SIMPLE PROBABILITY

Outcome: Any possible or actual result of the action considered, such as rolling a 5 on a standard number cube or getting tails when flipping a coin.

Event: A desired (or successful) outcome or group of outcomes from an experiment, such as rolling an even number on a standard number cube.

Sample space: All possible outcomes of a situation. For example, the sample space for flipping a coin is heads and tails; rolling a standard number cube has six possible outcomes (1, 2, 3, 4, 5, and 6).

Probability: The likelihood that an event will occur. Probabilities may be written as fractions, decimals, or percents. An event that is guaranteed to happen has a probability of 1, or 100%. An event that has no chance of happening has a probability of 0, or 0%. Events that "might happen" have probabilities between 0 and 1 or between 0% and 100%. In general, the more likely an event is to happen, the greater its probability.

Experimental probability: The probability based on data collected in experiments.

$$\text{Experimental probability} = \frac{\text{number of successful outcomes in the experiment}}{\text{total number of outcomes in the experiment}}$$

Theoretical probability is a calculated probability based on the possible outcomes when they all have the same chance of occurring.

$$\text{Theoretical probability} = \frac{\text{number of successful outcomes (events)}}{\text{total number of possible outcomes}}$$

In the context of probability, "successful" usually means a desired or specified outcome (event), such as rolling a 2 on a number cube (probability of $\frac{1}{6}$). To calculate the probability of rolling a 2, first figure out how many possible outcomes there are. Since there are six faces on the number cube, the number of possible outcomes is 6. Of the six faces, only one of the faces has a 2 on it. Thus, to find the probability of rolling a 2, you would write:

$$P(2) = \frac{\text{number of ways to roll 2}}{\text{number of possible outcomes}} = \frac{1}{6}. \text{ or } 0.1\overline{6} \text{.or approximately } 16.7\%$$

Example 1

If you roll a fair, 6-sided number cube, what is $P(3)$, that is, the probability that you will roll a 3?

Because the six sides are equally likely to come up, and there is only one 3, $P(3) = \frac{1}{6}$.

Example 2

There are 12 marbles in a bag: 2 clear, 4 green, 5 yellow, and 1 blue. If one marble is chosen randomly from the bag, what is the probability that it will be yellow?

$$P(\text{yellow}) = \frac{5 \text{ (yellow)}}{12 \text{ (outcomes)}} = \frac{5}{12}$$

Example 3

Joe flipped a coin 50 times. When he recorded his tosses, his result was 30 heads and 20 tails. Joe's activity provided data to calculate experimental probability for flipping a coin.

a. What is the theoretical probability of Joe flipping heads?

  The theoretical probability is 50% or $\frac{1}{2}$, because there are only two possibilities (heads and tails), and each is equally likely to occur.

b. What was the experimental probability of flipping a coin and getting heads based on Joe's activity?

  The experimental probability is $\frac{30}{50}$, $\frac{3}{5}$, or 60%. These are the results Joe actually got when he flipped the coin.

Example 4

Decide whether these statements describe theoretical or experimental probabilities.

a. The chance of rolling a 6 on a fair die is $\frac{1}{6}$.

  This statement is theoretical.

b. I rolled the die 12 times and 5 came up three times.

  This statement is experimental.

c. There are 15 marbles in a bag; 5 blue, 6 yellow, and 4 green. The probability of getting a blue marble is $\frac{1}{3}$.

  This statement is theoretical.

d. When Veronika pulled three marbles out of the bag she got 2 yellow and 1 blue, or $\frac{2}{3}$ yellow, $\frac{1}{3}$ blue.

  This statement is experimental.

Problems

1. There are 24 crayons in a box: 5 black, 3 white, 7 red, 2 yellow, 3 blue, and 4 green.
    What is the probability of randomly choosing a green? Did you respond with an
    experimental or theoretical probability?

2. A spinner is divided into four equal sections numbered 2, 4, 6, and 8. What is the
    probability of spinning an 8?

3. A fair number cube marked 1, 2, 3, 4, 5, and 6 is rolled. Tyler tossed the cube 40 times,
    and noted that 26 times an even number showed. What is the experimental probability that
    an even number will be rolled? What is the theoretical probability?

4. Sara is at a picnic and reaches into an ice chest, without looking, to grab a can of soda.
    If there are 14 cans of orange, 12 cans of fruit punch, and 10 cans of cola, what is the
    probability that she takes a can of fruit punch? Did you respond with an experimental
    probability or a theoretical one?

5. A baseball batting average is the probability a baseball player hits the ball when batting. If
    a baseball player has a batting average of 266, it means the player's probability of getting
    of getting a hit is 0.266. Is a batting average an experimental probability or theoretical?

6. In 2011, 39 people died by being struck by lightning, and 241 people were injured. There
    were 310,000,000 people in the United States. What is the probability of being one of the
    people struck by lightning?

7. In a medical study, 107 people were given a new vitamin pill. If a participant got sick, they
    were removed from the study. Ten of the participants caught a common cold, 2 came down
    with the flu, 18 got sick to their stomach, and 77 never got sick. What was the probability
    of getting sick if you participated in this study? Did you respond with an experimental
    probability or a theoretical one?

8. Insurance companies use probabilities to determine the rate they will charge for an
    insurance policy. In a study of 300 people that had life insurance policies, an insurance
    company found that 111 people were over 80 years old when they died, 82 people died
    when they were between 70 and 80 years old, 52 died between 60 and 70 years old, and 55
    died when they were younger than 60 years old. In this study what was the probability of
    dying younger than 70 years old? Did you respond with an experimental probability or a
    theoretical one?

Answers

1. $\frac{1}{6}$; theoretical   2. $\frac{1}{4}$ 3. $\frac{26}{40}$; $\frac{3}{6}$ 4. $\frac{1}{3}$; theoretical

5. experimental   6. $\frac{39+241}{310,000,000}$ 7. $\frac{10+2+18}{107} \approx 0.28$ 8. $\frac{55+52}{300} \approx 35.7\%$

$\approx 0.000\ 000\ 903$ experimental   experimental

INDEPENDENT AND DEPENDENT EVENTS

Two events are **independent** if the outcome of one event does not affect the outcome of the other event. For example, if you draw a card from a standard deck of playing cards but replace it before you draw again, the outcomes of the two draws are independent.

Two events are **dependent** if the outcome of one event affects the outcome of the other event. For example, if you draw a card from a standard deck of playing cards and do not replace it for the next draw, the outcomes of the two draws are dependent.

Example 1

Juan pulled a red card from the deck of regular playing cards. This probability is $\frac{26}{52}$ or $\frac{1}{2}$. He puts the card back into the deck. Will his chance of pulling a red card next time change?

No, his chance of pulling a red card next time will not change, because he replaced the card. There are still 26 red cards out of 52. This is an example of an independent event; his pulling out and replacing a red card does not affect any subsequent selections from the deck.

Example 2

Brett has a bag of 30 multi-colored candies. 15 are red, 6 are blue, 5 are green, 2 are yellow, and 2 are brown. If he pulls out a yellow candy and eats it, does this change his probability of pulling any other candy from the bag?

Yes, this changes the probability, because he now has only 29 candies in the bag and only 1 yellow candy. Originally, his probability of yellow was $\frac{2}{30}$ or $\frac{1}{15}$; it is now $\frac{1}{29}$. Similarly, red was $\frac{15}{30}$ or $\frac{1}{2}$ and now is $\frac{15}{29}$, better than $\frac{1}{2}$. This is an example of a dependent event.

Problems

Decide whether these events are independent or dependent events.

1. Flipping a coin, and then flipping it again.

2. Taking a black 7 out of a deck of cards and not returning it, then taking out another card.

3. Taking a red licorice from a bag and eating it, then taking out another piece of licorice.

Answers

1. independent   2. dependent   3. dependent

COMPOUND PROBABILITY AND COUNTING METHODS

COMPOUND PROBABILITY

Sometimes when you are finding a probability, you are interested in either of two outcomes taking place, but not both. For example, you may be interested in drawing a king or a queen from a deck of cards. At other times, you might be interested in one event followed by another event. For example, you might want to roll a one on a number cube and then roll a six. The probabilities of combinations of simple events are called **compound events**
.

To find the probability of *either* one event *or* another event that has nothing in common with the first, you can find the probability of each event separately and then add their probabilities. Using the example above of drawing a king or a queen from a deck of cards:

$P(\text{king}) = \frac{4}{52}$ and $P(\text{queen}) = \frac{4}{52}$ so $P(\text{king or queen}) = \frac{4}{52} + \frac{4}{52} = \frac{8}{52} = \frac{2}{13}$

For two independent events, to find the probability of *both* one *and* the other event occurring, you can find the probability of each event separately and then multiply their probabilities. Using the example of rolling a one followed by a six on a number cube:

$$P(1) = \frac{1}{6} \text{ and } P(6) = \frac{1}{6} \text{ so } P(1 \text{ then } 6) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

Note that you would carry out the same computation if you wanted to know the probability of rolling a one on a green cube, and a six on a red cube, if you rolled both of them at the same time.

Example 1

A spinner is divided into five equal sections numbered 1, 2, 3, 4, and 5. What is the probability of spinning *either* a 2 *or* a 5?
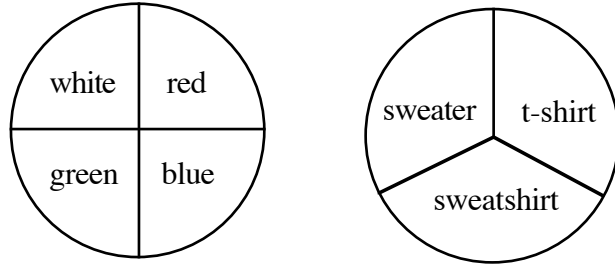
Step 1:   Determine both probabilities: $P(2) = \frac{1}{5}$ and $P(5) = \frac{1}{5}$

Step 2:   Since these are *either-or* compound events, add the fractions describing each probability: $\frac{1}{5} + \frac{1}{5} = \frac{2}{5}$

The probability of spinning a 2 or a 5 is $\frac{2}{5}$: $P(2 \text{ or } 5) = \frac{2}{5}$

Example 2

If each of the regions in each spinner at right is the same size, what is the probability of spinning each spinner and getting a green t-shirt?

Step 1: Determine both possibilities:
$P(\text{green}) = \frac{1}{4}$ and $P(\text{t-shirt}) = \frac{1}{3}$

Step 2: Since you are interested in the compound event of *both* green *and* a t-shirt, multiply both probabilities: $\frac{1}{4} \cdot \frac{1}{3} = \frac{1}{12}$

The probability of spinning a green t-shirt is $\frac{1}{12}$: $P(\text{green t-shirt}) = \frac{1}{12}$

Problems

Assume in each of the problems below that events are independent of each other.

1. One die, numbered 1, 2, 3, 4, 5, and 6, is rolled. What is the probability of rolling *either* a 1 *or* a 6?

2. Mary is playing a game in which she rolls one die and spins a spinner. What is the probability she will get *both* the 3 *and* black she needs to win the game?
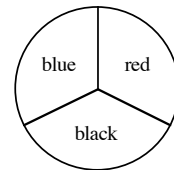
3. A spinner is divided into eight equal sections. The sections are numbered 1, 2, 3, 4, 5, 6, 7, and 8. What is the probability of spinning a 2, 3, or a 4?

4. Patty has a box of 12 colored pencils. There are 2 blue, 1 black, 1 gray, 3 red, 2 green, 1 orange, 1 purple, and 1 yellow in the box. Patty closes her eyes and chooses one pencil. She is hoping to choose a green or a red. What is the probability she will get her wish?

5. Use the spinners at right to tell Paul what his chances are of getting the silver truck he wants.
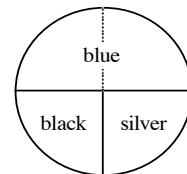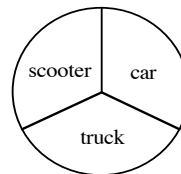
6. On the way to school, the school bus must go through two traffic signals. The first light is green for 25 seconds out of each minute, and the second light is green for 35 seconds out of each minute. What is the probability that both lights will be green on the way to school?

7. There are 250 students at South Lake Middle School. 125 enjoy swimming, 50 enjoy skateboarding, and 75 enjoy playing softball. What is the probability a student enjoys all three sports?

8. John has a bag of jellybeans. There are 100 beans in the bag. $\frac{1}{4}$ of the beans are cherry, $\frac{1}{4}$ of the beans are orange, $\frac{1}{4}$ of the beans are licorice, and $\frac{1}{4}$ of the beans are lemon. What is the probability that John will chose one of his favorite flavors, orange, or cherry?

9. A nationwide survey showed that only 4% of children liked eating lima beans. What is the probability that any two children will both like lima beans?

Answers

1. $\frac{2}{6}$ or $\frac{1}{3}$

2. $\frac{1}{18}$

3. $\frac{3}{8}$

4. $\frac{5}{12}$

5. $\frac{1}{12}$

6. $\frac{25}{60} \cdot \frac{35}{60} \approx 0.243$

7. $\frac{125}{250} \cdot \frac{50}{250} \cdot \frac{75}{250} = \frac{3}{100}$

8. $\frac{2}{4}$ or $\frac{1}{2}$

9. $\frac{1}{16}$

COUNTING METHODS

There are several different models you can use to determine all possible outcomes for compound events when *both* one event *and* the other occur: a systematic list, a probability table, and a probability tree. See the Math Notes box in Lesson 5.5.2 of the *Core Connections, Course 2* text for details on these three methods.

Not only can you use a **probability table** to help list all the outcomes, but you can also use it to help you determine probabilities of independent compound events when *both* one event *and* the other occur. For example, the following probability table (sometimes called an **area model**) helps determine the probabilities from Example 2 above:

	$\frac{1}{4}$ white	$\frac{1}{4}$ red	$\frac{1}{4}$ blue	$\frac{1}{4}$ green
$\frac{1}{3}$ sweater				
$\frac{1}{3}$ sweatshirt				
$\frac{1}{3}$ t-shirt				■

Each box in the rectangle represents the compound event of *both* a color *and* the type of clothing (sweater, sweatshirt, or t-shirt). The area of each box represents the probability of getting each combination. For example, the shaded region represents the probability of getting a green t-shirt: $\frac{1}{4} \cdot \frac{1}{3} = \frac{1}{12}$.

Example 3

At a class picnic Will and Jeff were playing a game where they would shoot a free throw and then flip a coin. Each boy only makes one free throw out of three attempts. Use a probability table (area model) to find the probability that one of the boys makes a free throw, and then flips a head. What is the probability that they miss the free throw and then flip tails?

	H	T
Make		
Miss		
Miss		

By finding the area of the small rectangles, the probabilities are:
P(make *and* heads) $= \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$, and P(miss *and* tails) $= \frac{2}{3} \cdot \frac{1}{2} = \frac{2}{6}$

Example 4

Chris owns a coffee cart that he parks outside the downtown courthouse each morning. 65% of his customers are lawyers; the rest are jury members. 60% of Chris's sales include a muffin, 10% include cereal, and the rest are coffee only. What is the probability a lawyer purchases a muffin or cereal?

The probabilities could be represented in an area model as follows:

	lawyer 0.65	jury 0.35
muffin 0.60		
cereal 0.10		
coffee only 0.30		

Probabilities can then be calculated:

	lawyer 0.65	jury 0.35
muffin 0.60	0.39	0.21
cereal 0.10	0.065	0.035
coffee only 0.30	0.195	0.105

The probability a lawyer purchases a muffin or cereal is
0.39 + 0.065 = 0.455 or 45.5%.

Example 5

The local ice cream store has choices of plain, sugar, or waffle cones. Their ice cream choices are vanilla, chocolate, bubble gum, or frozen strawberry yogurt. The following toppings are available for the ice cream cones: sprinkles, chocolate pieces, and chopped nuts.

What are all the possible outcomes for a cone *and* one scoop of ice cream *and* a topping? How many outcomes are possible?

Probability tables are useful only when there are two events. In this situation there are three events (cone, flavor, topping), so we will use a probability tree.

There are four possible flavors, each with three possible cones. Then each of those 12 outcomes can have three possible toppings. There are 36 outcomes for the compound event of choosing a flavor, cone, *and* topping.

Note that the list of outcomes, and the total number of outcomes, does not change if we change the order of events. We could just as easily have chosen the cone first.
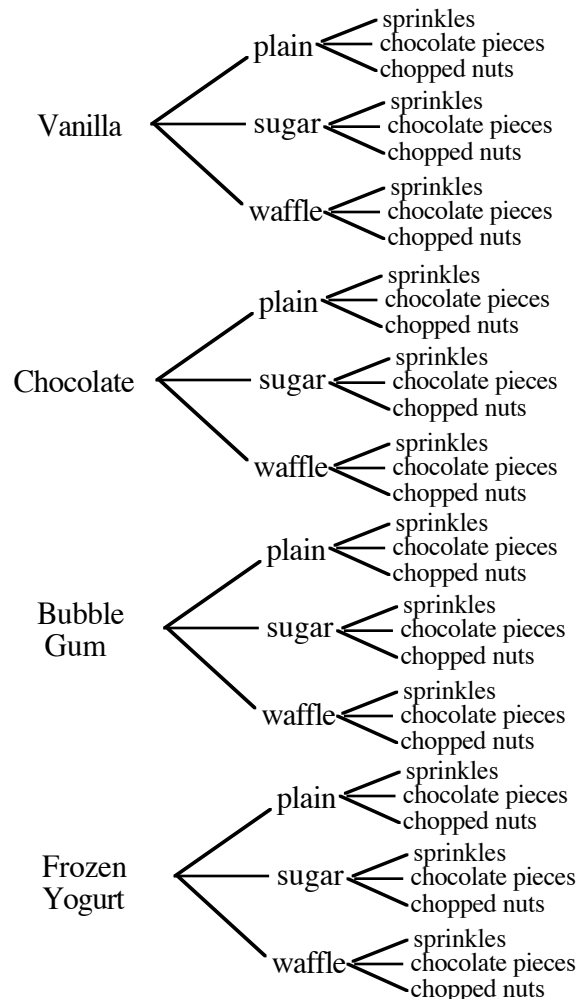
Problems

Use probability tables or tree diagrams to solve these problems.

1. How many different combinations are possible when buying a new bike if the following options are available:

    •   mountain bike or road bike

    •   black, red, yellow, or blue paint

    •   3–speed, 5–speed, or 10–speed

2. A new truck is available with:

    •   standard or automatic

                   transmission

    •   2–wheel or 4–wheel drive

    •   regular or king cab

    •   long or short bed

    How many combinations are possible?

3. A tax assessor categorizes 25% of the homes in how city as having a large backyard, 65% as having a small backyard, and 10% as having no backyard. 30% of the homes have a tile roof, the rest have some other kind of roof. What is the probability a home with a tile roof has a backyard?
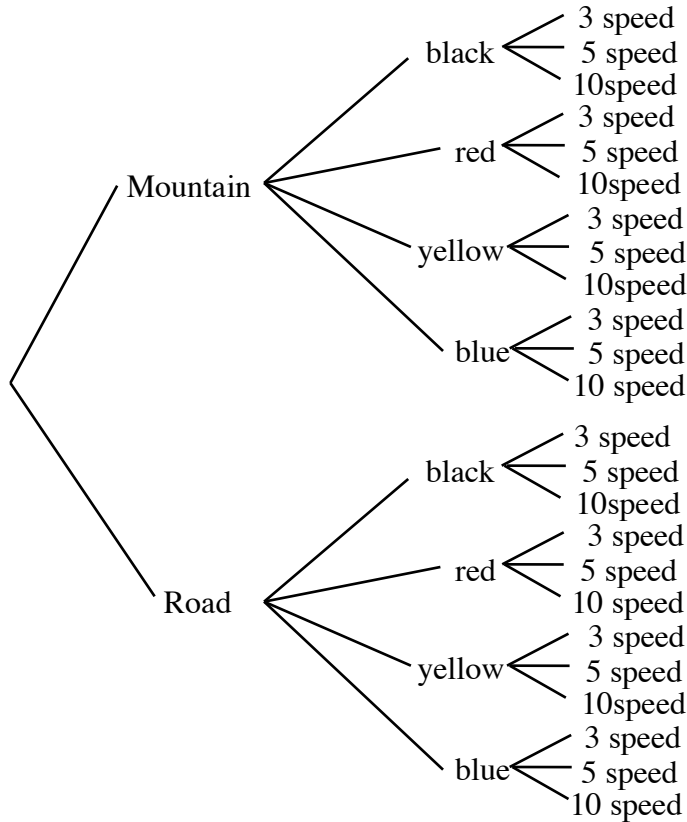
4. There is space for only 96 students at University High School to enroll in a "shop" class: 25 students in woodworking, 25 students in metalworking, and the rest in print shop. Three-fourths of the spaces are reserved for seniors, and one-fourth are for juniors. What is the probability that a student enrolled in shop class is a senior in print shop? What is the probability that a student enrolled in shop class is a junior in wood or metal shop?

5. Insurance companies use probabilities to determine the rate they will charge for an insurance policy. In a study of 3000 people that had life insurance policies, an insurance company collected the following data of how old people were when they died, compared to how tall they were. In this study, what was the probability of being tall (over 6ft) and dying young under 50 years old? What was the probability of being tall and dying under 70 years old? What was the probability of being between 50 and 70 years old?
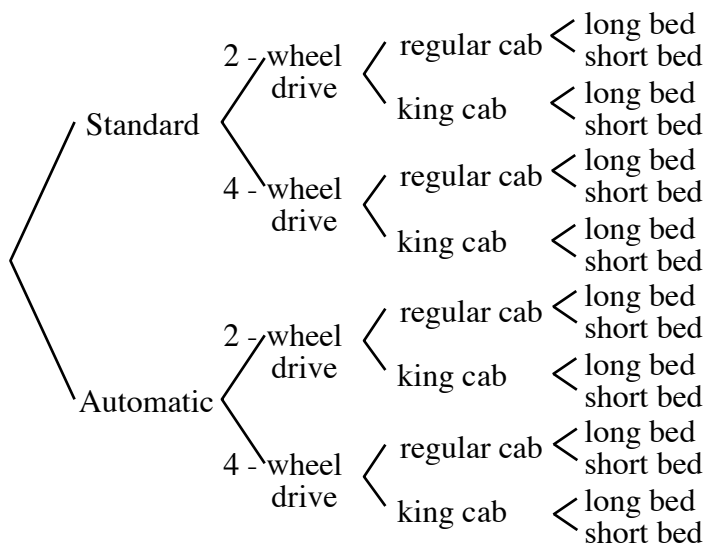
	<50 years old	50–60 yrs old	60–70 yrs old	70–80yrs old	>80 years old
over 6ft tall	30	25	52	82	111
under 6ft tall	270	225	468	738	999

Answers

1. There are 24 possible combinations as shown below.

```
                                              3 speed
                                  black       5 speed
                                              10speed
                                              3 speed
                                  red          5 speed
                 Mountain                     10speed
                                              3 speed
                                  yellow       5 speed
                                              10speed
                                              3 speed
                                  blue         5 speed
                                              10 speed

                                              3 speed
                                  black        5 speed
                                              10speed
                                              3 speed
                                  red          5 speed
                 Road                         10 speed
                                              3 speed
                                  yellow       5 speed
                                              10speed
                                              3 speed
                                  blue         5 speed
                                              10 speed
```

2. There are 16 possible combinations as shown below.

```
                                                    long bed
                          2 - wheel    regular cab  short bed
                          drive                     long bed
             Standard                   king cab    short bed
                                                    long bed
                          4 - wheel    regular cab  short bed
                          drive                     long bed
                                        king cab    short bed

                                                    long bed
                          2 - wheel    regular cab  short bed
                          drive                     long bed
             Automatic                  king cab    short bed
                                                    long bed
                          4 - wheel    regular cab  short bed
                          drive                     long bed
                                        king cab    short bed
```

3. The probability is $0.075 + 0.2275 = 0.3025$ or 30.25%.

	large yard 25%	small yard 65%	no yard 10%
tile roof 30%	0.075	0.2275	
other roof 70%			

4. The probability of a senior in print shop is about 0.359%. The probability of a junior in wood or metal shop is $0.065 + 0.065 \approx 0.13$.

	seniors $\frac{3}{4}$	juniors $\frac{1}{4}$	
woodworking $\frac{25}{96}$		≈ 0.065	
metalworking $\frac{25}{96}$		≈ 0.065	
print shop $\frac{46}{96}$	≈ 0.359		

5. The probability of being tall (over 6ft) and dying young under 50 years old is $\frac{30}{3000} = 0.01$. The probability of being tall and dying under 70 years old is $\frac{30+25+52}{3000} \approx 0.036$. The probability of being between 50 and 70 years old is $\frac{25+52+225+468}{3000} \approx 0.257$.

CIRCLE GRAPHS

A circle graph (or pie chart) is a diagram that represents proportions of categorized data as parts of a circle. Each sector or wedge represents a percent or fraction of the circle. The fractions or percents must total 1, or 100%. Since there are 360 degrees in a circle, the size of each sector (in degrees) is found by multiplying the fraction or percent by 360 degrees.

For additional information, see the Math Notes box in Lesson 7.1.1 of the *Core Connections, Course 3* text.

Example 1

Ms. Sallee's class of 30 students was surveyed about the number of hours of homework done each night and here are the results:

less than 1 hour	3 students
1 to 2 hours	9 students
2 to 3 hours	12 students
3 to 4 hours	4 students
more than 4 hours	2 students

The proper size for the sectors is found as follows:

< 1: $\frac{3}{30} \cdot 360° = 36°$; $1\text{–}2$: $\frac{9}{30} \cdot 360° = 108°$

$2\text{–}3$: $\frac{12}{30} \cdot 360° = 144°$; $3\text{–}4$: $\frac{4}{30} \cdot 360° = 40°$

> 4: $\frac{2}{30} \cdot 360° = 20°$
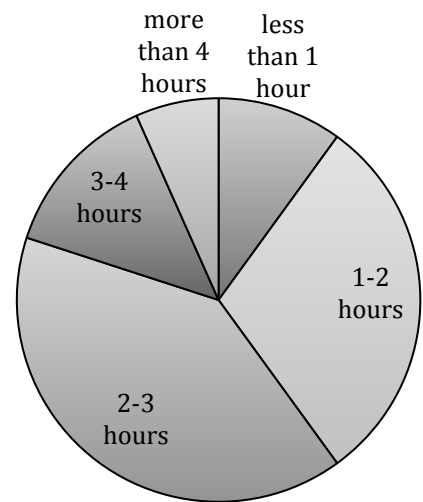
The circle graph is shown at right.

Example 2

The 800 students at Central Middle School were surveyed to determine their favorite school lunch item. The results are shown below.
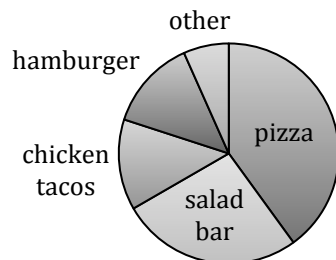
Use the circle graph at left to answer each question.

a.   Which lunch item was most popular?

b.   Approximately how many students voted for the salad bar?

c.   Which two lunch items appear to have equal popularity?

Answers: a. pizza is the largest sector; b. $\frac{1}{4} \cdot 800 = 200$;
c. hamburger and chicken tacos have the same size sectors

Problems

For problems 1 through 3 use the circle graph at right. The graph shows the results of the 1200 votes for prom queen.
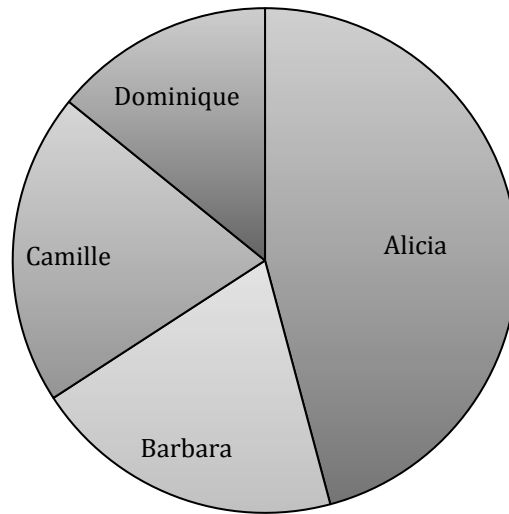
1. Who won the election?

2. Did the person who won the election get more than half of the votes?

3. Approximately how many votes did Camille receive?

4. Of the milk consumed in the United States, 30% is whole, 50% is low fat, and 20% is skim. Draw a circle graph to show this data.

5. On an average weekday, Sam's time is spent as follows: sleep 8 hours, school 6 hours, entertainment 2 hours, homework 3 hours, meals 1 hour, and job 4 hours. Draw a circle graph to show this data.

6. Records from a pizza parlor show the most popular one-item pizzas are: pepperoni 42%, sausage 25%, mushroom 10%, olive 9% and the rest were others. Draw a circle graph to show this data.

7. To pay for a 200 billion dollar state budget, the following monies were collected: income taxes 90 billion dollars, sales taxes 74 billion dollars, business taxes 20 billion dollars, and the rest were from miscellaneous sources. Draw a circle graph to show this data.

8. Greece was the host country for the 2004 Summer Olympics. The Greek medal count was 6 gold, 6 silver, and 4 bronze. Draw a circle graph to show this data.

Answers

1.   Alicia

2.   No Alicia's sector is less than half of a circle.

3.   Approximately 240 votes

4.

skim whole low fat

5.

job meals homework entertainment sleep school

6.

other olive mushroom sausage pepperoni

7.

business taxes misc. sales taxes income taxes

8.

bronze gold silver
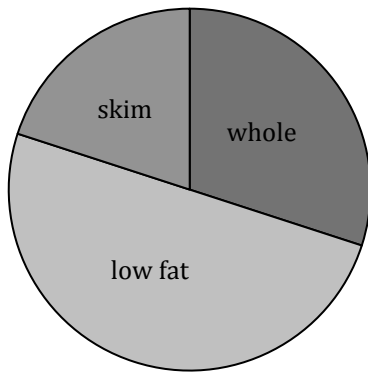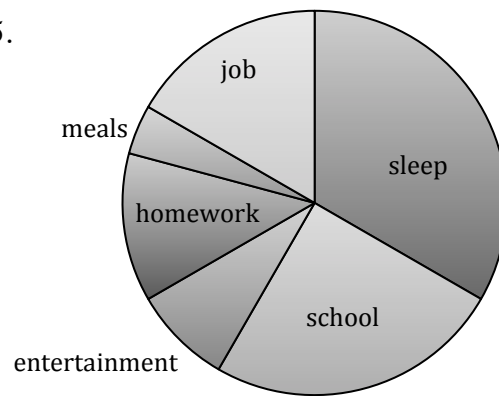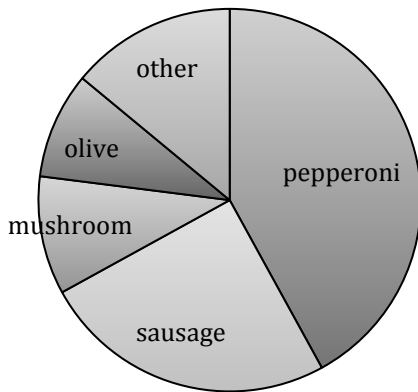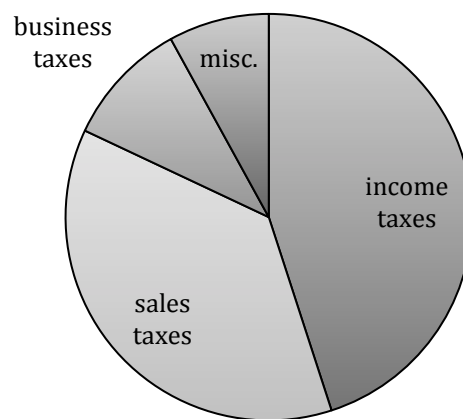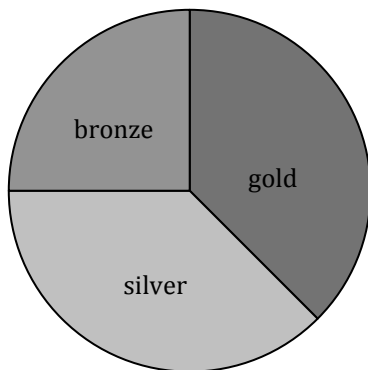
MEASURES OF CENTRAL TENDENCY

Measures of central tendency are numbers that locate or approximate the "center" of a set of data—that is, a "typical" value that describes the set of data. Mean and median are the most common measures of central tendency. (Mode will not be covered in this course.)

The mean is the arithmetic average of a data set. Add all the values in a set and divide this sum by the number of values in the set. The median is the middle number in a set of data arranged numerically. An outlier is a number that is much smaller or larger than most of the others in the data set. The range of a data set is the difference between the highest and lowest values of the data set.

For additional information, see the Math Notes boxes in Lesson 8.1.2 of the *Core Connections, Course 1* text or Lessons 1.1.3 and 1.1.4 of the *Core Connections, Course 2* text.

The mean is calculated by finding the sum of the data set and dividing it by the number of elements in the set.

Example 1

Find the mean of this set of data: 34, 31, 37, 44, 38, 34, 42, 34, 43, and 41.

- $34 + 31 + 37 + 44 + 38 + 34 + 42 + 34$
  $+ 43 + 41 = 378$

- $\frac{378}{10} = 37.8$

The mean of this set of data is 37.8.

Example 2

Find the mean of this set of data: 92, 82, 80, 92, 78, 75, 95, and 77.

- $92 + 82 + 80 + 92 + 78 + 75 + 95$
  $+ 77 + 77 = 748$

- $\frac{748}{9} = 83.1$

The mean of this set of data is 83.1.

Problems

Find the mean of each set of data.

1. 29, 28, 34, 30, 33, 26, and 34.

2. 25, 34, 35, 27, 31, and 30.

3. 80, 89, 79, 84, 95, 79, 78, 89, 76, 82, 76, 92, 89, 81, and 123.

4. 116, 104, 101, 111, 100, 107, 113, 118, 113, 101, 108, 109, 105, 103, and 91.

The median is the middle number in a set of data *arranged in numerical order*. If there is an even number of values, the median is the mean (average) of the two middle numbers.

Example 3

Find the median of this set of data: 34, 31, 37, 44, 38, 34, 43, and 41.

- Arrange the data in order: 31, 34, 34, 34, 37, 38, 41, 43, 44.

- Find the middle value(s): 37 and 38.

- Since there are two middle values, find their mean: $37 + 38 = 75$, $\frac{75}{2} = 37.5$. Therefore, the median of this data set is 37.5.

Example 4

Find the median of this set of data: 92, 82, 80, 92, 78, 75, 95, 77, and 77.

- Arrange the data in order: 75, 77, 77, 78, 80, 82, 92, 92, and 95.

- Find the middle value(s): 80. Therefore, the median of this data set is 80.

Problems

Find median of each set of data.

5. 29, 28, 34, 30, 33, 26, and 34.

6. 25, 34, 27, 25, 31, and 30.

7. 80, 89, 79, 84, 95, 79, 78, 89, 76, 82, 76, 92, 89, 81, and 123.

8. 116, 104, 101, 111, 100, 107, 113, 118, 113, 101, 108, 109, 105, 103, and 91.

The range of a set of data is the difference between the highest value and the lowest value.

Example 5

Find the range of this set of data: 114, 109, 131, 96, 140, and 128.

- The highest value is 140.

- The lowest value is 96.

- $140 - 96 = 44$.

- The range of this set of data is 44.

Example 6

Find the range of this set of data: 37, 44, 36, 29, 78, 15, 57, 54, 63, 27, and 48.

- The highest value is 78.

- The lowest value is 27.

- $78 - 27 = 51$.

- The range of this set of data is 51.

Problems

Find the range of each set of data in problems 5 through 8.

Outliers are numbers in a data set that are either much higher or much lower that the other numbers in the set.

Example 7

Find the outlier of this set of data: 88, 90 96, 93, 87, 12, 85, and 94.

- The outlier is 12.

Example 8

Find the outlier of this set of data: 67, 54, 49, 76, 64, 59, 60, 72, 123, 44, and 66.

- The outlier is 123.

Problems

Identify the outlier in each set of data.

9. 70, 77, 75, 68, 98, 70, 72, and 71.

10. 14, 22, 17, 61, 20, 16, and 15.

11. 1376, 1645, 1783, 1455, 3754, 1790, 1384, 1643, 1492, and 1776.

12. 62, 65, 93, 51, 55, 14, 79, 85, 55, 72, 78, 83, 91, and 76.

Answers

1. 30.57

2. $30.\overline{3}$

3. $86.\overline{13}$

4. $106.\overline{6}$

5. median 30; range 8

6. median 28.5; range 9

7. median 82; range 47

8. median 107; range 27

9. 98

10. 61

11. 3754

12. 14

GRAPHICAL REPRESENTATIONS OF DATA

Students represent distributions of single-variable data numerical data using dot plots, stem-and-leaf plots, box plots, and histograms. They represent categorical one-variable data on bar graphs. Each representation communicates information in a slightly different way.

STEM-AND-LEAF-PLOTS

A stem-and-leaf plot is a way to display data that shows the individual values from a set of data and how the values are distributed. The "stem" part on the graph represents all of the digits except the last one. The "leaf" part of the graph represents the last digit of each number.

Read more about stem-and-leaf plots, and how they compare to dot plots and histograms, in the Math Notes in Lesson 2.1.2 and 2.2.1 in the *Core Connections, Course 1* text, and in the Math Notes in Lesson 7.1.1 of the *Core Connections, Course 2* text.

Example 1

Make a stem-and-leaf plot of this set of data:
34, 31, 37, 44, 38, 29, 34, 42, 43, 34, 52, and 41.

```
2 | 9
3 | 1 4 4 4 7 8
4 | 1 2 3 4
5 | 2
```

Example 2

Make a stem-and-leaf plot of this set of data:
392, 382, 380, 392, 378, 375, 395, 377, and 377.

```
37 | 5 7 7 8
38 | 0 2
39 | 2 2 5
```

Problems

Make a stem-and-leaf plot of each set of data.

1. 29, 28, 34, 30, 33, 26, 18, and 34.

2. 25, 34, 27, 25, 19, 31, 42, and 30.

3. 80, 89, 79, 84, 95, 79, 89, 67, 82, 76, 92, 89, 81, and 123.

4. 116, 104, 101, 111, 100, 107, 113, 118, 113, 101, 108, 109, 105, 103, and 91.

HISTOGRAMS AND BAR GRAPHS

Histograms and bar graphs are visual ways to represent data. Both consist of vertical bars (called bins) with heights that represent the number of data points (called the frequency) in each bin.

Histograms are for displaying distributions of numerical data. In a histogram each bar represents the number of data elements within a certain range of values. All the bars touch each other. Values at the left side of a bin's range are included in that bin. Each range of values should have the same width.

Bar graphs are for displaying categorical data. In a bar graph each bar represents the number of data elements in a certain category. All the bars are the same width and are separated from each other.

For additional information and examples, see the Math Notes boxes in Lessons 2.1.2 and 2.2.1 of the *Core Connections, Course 1* text, or Lesson 7.1.1 of the *Core Connections, Course 2* text. For additional examples and practice, see the *Core Connections, Course 1* Checkpoint 9A materials or *Connections, Course 2* Checkpoint 7B materials in the back of those texts.

Example 3

The scores for a 25-point quiz are listed below arranged from least to greatest.

$7, 7, 12, 13, 15, 16, 16, 16, 18, 19, 20, 20, 20, 21, 21, 22, 23, 24$

Using intervals of five points, create a histogram for the class.

See histogram at right. Scores on the right end of the interval are included in the next interval. The interval between 10 and 15 only includes the two scores of 12 and 13. The interval between 15 and 20 only includes the six scores of 15, 16, 16, 16, 18, and 19.
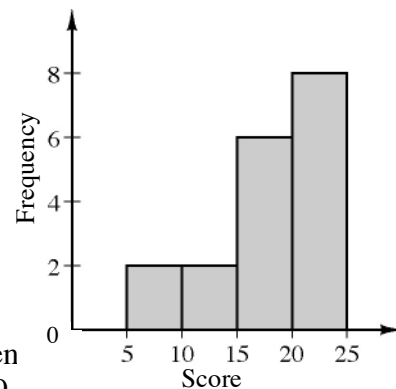
Example 4

Ms. Lim asked each of her students about their favorite kind of pet. Based on their responses, she drew the bar graph at right. Use the bar graph to answer each question.

   a. What is the favorite pet?

   b. How many students chose a bird as their favorite pet?

   c. What was the least favorite pet?

   d. If every student voted once, how many students are in the class?
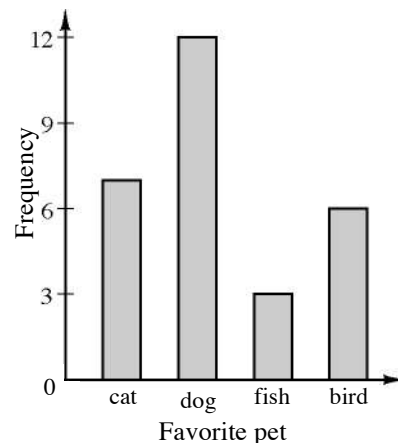
Answers: a. dog   b. 6   c. fish   d. 28

                                             Core Connections, Courses 1–3

Problems

5. Mr. Diaz surveyed his employees about the time it takes them to get to work. The results are shown in the histogram at right.

   a. How many employees completed the survey?

   b. How many employees get to work in less than 20 minutes?

   c. How many employees get to work in less than 40 minutes?

   d. How many employees take 60 minutes to get to work?
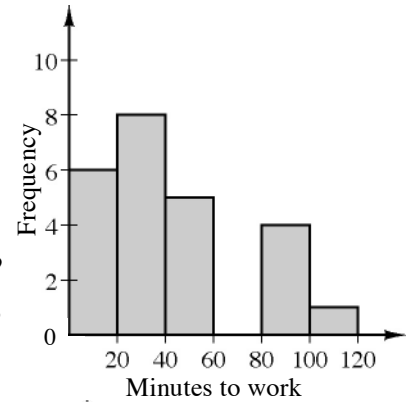
6. The two sixth grade classes at Vista Middle School voted for their favorite dessert. The results are shown in the bar graph at right for the five favorite choices.

   a. What was the favorite dessert and how many students made that choice?

   b. How many students selected cake as their favorite dessert?

   c. How many students selected yogurt as their favorite?

   d. How many more students selected ice cream than pudding?
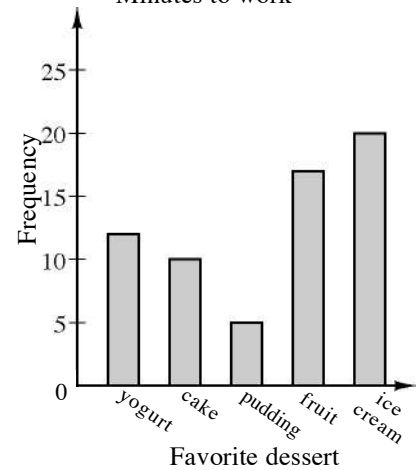
7. Mr. Fernandez asked 30 people at work how many pets they owned. The results are shown at right. Make a histogram to display this data. Use intervals of one pet.

0 pets	5 people
1 pet	8 people
2 pets	10 people
3 pets	3 people
4 pets	2 people
5 pets	1 person
9 pets	1 person

8. During the fist week of school Ms. Chan asked her students to name the county where they were born. There were so many different countries she grouped them by continent:

   North America: 14 students, South America: 2 students, Europe: 3 students, Asia: 10 students, Africa: 1 student, Australia: 0 students.

   Make a bar graph to display this information.

9. Three coins were tossed 20 times and the number of results that were "heads" each time is shown below:

   1, 1, 2, 0, 2, 3, 1, 2, 1, 2, 2, 1, 3, 2, 0, 1, 2, 0, 2, 1

   Make a histogram to show the results.

10. The physical education teacher at West Middle School asked the class about their favorite winter activity. Here were the results:

    reading: 8 students, ice skating: 4 students, skiing: 6 student, snowboarding: 11 students, computer activities: 14 students.

    Make a bar graph to show the results.

BOX PLOTS

Another way to display a distribution of one-variable numerical data is with a box plot. A box plot is the only display of data that clearly shows the median, quartiles, range, and outliers of a data set.

For additional information, see the Math Notes boxes in Lessons 8.1.4 and 8.1.5 of the *Core Connections, Course 1* text, or Lessons 7.1.2 and 7.1.5 of the *Core Connections, Course 2* text. For additional examples and practice, see the *Core Connections, Course 1* Checkpoint 9A materials or the *Core Connections, Course 2* Checkpoint 7B materials in the back of the text.

Example 5

Display this data in a box plot:
51, 55, 55, 62, 65, 72, 76, 78, 79, 82, 83, 85, 91, and 93.

- Since this data is already in order from least to greatest, the range is $93 - 51 = 42$. Thus you start with a number line with equal intervals from 50 to 100.

- The median of the set of data is 77. A vertical segment is drawn at this value above the number line.

- The median of the lower half of the data (the first quartile) is 62. A vertical segment is drawn at this value above the number line.

- The median of the upper half of the data (the third quartile) is 83. A vertical segment is drawn at this value above the number line.

- A box is drawn between the first and third quartiles.

- Place a vertical segment at the minimum value (51) and at the maximum value (93). Use a line segment to connect the minimum to the box and the maximum to the box.
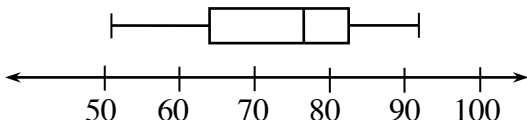
Example 6

Display this data in a box plot:
62, 65, 93, 51, 12, 79, 85, 55, 72, 78, 83, 91, and 76.

- Place the data in order from least to greatest: 12, 51, 55, 62, 65, 72, 76, 78, 79, 83, 85, 91, 93. The range is $93 - 12 = 81$. Thus you want a number line with equal intervals from 10 to 100.

- Find the median of the set of data: 76. Draw the line segment.

- Find the first quartile: $55 + 62 = 117$; $117 \div 2 = 58.5$. Draw the line segment.

- Find the third quartile: $83 + 85 = 168$; $168 \div 2 = 84$. Draw the line segment.

- Draw the box connecting the first and third quartiles. Place a line segment at the minimum value (12) and a line segment at the maximum value (93). Connect the minimum and maximum values to the box.
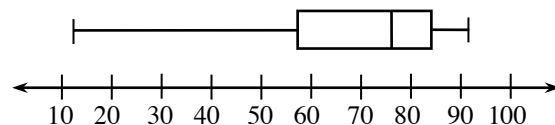
Problems

Create a stem-and-leaf plot and a box plot for each set of data in problems 5 through 8.

11. 45, 47, 52, 85, 46, 32, 83, 80, and 75.

12. 75, 62, 56, 80, 72, 55, 54, and 80.

13. 49, 54, 52, 58, 61, 72, 73, 78, 73, 82, 83, 73, 61, 67, and 68.

14. 65, 35, 48, 29, 57, 87, 94, 68, 86, 73, 58, 74, 85, 91, 88, and 97.

15. Given a set of data: 265, 263, 269, 259, 267, 264, 253, 275, 264, 260, 273, 257, and 291.

    a.   Make a stem-and-leaf plot of this data.

    b.   Find the mean and median of this data.

    c.   Find the range of this data.

    d.   Make a box plot for this data.

16. Given a set of data: 48, 42, 37, 29, 49, 46, 38, 28, 45, 45, 35, 46.25, 34, 46, 46.5, 43, 46.5, 48, 41.25, 29, and 47.75.

    a.   Make a stem-and-leaf plot of this data.

    b.   Find the mean and median of this data.

    c.   Find the range of the data.

    d.   Make a box plot for this data.

Answers

1.

1	8
2	6 8 9
3	0 3 4 4

2.

1	9
2	5 5 7
3	0 1 4 5
4	2

3.

6	7
7	6 9 9
8	0 1 2 4 9 9 9
9	2 5
10	
11	
12	3

4.

9	1
10	0 1 1 3 4 5 7 8 9
11	1 3 3 6 8

5. a. 24   b. 6   c. 14   d. 0
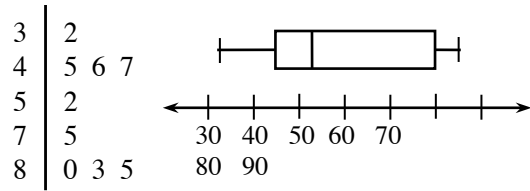
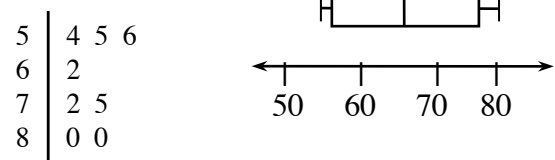**6.**   a. ice cream 20   b. 10

c. 12   d. 15

7.

8.

9.

10.

11.

3	2
4	5 6 7
5	2
7	5
8	0 3 5

30 40 50 60 70
80 90

12.

5	4 5 6
6	2
7	2 5
8	0 0

50 60 70 80

13.

4	9
5	2 4 8
6	1 1 7 8
7	2 3 3 3 8
8	2 3

40 50 60 70 80 90

14.

2	9
3	5
4	8
5	7 8
6	5 8
7	3 4
8	5 6 7 8
9	1 4 7

20 30 40 50 60 70 80 90
100

15.

25	3 7 9
26	0 3 4 4 5 7 9
27	3 5
28	
29	1
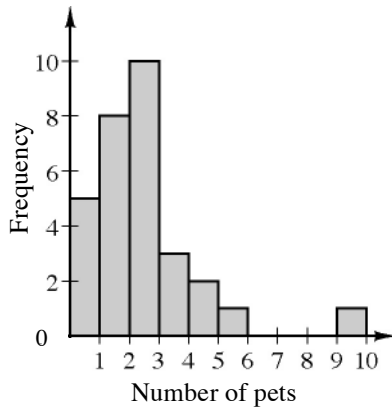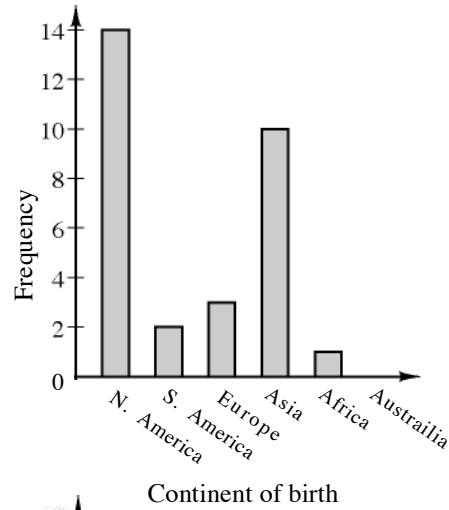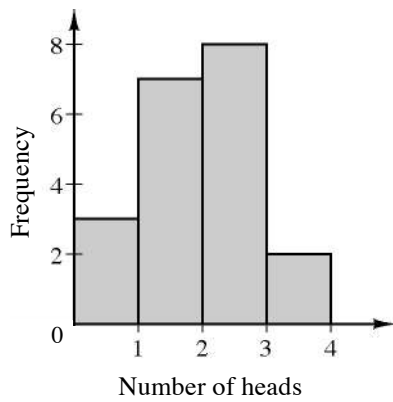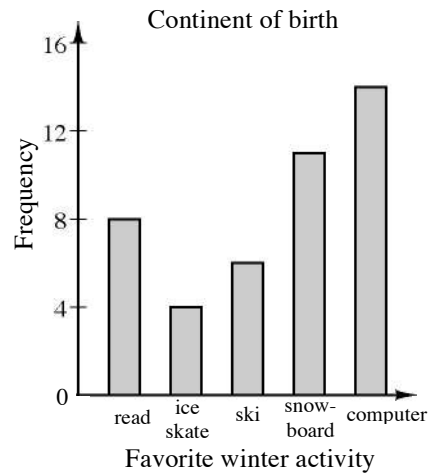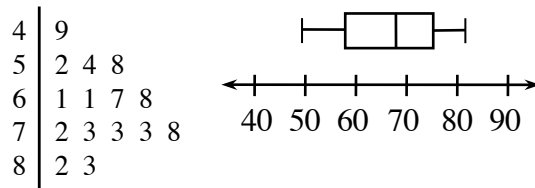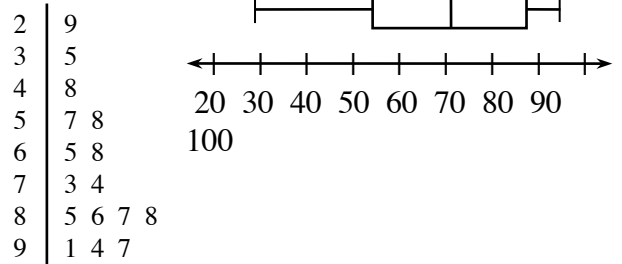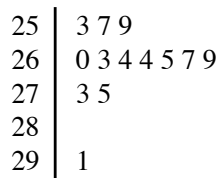
Mean: 266.15
Median: 264
Range: 38

250 270 300

16.

2	8 9 9
3	4 5 7 8
4	1.25 2 3 5 5 6 6 6.25 6.5 6.5 7.75 8 8 9
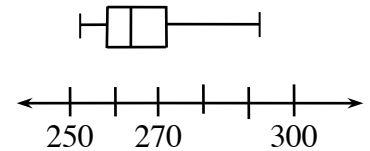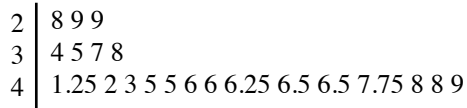
Mean: 41.4405
Median: 43
Range: 21

25 30 35 40 45 50
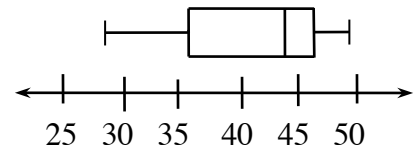
SCATTERPLOTS, ASSOCIATION, AND LINE OF BEST FIT

Data that is collected by measuring or observing naturally varies. A scatterplot helps students decide is there is a relationship (an **association**) between two numerical variables.

If there is a possible linear relationship, the trend can be shown graphically with a line of best fit on the scatterplot. In this course, students use a ruler to "eyeball" a line of best fit. The equation of the best-fit line can be determined from the slope and the y-intercept.

An association is often described by its form, direction, strength, and outliers. See the Math Notes boxes in Lessons 7.1.2, 7.1.3, and 7.3.2 of the *Core Connections, Course 3* text.

For additional examples and practice, see the *Core Connections, Course 3* Checkpoint 9 materials.

Example 1

Sam collected data by measuring the pencils of her classmates. She recorded the length of the painted part of each pencil and its weight. Her data is shown on the graph at right.

a. Describe the association between weight and length of the pencil.

b. Create a line of best fit where y is the weight of the pencil in grams and x is the length of the paint on the pencil in centimeters.

c. Sam's teacher has a pencil with 11.5 cm of paint. Predict the weight of the teacher's pencil using the equation found in part (b).
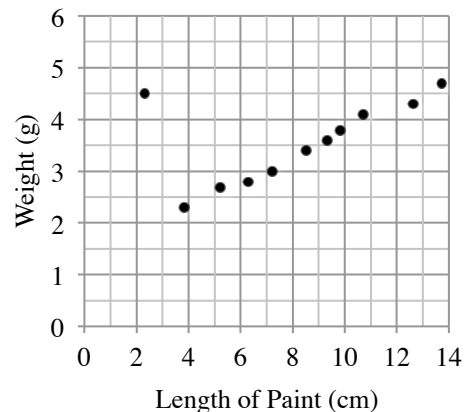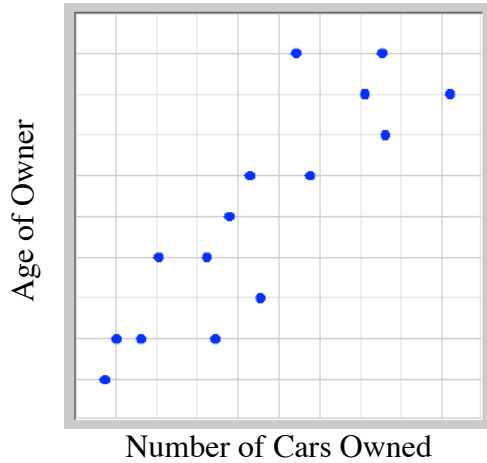
Answer:

a. There is a strong positive linear association with one apparent outlier at 2.3cm.

b. The equation of the line of best fit is approximately: $y = \frac{1}{4}x + 1.5$. See graph at right.
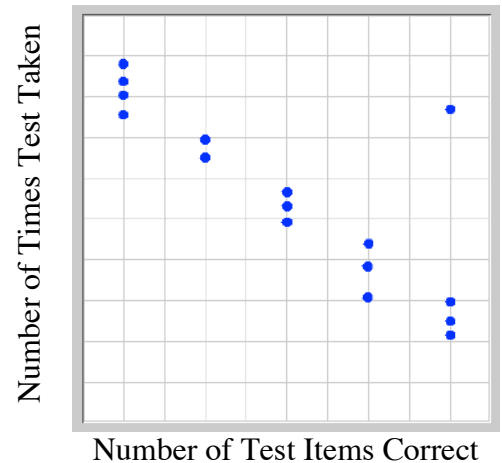
c. $\frac{1}{4}(11.5) + 1.5 \approx 4.4$ g.

Core Connections, Courses 1–3
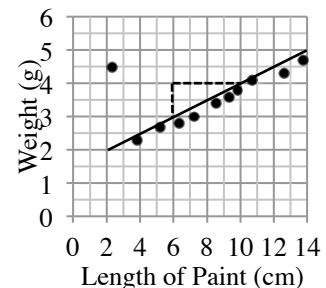
Problems

In problems 1 through 4 describe (if they exist), the form, direction, strength, and outliers of the scatterplot.

1.

Age of Owner (y-axis)

Number of Cars Owned (x-axis)

2.

Number of Times Test Taken (y-axis)

Number of Test Items Correct (x-axis)

3.

Chapter 5 Test Score (y-axis)

Height (x-axis)

4.
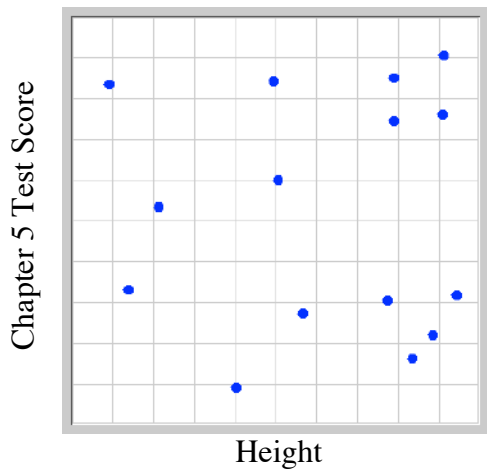
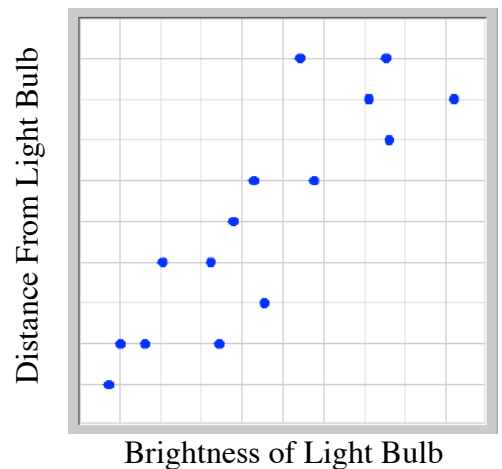Distance From Light Bulb (y-axis)

Brightness of Light Bulb (x-axis)

5. Dry ice (frozen carbon dioxide) evaporates at room temperature. Giulia's father uses dry ice to keep the glasses in the restaurant cold. Since dry ice evaporates in the restaurant cooler, Giulia was curious how long a piece of dry ice would last. She collected the data shown in the table at right.

Draw a scatterplot and a line of best fit. What is the approximate equation of the line of best fit?

# of hours after noon	Weight of dry ice (oz)
0	15.3
1	14.7
2	14.3
3	13.6
4	13.1
5	12.5
6	11.9
7	11.5
8	11.0
9	10.6
10	10.2

6.	Ranger Scott is responsible for monitoring the population of the elusive robins in McNeil State Park. He would like to find a relationship between the elm trees (their preferred nesting site) and the number of robins in the park. He randomly selects 7 different areas in the park and painstakingly counts the elms and robins in each area.

| Elms	8	13	4	5	10	9	4
Robins	5	9	3	5	7	7	5

a.	Make a scatterplot on graph paper and describe the association.

b.	Sketch the line of best fit on your scatterplot. Find the equation of the line of best fit.

c.	Based on the equation, how many robins should Ranger Scott expect to find in an area with 6 elm trees?

7.	A study was done for a vitamin supplement that claims to shorten the length of the common cold. The data the scientists collected from ten patients in an early study are shown in the table below.

| Number of months taking supplement	0.5	2.5	1	2	0.5	1	2	1	1.5	2.5
Number of days cold lasted	4.5	1.6	3	1.8	5	4.2	2.4	3.6	3.3	1.4

a.	Create a scatterplot and describe the association.

b.	Model the data with a line of best fit. Use x to represent the number of months taking the supplement and y to represent the length of the cold.

c.	According to your model, how many days do you expect a cold to last for patient taking the supplement for 1.5 months?

Answers

1.   Moderate, positive, linear association with no outliers.

2.   Strong, negative, linear association with an outlier.

3.   No association

4.   Strong, positive, curved association.

5. $y = -\frac{1}{2}x + 15.3$
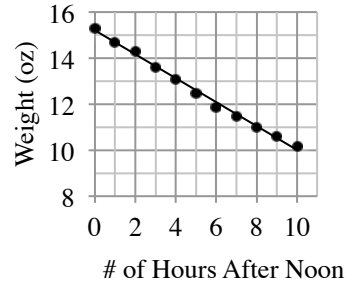
6.   Strong, positive linear association and no outliers
      $y = \frac{1}{2}x + 2$ 5 robins
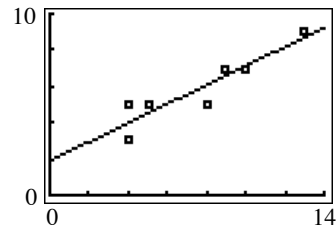
7.   a. The form is linear, the direction is negative, the strength
         is moderate, and there are no apparent outliers.

      b. $y \approx -\frac{5}{3}x + 5$

      c. $-\frac{5}{3}\left(\frac{3}{2}\right) + 5 = 2\frac{1}{2}$ days