a big group of data can be summarized using an average. You can usually think of an average as the "middle" space or as a fulcrum on a seesaw. It's the point in a range of values that seems to most fairly represent all the values.

Averages, also called **measures of central tendency**, come in three flavors: the mean, the median, and the mode. Each provides you with a different type of information about a distribution of scores and is simple to compute and interpret.

COMPUTING THE MEAN

The **mean** is the most common type of average that is computed. It is so popular that scientists sometimes sloppily treat the word *average* as if it means *mean* when it only sometimes means *mean*. The mean is simply the sum of all the values in a group, divided by the number of values in that group. So, if you had the spelling scores for 30 fifth graders, you would simply add up all the scores to get a total and then divide by the number of students, which is 30.

We are about to show a formula or equation for the first time in this book. Don't panic. Equations are just statements or sentences that use symbols instead of words. We will always tell you what words the symbols stand for. The formula for computing the mean is shown in Formula 2.1:

$$\bar{X} = \frac{\sum X}{n}, \quad (2.1)$$

where

- the letter X with a line above it (also sometimes called "X bar") is the mean value of the group of scores;
- the \sum , or the Greek letter sigma, is the summation sign, which tells you to "sum up" or add together whatever follows it;
- the X is each individual score in a group of scores; and
- the n is the size of the sample from which you are computing the mean, the number of scores.



To compute the mean, follow these steps:

1. List the entire set of values in one or more columns. These are all the X s.
2. Compute the sum or total of all the values.
3. Divide the total or sum by the number of values.

For example, if you needed to compute the average number of shoppers at three different locations, you would compute a mean for that value.

Location	Number of Shoppers Last Year
Lanham Park Store	2,150
Williamsburg Store	1,534
Downtown Store	3,564

The mean or average number of shoppers in each store is 2,416. Formula 2.2 shows how this average was computed using the formula you saw in Formula 2.1:

$$\bar{X} = \frac{\sum X}{n} = \frac{2,150 + 1,534 + 3,564}{3} = \frac{7,248}{3} = 2,416. \quad (2.2)$$

Or, if you needed to compute the average number of students in each grade in a school building, you would follow the same procedure.

Grade	Number of Students
Kindergarten	18
1	21
2	24
3	23
4	22
5	24
6	25

The mean or average number of students in each class is 22.43. Formula 2.3 shows how this average was computed using the formula you saw in Formula 2.1:

$$\bar{X} = \frac{\sum X}{n} = \frac{18 + 21 + 24 + 23 + 22 + 24 + 25}{7} = \frac{157}{7} = 22.43. \quad (2.3)$$

See, we told you it was easy. No big deal. By the way, when you calculated that mean just now, you may have gotten a number with lots more digits in it: 22.42857143 or something like that. Statisticians are usually okay with you shortening numbers to just a couple digits past the decimal. So, we felt fine reporting the mean as 22.43 (rounding up for that last digit).

- The mean is sometimes represented by the letter M and is also called the typical, average, or most central score. If you are reading another statistics book or a research report and you see something like $M = 45.87$, it probably means that the mean is equal to 45.87. Technically, that capital letter M is used when you are talking about the mean of the larger population represented by the sample in front of you. Those sorts of distinctions aren't important right now but might be interesting later on.

- In the formula, a small n represents the sample size for which the mean is being computed. A large N (← like this) would represent the population size. In some books and in some journal articles, no distinction is made between the two. Notice, as with the capital M when talking about the mean of a population, statistical types often capitalize a letter symbol to refer to a population and keep the letter as lowercase when talking about samples.
- The mean is like the fulcrum on a seesaw. It's the centermost point where all the values on one side of the mean are equal in weight to all the values on the other side of the mean.
- Finally, for better or worse, the mean is very sensitive to extreme scores. An extreme score can pull the mean in one or the other direction and make it less representative of the set of scores and less useful as a measure of central tendency. This, of course, all depends on the values for which the mean is being computed. And, if you have extreme scores and the mean won't work as well as you want, we have a solution! More about that later.

The mean is also referred to as the **arithmetic mean**, and there are other types of means that you may read about, such as the harmonic mean. Those are used in special circumstances and need not concern you here. And if you want to be technical about it, the arithmetic mean (which is the one that we have discussed up to now) is also defined as the point about which the sum of the deviations is equal to zero (whew!). Each score in a sample is some distance from the mean. If you add up all those distances, it will equal zero. Always. Every time. That's why we like the mean. For instance, if you have scores like 3, 4, and 5 (whose mean is 4), the sum of the deviations about the mean (-1 , 0 , and $+1$) is 0 .

Remember that the word *average* means only the one measure that best represents a set of scores and that there are many different types of averages. Which type of average you use depends on the question that you are asking and the type of data that you are trying to summarize. This is a levels of measurement issue that we will cover later in this chapter when we talk about when to use which statistic.

In basic statistics, an important distinction is made between those values associated with samples (a part of a population) and those associated with populations. To do this, statisticians use the following conventions. For a sample statistic (such as the mean of a sample), Roman letters are used. For a population parameter (such as the mean of a population), Greek letters are used. So, for example, the mean for the spelling score for a sample of 100 fifth graders is represented as $\bar{X} = 5$, whereas the mean for the spelling score for the entire population of fifth graders is represented, using the Greek letter μ , as μ_5 .

COMPUTING THE MEDIAN

The median is also an average, but of a very different kind. The **median** is defined as the midpoint in a set of scores. It's the point at which one half, or 50%, of the scores fall above and one half, or 50%, fall below. It's got some special qualities that we will talk about later in this section, but for now, let's concentrate on how it is computed. There's not really a formula for computing the median but instead a set of steps.

To compute the median, follow these steps:

1. List the values in order, from either highest to lowest or lowest to highest.
2. Find the middle-most score. That's the median.



For example, here are the annual incomes from five different households:

\$135,456
\$45,500
\$62,456
\$54,365
\$37,668

Here is the list ordered from highest to lowest:

\$135,456
\$62,456
\$54,365
\$45,500
\$37,668

There are five values. The middle-most value is \$54,365, and that's the median.

Now, what if the number of values is even? An even number of scores means there is no middle value. Let's add a value (\$64,500) to the list so there are six income levels. Here they are sorted with the largest value first:

\$135,456
\$64,500

\$62,456

\$54,365

\$45,500

\$37,668

When there is an even number of values, the median is simply the mean of the two middle values. In this case, the middle two cases are \$54,365 and \$62,456. The mean of those two values is \$58,410.50. That's the median for that set of six values.

What if the two middle-most values are the same, such as in the following set of data?

\$45,678

\$25,567

\$25,567

\$13,234

Then the median is same as both of those middle-most values. In this example, it's \$25,567.

If we had a series of values that was the number of days spent in rehabilitation for a sports-related injury for seven different patients, the numbers might look like this:

43

34

32

12

51

6

27

As we did before, we can order the values (51, 43, 34, 32, 27, 12, 6) and then select the middle value as the median, which in this case is 32. So, the median number of days spent rehabilitating an injury is 32.

If you know about medians, you should also know about **percentile ranks**. Percentile ranks are used to define the percentage of cases equal to or below a certain point in a distribution or set of scores. For example, if a score is "at the

75th percentile,” it means that the score is at or above 75% of the other scores in the distribution. The median is also known as the 50th percentile, because it’s the point at or below which 50% of the cases in the distribution fall. Other percentiles are useful as well, such as the 25th percentile, often called Q_1 , and the 75th percentile, referred to as Q_3 . So what’s Q_2 ? The median, of course.

Here comes the answer to the question you’ve probably had in the back of your mind since we started talking about the median. Why use the median instead of the mean? For one very good reason. The median is insensitive to extreme scores, whereas the mean is not. When you have a set of scores in which one or more scores are extreme, the median better represents the centermost value of that set of scores than any other measure of central tendency. Yes, even better than the mean.

What do we mean by *extreme*? It’s probably easiest to think of an extreme score as one that is very different from the group to which it belongs. For example, consider the list of five incomes that we worked with earlier (shown again here):

\$135,456
 \$54,365
 \$37,668
 \$32,456
 \$25,500

The value \$135,456 is more different from the other four than is any other value in the set. We would consider that an extreme score.

The best way to illustrate how useful the median is as a measure of central tendency is to compute both the mean and the median for a set of data that contains one or more extreme scores and then compare them to see which one best represents the group. Here goes.

The mean of the set of five scores you see above is the sum of the set of five divided by 5, which turns out to be \$57,089. On the other hand, the median for this set of five scores is \$37,668. Which is more representative of the group? The value \$37,668, because it clearly lies closer to most of the scores, and we like to think about “the average” (in this case, we are using the median as a measure of average) as being representative or assuming a central position. In fact, the mean value of \$57,089 falls above the fourth highest value (\$54,365) and is not very central or representative of the distribution.

It’s for this reason that certain social and economic indicators (often involving income) are reported using a median as a measure of central tendency—“The

median income of the average American family is . . .”—rather than using the mean to summarize the values. There are just too many extreme scores that would **skew**, or significantly distort, what is actually a central point in the set or distribution of scores. As an example, the median annual family income in the United States for 2014 (the most recent year we could find data for) was about \$54,000, while the mean annual family income was about \$73,000. Which is closer to your family’s income?

You learned earlier that sometimes the mean is represented by the capital letter M instead of \bar{X} . Well, other symbols are used for the median as well. We like the letter M , but some people confuse it with the mean, so they use *Med* or *Mdn* for median. Don’t let that throw you—just remember what the median is and what it represents, and you’ll have no trouble adapting to different symbols.

Here are some interesting and important things to remember about the median:

- We use the word *median* to describe other middle things, like the median on a highway, that stripe down the middle of a road.
- Because the median is based on how many cases there are, and not the values of those cases, extreme scores (sometimes called **outliers**) only count a little.

COMPUTING THE MODE

The third and last measure of central tendency that we’ll cover, the mode, is the most general and least precise measure of central tendency, but it plays a very important part in understanding the characteristics of a sample of scores. The **mode** is the value that occurs most frequently. Like the median, there is no formula for computing the mode.



To compute the mode, follow these steps:

1. List all the values in a distribution but list each value only once.
2. Tally the number of times that each value occurs.
3. The value that occurs most often is the mode.

For example, an examination of the political party affiliation of 300 people might result in the following distribution of scores:

Party Affiliation	Number or Frequency
Democrats	90
Republicans	70
Independents	140

The mode is the value that occurs most frequently, which in the preceding example is Independents. That's the mode for this distribution.

If we were looking at the modal response on a 100-item multiple-choice test, we might find that Alternative A was chosen more frequently than any other. The data might look like this:

Item Alternative Selected	A	B	C	D
Number of times	57	20	12	11

On this 100-item multiple-choice test where each item has four choices (A, B, C, and D), A was the answer selected 57 times. It's the modal response.

Want to know the easiest and most commonly made mistake made when computing the mode? It's selecting the number of *times* a category occurs rather than the *label* of the category itself. Instead of the mode being Independents (in our first example), it's easy for someone to conclude the mode is 140. Why? Because he or she is looking at the number of times the value occurred, not the value that occurred most often! This is a simple mistake to make, so be on your toes when you are asked about these things.

Apple Pie à la Bimodal

If every value in a distribution contains the same number of occurrences, then there really isn't a single mode. But if more than one value appears with equal frequency, the distribution is multimodal. The set of scores can be bimodal (with two modes), as the following set of data using hair color illustrates:

Hair Color	Number or Frequency
Red	7
Blond	12
Black	45
Brown	45

In the previous example, the distribution is bimodal because the frequency of the values of black and brown hair occurs equally. You can even have a bimodal distribution when the modes are relatively close together but not exactly the same, such as 45 people with black hair and 44 with brown hair. The question becomes, How much does one class of occurrences stand apart from another?

Can you have a trimodal distribution? Sure—where three values have the same frequency. It's unlikely, especially when you are dealing with a large set of **data points**, or observations, but certainly possible. The real answer to the above stand-apart question is that categories have to be mutually exclusive—you simply cannot have both black and red hair (although if you look around the classroom, you may think differently). Of course, you can have those two colors, but each person's hair color is forced into only one category.

WHEN TO USE WHAT MEASURE OF CENTRAL TENDENCY (AND ALL YOU NEED TO KNOW ABOUT SCALES OF MEASUREMENT FOR NOW)

Which measure of central tendency you use depends on certain characteristics of the data you are working with—specifically the **scale of measurement** at which those data occur. And that scale or level dictates the specific measure of central tendency you will use.

But let's step back for just a minute and make sure that we have some vocabulary straight, beginning with the idea of what measurement is.

Measurement is the assignment of values to outcomes following a set of rules—simple. The results are the different scales we'll define in a moment, and an outcome is anything we are interested in measuring, such as hair color, gender, test score, or height.

These scales of measurement, or rules, are the particular levels at which outcomes are observed. Each level has a particular set of characteristics, and scales of measurement come in four flavors (there are four types): nominal, ordinal, interval, and ratio.

Let's move on to a brief discussion and examples of the four scales of measurement and then discuss how these levels of scales fit with the different measures of central tendency discussed earlier.

A Rose by Any Other Name: The Nominal Level of Measurement

The **nominal level of measurement** is defined by the characteristics of an outcome that fit into one and only one class or category. For example, gender can

be a nominal variable (female and male), as can ethnicity (Caucasian or African American), as can political affiliation (Republican, Democrat, or Independent). Nominal-level variables are “names” (*nominal* in Latin), and the nominal level is the least precise level of measurement. Nominal levels of measurement have categories that are mutually exclusive; for example, political affiliation cannot be both Republican and Democrat.

Any Order Is Fine With Me: The Ordinal Level of Measurement

The *ord* in **ordinal level of measurement** stands for *order*, and the characteristic of things being measured here is that they are ordered. The perfect example is a rank of candidates for a job. If we know that Russ is ranked 1, Marquis is ranked 2, and Hannah is ranked 3, then this is an ordinal arrangement. We have no idea how much higher on this scale Russ is relative to Marquis than Marquis is relative to Hannah. We just know that it’s “better” to be 1 than 2 than 3 but not by how much.

1 + 1 = 2: The Interval Level of Measurement

Now we’re getting somewhere. When we talk about the **interval level of measurement**, a test or an assessment tool is based on some underlying continuum such that we can talk about how much more a higher performance is than a lesser one. For example, if you get 10 words correct on a vocabulary test, that is 5 more than getting 5 words correct. A distinguishing characteristic of interval-level scales is that the intervals or spaces or points along the scale are equal to one another. Ten words correct is 2 more than 8 correct, and 6 is 2 more than 4 correct. And it’s not just the actual math that determines whether there are equal intervals; it is assumptions about the underlying concept being measured. If two more spelling words correct always represents two more units of spelling ability (whatever that means), then the measurement is at the interval level.

Can Anyone Have Nothing of Anything? The Ratio Level of Measurement

Well, here’s a little conundrum for you. An assessment tool at the **ratio level of measurement** is characterized by the presence of an absolute zero on the scale. What that zero means is the absence of any of the trait that is being measured. The conundrum? Are there outcomes we measure where it is possible to have nothing of what is being measured? In some disciplines, that can be the case. For example, in the physical and biological sciences, you can have the absence of a characteristic, such as absolute zero (no molecular movement) or zero light. In the social and behavioral sciences, it’s a bit harder. Even if you score zero on that spelling test or miss every item of an intelligence test (written in Martian), that does not mean that you have no spelling ability or no intelligence, right?

In Sum . . .

These scales of measurement, or rules, represent particular levels at which outcomes are measured. And, in sum, we can say the following:

- Any outcome can be assigned to one of four levels of measurement.
- Levels of measurement have an order, from the least precise being nominal to the most precise being ratio.
- The “higher up” the scale of measurement, the more precise the data being collected, and the more detailed and informative the data are. It may be enough to know that some people are rich and some poor (and that’s a nominal or categorical distinction), but it’s much better to know exactly how much money they make (ratio). We can always make the “rich” versus “poor” distinction if we want to once we have all the information.
- Finally, the more precise scales contain all the qualities of the scales below them; for example, the interval scale includes the characteristics of the ordinal and nominal scales. If you know that the Cubs’ batting average is .350, you know it is better than that of the Tigers (who hit .250) by 100 points, but you also know that the Cubs are better than the Tigers (but not by how much) and that the Cubs are different from the Tigers (but there’s no direction to the difference).

Okay, we’ve defined levels of measurement and discussed three different measures of central tendency and given you good examples of each. But the most important question remains unanswered. That is, “When do you use which measure?”

In general, which measure of central tendency you use depends on the type of data that you are describing, which in turn means at what level of measurement the data occur. Unquestionably, a measure of central tendency for qualitative, categorical, or nominal data (such as racial group, eye color, income bracket, voting preference, and neighborhood location) can be described using only the mode. It’s not meaningful to talk about the mean eye color in a classroom.

For example, you can’t be interested in the most central measure that describes which political affiliation is most predominant in a group and use the mean—what in the world could you conclude, that everyone is half Republican? Rather, saying that out of 300 people, almost half (140) are Independent seems to be the best way to describe the value of this variable. In general, the median and the mean are best used with quantitative data, such as height, income level in dollars (not categories), age, test score, reaction time, and number of hours completed toward a degree.

It’s also fair to say that the mean is a more precise measure than the median and that the median is a more precise measure than the mode. The best strategy, then, if you are measuring at a high level, such as interval or ratio, is to use the mean, and

indeed, the mean is the most often used measure of central tendency. However, we do have occasions when the mean would not be appropriate as a measure of central tendency—for example, when we have categorical or nominal data, such as hospitalized versus nonhospitalized people. Then we use the mode.

So, here is a set of three guidelines that may be of some help. And remember, there can always be exceptions.

1. Use the mode when the data are categorical (“nominal”) in nature and the people or things can fit into only one class, such as hair color, political affiliation, neighborhood location, and religion. When this is the case, these categories are called mutually exclusive.
2. Use the median when you have extreme scores and you don’t want an average that is misleading, such as when the variable of interest is income expressed in dollars.
3. Finally, use the mean when you have data that do not include extreme scores and are not categorical, such as the numerical score on a test or the number of seconds it takes to swim 50 meters.

USING SPSS TO COMPUTE DESCRIPTIVE STATISTICS

If you haven’t already, now would be a good time to check out Appendix A so you can become familiar with the basics of using SPSS. Then come back here.

Let’s use SPSS to compute some descriptive statistics. The data set we are using is named Chapter 2 Data Set 1, and it is a set of 20 scores on a test of prejudice. All of the data sets are available in Appendix C and from the SAGE website edge.sagepub.com/salkindfrey7e. There is one variable in this data set:

Variable	Definition
Prejudice	The value on a test of prejudice as measured on a scale from 1 to 100

Here are the steps to compute the measures of central tendency that we discussed in this chapter. Follow along and do it yourself. With this and all exercises, including data that you enter or download, we’ll assume that the data set is already open in SPSS.

1. Click Analyze → Descriptive Statistics → Frequencies . . .
2. Double-click on the variable named Prejudice to move it to the Variable(s) box.

(Continued)



(Continued)

3. Click Statistics, and you will see the Frequencies: Statistics dialog box shown in Figure 2.1.
4. Under Central Tendency, click the Mean, Median, and Mode boxes.
5. Click Continue.
6. Click OK.

The SPSS Output

Figure 2.2 shows you selected output from the SPSS procedure for the variable named Prejudice.

In the Statistics part of the output, you can see how the mean, the median, and the mode are all computed along with the sample size and the fact that there were no missing data. SPSS does not use symbols such as X in its output. Also listed in the output are the frequency of each value and the percentage of times it occurs, all useful descriptive information.

It's a bit strange, but if you select Analyze → Descriptive Statistics → Descriptives (instead of clicking Frequencies last) in SPSS and then click Options, there's no median or mode option, which you might expect because they are basic descriptive statistics. The lesson here? Statistical analysis programs are usually quite different from one another, use different names for the same things, and make different assumptions about what's where. If you can't find what you want, it's probably there. Just keep hunting. Also, be sure to use the Help feature to get help navigating through all this new information until you find what you need.

FIGURE 2.1 The Frequencies: Statistics dialog box from SPSS

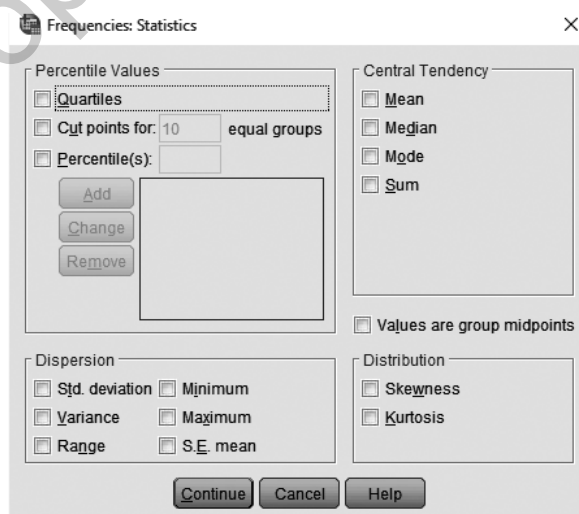


FIGURE 2.2 ● Descriptive statistics from SPSS→ **Frequencies**

Statistics

Prejudice

N	Valid	20
	Missing	0
Mean		84.70
Median		87.00
Mode		87

Prejudice

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 55	1	5.0	5.0	5.0
64	1	5.0	5.0	10.0
67	1	5.0	5.0	15.0
76	1	5.0	5.0	20.0
77	1	5.0	5.0	25.0
81	2	10.0	10.0	35.0
82	1	5.0	5.0	40.0
87	4	20.0	20.0	60.0
89	1	5.0	5.0	65.0
93	1	5.0	5.0	70.0
94	2	10.0	10.0	80.0
96	1	5.0	5.0	85.0
99	3	15.0	15.0	100.0
Total	20	100.0	100.0	

Understanding the SPSS Output

This SPSS output is pretty straightforward and easy to interpret.

The average mean score for the 20 scores is 84.70 (and remember that the span of possible scores is from 0 to 100). The median, or the point at which 50% of the scores fall above and 50% fall below, is 87 (which is pretty close to the mean), and the most frequently occurring score, or the mode, is 87.

SPSS output can be full of information or just give you the basics. It all depends on the type of analysis that you are conducting. In the preceding example, we have just the basics and, frankly, just what we need. Throughout this book, you will be seeing output and then learning about what it means, but in some cases, discussing the entire collection of output information is far beyond the scope of the book. We focus on output that is directly related to what you learned in the chapter.

REAL-WORLD STATS

The Pew Research Center reported a recent study of who uses the big social media sites, like Facebook, Instagram, LinkedIn, and Twitter. About 70% of all Americans use one or more of these sites! Women tend to use Facebook and Instagram more than men do, but there wasn't really a gender difference for LinkedIn

or Twitter. The researchers also looked at the relationship between age and the percentage of people who used each site. In the following table, notice that, as you might suspect, younger adults use social networks more than older adults.

Age	Facebook	Instagram	LinkedIn	Twitter
18 to 29	81%	64%	29%	40%
30 to 49	78%	40%	33%	27%
50 to 64	65%	21%	24%	19%
65 and over	41%	10%	9%	8%

Want to know more? Go online and read about these and other findings at

<https://www.pewinternet.org/fact-sheet/social-media/>.

Summary

No matter how fancy-schmancy your statistical techniques are, you will almost always start by simply describing what's there—hence the importance of understanding the simple notion of central tendency. From here, we go to another important descriptive characteristic: variability, or how different scores are from one another. That's what we'll explore in Chapter 3!

Time to Practice

1. By hand (which means using a calculator if you'd like), compute the mean, the median, and the mode for the following set of 40 chemistry final scores.

93	85	99	77
94	99	86	76
95	99	97	84
91	89	77	87

97	83	80	98
75	94	81	85
78	92	89	94
76	94	96	94
90	79	80	92
77	86	83	81

2. Compute the mean, the median, and the mode for the following three sets of scores saved as Chapter 2 Data Set 2. Do it by hand or use a computer program such as SPSS. Show your work, and if you use SPSS, print out a copy of the output.

Score 1	Score 2	Score 3
3	34	154
7	54	167
5	17	132
4	26	145
5	34	154
6	25	145
7	14	113
8	24	156
6	25	154
5	23	123

3. Compute the means for the following set of scores saved as Chapter 2 Data Set 3 using SPSS. Print out a copy of the output.

Number of Beds (infection rate)	Infection Rate (per 1,000 admissions)
234	1.7
214	2.4
165	3.1
436	5.6
432	4.9
342	5.3

(Continued)

(Continued)

Number of Beds (infection rate)	Infection Rate (per 1,000 admissions)
276	5.6
187	1.2
512	3.3
553	4.1

4. You are the manager of a fast-food restaurant. Part of your job is to report to the boss at the end of each day which item is selling best. Use your vast knowledge of descriptive statistics and write one paragraph to let the boss know what happened today. Here are the data. Don't use SPSS to compute important values; rather, do it by hand. Be sure to include a copy of your work.

Special	Number Sold	Cost
Huge Burger	20	\$2.95
Baby Burger	18	\$1.49
Chicken Littles	25	\$3.50
Porker Burger	19	\$2.95
Yummy Burger	17	\$1.99
Coney Dog	20	\$1.99
Total specials sold	119	

5. Imagine you are the CEO of a small group of clothing stores and you are planning an expansion. You'd like your new store to post similar numbers as the other three that are in your empire. By hand, provide some idea of what you want the store's financial performance to look like. And remember that you have to select whether to use the mean, the median, or the mode as an average. Good luck, young Jedi.

Average	Store 1	Store 2	Store 3	New Store
Sales (in thousands of dollars)	323.6	234.6	308.3	
Number of items purchased	3,454	5,645	4,565	
Number of visitors	4,534	6,765	6,654	

6. Here are ratings (on a scale from 1 through 5) for various Super Bowl party foods. You have to decide which food is rated highest (5 is a winner and 1 a loser). Decide what type of average you will use and why. Do this by hand or use SPSS.

Snack Food	North Fans	East Fans	South Fans	West Fans
Loaded Nachos	4	4	5	4
Fruit Cup	2	1	2	1
Spicy Wings	4	3	3	3
Gargantuan Overstuffed Pizza	3	4	4	5
Beer Chicken	5	5	5	4

- Under what conditions would you use the median rather than the mean as a measure of central tendency? Why? Provide an example of two situations in which the median might be more useful than the mean as a measure of central tendency.
- Suppose you are working with a data set that has some very “different” (much larger or much smaller than the rest of the data) scores. What measure of central tendency would you use and why?
- For this exercise, use the following set of 16 scores (ranked) that consists of income levels ranging from about \$50,000 to about \$200,000. What is the best measure of central tendency, and why?

\$199,999

\$98,789

\$90,878

\$87,678

\$87,245

\$83,675

\$77,876

\$77,743

\$76,564

\$76,465

\$75,643

\$66,768

\$65,654

\$58,768

\$54,678

\$51,354

- Use the data in Chapter 2 Data Set 4 and, manually, compute the average attitude scores (with a score of 10 meaning a very positive attitude and a 1 meaning a very negative attitude) for three groups of individuals' attitudes reflecting their experience with urban transportation. Choose the mean as the average to use.

(Continued)

(Continued)

11. Take a look at the following number of pie orders from the Lady Bird diner and determine the average number of orders for each week.

Week	Chocolate Silk	Apple	Douglas County Pie
1	12	21	7
2	14	15	12
3	18	14	21
4	27	12	15

Student Study Site



Get the tools you need to sharpen your study skills! Visit edge.sagepub.com/salkindfrey7e to access practice quizzes, eFlashcards, original and curated videos, data sets, and more!