

a fairly sharp peak. The variation is higher when the data are distributed more widely around the center, which makes the peak broader. Figure 6.5b shows a distribution with moderate variation and Figure 6.5c shows a distribution with high variation. We'll discuss methods for quantitatively describing variation in Unit 6B.

DEFINITION

Variation describes how widely data values are spread out about the center of a distribution.

Example 7 Variation in Marathon Times

How would you expect the variation to differ between times in the Olympic marathon and times in the New York marathon? Explain.

Solution The Olympic marathon invites only elite runners, whose times are likely to be clustered relatively near world record times. The New York marathon allows runners of all abilities, whose times are spread over a very wide range (from near the world record to many hours). Therefore, the variation among the times should be greater in the New York marathon than in the Olympic marathon.

Now try Exercises 27–34, part c.

QUICK QUIZ

Choose the best answer to each of the following questions. Explain your reasoning with one or more complete sentences.

- You want to find the *mean* weight of the apples in a barrel. What do you need to do?
 - Lay the apples out in order of increasing weight and find the weight of the apple in the middle.
 - Find the total weight of all the apples and divide by the number of apples.
 - Divide the number of apples in the barrel by the weight of the barrel.
- On a math exam, one student scores 79 while 25 students score below 79 and 25 students score above 79. The score of 79 therefore represents the
 - mean.
 - median.
 - mode.
- One hundred students take a chemistry exam. All but two of the students score between 50 and 80 points, but one student gets a 31 and one student gets a 98. The scores of 31 and 98 represent
 - mean scores.
 - skewed scores.
 - outliers.
- Twenty students take a political science exam. Eighteen score between 70 and 75, and two score 100. What can you conclude?
 - The mean is higher than the median.
 - The median is higher than the mean.
 - The median is higher than the mode.
- A survey asks students at your school to state how many sodas they drink per week. The results show a mean of 12 sodas per week and a median of 8 sodas per week. What can you conclude?
 - Something must have been done wrong in computing the mean and median.
 - Most students drink between 8 and 12 sodas per week.
 - At least one student drinks more than 16 sodas per week.
- Among professional actors, a small number of superstars earn much more money than most other actors. A distribution of actor salaries is therefore
 - symmetric.
 - left-skewed, with outliers at low values.
 - right-skewed, with outliers at high values.
- The distribution of wages at the company you work for is right-skewed, with outliers at high values. Assuming you would like a high wage, you would hope that your wage was closer to the
 - mean.
 - median.
 - mode.
- Compared to a distribution with a broad central peak, a distribution with a sharp central peak
 - has low variation.
 - has high variation.
 - is symmetric.

9. Consider the time it takes to drive from an outlying suburb to downtown. At which time would you expect the driving time to have the highest variation?
- at 3:00 in the morning
 - at 7:00 Sunday morning
 - during rush hour
10. The mayor of a town is considering a run for governor. She conducts a poll asking registered voters to rate their likelihood of voting for her on a scale of 1 to 5, where 1 means "definitely would not vote for her" and 5 means "definitely would vote for her." The most encouraging result would be
- low median, high variation.
 - high median, low variation.
 - high median, high variation.

Exercises

REVIEW QUESTIONS

- Define and distinguish among mean, median, and mode.
- What are outliers? Describe the effects of outliers on the mean, median, and mode.
- Briefly describe at least two possible sources of confusion about the "average."
- Give simple examples of a single-peaked distribution and a double-peaked distribution.
- What do we mean when we say that a distribution is symmetric? Give simple examples of a symmetric distribution, a left-skewed distribution, and a right-skewed distribution.
- What do we mean by the variation of a distribution? Give simple examples of distributions with different amounts of variation.

DOES IT MAKE SENSE?

Decide whether each of the following statements makes sense (or is clearly true) or does not make sense (or is clearly false). Explain your reasoning.

- In my data set of 10 exam scores, the mean turned out to be the score of the person with the third highest grade. No two people got the same score.
- In my data set of 10 exam scores, the median turned out to be the score of the person with the third highest grade. No two people got the same score.
- I made a distribution of 15 apartment rents in my neighborhood. One apartment had a much higher rent than all of the others, and this outlier caused the mean rent to be higher than the median rent.
- If management and employees use the same data and do the calculations properly, they will always agree on the average wage.
- The distribution of grades was left-skewed, but the mean, median, and mode were all the same.
- There's much more variation in the ages of the general population than in the ages of students in my college extension course, but both turn out to have the same mean.

BASIC SKILLS & CONCEPTS

13–18: Mean, Median, and Mode. Compute the mean, median, and mode of the following data sets.

13. Weights (in grams) of copper-nickel quarters made after 1965:
- 5.67 5.72 5.63 5.70 5.73 5.64

14. Body temperature (in degrees Fahrenheit) of randomly selected normal and healthy adults:

98.6 98.6 98.0 98.0 99.0
98.4 98.4 98.4 98.4 98.6

15. Total points scored in the games of an NBA (basketball) play-off series:

201 195 232 176 189 215 200

16. Golf scores from a recent WPGA tournament:

69 70 70 71 73 73 75 76 76 78
80 81 81 82 83 84 85 86 87 88

17. Rainfall amounts (in inches) for six consecutive months in a particular city:

2.5 0.8 1.3 1.2 0.4 0.9

18. Weights (in grams) of randomly selected M&M plain candies:

0.957 0.912 0.842 0.925 0.939 0.886
0.914 0.913 0.958 0.947 0.920

19. Outlier Coke. Cans of Coca-Cola vary slightly in weight. Here are the measured weights of seven cans, in pounds:

0.8161 0.8194 0.8165 0.8176
0.7901 0.8143 0.8126

Find the mean and median of these weights. Which, if any, of these weights would you consider to be an outlier? What are the mean and median weights if the outlier is excluded?

20. Margin of Victory. The data set below gives the margin of victory in the NFL Superbowl games for 2002–2009.

3 12 11 3 3 27 3 27

- Find the mean and median margin of victory.
- Identify the outliers in the set. If you eliminate the outliers on the high side, what are the new mean and median?

21–26: Appropriate Average. State, with an explanation, whether the mean, median, or mode gives the best description of the following averages.

- The average household income in Missouri
- The average age at first marriage for men in America
- The average number of times that people change jobs during their careers
- The average number of pieces of lost luggage per flight from an airline company's perspective.
- The average weight of potatoes in a 10-pound bag