

Lesson

1-7

Data and Spread

► **BIG IDEA** Two measures of the spread of a data set are *range* and *mean absolute deviation*.

In this chapter you have investigated number sequences that fit a pattern exactly and are described by an expression. However, many collections of numbers are “messy” and do not have an exact algebraic description. Statistics can help analyze and summarize these kinds of data.

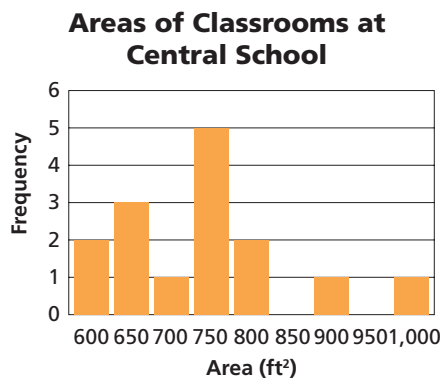
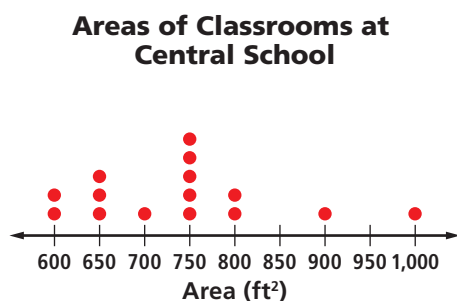
Picturing the Mean of a Data Set

Suppose that a school board has received complaints that classrooms at Central School are small and crowded. To decide if this is the case, data about room size were gathered. Areas of 15 rooms were calculated in square feet, then rounded to the nearest 50 square feet. Here are the areas they found.

750; 700; 800; 750; 650; 800; 750; 750; 650; 750; 600; 900; 650; 600; 1,000

You can represent this data in a frequency table at the right.

Two other ways are in a *dot plot* (below at the left) and a *bar graph* (below at the right).



Vocabulary

range
mean absolute deviation
symmetric
skewed right
skewed left
uniform

Mental Math

- What is the mean of 2, 3, and 4?
- What is the mean of -2, -3, and -4?
- What is the mean of -1, 0, and 1?
- What is the mean of $x - 1$, x , and $x + 1$?

Room Area (ft ²)	Frequency
600	2
650	3
700	1
750	5
800	2
900	1
1,000	1

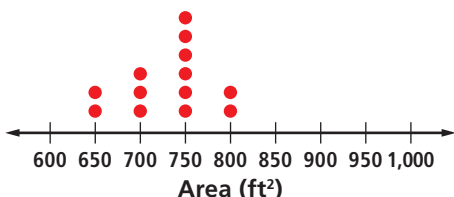
It is customary to use the Greek letter μ (*mu*, pronounced “mew”) to stand for the mean. The mean classroom area is shown below.

$$\begin{aligned}\mu &= \frac{2 \cdot 600 + 3 \cdot 650 + 1 \cdot 700 + 5 \cdot 750 + 2 \cdot 800 + 1 \cdot 900 + 1 \cdot 1,000}{15} \\ &= 740 \text{ ft}^2\end{aligned}$$

The school board at Central decided to compare the areas of classrooms at Central with those of Whittier School. Here is the list of classroom areas and a dot plot for Whittier.

750; 750; 750; 700; 800; 750; 700; 750; 800; 650; 700; 650; 750

**Areas of Classrooms at
Whittier School**



The Range of a Data Set

The mean of the classroom areas for Whittier is about 730 square feet, which is less than the mean of 740 square feet for Central. But notice that the room areas at Central are more spread out than those at Whittier. The data values for Whittier lie in the interval $650 \leq A \leq 800$, where A is the area of a room. For Central, A is in the interval $600 \leq A \leq 1,000$. The **range** r of a collection of data is the difference between the maximum value M and minimum value m , so $r = M - m$. The range of classroom areas for Whittier is $800 - 650$, or 150 square feet.

STOP See Quiz Yourself 1 at the right.

▶ QUIZ YOURSELF 1

What is the range of classroom areas for Central School?

The Mean Absolute Deviation of a Data Set

The range is a measure of the spread of a data set. Another measure of spread is the average difference between the areas of the rooms and the mean area. This is the **mean absolute deviation** of the data set.

Example

Use the information about the classroom areas at Central School on page 47 to find the mean absolute deviation.

Solution We show how to find the mean absolute deviation for the classroom areas at Central School using spreadsheets or lists, although the same computations can be done by hand or on a calculator without lists. See the table on the next page.

Step	Spreadsheet	Graphing Calculator
1. Organize the data.	After naming column A in cell A2, enter the 15 room areas in cells A3 through A17.	Enter the 15 room areas into L1, the first list on your calculator.
2. Calculate the mean.	As shown below, it is 740 square feet. Find the mean from cells B3 through B17.	Keep the mean in mind.
3. Calculate the deviation of each area from the mean.	Enter " $=A3-B3$ " in cell C3. Copy down to C17.	Enter L1-740 in place of L2 at the top of the screen. L2 will contain the deviations.
4. Calculate the mean of the deviations.	In cell C19 use the formula " $=average(C3:C17)$ ". The mean of the deviations should be 0 because the positive and negative differences balance.	Leave the LIST screen to calculate mean (L2).
5. Put the absolute value of each deviation in column D.	Enter " $=abs(C3)$ " in cell D3 and copy the formula down the rest of the column.	Return to LIST. Enter abs (L2) for L3.
6. Calculate the mean absolute deviation.	In cell D19 enter " $=average(D3:D17)$ ". The mean of the absolute deviations is approximately 78.7.	Leave LIST again. Calculate mean (L3).

	A	B	C	D
1	Central School			
2	classroom area	mean area	Deviation (area - mean area)	Absolute Deviation (area - mean area)
3	750	740	10	10
4	700	740	-40	40
5	800	740	60	60
6	750	740	10	10
7	650	740	-90	90
8	800	740	60	60
9	750	740	10	10
10	750	740	10	10
11	650	740	-90	90
12	750	740	10	10
13	600	740	-140	140
14	900	740	160	160
15	650	740	-90	90
16	600	740	-140	140
17	1000	740	260	260
18				
19	740		0	78.66667
20				

L1	L2	L3
750	10	
700	-40	
800	60	
750	10	
650	-90	
800	60	
750	10	

L2 = L1 - 740

L1	L2	L3
750	10	10
700	-40	40
800	60	60
750	10	10
650	-90	90
800	60	60
750	10	10

L3 = abs(L2)

mean(L2)
0

mean(L3)
78.66666667

STOP See Quiz Yourself 2 at the right.

An abbreviation for mean absolute deviation is m.a.d. You should have found that the m.a.d. for Whittier is much smaller than the m.a.d. for Central. The rooms at Central are larger, on average, than those at Whittier. The range of the rooms at Central is also larger than the range of the rooms at Whittier. Do you see what is probably causing the complaints that the classrooms at Central are too small?

An Algorithm for Finding the Mean Absolute Deviation

As its name indicates, the mean absolute deviation is found using the following three steps:

Step 1 Find the mean, μ , of the data values.

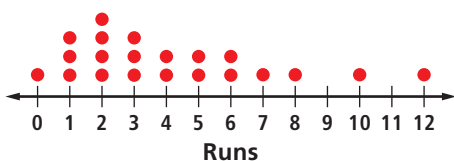
Step 2 For each data value v , find the absolute value of its deviation from the mean, which is represented by $|v - \mu|$.

Step 3 Take the mean of the absolute deviations.

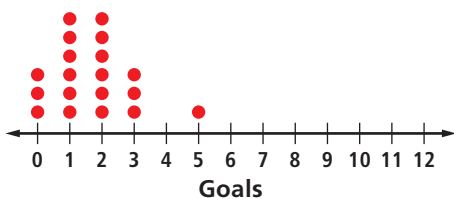
The Shape of a Distribution

Here are dot plots of the number of runs scored by Westview's softball team and the number of goals scored by its soccer team.

**Softball Runs Westview
Scored in Games**



**Soccer Goals Westview
Scored in Games**



▶ QUIZ YOURSELF 2

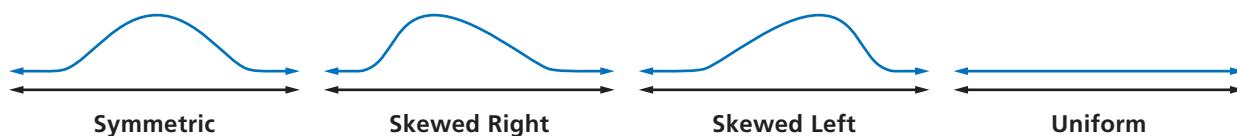
Calculate the mean absolute deviation for the areas of the classrooms at Whittier on page 48.



There was a 67.3% increase in the number of female high school softball players from the 1980–1981 season to the 2000–2001 season.

Source: profaspitch.com

Some shapes of distributions are given special names. When most of the data are centered around one point and the values on the left and right sides are roughly mirror images, the distribution is called **symmetric**. The distribution of goals scored by the soccer team is symmetric. When the upper half of the values extends much farther to the right than the lower half, leaving a tail on the right, the shape is said to be **skewed right**. The softball team's run distribution has this kind of shape because most games are relatively low-scoring, but a few have high scores. Likewise, when the lower half of the data is much farther out, leaving a tail to the left, the shape is said to be **skewed left**. If the distribution has roughly the same height for all values it is called **uniform**.



Questions

COVERING THE IDEAS

1. Calculate the mean absolute deviation of the runs scored by Westview's softball team on page 50.
2. Calculate the m.a.d. of the goals scored by Westview's soccer team on page 50.
3. The table below gives the quiz scores of a group of students. Copy the table and fill in the columns. Then find the mean absolute deviation of the quiz scores.

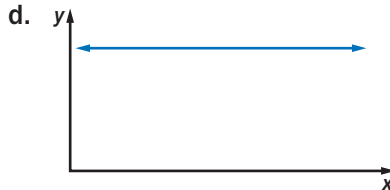
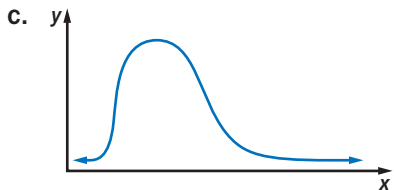
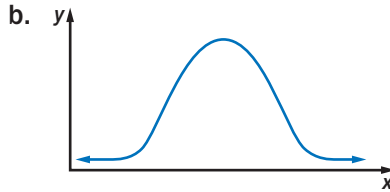
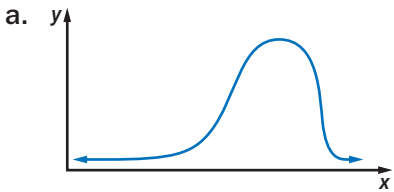
Score	Mean	Deviation	Absolute Deviation
21	?	?	?
15	?	?	?
25	?	?	?
22	?	?	?
27	?	?	?

APPLYING THE MATHEMATICS

In 4 and 5, find the mean and mean absolute deviation in your head.

4. 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7
5. 2, 2, 2, 2, 2, 2, 4, 4, 4, 4, 4, 4

For 6–9, use the following sketches.



6. Which distribution is symmetric?
7. Which distribution is skewed left?
8. Which distribution is skewed right?
9. Which of the distributions most closely describes the scores of all baseball games of one team in a season?
10. In this lesson, a story is told in which people at Central School complain about the classrooms being too small. Whittier School has a mean classroom size that is lower than Central's, but there are no complaints at Whittier. What seems to be the cause of the complaints at Central?
11. These two data sets give the lengths of terms of Presidents of the United States and the reigns of kings and queens of England.

Years served by the first 42 American presidents: 8, 4, 8, 8, 8, 4, 8, 4, 0, 4, 4, 1, 3, 4, 4, 4, 4, 8, 4, 1, 3, 4, 4, 4, 5, 7, 4, 8, 2, 6, 4, 12, 8, 8, 3, 5, 6, 2, 4, 8, 4, 8

Years reigned by 39 English rulers: 21, 13, 35, 19, 35, 10, 17, 56, 35, 20, 50, 22, 13, 9, 39, 22, 0, 24, 38, 6, 5, 44, 22, 24, 25, 3, 13, 6, 12, 13, 33, 59, 10, 7, 63, 9, 25, 1, 15

 - a. Use a spreadsheet or graphing calculator to calculate the mean, mean absolute deviation, and range of each collection of data.
 - b. Write a few sentences explaining why the two distributions have such different values for their means, mean absolute deviations, and ranges.
12.
 - a. Construct a data set of 8 items that has a mean absolute deviation of 0.5 and for which $\mu = 8$.
 - b. What are the M , m , and r for this data set?



The first Oval Office was built in 1909 in the center of the south side of the West Wing of the White House.

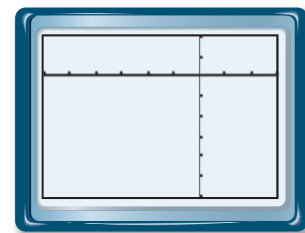
Source: www.whitehouse.gov

13. Construct a data set of 7 items that has a median of 9 and a mean of 8.
14. Construct a data set of 12 items that has a mode of 6, and for which $\mu = 9$.
15. The mean absolute deviation can be used to measure consistency. The more consistent data set is the one with the smaller m.a.d. Here are the mean maximum daily temperatures for each month in San Diego and Miami. In which city are the year's high temperatures more consistent?

Month	Miami (°C)	San Diego (°C)
January	24.0	18.8
February	24.7	19.2
March	26.2	19.1
April	28.0	20.2
May	29.6	20.6
June	30.9	22.0
July	31.7	24.6
August	31.7	25.4
September	31.0	25.1
October	29.2	23.7
November	26.9	21.1
December	24.8	18.9

REVIEW

16. Evaluate each expression. (Lesson 1-6)
- $|3 - |-3 + 9||$
 - $|2x - 10|$ when $x = 3.5$
 - $-4|x - y|$ when $x = 0.5$ and $y = 3$
17. Use the window at the right. The tick marks on the x -axis occur every 20 units. On the y -axis, tick marks occur every 0.25 unit. Find the values of each. (Lesson 1-5)
- | | | | |
|----------|---|----------|---|
| x -min | ? | x -max | ? |
| y -min | ? | y -max | ? |
18. On a graphing calculator, graph $y = x^3 - 5x$ using the window $-10 \leq x \leq 10$ and $-10 \leq y \leq 10$. Sketch the graph that results. (Lesson 1-5)
19. Make a table and scatterplot for the ordered pair $(n, n^2 - n)$ using the following values of n : $-3, -2, 0, 2, \text{ and } 3$. (Lesson 1-4)

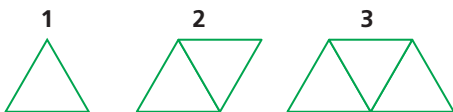


20. Complete the table of values for each expression. Use the tables to determine whether the expressions appear to be equivalent. Explain your reasoning. (Lesson 1-3)

n	$(n + 2) + 7 + (n + 2)$
-5	?
-3	?
1	?
2	?
5	?

n	$3n + 3 - n + 8$
-5	?
-3	?
1	?
2	?
5	?

21. Consider the sequence of triangles. Assume the designs are made with toothpicks. (Lesson 1-2)



- Draw a table describing the number of toothpicks needed for the first 5 terms.
- Write an expression for the number of toothpicks in the n th term.
- How many toothpicks will be needed to make the 100th term?

EXPLORATION

22. Choose the data set that would probably have the smaller m.a.d. and explain why.
- incomes of 10 union members from a factory; incomes of 10 management members from the same factory
 - temperatures in one place on the moon for a month; temperatures in one place on Earth for a month
 - recovery times for 6 people from an appendectomy; recovery times for 6 basketball players from knee operations



Two unionized workers inspecting air-bag canisters.

QUIZ YOURSELF ANSWERS

- $r = 1,000 - 600 = 400 \text{ ft}^2$
- 39