12

Working With Data

The ability to interpret data is an important skill in society.

Each of the following statements involves data.
- Jerry has a size 12 foot. Mark, Vince, and Peter all have size 10 feet.
- Kim had an average of 98% in math this term. The class average was 72%.
- A group of students with summer jobs earned between $100 and $1500 each last summer.
- In a snowboarding Big Air competition, Sean finished eighth. Five judges gave him the following scores out of ten: 8.2, 7.5, 7.0, 7.2, and 8.0.

In this chapter, you will learn ways of determining and reporting on the centre of a set of data.

What You Will Learn

❏ to determine the centre of a set of data in several ways
❏ to identify unusual values in a set of data
❏ to decide when an unusual value should be included when measuring the centre of a set of data
❏ to choose the best measure of the centre of a set of data

Key Words

measure of
  central tendency
median
mode
mean
range
outlier

MATH LINK

Think of situations in which you might want to gather data.
For example,
- What size of running shoes do people wear?
- How well did the class do on the last math test?
- How well do players score in various sports competitions?
- How much do teenagers in your community earn per hour?
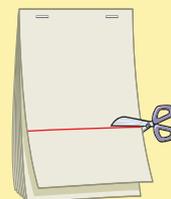By the end of the chapter, you will know how to analyse the data you would like to collect.

Make the following Foldable to organize what you learn in Chapter 12.

Step 1 Staple eight sheets of paper together along the top edge.

Step 2 Make a line 10.5 cm up from the bottom of the top page. Cut across the entire page at this mark.

Step 3 Make a line 9 cm up from the bottom of the second page. Cut across the entire page at this mark.

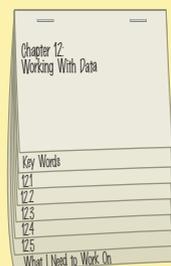**Step 4** Cut across a line 7.5 cm up from the bottom of the third page.

Step 5 Cut across a line 6 cm up from the bottom of the fourth page.

Step 6 Cut across a line 4.5 cm up from the bottom of the fifth page.

Step 7 Cut across a line 3 cm up from the bottom of the sixth page.

Step 8 Cut across a line 1.5 cm up from the bottom of the seventh page.

Step 9 Label the tabs formed.

Literacy ⊗ Link

As you work through Chapter 12, take notes under the appropriate tab. Include information about the key words, examples, and key ideas.
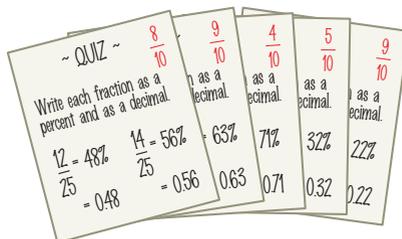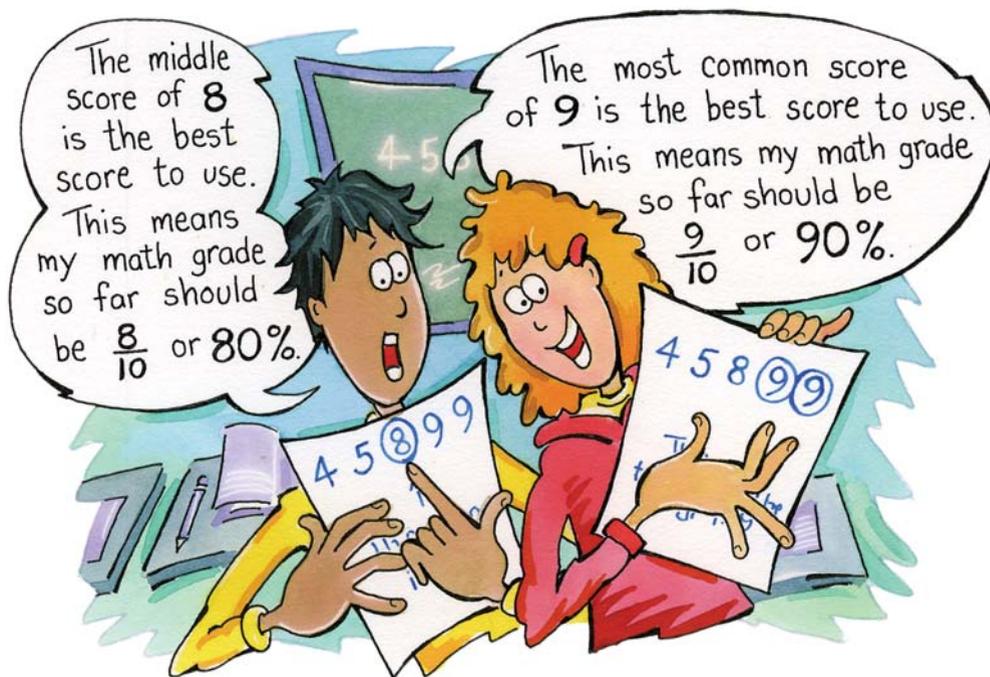
Median and Mode

Early in the school year, Melanie and Amir had the exact same scores on five weekly math quizzes. Their scores out of 10 were 8, 9, 4, 5, 9. Their teacher asked them to report on a single score that best summarizes their progress in math.

Discuss the Math

How do you determine median and mode?

The middle score of **8** is the best score to use. This means my math grade so far should be $\frac{8}{10}$ or 80%.

The most common score of **9** is the best score to use. This means my math grade so far should be $\frac{9}{10}$ or 90%.

1. Which student do you think is right? Discuss your response with a classmate. Do you agree with your classmate?

2. Both Amir and Melanie have justified their decisions. They have each found a **measure of central tendency** to represent the "centre" of the data.

a) Amir has calculated the **median** of the five scores. Describe how he found the median.

b) Melanie has calculated the **mode** of the five scores. Describe how she found the mode.

Reflect on Your Findings

3. a) Which measure, median or mode, do you think better represents how these students are doing in math this term? Explain why.

b) Give a set of ten numbers where you think the median is a more accurate measure of central tendency.

c) Give a set of ten numbers where you think the mode is a more accurate measure of central tendency.

Example 1: Determine the Mode and the Median

One month, the Edmonton Oilers won 7 out of 10 games.

As a team they scored the following number of goals in these 10 games:

5, 4, 3, 2, 1, 5, 3, 3, 7, 3

a) What is the mode for the number of goals scored?

b) What is the median number of goals scored?

Solution

a) Arrange the numbers in increasing order.
The mode is the most frequently occurring number in the list.
1 2 3 3 3 3 4 5 5 7
The mode is 3 since it occurs four times.

b) Arrange the numbers in increasing order. The median is the middle value. Cross off pairs of least and greatest values until you reach the middle one or two numbers.
1̸ 2̸ 3̸ 3̸ 3 3 4̸ 5̸ 5̸ 7̸
If one number remains, then it is the median value.
If two numbers remain then you must find the value halfway between them. In this case, since the two numbers are both 3, the median is 3.
Notice that both the mode and median are 3 in this example. The mode and median will sometimes be the same value, but not always.

measure of central tendency
- a value that represents the centre of a set of data
- can be the mean, median, or mode

median
- the middle number in a set of data after the data have been arranged in order
- median of 2, 5, 6, 8, and 9 is 6
- median of 1, 3, 6, 8, 9, and 10 is 7

mode
- the most frequently occurring number in a set of data
- mode of 3, 5, 7, 7, and 9 is 7
- modes of 2, 2, 4, 6, 6, 8, and 11 are 2 and 6
- the data 1, 3, 4, and 5 has no mode

Example 2: Determine the Mode and Median From a Frequency Table

The frequency table shows the hourly wages of employees at a recycling depot.

Hourly Wages ($)	Number of Employees
8	3
10	2
11	3
14	2

a) What is the mode for the hourly wages?

b) What is the median wage?

c) How would the mode and median be affected if one employee gets a raise from $10 per hour to $11 per hour?

Solution

a) *Method 1: List the Values in Order*

Record the hourly wages, in dollars, for each employee in increasing order.

8   8   8   10   10   11   11   11   14   14

There are three 8s and three 11s.

So, there are two modes: $8 and $11.

> Since there are three people who earn $8 per hour, record three 8s.

> If no number is repeated in a set of data then there is no mode.

Method 2: Use the Frequency Table

From the frequency table, you can determine the mode by looking at the numbers in the second column. The highest number is 3. It corresponds to 3 people who earn $8 and 3 people who earn $11. So, there are two modes: $8 and $11.

> **Literacy Link**
>
> When there are two modes the data are said to be *bimodal*. The prefix "bi" means two. For example, a bicycle has two wheels.

b) *Method 1: List the Values in Order*

Record the hourly wages, in dollars, for each employee in increasing order. The median is the middle value. Cross off pairs of least and greatest values until you reach the middle one or two numbers.

8̶   8̶   8̶   1̶0̶   10   11   1̶1̶   1̶1̶   1̶4̶   1̶4̶

Since two numbers remain, find the middle value between 10 and 11.

The median hourly wage is $10.50.

> The median does not have to be a number in a set of data.

> To calculate the middle value, you could add the numbers and then divide by two.
> $$\frac{10 + 11}{2} = \frac{21}{2}$$
> $$= 10.5$$

Method 2: Use the Frequency Table

Since there are ten employees, the median must be halfway between the wages of the fifth and sixth employees when the salaries are arranged in order. From the frequency table, five people earn $8 or $10 and five people earn $11 or $14.

The median must be halfway between $10 and $11.

The median hourly wage is $10.50.

c) By changing one $10 to $11 the ordered values would now be

8 8 8 10 11 11 11 11 14 14

The only mode is $11, since $11 now occurs four times.

The fifth number is $11 and sixth number is $11, so the median is $11.

Show You Know

Find the mode and median prices of the baseball caps sold in the last week.

Baseball Cap Price ($)	Number of Sales
7	5
9	5
10	6
12	4

Key Ideas

- The mode is the most frequently occurring number in a set of data.
  - If no number repeats, there is no mode.
    1 2 3 4 5   No mode
  - There can be more than one mode.
    1 1 2 2 3   Two modes: 1 and 2 (bimodal)
- The median is the middle value in a set of data after the numbers have been arranged in order.
  1̶ 2̶ 3 4̶ 5̶   Median is 3
- If there is an even number of data values, then the median is the value halfway between the two middle numbers.
  2̶ 4̶ 6 8 1̶0̶ 1̶2̶

  The median here is the value halfway between 6 and 8, or 7.

  $$\frac{6+8}{2} = \frac{14}{2}$$
  $$= 7$$

- The median does not have to be a number in the set of data.

Communicate the Ideas

1. Over ten days, the following numbers of juice cans were collected by the school recycling team: 15, 20, 12, 16, 24, 20, 12, 20, 23, 17.
Which value do you think is easier to determine, mode or median? Why?

2. Create a set of five numbers where the median and mode are the same. Explain why you chose the numbers you did.

3. Dana was asked to find the median of the following numbers: 3, 6, 5, 4, 2. She decided to order the numbers from greatest to least: 6, 5, 4, 3, 2. She removed outer pairs of numbers until she was left with the middle value, 4. Has she correctly found the median? Explain.

Practise

For help with #4 to #7, refer to Example 1 on page 423.

4. What are the median and mode of each set of data?
   a) 4, 2, 9, 6, 4
   b) 21, 15, 18, 21, 20, 18
   c) 3, 8, 5, 12, 10, 8, 2

5. Determine the median and mode of each set of data.
   a) 6, 4, 8, 6, 2, 9
   b) 14, 5, 8, 11, 10
   c) 18, 24, 16, 18, 24, 16, 18, 18

6. In one week, a store in the mall sold the following numbers of Nickelback CDs:
34, 42, 37, 44, 46, 42, 51
What were the mode and median for the CD sales that week?

7. In 12 hockey games, the Calgary Flames scored the following numbers of goals:
4, 4, 0, 2, 3, 3, 4, 1, 3, 0, 3, 2
What were the mode and median?

For help with #8 and #9, refer to Example 2 on pages 424–425.

8. A new T-shirt company sold 26 shirts on their first day. The table shows the number of T-shirts sold

Price ($)	Number Sold
8	6
14	7
17	5
20	4
25	4

according to price. What were the mode and median for the T-shirt prices?

9. A coffee shop sold 36 beverages one hour. The prices of the beverages sold are shown in the table. What were the mode and median prices?

Price	Number Sold
$2	12
$3	10
$3.50	9
$4	5

10. A cat gives birth to a litter of six kittens. The masses of the kittens are:

95 g, 100 g, 100 g, 105 g, 110 g, 110 g

a) What is the mode?

b) What is the median mass?

c) The mother cat has a second litter. The masses of these kittens are:

90 g, 95 g, 100 g, 105 g, 110 g, 115 g

What are the median and mode of the masses of all 12 kittens?

11. The following tally chart shows the heights of a grade 7 boys' basketball team.

Height (cm)	Number of Players
150	II
155	I
160	II
165	IIII
170	₩
180	I

a) What is the mode for the heights? How can you determine the mode(s) quickly from the chart?

b) What is the median height?

12. What is one possible set of four numbers that have a mode of 7 and a median of 11?

Extend

13. Five whole numbers have a mode of 4 and a median of 3. What are the five numbers?

14. If the only mode is 4, describe all possible whole number values for n.

a) 5, 6, 7, 4, 2, n

b) 5, 6, 4, 4, n

15. If the median is 4, describe all possible whole number values for n.

a) 3, 2, 6, 7, n

b) 1, 3, 4, 5, n

16. The median of the set of numbers 3, 4, 5, x, and y is 5. What are the possible whole number values for x and y?

17. Five classmates have birthdays in March. The median of their birth dates is 12 (March 12) and the mode is 8. If the sum of the five birth dates is 56, find all possible sets of birth dates.

MATH LINK

David surveyed ten friends about their shoe size. He recorded the following sizes:

6, 7, 5, 8, 8, 7, 7, 6, 9, 8
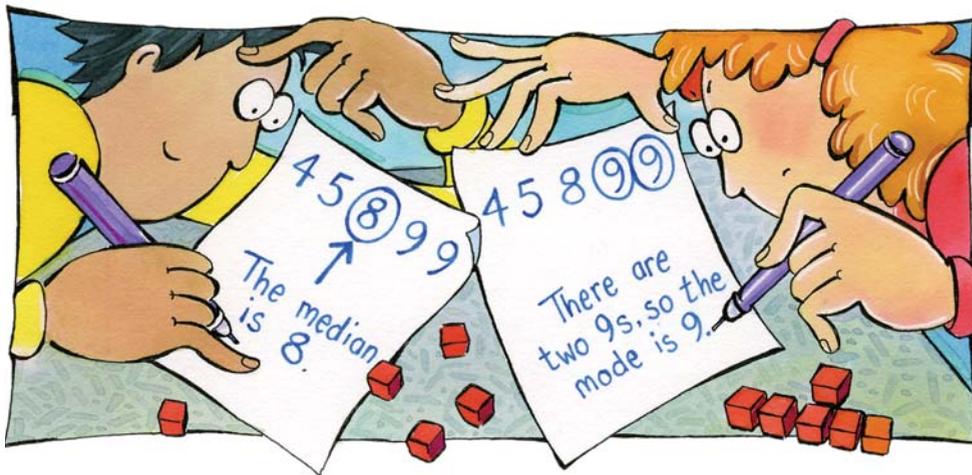
a) What is the median shoe size?

b) What is the mode?

Mean

Ms. Fermat was not satisfied with the way Amir and Melanie calculated their math midterm reports. She did not feel that the median and mode provided a correct view of their performances. Ms. Fermat asked the students to explore another way of representing the centre of the data.

Explore the Math

Materials

• 35 centimetre linking cubes

How do you calculate the mean of a set of data?

We need to find a value that the test scores are centred around.

Let's try levelling out the towers without adding any more blocks.

1. Build a tower that represents each score that Amir and Melanie received on their weekly math quizzes: 4, 5, 8, 9, 9.
A tower 4 cubes high represents a score of 4 out of 10.

2. Sketch a picture of your towers for each of the five scores.

3. Move cubes from the taller towers to the shorter towers to create five identical towers with the same height. Use only the cubes that you used in #1.
  a) What is the new height of each tower?
  b) How does this value represent the centre of the data?

4. a) What is the sum of the five original scores?

b) Divide the sum you found in part a) by the number of quizzes.

c) How does your answer compare to the height of each tower in #3?

Reflect on Your Findings

5. The value you determined in #3 and #4 above is called the mean .

a) How does this value compare with the median and the mode?

b) Do you think the mean is the best measure of how Amir and Melanie are doing in math this term? Explain why.

mean
- a measure of central tendency
- the sum of a set of values divided by the number of values in the set
- for example,

$$\text{Mean} = \frac{6 + 8 + 4}{3}$$
$$= 6$$

Example 1: Calculate the Mean

The daily sales at More Movies for the first week after the DVD release of *The Threat of the Brain Snatchers* are shown in the table.

a) What is the daily mean number of sales from Monday to Saturday?

b) How many sales will be needed on Sunday if More Movies hopes to have a daily mean of 100 sales for the first week?

Day	Number of Daily Sales
Monday	140
Tuesday	90
Wednesday	80
Thursday	90
Friday	110
Saturday	120

Solution

a) Calculate the sum of the six numbers.
140 + 90 + 80 + 90 + 110 + 120 = 630
Divide the sum by the number of days, 6.
630 ÷ 6 = 105
The daily mean number of sales is 105 from Monday to Saturday.

> Look for numbers that are easy to add.
> 90 + 110 = 200
> 80 + 120 = 200
> So, 140 + 90 + 200 + 200 = 630

b) Calculate the total number of sales that will be necessary in order to have a daily mean number of sales of 100 for 7 days (one week).
Since the mean needs to be 100, multiply 100 by 7.
100 × 7 = 700
From part a), the sum of the sales for the first 6 days was 630.
Subtract to calculate the number of sales needed on Sunday.
Total Sales = 700 − 630
= 70
70 sales need to be made on Sunday.

Strategies
Work Backwards
Refer to page xvi.

Show You Know

What is the mean of each set of values?

a) 7, 8, 6, 9, 9, 5, 7, 7, 8, 4   **b)** 300, 250, 400, 300, 250

Literacy Link
You often see the word "average" used instead of the word "mean."

Example 2: Use the Mean to Make Predictions

A scientist recorded the daily distance travelled by a spawning salmon for five days in a row.

Day	Distance Travelled (km)
Monday	44
Tuesday	52
Wednesday	51
Thursday	46
Friday	57

a) What is the mean distance travelled by the salmon each day?

b) How far would you expect the salmon to travel in the next ten days? Assume the salmon has not yet reached its spawning destination.

c) A Chinook salmon may travel 8000 km from its feeding grounds in the North Pacific Ocean! How many days would you predict this journey to take based on the mean daily travel distance you calculated in part a)?

Solution

a) Calculate the sum of the five distances.
$44 + 52 + 51 + 46 + 57 = 250$
Divide the sum by the number of days.
$250 \div 5 = 50$

C **(** **44** **+** **52** **+** **51** **+** **46** **+** **57** **)** **÷** **5** **=** 50.

The mean distance travelled each day is 50 km.

> **M E**
> Add the tens.
> $40 + 50 + 50 + 40 + 50 = 230$
> Add the ones.
> $4 + 2 + 1 + 6 + 7 = 20$
> Add the subtotals.
> $230 + 20 = 250$

b) The mean distance travelled each day is 50 km. Multiply to calculate the expected distance travelled over the next ten days.
$50 \times 10 = 500$
You would expect the salmon to travel about 500 km over the next ten days.

c) The total distance travelled is 8000 km.
The mean distance travelled each day is 50 km.
Divide to find the length of time the total journey would take.
$8000 \div 50 = 160$

C **8000** **÷** **50** **=** 160.

The total journey would take about 160 days.

Key Ideas

- To calculate the mean, add all of the numbers in a set of data and then divide by the number of numbers.
- The mean does not have to be a number in the set of data.

$$\text{Mean} = \frac{4 + 6 + 8 + 10 + 6 + 7 + 15}{7}$$
$$= \frac{56}{7}$$
$$= 8$$

Communicate the Ideas

1. Describe to a classmate who missed the lesson how to calculate the mean of the following set of five numbers:

   2, 6, 8, 9, 10

2. A toy store has six bins of stuffed animals. These bins contain 8, 7, 4, 5, 3, and 9 stuffed animals each.

   a) What is the mean number of stuffed animals?

   b) How could the vertical towers of linking cubes be levelled to determine the mean number of stuffed animals in a bin?
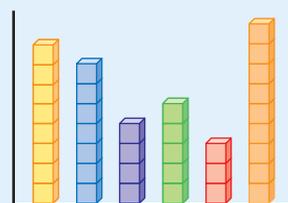
Practise

Round all answers to the nearest tenth where necessary.

For help with #3 to #6, refer to Example 1 on page 429.

3. What is the mean of each set of data?

   a) 5, 4, 10, 5, 6

   b) 2.2, 1.6, 1.9, 2.3, 2.1, 1.9

   c) 30, 85, 50, 105, 100, 65, 20, 25

4. Determine the mean of each set of data.

   a) 6, 7, 8, 9, 4, 11

   b) 3.4, 2.2, 1.4, 4.6, 2.2, 1.4, 1.6, 1.6

   c) 120, 72, 100, 110, 150, 75, 73

5. A store's sales of projection TVs on four Saturdays in February were 8, 7, 9, and 10. What was the mean number of Saturday sales in February?

6. Faith picked the following amounts of cranberries over seven days:

   2 L, 1 L, 1.5 L, 3 L, 1 L, 0.5 L, 1.5 L

   What is the mean volume of cranberries she picked?

For help with #7 and #8, refer to Example 2 on page 430.

7. Juanita scored the following points in her first six basketball games:

    12, 10, 11, 12, 14, 13

   a) What is the mean number of points scored?

   b) How many points would she need to score in her next game to increase her mean by 1 point for the seven games?

8. The chart shows the growth of a seed planted indoors in January.

Month	Height (cm)
Jan	3
Feb	4
Mar	4
Apr	3
May	5
June	5

a) What is the mean monthly growth?

b) How much will the plant have to grow in July for the mean monthly growth to be 5 cm for the seven-month period?

**c)** Predict the height of the plant after one year.

Apply

9. A survey company collected information about the amount of television teens watch each week.

Province	Television Viewing (hours per week)
British Columbia	12.4
Alberta	13.4
Saskatchewan	14.6
Manitoba	14.7

a) What is the mean for the four provinces listed? Round your answer to the nearest tenth of an hour.

b) The number of hours per week of television watched by teens for all of Canada was given as 14.0. How does the mean for the four provinces compare with the value given for all of Canada?

c) Would you predict the mean for the provinces not listed to be more or less than 14.0? Explain your reasoning.

d) How many hours of TV would you expect a typical Canadian teen to watch in one day?

e) How many hours of TV would you expect a typical Canadian teen to watch in ten weeks?

10. The graph shows the number of homes cleaned by Quick & Clean Housecleaning. What is the mean number of homes cleaned for the months shown?

Quick & Clean Housecleaning
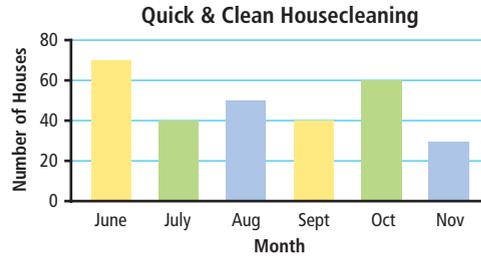
**11.** Here are Sasha's test scores in math this term:

78%, 68.5%, 82%, 87%, 91%, 88%, 74.5%, 71%

a) What is the mean of Sasha's test scores?

b) Sasha wants to raise her mean mark by 1% after the next test. What mark does she need to get on the next test?

12. The table gives the maximum daytime temperatures for five cities in Saskatchewan during August.

City	Maximum Temperature (°C)
La Ronge	22
North Battleford	24
Regina	26
Uranium City	19
Yorkton	24

a) What is the mean daytime temperature?

b) Predict the maximum daily temperature for Saskatoon, SK, in August. Explain your reasoning.
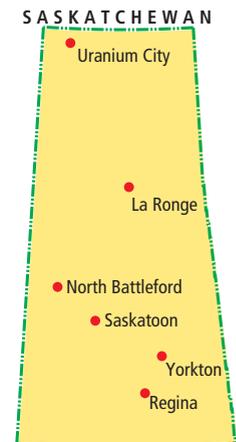
13. Brennan recorded his test scores for his English course in a table.

Test	My Score	Out of (Total)
Grammar	6	10
Spelling	11	15
Novel Study	27	35
Poetry	31	40

a) What is the sum of Brennan's four scores?

b) What is the total possible score Brennan could have received?

c) Calculate Brennan's overall mean score, to the nearest whole percent.

14. The table shows the percent of people who take their own bags shopping in order to minimize the waste of plastic and paper. What is the mean percent? Round your answer to the nearest tenth of a percent.

Province	People Who Bring Their Own Shopping Bags
Alberta	31%
British Columbia	31%
Manitoba	35%
Ontario	25%
Québec	18%
Saskatchewan	33%

15. The table gives the ticket prices and number of seats available at each seat location for a soccer stadium.

Seat Location	Ticket Price	Number of Seats
Lower Centre	$12	500
Upper Centre	$10	1000
End Zone	$6	500

What is the mean ticket price for a sell-out?

16. During one month, major league baseball players Joe McEwing and Mike Lieberthal each had 45 hits in 132 attempts.

a) Determine the *batting average* for each player by dividing the number of hits by the number of attempts. A *batting average* is always expressed as a decimal to 3 places.

b) Suppose these players played against each other. Joe hit .800 by hitting 4 times in 5 attempts. Mike hit 1.000 after successfully hitting 3 times in 3 attempts. Who do you think has the better updated *batting average* for the year? Explain your prediction.

c) Calculate each player's new *batting average*. Was your prediction correct?

MATH LINK

Leah interviewed ten friends about the number of cousins they have.

Name	Number of Cousins	Name	Number of Cousins
Danika	18	Kyle	20
Jerome	3	Nicole	8
Paula	9	Vishal	22
Sam	14	Michelle	6
Janice	12	Jonah	10

What is the mean number of cousins among Leah's friends? Round your answer to the nearest whole number.
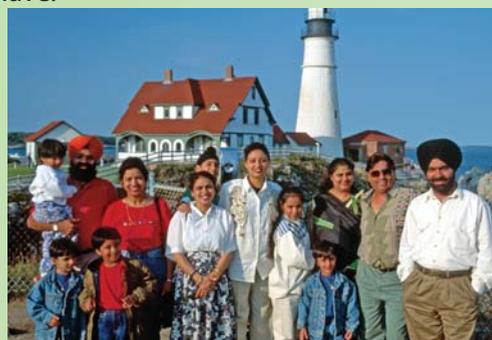
Range and Outliers

The wooden roller coaster at Playland in Vancouver was built in 1958. It is one of the oldest wooden roller coasters that is still in use. Most newer roller coasters are made of steel.

Discuss the Math

How can the largest and smallest values be used to describe a set of data?

The following diagram shows the elevation changes for the Stomach Bender roller coaster. The measurements represent the heights of the roller coaster relative to the starting point of the ride.
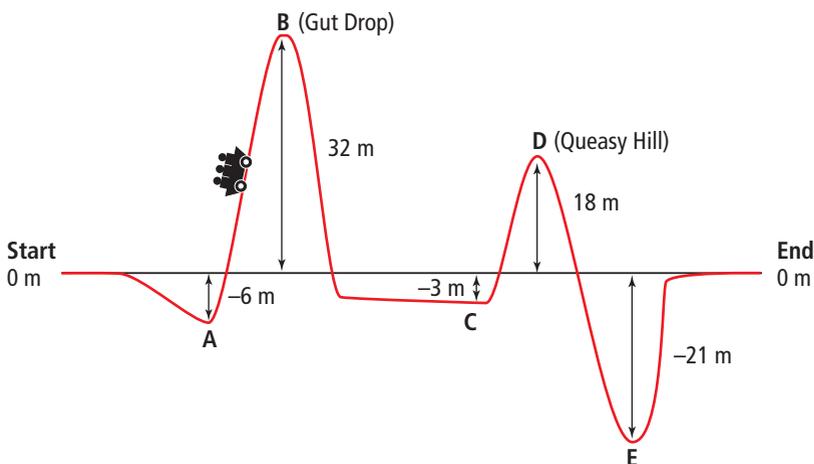
1. Copy the table below into your notebook.

Location Along Ride	Start	A	B	C	D	E	End
Elevation Relative to Starting Point (m)							

2. Fill in the elevations at each point.

3. a) What is the largest value?

b) What is the smallest value?

c) What is the **range** of heights for the roller coaster?

4. Do you think there are any **outliers** in the set of 7 locations on the roller coaster? If so, what are they?

Reflect on Your Findings

5. How would you use the highest and lowest values of a data set to help you describe the data?

Example 1: Calculate the Range

A hospital recorded the following number of births one week.

Day	Mon	Tues	Wed	Thurs	Fri	Sat	Sun
Number of Births	7	10	4	6	7	8	3

a) What is the highest number of births?
b) What is the lowest number of births?
c) What is the range of the data?

Solution

a) The highest number of births is 10 births on Tuesday.

b) The lowest number of births is 3 births on Sunday.

c) The range is calculated by subtracting the lowest value from the highest value.

Range = highest value − lowest value
$$= 10 - 3$$
$$= 7$$

Show You Know

What is the range of each set of values?
a) 7, 4, 3, 6, 2, 5, 8, 3
b) 11, 8, 8, 5, 7, 6, 7, 6, 6
c) −4, 3, −1, 0, 2

Example 2: Identify Outliers in a Data Set

In the 1985–1986 NHL hockey season, Wayne Gretzky set a record for the most points (goals + assists) recorded during a single regular season. The following table shows the top five point leaders for the 1985–1986 season and the 2005–2006 season.

NHL Point Leaders			
1985–1986 Season	**Points**	**2005–2006 Season**	**Points**
Wayne Gretzky	215	Joe Thornton	125
Mario Lemieux	141	Jaromir Jagr	123
Paul Coffey	138	Alexander Ovechkin	106
Jari Kurri	131	Dany Heatley	103
Mike Bossy	123	Daniel Alfredsson	103

a) Determine the range for each season. Compare the two ranges.
b) Which data value appears to be an outlier in one of the two seasons?
c) Remove the outlier and determine the new range for that season. Compare the new range with the range from the other season.

Solution

a) 1985–1986 Season:
215 is the highest value, and 123 is the lowest value
Range = 215 − 123
= 92

2005–2006 Season:
125 is the highest value, and 103 is the lowest value
Range = 125 − 103
= 22

The ranges are very different: 92 and 22.

b) The outlier appears to be Gretzky's 215 points in the 1985–1986 season.

c) The new highest value is 141.
The new range = 141 − 123
= 18
The two ranges are more comparable in size: 22 and 18.

Show You Know

What value(s) appear to be outliers in each set of data?
a) 18, 19, 79, 17, 20, 12
b) 14, 14, 16, 15, 1, 16
c) 98, 202, 99, 91, 2, 95, 89, 88, 94

- The range provides information about the spread of the data.
  Range = highest value − lowest value
- Outliers are values that are very different from the rest of the numbers in a set of data. A data set can have no outliers, or one or more outliers.

Communicate the Ideas

1. In your own words, explain the meaning of range and outliers.

2. How can you determine the smallest value in a data set if you are given the range and the largest value? Use an example to explain your response.

Practise

For help with #3 to #5, refer to Example 1 on page 435.

3. What is the range of each set of data?
   a) 11, 8, 7, 10, 6, 5, 18, 13
   b) 4, −5, 3, −2, 0, 7

4. Determine the range of each set of data.
   a) 16, 11, 7, 29, 31, 18, 21, 18
   b) 7, 1, −1, −2, 9, −7, 4

5. Leanne timed how long she could hold her breath while swimming under water. She recorded the following timings, in seconds:
47, 55, 42, 65, 61
What is the range of the data?

For help with #6 to #8, refer to Example 2 on page 436.

6. What value(s) appear to be outliers in each set of data?
   a) 6, 9, 9, 37, 8, 7
   b) 24, 34, 46, 26, 32, 43, 115
   c) 48, 32, 67, 61, 47, 95, 89, 888, 1

7. Identify any possible outlier(s) in each set of data.
   a) 666, 11, 9, 12, 8, 13, 10
   b) 43, 54, 62, 64, 0, 211, 45
   c) 82, 75, 76, 85, 89, 95, 92, 88, 80

8. Curb-cut ramps are often built at intersections to allow wheelchairs and strollers to cross streets more easily. In one town, a study was done to count the number of curb-cuts at ten random intersections. The results for five different neighbourhoods are given:
9, 8, 8, 1, 7
   a) What is the range?
   b) Which data value may be an outlier?
   c) If you remove the outlier, what is the new range?

9. Vincent received a 200-piece jigsaw puzzle for his birthday. The following times indicate how long it took him to complete the puzzle the first six times he did it:

54 min, 36 min, 34 min, 29 min, 27 min, 28 min

a) What is the range?

b) Which time may be an outlier?

c) Why might this particular value be so different from the others?

d) If you remove the outlier, what is the new range?

10. The following table shows the mean high temperature for each month in Whitehorse, Yukon Territory.

Month	Mean Temp.
January	−13°C
February	−7°C
March	−1°C
April	6°C
May	13°C
June	18°C
July	20°C
August	18°C
September	12°C
October	4°C
November	−5°C
December	−11°C

a) What is the highest value?

b) What is the lowest value?

c) What is the range of the data?

11. The table gives the mean number of wet days per year for several locations across Canada. What is the range of wet days?

Location	Mean Number of Wet Days
St. John's	217
Halifax	170
Fredericton	156
Toronto	139
Winnipeg	119
Regina	109
Edmonton	123
Vancouver	164
Whitehorse	122
Yellowknife	118

12. The table gives the magnitudes of five of the largest earthquakes that have occurred in western Canada.

Location	Date	Magnitude
West of Vancouver Island, BC	Jan 26, 1700	9.0
South of Hope, BC	Dec 14, 1872	7.4
West Coast of Vancouver Island	Dec 6, 1918	7.0
Central Vancouver Island	June 23, 1946	7.3
Queen Charlotte Islands, BC	Aug 22, 1949	8.1

Magnitude refers to the relative size of an earthquake. Magnitude measurements are recorded using the Richter scale, which is a special scale that goes from 0 to 10. What is the range of magnitudes of the five earthquakes in the table?

WWW Web Link

To learn more about earthquakes in Canada, go to www.mathlinks7.ca and follow the links.

13. Final times, in minutes, for the top ten contestants in a race were 55.2, 56.7, 56.9, 57.3, 58.8, 58.9, 59.2, 59.3, 59.3, and 59.7.

  a) What is the range of times?

  b) Do you think this set of data contains any outliers? Explain.

Extend

14. A data set contains five numbers:

  12, 20, 18, 15, 5

  a) Determine the range, median, mode, and mean of the data set.

  b) Add 10 to each of the five numbers and recalculate the range, median, mode, and mean. How do your answers compare with your answers to part a)?

  c) Multiply the numbers in the original data set by 10 and calculate the new measures of central tendency. How do your answers compare with your answers to part a)?

15. Five students competed in a pie-eating contest.

  • The range of times for the group to finish eating their pies was 9 min.
  • The least amount of time it took for someone to finish their pie was 1 min.
  • The mode was 4 min.
  • One boy took 2 min more than the median.

  a) What are the finishing times for the five contestants?

  b) What was the mean number of minutes it took the five contestants to eat their pies?

  c) What strategy did you use to determine the mean?

16. Conduct the following experiment to simulate rolling two 10-sided dice each numbered from 0 to 9.

  • Randomly choose a page from your textbook.
  • Calculate the sum of the last two digits of the page number. For example, for page 108 you would add 0 and 8 to get 8.
  • Randomly choose 49 more pages for a total of 50 pages.
  • Copy and complete the frequency table.

Sum	Tally	Freq.	Sum	Tally	Freq.
0			10		
1			11		
2			12		
3			13		
4			14		
5			15		
6			16		
7			17		
8			18		
9			19		

  a) What is the mode?

  b) What is the mean?

  c) What is the median?

  d) What is the range?

  e) Why are there no tally marks on 19?

MATH LINK

Measure and record the heights of ten people in your class, including your teacher.

a) What is the range of heights?
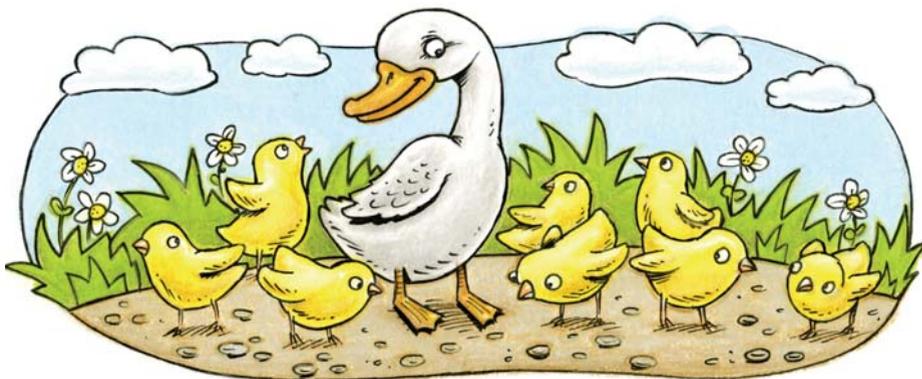
b) Identify any possible outliers.

The Effects of Outliers

Can you spot the outlier in the cartoon shown?
Suppose you are asked to determine the mean mass of these babies. Should this outlier be removed from the data set?

Some outliers are caused by mistakes in data collection. Other outliers are just as important as the other data values. When there are outliers in a data set, the mean, median, and range can be different from what they are when the outliers are removed. People who work with data need to decide when outliers should and should not be used when calculating measures of central tendency.

Explore the Math

Materials
• ruler

How do outliers affect measures of central tendency?

For a grade 7 science project, students were asked to plant and observe the growth of six bean seedlings. Students were given specific instructions to give their planted seeds 1 h of light and 30 mL of water per day. After two weeks the students brought their plants back to school.

1. With a ruler, measure the heights of the six bean seedlings shown. What are the heights?

2. Copy the following table into your notebook.

Plant Height	Mean (cm)	Median (cm)	Largest Value	Smallest Value	Range
With Outlier					
Without Outlier					

3. Complete the following calculations. Record your answers in the first row of your table.
   a) Determine the mean and median heights for the plants. Round your answers to the nearest tenth of a centimetre.
   b) What is the highest seedling height?
   c) What is the lowest seedling height?
   d) What is the range in heights?

4. Identify a possible outlier value.

5. Remove the outlier from your data. Repeat the calculations from #3. Record these answers in the second row of your table.

6. a) How has the median changed by removing the outlier?
   b) How has the mean changed by removing the outlier?

7. What are some possible reasons why the one plant grew so much more than the other five? Compare your reasons with those of a classmate.

Reflect on Your Findings

8. a) Which value is affected more by the presence of an outlier, the median or the mean? Explain.
   b) Should the outlier value be included in the data for the science experiment? Explain why or why not.

Example 1: Identify Outliers

Shannon practised shooting baskets every day last week to prepare for a basketball tournament. She recorded the number of baskets she made each day out of 25 shots.

a) What is the range of baskets scored?
b) What are the median and mean numbers of baskets scored?
c) Identify any possible outliers. Should the outlier(s) be removed from the data set? Explain why or why not.

Day	Number of Baskets
Sunday	14
Monday	17
Tuesday	16
Wednesday	20
Thursday	5
Friday	22
Saturday	18

Solution

a) The highest and lowest values are 22 and 5.

$$\text{Range} = 22 - 5$$
$$= 17$$

b) Arrange the data in order: 5̶, 1̶4̶, 1̶6̶, 17, 1̶8̶, 20, 2̶2̶
   The median number of baskets scored is 17.

$$\text{Mean} = \frac{5 + 14 + 16 + 17 + 18 + 20 + 22}{7}$$
$$= \frac{112}{7}$$
$$= 16$$

The mean number of baskets scored is 16.

c) The value 5 could be considered an outlier. This value is significantly different from the other values. But this value should not be removed from the data set because it is just as important as the other data values. It is probably not an error in measurement. It may simply represent a poorer performance that day.

Example 2: Identify Outliers and Determine Their Effects

In a science experiment, students were asked to measure the length of their right thumb from the first knuckle to the end of their thumb. The table shows the lengths that were measured by ten different students.

Student	A	B	C	D	E	F	G	H	I	J
Length (cm)	3.2	2.7	3.1	2.8	2.8	2.7	31	3.3	2.6	3.0

a) What is the range?
b) What are the median and the mean?
c) Identify any possible outlier(s). Should the outlier(s) be removed from the data set? Explain why or why not.
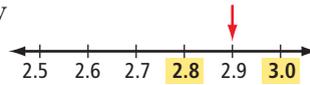d) How would removing the outlier affect the median and the mean?
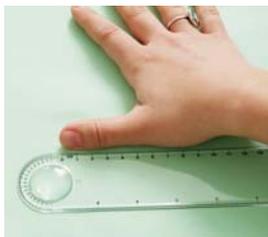
Solution

a) The highest value is 31 cm. Range = 31 − 2.6
   The lowest value is 2.6 cm.   = 28.4
   The range 28.4 cm.

b) To find the median, arrange the numbers in order.
   2.6, 2.7, 2.7, 2.8, 2.8, 3.0, 3.1, 3.2, 3.3, 31
   With ten values, the median will be halfway
   between the fifth and sixth values.
   The median is 2.9 cm.

$$\text{Mean} = \frac{3.2 + 2.7 + 3.1 + 2.8 + 2.8 + 2.7 + 31 + 3.3 + 2.6 + 3.0}{10}$$
$$= \frac{57.2}{10}$$
$$= 5.72$$

The mean is 5.72 cm.

c) The value of 31 is an outlier. It is about ten times greater than the
   other measurements and is most likely an error in either measurement
   or recording. The student may have measured in millimetres instead of
   centimetres. Or, it is possible that the student forgot to place the
   decimal point between the digits 3 and 1. In this case, the outlier
   should be removed.

d) Remove the outlier value of 31 from the ordered list of values from
   part b). Recalculate the median and the mean.
   2.6, 2.7, 2.7, 2.8, 2.8, 3.0, 3.1, 3.2, 3.3
   Since there are only nine values now, the median will be the fifth
   value. The median is 2.8 cm.

$$\text{Mean} = \frac{2.6 + 2.7 + 2.7 + 2.8 + 2.8 + 3.0 + 3.1 + 3.2 + 3.3}{9}$$
$$\approx 2.9$$

The mean is approximately 2.9.

The median changes from 2.9 to 2.8. The mean changes from 5.7
to 2.9. The mean is affected more by removing the outlier.

Show You Know

The following times were recorded, in seconds, for the runners in a race:
20.2, 16.5, 40.4, 18.5, 21.4, 20.5, 17.1, 24.5, 19.0

a) What is the range of times?

b) What are the median and mean times?

c) Identify any possible outliers. Should the outlier(s) be removed
   from the data set? Explain why or why not.

Key Ideas

- Outliers can affect all measures of central tendency.
- When a small set of data has an outlier, the mean is usually affected more by the outlier than the median.
- Some outliers are just as important as the other data values, while others are better removed from the data set.

Communicate the Ideas

1. Brian's bowling scores are 135, 132, 128, 316, 135, and 138. Identify a possible outlier in his scores. Should you remove it from the data set? Explain your reasoning.

2. a) Give an example of a situation where an outlier would be used when reporting on measures of central tendency.

   b) Give an example of a situation where an outlier would not be used when reporting on measures of central tendency.

Practise

3. The table shows the percent of students surveyed that had at least one song on their MP3 players by the musicians listed.

Musician	Students With at Least One Song
Snoop Dogg	42%
Shania Twain	38%
Britney Spears	6%
Kanye West	40%
Led Zeppelin	41%
Avril Lavigne	38%
U2	88%

   a) What is the range?

   b) What are the median and mean?

   c) Identify any possible outliers. Should the outlier(s) be removed from the data set? Explain why or why not.

4. Two grade 7 students randomly ask the ages of the first eight people to pass them in the hallway. They record the following ages:

7, 11, 8, 8, 52, 9, 9, 10

   a) What is the range?

   b) What are the median and mean age?

   c) Identify any possible outliers. Should the outlier(s) be removed from the data set? Explain why or why not.

5. Sharon recorded the following prices for five different brands of canned tomatoes on the grocery store shelf:

$1.29, $1.69, $1.59, $9.61, $1.39

   a) What is the range?

   b) What are the median and the mean?

   c) Identify any possible outlier(s). Should the outlier(s) be removed from the data set? Explain why or why not.

   d) How would removing the outlier(s) affect the median and the mean?

Apply

6. A medical study was conducted to learn the effect of caffeine on heart rate. Participants were asked to drink one 250-mL cup of coffee and record their number of heartbeats over a 15-second interval. The following data were collected.

Heartbeats in 15 seconds:
33, 35, 30, 70, 33, 31, 36, 40, 37, 29

a) What is the range?

b) What are the median and the mean?

c) Identify any possible outlier(s). Should the outlier(s) be removed from the data set? Explain why or why not.

d) How would removing the outlier(s) affect the median and the mean?

7. David had the following scores on his eight weekly spelling tests:

70%, 80%, 80%, 70%, 100%, 80%, 20%, 90%

a) What is the range?

b) What are his median and mean scores?

c) Identify any possible outlier(s). Should the outlier(s) be removed from the data set? Explain why or why not.

d) How would removing the outlier(s) affect the median and the mean?

e) Describe two different ways that you could determine David's overall mark to be between 75% and 85%.

Extend

8. A set of nine numbers has two outliers. The mean and median both equal 50 if you include or exclude the outliers. What are the possible nine numbers?

MATH LINK

In a gymnastics competition, each performance was judged by eight judges on a scale from 0.25 to 10.00. In order to calculate the gymnast's overall performance, the top score and bottom score were removed and the mean of the remaining scores was determined. This value is called the trimmed mean.

Jordan recorded the following scores for her friend's performance.

Judge	A	B	C	D	E	F	G	H
Score	8.25	7.50	9.75	8.50	6.50	7.75	8.00	8.25

Round your answers to 2 decimal places.

a) Using all the scores, what is the median? mean? highest score? lowest score? range?

b) Remove the top and bottom scores. What is the new median? mean? range?

c) Which value in part b) has changed the most?

d) Would you consider the highest and lowest scores to be outliers in this example? Explain.

Choose the Best Measure of Central Tendency

Focus on...

After this lesson, you will be able to...

❑ determine when to use the mean, median, or mode to best describe a set of data

❑ solve problems using mean, median, and mode

The table below shows school T-shirt sales for the past ten weeks. The school wants to make one more order for the next 30 weeks. How could the school decide how many T-shirts to order?

Date	Sept 10	Sept 17	Sept 24	Oct 1	Oct 8	Oct 15	Oct 22	Oct 29	Nov 5	Nov 12
Sales	7	50	8	9	10	12	7	7	9	11

Discuss the Math

Which measure of central tendency is most appropriate?

1. What are the mean, median, mode, and range of the T-shirt data shown above?

2. Compare the mean, median, and mode. Which measure seems to best represent the ten numbers? Explain.

3. Share your answer from #2 with a classmate. Do your answers agree? If not, explain the reasons for your selection.

4. How many T-shirts should be purchased for the next 30 weeks? Explain your thinking.

Reflect on Your Findings

5. a) Why do you think the sales in the second week were so high compared to the other nine weeks?

b) Which measure of central tendency was most affected by the large number of sales in the second week?

c) Which measure of central tendency do you think is not a good measure of the centre of these data? Explain why.

Example 1: Compare Measures of Central Tendency

Amir and Melanie's weekly quiz scores were 4, 5, 8, 9, 9.

Which measure(s) of central tendency best describe the data? Explain why.

The median is 8.
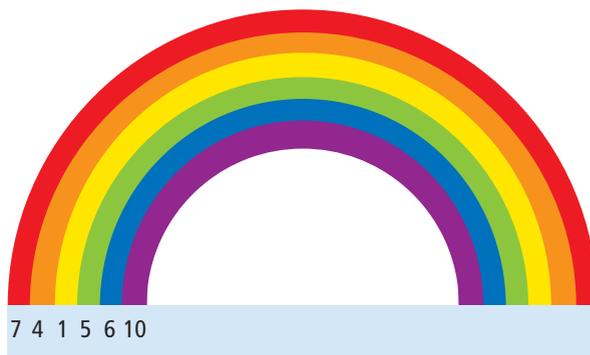
The mode is 9.

The mean is 7.
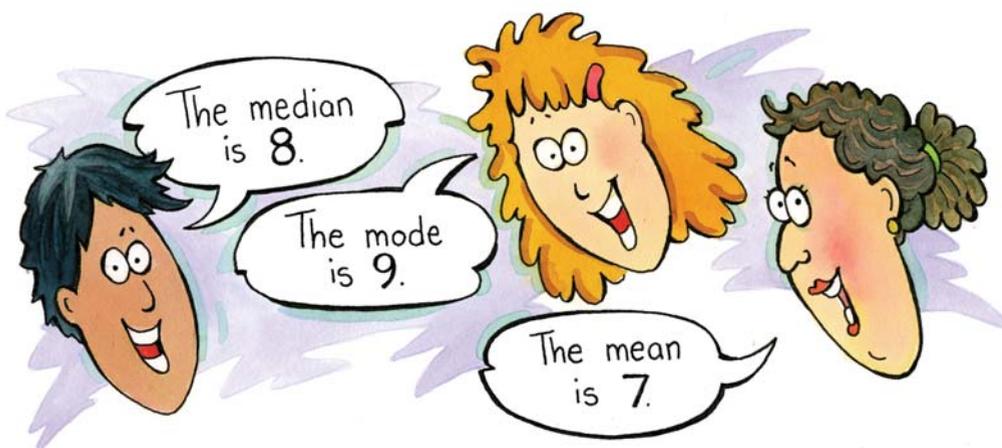
Solution

Since the mode represents the highest score, it is not the best representation of the five scores. The other two measures, median and mean, are both acceptable.

Show You Know

a) What are the mean, median, and mode for the following data set? Round your answers to the nearest tenth, if necessary.

16, 53, 14, 16, 11, 11, 12, 13, 11

b) Which measure(s) of central tendency best describe the data? Explain.

Example 2: Choose Mode as the Best Measure of Central Tendency

Students have taken a vote on the new official school colours for sports uniforms. The number of votes for each colour is shown. Which measure of central tendency would you use to decide the winning colour?

7 4 1 5 6 10

Solution

The data collected involve the frequency of colour choices. The most popular choice wins.

In this case, the median and mean do not provide any meaningful information about colour choice. The best measure to use is the mode. The mode is purple since purple is the most popular choice.

Example 3: Compare Median and Mean

Suki paid the following amounts for her last six pairs of jeans before tax:

$44, $38, $45, $49, $125, $50

a) What are the median and the mean jean prices?
b) Which measure of central tendency best describes these data?

Solution

a) Arrange the numbers in order. The median is the middle value.
    38̶ 4̶4 45 49 5̶0̶ 1̶2̶5̶
    The median is halfway between the values of 45 and 49 at 47.
    The median price is $47.00.

$$\text{Mean} = \frac{38 + 44 + 45 + 49 + 50 + 125}{6}$$
$$= 58.50$$

The mean price is $58.50.

> The number of pairs of jeans is 6.

b) The value of $125 is very different from the other five values. The single value, $125, alters the mean much more than the median. The median is a better measure of central tendency for the six prices.

Key Ideas

- The mode is the best measure of central tendency for data that represent frequency of choice such as favourite colour, clothing and shoe sizes, or most popular musical group.
- If all the numbers in a set of data are relatively close together, either the median or mean can be used as a measure of central tendency.
- If a data set contains unusually large or small numbers relative to the rest of the data, the median is usually the best measure of central tendency.

Communicate the Ideas

1. A group of adults were asked to choose their favourite female singer: Sarah McLachlan, Hillary Duff, Avril Lavigne, Shania Twain, or Celine Dion. Which measure of central tendency would best represent the data? Explain why.

2. A student recorded the following hours of volunteer work for each of the past ten months:

23, 18, 21, 19, 23, 24, 84, 22, 20, 16

Determine the median and mean. Which measure of central tendency better represents the data? Explain why.

Practise

Round all answers to one decimal place unless otherwise specified.

For help with #3, refer to Example 1 on page 447.

3. Min recorded the number of baskets she made out of ten attempts during each basketball practice. After nine practices, her results were:

4, 7, 5, 6, 3, 7, 2, 3, 7

a) What are the median, mode, and mean?

b) Which measure of central tendency best describes these data? Explain why.

4. Which measure of central tendency best represents the following data? What is its value?

Favourite Hockey Player	Number of Votes
Sidney Crosby	8
Jarome Iginla	5
Alexander Ovechkin	8
Ryan Smyth	7

5. The following tally chart represents the sizes of running shoes that were sold last Saturday.

Size	7	8	9	10
Number Sold	⊞	⊞		‖

a) What are the mean and the mode size of shoe?

b) If you are restocking the shoes at the end of the day, which measure of central tendency is more meaningful? Why?

6. A realtor in Rainbow Town sold the following houses in the past month.

House Description	Selling Price
Red starter house	$80 000
Blue house	$140 000
Green house	$145 000
Grey house	$150 000
Pink mansion	$2 100 000

a) What are the median and mean?

b) Which measure of central tendency is more representative of the house prices in Rainbow Town?

7. In a grade 5 class, 16 students are 10 years old, 2 students are 9 years old, and 2 students are 11 years old. The teacher and her assistant are both 50 years old.

a) What are the mean, median, and mode ages of all students and teachers?

b) Which measure(s) of central tendency best describes the data? Explain why.

Apply

8. The following table shows survey results for the percent of radio listening time by music type among 100 Canadian teens.

Music Type	Listening Time (%)
Pop	19.0
Contemporary rock	31.0
Rap	14.7
Album rock	10.6
Country	8.7
Other	16.0

Which single music type best represents Canadian teenagers? Which measure of central tendency did you use to find your answer? Explain why.

9. Juan's Cleaners had developed a new disinfectant to kill germs. Ten tests were performed with the following results.

Percent of germs eliminated:
67, 99, 91, 87, 99, 70, 99, 69, 92, 61

a) If you were the owner of the company, which measure of central tendency would you use for advertising? Why?

b) If you were working for the Centre for Disease Control, which measure of central tendency would be best for the public to use in evaluating the product? Why?

10. A school held a canned food drive for holiday season food baskets in their community. The following numbers of cans were collected.

Grade	Number of Students	Total Cans Collected
1	20	100
2	24	96
3	22	99
4	20	120
5	18	90
6	26	130
7	25	125

a) Which grade collected the most cans?

b) What was the mean number of cans collected for each student in each grade?

c) Which grade should win a prize for most cans collected? Explain your choice.

Extend

11. A set of five distinct natural numbers has a mean of 6 and a median of 6. What is the largest possible number in the set?

12. The mean of four natural numbers is 28. If there is only one mode, what is a possible set of numbers?

13. A set of three whole numbers has a median of 17. If the smallest number is 10 and the mean is 15, what is the largest number?

14. A set of four whole numbers has a mode of 7. The smallest number is 2. What is the mean if the largest number is 8?

15. On a recent class test, the mean score was 80%, the mode score was 70%, and the median score was 90%. Kyle scored 80%. Which measure of central tendency would you use to make Kyle feel

a) confident about himself? Why?

b) like he could do better? Why?

16. Work with a partner to create a scenario where there are two groups that take opposite positions (labour and management, children and parents, etc.).

a) Provide data that supports using the mode as the best measure of central tendency.

b) Provide data that supports using the median as the best measure of central tendency.

MATH LINK

A set of seven judges gave the following scores to Susan's diving performance:
7.2, 6.8, 7.3, 8.0, 8.5, 8.2, 6.8

a) What is the mean? Round your answer to the nearest tenth.

b) What is the median?

c) What is the mode?

d) Which measure(s) of central tendency best represent the centre of the data? Explain why.

Key Words

For #1 to #4, unscramble the letters to identify the appropriate key word.

1. A E I M N D
the middle number in a set of data arranged from smallest to largest

2. N A M E
the sum of a data set divided by the number of values in the set

3. L R U O I T E
a value that is significantly smaller or larger than the other values in a data set

4. A G R E N
the positive difference between the largest and smallest values in a data set

12.1 Median and Mode, pages 422–427

5. The number of phone calls received at Jessica's house in the last seven days were 5, 9, 8, 5, 6, 3, and 4. What is the mode and median?

6. The table shows the mean wind speed in kilometres per hour from weather readings taken at the Blood Reserve near Lethbridge, Alberta.

Month	Wind Speed (km/h)	Month	Wind Speed (km/h)
Jan	22	July	16
Feb	20	Aug	15
Mar	19	Sept	17
Apr	21	Oct	21
May	20	Nov	21
June	19	Dec	22

What is the mode and median?

7. Make a data set with six numbers for which the values of the median and mode differ by 2. Explain how you found your answer.

12.2 Mean, pages 428–433

8. Between January and March one year, school was cancelled in Chesterfield Inlet seven times due to blizzards. The following data give the number of days each blizzard lasted:

1, 3, 2, 3, 5, 2, 3

What was the mean length of a blizzard that year?

9. The Wong family took a family vacation to California. The one-way driving trip was 2100 km. It took them 3 days to drive there.

a) What was the mean distance they travelled each day?

b) Predict how many days it would take them next summer to drive one way to eastern Canada, a distance of 4900 km. Assume they can travel the same mean distance each day.

10. A set of five numbers has a sum of 35.

a) What is the mean of the five numbers?

b) If one of the numbers is removed, the mean is 6. What is the sum of the remaining four numbers?

c) Which number was removed?

12.3 Range and Outliers, pages 434–439

11. The following table shows the number of forest fires that were caused by humans in one year.

Region	Fires Caused by Humans in One Year
Manitoba	203
Saskatchewan	239
Alberta	336
British Columbia	644
Yukon Territory	25
Northwest Territories	17

a) What are the highest and lowest values?

b) What is the range?

c) Identify any possible outlier(s). What might be a possible explanation for the outlier(s)?

12. The following table shows the mean hourly wages of six different categories of employment in British Columbia, Manitoba, and Alberta in 2006.

Type of Occupation	Hourly Wages ($)		
	MB	AB	BC
processing and manufacturing	13.62	17.68	17.53
health related	20.41	22.72	25.16
trades, transport, and equipment	16.87	19.98	20.73
science related	23.99	28.84	25.71
business, finance, and administration	16.29	18.77	18.94
sales and service	11.37	12.68	13.60

a) What is the range of wages in each province?

b) What is the range for the three highest paying categories in each province?

12.4 The Effects of Outliers, pages 440–445

13. The following table shows the October rainfall in several locations.

Location	October Rainfall (mm)
La Ronge, SK	22
Regina, SK	14
Jasper, AB	24
Edmonton, AB	10
Slave Lake, AB	18
Terrace, BC	197
Kelowna, BC	21

a) What are the median and mean October rainfalls?

b) Identify any possible outlier(s). Should the outlier(s) be removed from the data set? Explain why or why not.

c) How would removing the outlier(s) affect the median and the mean?

> **Geography 🔗 Link**
>
> The rainfall recorded for Terrace, BC is correct. You may wish to learn more about the geography of Terrace that makes it such a wet place.

12.5 Choose the Best Measure of Central Tendency, pages 446–451

14. Two grade 7 classes collected cans of food for a local food bank. The table shows the number of cans of food that were collected by each student.

	Number of Cans Per Student						
	0	1	2	3	4	5	37
Class A	IIII	HH II	HH III				I
Class B		IIII	HH	HH II	II	I	

a) What is the median for each class?

b) What is the mean for each class?

c) If a tie is not allowed, which class should get the prize? Explain.

For #1 to #3, choose the best answer.

Use the following information to answer #1 to #3.

In April, 2003, Canadian Mike Weir won the Masters golf tournament. His scores for the four-day event appear in the table.

Day	Score
Thursday	70
Friday	68
Saturday	75
Sunday	68

1. What is the mode of the set of golf scores?

A 7 **B** 68 **C** 69 **D** 70.25

2. What is the mean of the set of golf scores?

A 7 **B** 68 **C** 69 **D** 70.25

3. What is the range of the set of golf scores?

A 7 **B** 68 **C** 69 **D** 70.25

Use the following information to answer #4 to #6.

In the past year the following sizes of litters of kittens were born in one neighbourhood:

5, 6, 3, 5, 6, 3, 4, 5, 5, 4

4. The mode is ■.

5. The median is ■.

6. The mean is ■.

Short Answer

7. Suppose every piece of data in a set is increased by 20. How does this affect the values of the mean, median, and mode?

8. a) Give an example of a situation where the mode is the best measure of central tendency.

b) Give an example of a situation where the median is the best measure of central tendency.

c) Give an example of a situation where the mean is the best measure of central tendency.

9. The table shows the percent of homes in four provinces that have water-saving, low-flow shower heads installed.

| Province	Homes With Low-Flow Shower Heads (%)
Manitoba	34.0
Saskatchewan	26.6
Alberta	32.2
British Columbia	43.4

a) What is the range of percents?

b) What is the mean of percents?

10. Mary and Iblauk went sledding after school. They both went down the hill six times on the same sled. The following list gives the distances they slid, in metres:

25, 42, 53, 47, 41, 51

a) What is the mean distance travelled, to the nearest tenth of a metre?

b) Identify any outlier(s). Should the outlier(s) be included in the calculation of the mean? Explain why.

11. The following times, in seconds, were recorded for a mouse to run through a maze:

34, 12, 10, 11, 12, 13, 9

a) What are the median and mean times?

b) Which measure of central tendency best represents the data?

c) Identify any possible outlier(s). Should the outlier(s) be removed from the data set? Explain why or why not.

d) How would removing the outlier(s) affect the median and the mean?

Extended Response

12. Laura tracks the number of hours she spends playing games on the computer each night for one week:

2.5, 2.0, 1.0, 2.5, 2.0, 15, 3.5

a) What is the range?

b) What are the mode, median, and mean?

c) Identify any possible outlier(s). Should the outlier(s) be removed from the data set? Explain why or why not.

d) Which measure of central tendency changes the most when the outlier is removed? By how much does it change?

13. For every soccer practice, Cam keeps track of the number of goals he makes in ten attempts. In the last ten practices, he made the following numbers of goals:

2, 6, 8, 7, 8, 9, 7, 8, 6, 8

a) What is the range?

b) What is the median as a percent?

c) What is the mean as a percent?

d) What is a possible explanation for Cam's low score on the first day?

e) How would the mean and median change if Cam did not include his low score of 2?

f) If Cam wishes to use the mean to describe his performance, should he include his low score? Explain.
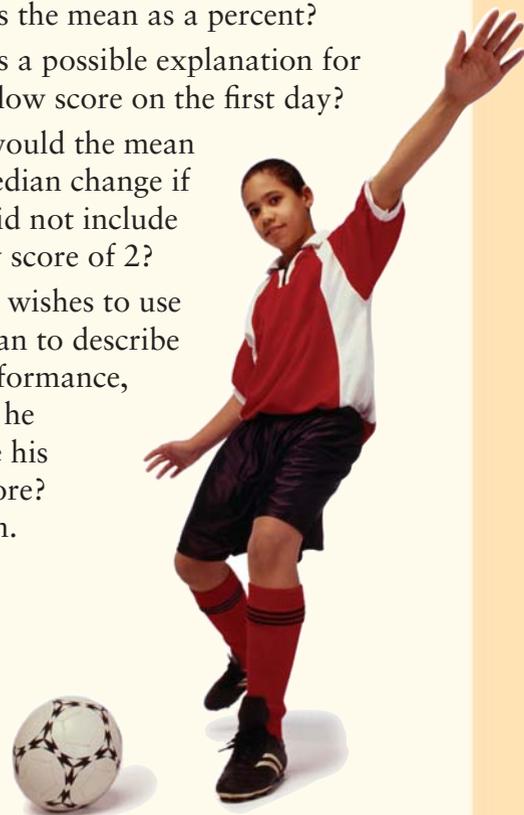
WRAP IT UP!

Collect at least ten pieces of numerical data about a topic of your choice. Analyse the data to show what you know about working with data. Your report should include:
- a description of why you collected this data set and how you collected it
- the range, mean, median, and mode
- an explanation of which measure of central tendency best represents the data
- a description of any possible outliers
- an explanation for whether any outliers should be used in determining the measures of central tendency

Math Games

Data Duel

Play Data Duel with a partner. These are the rules:

- Each player rolls two dice, adds the results, and records the total. Each player rolls 9 more times to make a list of 10 totals.
- Either player spins the Data Duel spinner.
- Both players use their list of 10 totals to determine the measure shown on the spinner.
- The players compare the measures and check each other's reasoning. The higher measure scores a point for this round.
- If the measures are tied, both players score a point for the round.
- There are some extra rules if players are finding the mode:
  - If both players find that there is no mode, the round is tied and both players score a point.
  - A player who finds two or more modes should compare the highest mode with the other player's mode.
  - Any mode value beats finding no mode.
- Play more rounds using a new list of 10 totals for each round. The first player to reach 5 points wins. If both players reach 5 points in the same round, play more rounds until one player pulls ahead. This player wins.

Materials
- 4 six-sided dice
- Data Duel spinner
- paper clip
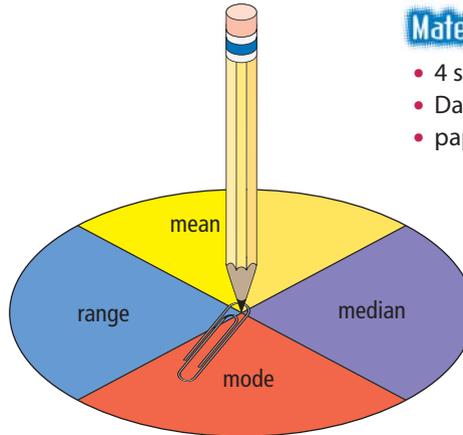
The spinner stopped on "mode," so I find the mode of my 10 totals.

We found the mode. My mode is 9 and my partner's mode is 6, so I score a point.

Managing a Business

Before buying items to sell in their stores, store owners collect and analyse data about their customers and their needs in order to predict sales and place their orders.

You be the store owner!

You are going to open a shoe store for teens and need the following information about the teens in your area:
• What sizes of shoes are worn by teens between 12 and 14 years old?
• What styles of shoes are most popular? Consider basketball, trainer, skateboarding, tennis, and so on.
• How much are teens willing to pay for a pair of shoes?

a) Collect and organize data from your class on shoe sizes, style preferences, and prices.

b) Analyse the data to help you decide what you need to order. Include measures of central tendency, such as mean, median, and mode, as appropriate.

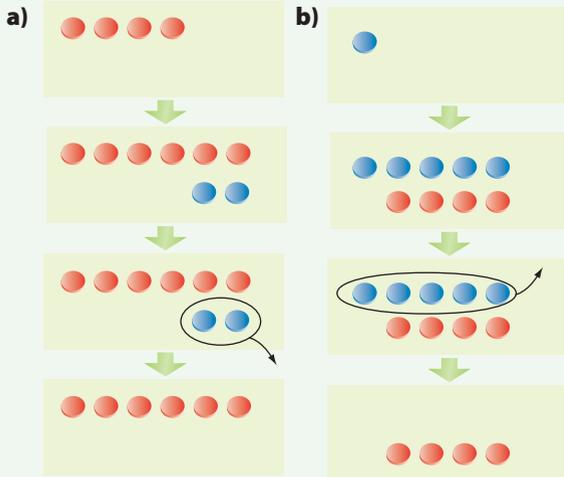**c)** Choose which measure of central tendency is best for each set of results. Explain your choices.

Chapters 9–12 Review

Chapter 9 Add and Subtract Integers

1. What addition statement does each diagram represent?

a)

b)
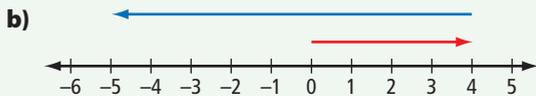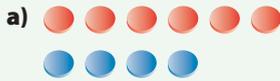
2. What subtraction statement does each diagram represent?

a) **b)**

3. What is each sum or difference?

a) $(+5) + (-5)$ **b)** $(+7) + (-4)$

c) $(-9) - (-7)$ **d)** $(+8) - (-8)$

e) $(+3) - (+7)$ **f)** $(-2) - (-9)$

4. One October morning in Prince Edward, Saskatchewan, the temperature was $-3°C$. In the late afternoon, the temperature was $12°C$ higher. What was the temperature in the late afternoon?

5. The surface of Great Bear Lake is 156 m above sea level. The bottom of the lake is 257 m below sea level. Estimate and calculate the depth of the lake.

Chapter 10 Patterns and Expressions

6. Describe each pattern. What are the next two numbers in each pattern?

a) 1, 4, 7, 10, …

b) 8, 13, 18, 23, 26 …

c) 17, 14, 11, 8, …

7. Look at the following number pattern.

$$\frac{2}{30} = 0.0666..., \text{ or } 0.0\overline{6}$$

$$\frac{5}{30} = 0.1666..., \text{ or } 0.1\overline{6}$$

$$\frac{8}{30} = 0.2666..., \text{ or } 0.2\overline{6}$$

a) Describe the pattern.

b) What is the next fraction in the pattern?

c) What is $\frac{17}{30}$ written as a repeating decimal using bar notation?

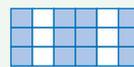**8.** The tile pattern shown is being used to cover a bathroom floor.

Figure 1   Figure 2   Figure 3

a) Make a table showing the number of white and blue tiles in the first five designs.

b) Describe the pattern for the number of blue tiles in relation to the number of white tiles.

c) Choose a variable and tell what it represents. Then use the variable to write an algebraic expression for the number of blue tiles.

d) How many blue tiles will there be if there are 24 white tiles?

9. The expression $2n + 4$ represents the perimeter of any figure in this pattern, where n is the figure number.

Figure 1   Figure 2   Figure 3

a) What would be the perimeter of Figure 12?

b) Make a table of values for the first six figures in the pattern.

10. The graph shows the number of roller coaster riders allowed on a roller coaster train, depending on the number of cars that make up the train.
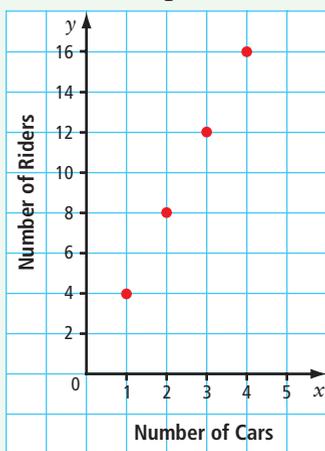

a) Make a table of values for the first five values of x starting at $x = 1$.

b) What is an algebraic expression for the number of riders in relation to the number of cars?

c) Describe the pattern of points on the graph in two different ways.

d) If there are 10 cars in the roller coaster train, how many riders are allowed on the train?

Chapter 11 Solving Equations

11. The diagram represents an equation.
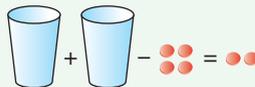
a) What are the two expressions that make up this equation?

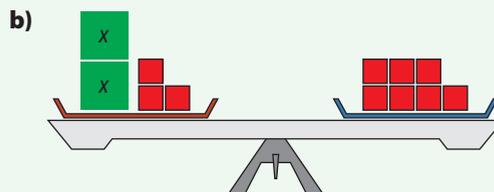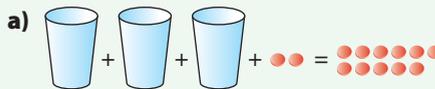**b)** What is the equation?

12. Solve by inspection. Verify your answer.

a) $k - 7 = 19$ **b)** $p + 12 = 12$

c) $2n = 18$ **d)** $\frac{c}{8} = 3$

13. Solve the equation modelled by each diagram. Check your solution.

a)

b)

14. The formula for the perimeter of an equilateral triangle is $P = 3s$. What side length is needed to make an equilateral triangle with a perimeter of 48 cm?

15. An adventure company charges $95 per day for canoeing equipment plus $10 per student for food. The total cost for one day can be modelled using the equation $C = 10n + 95$.

a) What do the variables C and n represent?

b) Students in one class raised $345 for a one-day trip. How many students can go?

Chapter 12 Working With Data

16. The following numbers of haircuts were given at a salon in the last five days:

25, 29, 25, 26, 35.

  a) What is the mode?

  b) What is the median?

  c) What is the mean?

17. The mode is 6 for the set of numbers 3, 4, 4, 5, 6, 6, 7, 8, 9, 10, x, and y. What are possible whole number values for x and y?

18. Robert takes his dog for a walk six days a week. The following times indicate how long they walked last week:

54 min, 56 min, 60 min, 58 min, 55 min, 28 min

  a) What is the range?

  b) Which time may be an outlier?

  c) Why might this value be so different from the others?

  d) If you remove the outlier, what is the new range?

19. Steven scored 338 points in 26 basketball games this fall. Micha played in 8 fewer games than Steven. Her mean score per game was 0.5 points higher than Steven's. How many total points did Micha score in her basketball season?

20. Thirty students wrote a science test. Twenty-five students achieved a mean mark of 74%. The other five students had a mean mark of 45%. What was the class mean mark? Express your answer as a percent.

21. Melissa found the following prices for five different brands of orange juice in the refrigerated section at the grocery store:

$3.29, $2.99, $3.49, $6.98, and $3.79.

  a) What is the range?

  b) What are the median and the mean?

  c) Which is the best measure of central tendency for the data?

  d) Identify any possible outlier(s). Should the outlier(s) be removed from the data set? Explain why or why not.

  e) How would removing the outlier(s) affect the median and the mean?

22. Roach-Away developed a new chemical to destroy cockroaches. Ten tests were performed. The percent of roaches destroyed were:

60, 99, 90, 99, 70, 91, 88, 71, 69, 99

If you were the owner of the company, which measure of central tendency would you use for advertising? Explain your choice.

The Birthday Party

1. Students in two classes were asked how many friends were invited to their last birthday party.

Class A: 5, 10, 10, 6, 8, 9, 6, 5, 4, 6, 8
Class B: 11, 9, 13, 11, 18, 12, 14, 18, 16, 15, 11, 14

  a) Find the mean, median, and mode for each class.

  b) What is the range for each class?

2. You are planning a birthday party. You can choose from the following three activities. Costs include all taxes.

  A Bowling: Cost is $12 per person. This includes the cost of food, drinks, shoe rental, and 1 h of bowling.

  B Swimming: Cost is $8 per person plus $55 for renting the pool.

  C Movie: Cost is $10 per person plus $30 for the party room.

In all cases, the birthday person is free.

You want to invite at least 8 friends. The total cost cannot be more than $200.

Model the cost of each activity using an algebraic equation.

3. Which party would you choose? How many friends can you take? What is the total cost?

4. Explain the decisions you made in #3. Justify your response using at least two of a table of values, equations, or graphs.