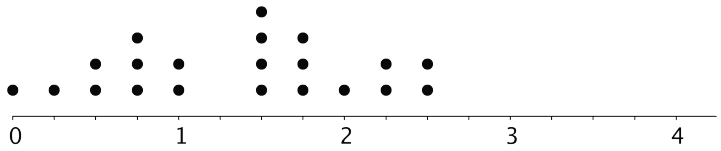
NAME DATE PERIOD

Unit 8, Lesson 1: Got Data?

Let's explore different kinds of data.

1.1: Dots of Data

Here is a **dot plot** for a data set.



1. Determine if each of the following would be an appropriate label to represent the data in the dot plot? Be prepared to explain your reasoning.

   a. Number of children per class.

   b. Distance between home and school, in miles.

   c. Hours spent watching TV each day.

   d. Weight of elephants, in pounds.

   e. Points received on a homework assignment.

2. Think of another label that can be used with the dot plot.

   a. Write it below the scale of the dot plot. Be sure to include the unit of measurement.

   b. In your scenario, what does one dot represent?

   c. In your scenario, what would a data point of 0 mean? What would a data point of $3\frac{1}{4}$ mean?

NAME DATE PERIOD

1.2: Surveying the Class

Here are some survey questions. Your teacher will explain which questions can be used to learn more about the students in your class and how the responses will be collected. The data that your class collects will be used in upcoming activities.

1. How long does it usually take you to travel to school? Answer to the nearest minute.

2. How do you travel to school on most days? Choose one.

   - ○ Walk
   - ○ Bike
   - ○ Scooter or skateboard
   - ○ Car
   - ○ School bus
   - ○ Public transport
   - ○ Other

3. How tall are you without your shoes on? Answer to the nearest centimeter.

4. What is the length of your right foot without your shoe on? Answer to the nearest centimeter.

5. What is your arm span? Stretch your arms open, and measure the distance from the tip of your right hand's middle finger to the tip of your left hand's middle finger, across your back. Answer to the nearest centimeter.

6. How important are the following issues to you? Rate each on a scale from 0 (not important) to 10 (very important).

   a. Reducing pollution b. Recycling c. Conserving water

7. Do you have any siblings? _____ Yes _____ No

8. How many hours of sleep per night do you usually get when you have school the next day? Answer to the nearest half hour.

9. How many hours of sleep per night do you usually get when you do not have school the next day? Answer to the nearest half hour.

NAME DATE PERIOD

10. Other than traveling from school, what do you do right after school on most days?

- ◦ Have a snack
- ◦ Do homework
- ◦ Read a book
- ◦ Talk on the phone
- ◦ Practice a sport
- ◦ Do chores
- ◦ Use the computer
- ◦ Participate in an extracurricular activity

11. If you could meet one of these celebrities, who would you choose?

- ◦ A city or state leader
- ◦ A champion athlete
- ◦ A movie star
- ◦ A musical artist
- ◦ A best-selling author

12. Estimate how much time per week you usually spend on each of these activities. Answer to the nearest quarter of an hour.

a. Playing sports or doing outdoor activities

b. Using a screen for fun (watching TV, playing computer games, etc.)

c. Doing homework

d. Reading

NAME DATE PERIOD

1.3: Numerical and Categorical Data

The list of survey questions in the activity earlier can help you complete these exercises.

1. The first survey question about travel *time* produces **numerical data**. Identify two other questions that produce numerical data. For each, describe what was measured and its unit of measurement.

   a. Question #: _____ What was measured:

   Unit of measurement:

   b. Question #: _____ What was measured:

   Unit of measurement:

2. The second survey question about travel *method* produces **categorical data**. Identify two other questions that produce categorical data. For each, describe what characteristic or feature was being studied.

   a. Question #: _____ Characteristic being studied:

   b. Question #: _____ Characteristic being studied:

3. Think about the responses to these survey questions. Do they produce numerical or categorical data? Be prepared to explain how you know.

   a. How many pets do you have?

   b. How many years have you lived in this state?

   c. What is your favorite band?

   d. What kind of music do you like best?

   e. What is the area code of your school's phone number?

   f. Where were you born?

   g. How much does your backpack weigh?

NAME DATE PERIOD

4. Name two characteristics you could investigate to learn more about your classmates. Make sure one would give categorical data and the other would give numerical data.

Are you ready for more?

Priya and Han collected data on the birth months of students in their class. Here are tables showing their records for the same group of students.

This table shows Priya's records.

Jan	Apr	Jan	Feb	Oct	May	June	July	Aug	Aug
Sep	Jan	Feb	Mar	Apr	Nov	Nov	Dec	Feb	Mar

This table shows Han's records.

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

1. How are their records alike? How are they different?

2. What kind of data—categorical or numerical—do you think the variable "birth month" produces? Explain how you know.

NAME DATE PERIOD

Lesson 1 Summary

The table contains data about 10 dogs.

dog name	weight (kg)	breed
Duke	36	German shepherd
Coco	6	pug
Pierre	7	pug
Ginger	35	German shepherd
Lucky	10	beagle
Daisy	10	beagle
Buster	35	German shepherd
Pepper	7	pug
Rocky	7	beagle
Lady	32	German shepherd

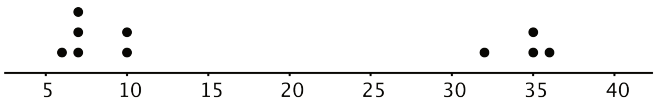
- The weights of the dogs are an example of **numerical data**, which is data that are numbers, quantities, or measurements. The weights of the dogs are measurements in kilograms.

- The dog breeds are an example of **categorical data**, which is data containing values that can be sorted into categories. In this case, there are three categories for dog breeds: pug, beagle, and German shepherd.

Some data with numbers are categorical because the numbers are *not* quantities or measurements. For example, telephone area codes are categorical data, because the numbers are labels rather than quantities or measurements.

NAME DATE PERIOD

Numerical data can be represented with a **dot plot** (sometimes called a line plot). Here is a dot plot that shows the weights of the dogs.



dog weights in kilograms

We can collect and study both kinds of data by doing surveys or taking measurements. When we do, it is important to think about what feature we are studying (for example, breeds of dogs or weights of dogs) and what units of measurement are used.

Lesson 1 Glossary Terms

- categorical data
- numerical data
- dot plot

NAME DATE PERIOD

Unit 8, Lesson 2: Statistical Questions

Let's look more closely at data and the questions they can help to answer.

2.1: Pencils on A Plot

1. Measure your pencil to the nearest $\frac{1}{4}$-inch. Then, plot your measurement on the class dot plot.

2. What is the difference between the longest and shortest pencil lengths in the class?

3. What is the most common pencil length?

4. Find the difference in lengths between the most common length and the shortest pencil.

NAME DATE PERIOD

2.2: What's in the Data?

Ten sixth-grade students at a school were each asked five survey questions. Their answers to each question are shown here.

data set A	0	1	1	3	0	0	0	2	1	1
data set B	12	12	12	12	12	12	12	12	12	12
data set C	6	5	7	6	4	5	3	4	6	8
data set D	6	6	6	6	6	6	6	6	6	6
data set E	3	7	9	11	6	4	2	16	6	10

1. Here are the five survey questions. Match each question to a data set that could represent the students' answers. Explain your reasoning.
    - Question 1: Flip a coin 10 times. How many heads did you get?

      Data set _____ Reason:
    - Question 2: How many books did you read in the last year?

      Data set _____ Reason:
    - Question 3: What grade are you in?

      Data set _____ Reason:
    - Question 4: How many dogs and cats do you have?

      Data set _____ Reason:
    - Question 5: How many inches are in 1 foot?

      Data set _____ Reason:
2. Discuss with a partner: How are Question 3 and Question 5 different from the other questions?

NAME DATE PERIOD

2.3: What Makes a Statistical Question?

These three questions are examples of **statistical questions**:

- What is the most common color of the cars in the school parking lot?
- What percentage of students in the school have a cell phone?
- Which kind of literature—fiction or nonfiction—is more popular among students in the school?

These three questions are not examples of statistical questions:

- What color is the principal's car?
- Does Elena have a cell phone?
- What kind of literature—fiction or nonfiction—does Diego prefer?

1. Study the examples and non-examples. Discuss with your partner:
    - How are the three statistical questions alike? What do they have in common?
    - How are the three non-statistical questions alike? What do they have in common?
    - How can you find answers to the statistical questions? How about answers to non-statistical questions?
    - What makes a question a statistical question?

    Pause here for a class discussion.

2. Read each question. Think about the data you might collect to answer it and whether you expect to see **variability** in the data. Complete each blank with "Yes" or "No."

    a. How many cups of water do my classmates drink each day?
    - Is variability expected in the data? _____ Is the question statistical? _____

    b. Where in town does our math teacher live?
    - Is variability expected in the data? _____ Is the question statistical? _____

NAME DATE PERIOD

c. How many minutes does it take students in my class to get ready for school in the morning?
   - Is variability expected in the data? _____ Is the question statistical? _____

d. How many minutes of recess do sixth-grade students have each day?
   - Is variability expected in the data? _____ Is the question statistical? _____

e. Do all students in my class know what month it is?
   - Is variability expected in the data? _____ Is the question statistical? _____

2.4: Sifting for Statistical Questions

1. Your teacher will give you and your partner a set of cards with questions. Sort them into three piles: Statistical Questions, Not Statistical Questions, and Unsure.

2. Compare your sorting decisions with another group of students. Start by discussing the two piles that your group sorted into the Statistical Questions and Not Statistical Questions piles. Then, review the cards in the Unsure pile. Discuss the questions until both groups reach an agreement and have no cards left in the Unsure pile. If you get stuck, think about whether the question could be answered by collecting data and if there would be variability in that data.

3. Record the letter names of the questions in each pile.

   - Statistical questions:
   - Non-statistical questions:

Are you ready for more?

Tyler and Han are discussing the question, "Which sixth-grade student lives the farthest from school?"

- Tyler says, "I don't think the question is a statistical question. There is only one person who lives the farthest from school, so there would not be variability in the data we collect."

- Han says: "I think it is a statistical question. To answer the question, we wouldn't

NAME DATE PERIOD

actually be asking everyone, 'Which student lives the farthest from school?' We would have to ask each student how far away from school they live, and we can expect their responses to have variability."

Do you agree with either one of them? Explain your reasoning.

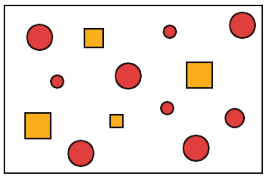
NAME DATE PERIOD

Lesson 2 Summary

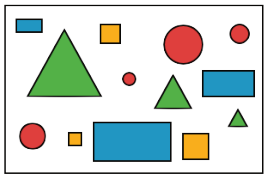
We often collect data to answer questions about something. The data we collect may show **variability**, which means the data values are not all the same.

Some data sets have more variability than others. Here are two sets of figures.

A



B



Set A has more figures with the same shape, color, or size. Set B shows more figures with different shapes, colors, or sizes, so set B has greater variability than set A.

Both numerical and categorical data can show variability. Numerical sets can contain different numbers, and categorical sets can contain different categories or types.

When a question can only be answered by using data and we expect that data to have variability, we call it a **statistical question**. Here are some examples.

- Who is the most popular musical artist at your school?
- When do students in your class typically eat dinner?
- Which classroom in your school has the most books?

To answer the question about books, we may need to count all of the books in each classroom of a school. The data we collect would likely show variability because we would expect each classroom to have a different number of books.

In contrast, the question "How many books are in your classroom?" is *not* a statistical question. If we collect data to answer the question (for example, by asking everyone in the class to count books), the data can be expected to show the same value. Likewise, if we ask all of the students at a school where they go to school, that question is not a statistical question because the responses will all be the same.

Lesson 2 Glossary Terms

- variability

- statistical question

NAME DATE PERIOD

Unit 8, Lesson 3: Representing Data Graphically

Let's represent data with dot plots and bar graphs.

3.1: Curious about Caps

Clare collects bottle caps and keeps them in plastic containers.

Write one statistical question that someone could ask Clare about her collection. Be prepared to explain your reasoning.

3.2: Estimating Caps

1. Write down the statistical question your class is trying to answer.

2. Look at the dot plot that shows the data from your class. Write down one thing you notice and one thing you wonder about the dot plot.

3. Use the dot plot to answer the statistical question. Be prepared to explain your reasoning.

NAME DATE PERIOD

3.3: Been There, Done That!

Priya wants to know if basketball players on a men's team and a women's team have had prior experience in international competitions. She gathered data on the number of times the players were on a team before 2016.

men's team	3	0	0	0	0	1	0	0	0	0	0	0
women's team	2	3	3	1	0	2	0	1	1	0	3	1

1. Did Priya collect categorical or numerical data?

2. Organize the information on the two basketball teams into these tables.

Men's Basketball Team Players

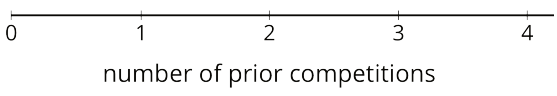
number of prior competitions	frequency (number)
0	
1	
2	
3	
4	

Women's Basketball Team Players

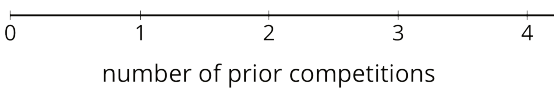
number of prior competitions	frequency (number)
0	
1	
2	
3	
4	

3. Make a dot plot for each table.

Men's Basketball Team Players



Women's Basketball Team Players



NAME DATE PERIOD

4. Study your dot plots. What do they tell you about the competition participation of:

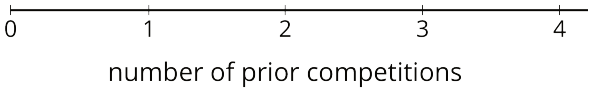
   a. the players on the men's basketball team?

   b. the players on the women's basketball team?

5. Explain why a dot plot is an appropriate representation for Priya's data.

Are you ready for more?

Combine the data for the players on the men's and women's teams and represent it as a single dot plot. What can you say about the repeat participation of the basketball players?



NAME DATE PERIOD

3.4: Favorite Summer Sports

Kiran wants to know which three summer sports are most popular in his class. He surveyed his classmates on their favorite summer sport and collected these responses:

swimming	gymnastics	track and field	volleyball	swimming	swimming
diving	track and field	gymnastics	basketball	basketball	volleyball
track and field	track and field	volleyball	gymnastics	diving	gymnastics
volleyball	rowing	track and field	track and field	soccer	swimming
gymnastics	track and field	swimming	rowing	diving	soccer

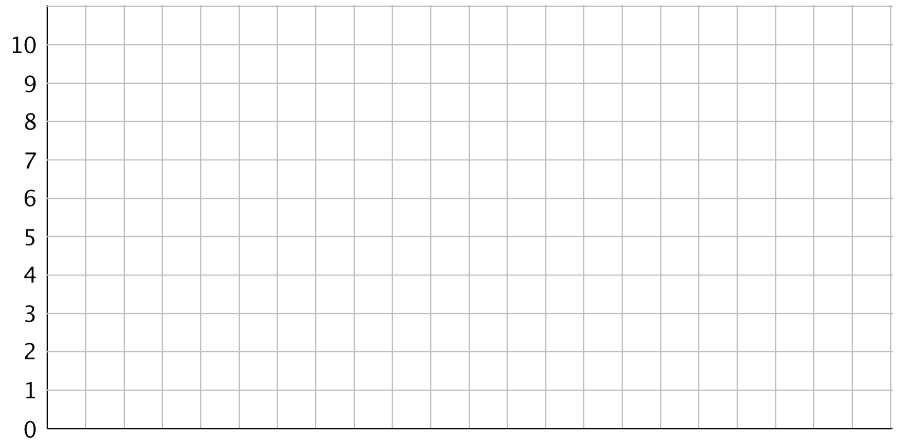
1. Did Kiran collect categorical or numerical data?

2. Organize the responses in a table to help him find which summer sports are most popular in his class.

sport	frequency

NAME DATE PERIOD

3. Represent the information in the table as a bar graph.



4. a. How can you use the bar graph to find how many classmates Kiran surveyed?

b. Study your bar graph, and answer Kiran's question about the top three summer sports in his class. Then, make at least one other observation, based on your bar graph, about his classmates' preferred summer sports.

5. Could a dot plot be used to represent Kiran's data? Explain your reasoning.

NAME DATE PERIOD

Lesson 3 Summary

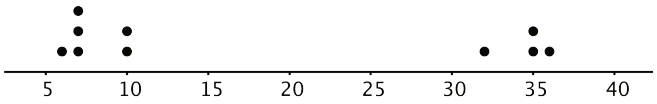
When we analyze data, we are often interested in the **distribution**, which is information that shows all the data values and how often they occur.

In a previous lesson, we saw data about 10 dogs. We can see the distribution of the dog weights in a table such as this one.

weight in kilograms	frequency
6	1
7	3
10	2
32	1
35	2
36	1

The term **frequency** refers to the number of times a data value occurs. In this case, we see that there are three dogs that weigh 7 kilograms, so “3” is the frequency for the value “7 kilograms.”

Recall that dot plots are often used to to represent numerical data. Like a frequency table, a dot plot also shows the distribution of a data set. This dot plot, which you saw in an earlier lesson, shows the distribution of dog weights.



dog weights in kilograms

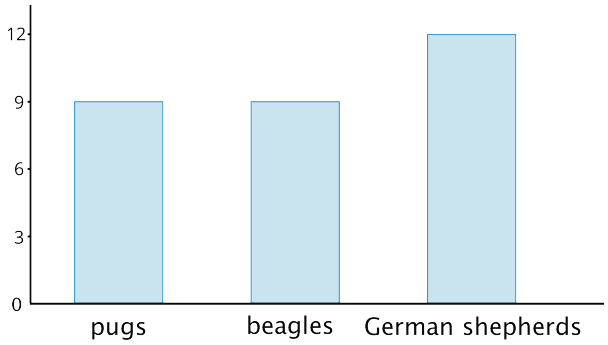
A dot plot uses a horizontal number line. We show the frequency of a value by the number of dots drawn above that value. Here, the two dots above the number 35 tell us that there are two dogs weighing 35 kilograms.

NAME DATE PERIOD

The distribution of categorical data can also be shown in a table. This table shows the distribution of dog breeds.

breed	frequency
pug	9
beagle	9
German shepherd	12

We often represent the distribution of categorical data using a bar graph.



A bar graph also uses a horizontal line. Above it we draw a rectangle (or "bar") to represent each category in the data set. The height of a bar tells us the frequency of the category. There are four German shepherds in the data set, so the bar for this category is 4 units tall. Below the line we write the labels for the categories.

In a dot plot, a data value is placed according to its position on the number line. A weight of 10 kilograms must be shown as a dot above 10 on the number line.

In a bar graph, however, the categories can be listed in any order. The bar that shows the frequency of pugs can be placed anywhere along the horizontal line.

Lesson 3 Glossary Terms

- frequency
- distribution

NAME DATE PERIOD

Unit 8, Lesson 4: Dot Plots

Let's investigate what dot plots and bar graphs can tell us.

4.1: Pizza Toppings (Part 1)

Fifteen customers in a pizza shop were asked, “How many toppings did you add to your cheese pizza?” Their responses are shown in the table.

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

1. Could you use a dot plot to represent the data? Explain your reasoning.

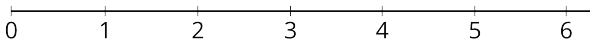
2. Complete the table.

number of toppings	frequency (number)
0	
1	
2	
3	

NAME DATE PERIOD

4.2: Pizza Toppings (Part 2)

1. Use the tables from the warm-up to display the number of toppings as a dot plot. Label your drawing clearly.



2. Use your dot plot to study the distribution for number of toppings. What do you notice about the number of toppings that this group of customers ordered? Write 2–3 sentences summarizing your observations.

Are you ready for more?

Think of a statistical question that can be answered with the data about the number of toppings ordered, as displayed on the dot plot. Then answer this question.

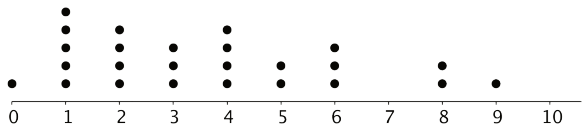
4.3: Homework Time

Twenty-five sixth-grade students answered the question: "How many hours do you generally spend on homework each week?"

1. Why is this question a statistical question?

NAME DATE PERIOD

2. This dot plot shows the number of hours per week that these 25 students reported spending on homework.



hours spent on homework per week

Use the dot plot to answer the following questions. For each, show or explain your reasoning.

   a. What percentage of the students reported spending 1 hour on homework each week?

   b. What percentage of the students reported spending 4 or fewer hours on homework each week?

3. Would 6 hours per week be a good description of the number of hours this group of students spends on homework per week? What about 1 hour per week? Explain your reasoning.

4. What value do you think would be a good description of the homework time of the students in this group? Explain your reasoning.

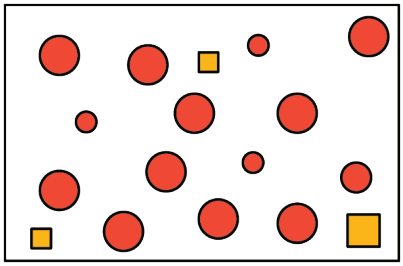
5. Someone said, “In general, these students spend roughly the same number of hours doing homework.” Do you agree? Explain your reasoning.

NAME DATE PERIOD

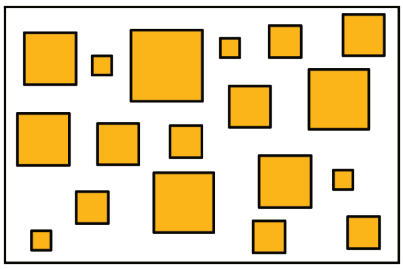
Lesson 4 Summary

We often collect and analyze data because we are interested in learning what is "typical," or what is common and can be expected in a group.

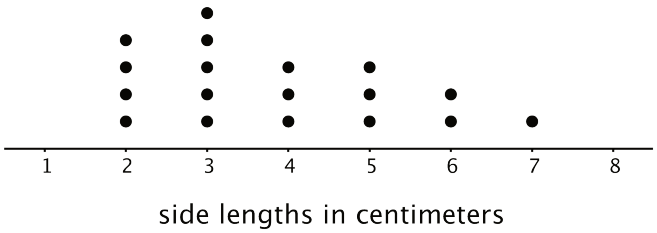
Sometimes it is easy to tell what a typical member of the group is. For example, we can say that a typical shape in this set is a large circle.



Just looking at the members of a group doesn't always tell us what is typical, however. For example, if we are interested in the side length typical of squares in this set, it isn't easy to do so just by studying the set visually.



In a situation like this, it is helpful to gather the side lengths of the squares in the set and look at their distribution, as shown in this dot plot.



We can see that many of the data points are between 2 and 4, so we could say that side lengths between 2 and 4 centimeters or close to these lengths are typical of squares in this set.

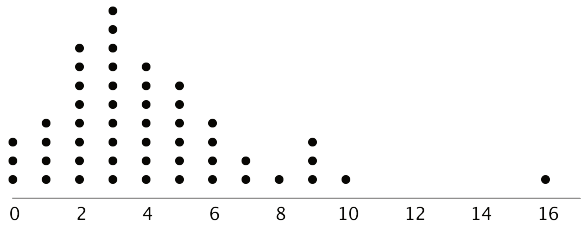
NAME DATE PERIOD

Unit 8, Lesson 5: Using Dot Plots to Answer Statistical Questions

Let's use dot plots to describe distributions and answer questions.

5.1: Packs on Backs

This dot plot shows the weights of backpacks, in kilograms, of 50 sixth-grade students at a school in New Zealand.



1. The dot plot shows several dots at 0 kilograms. What could a value of 0 mean in this context?

2. Clare and Tyler studied the dot plot.

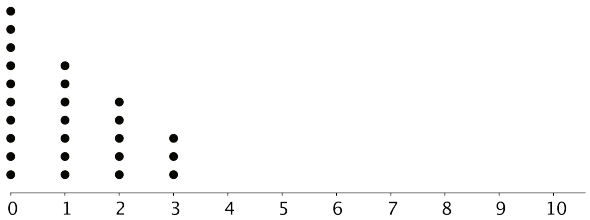
   - Clare said, “I think we can use 3 kilograms to describe a typical backpack weight of the group because it represents 20%—or the largest portion—of the data.”
   - Tyler disagreed and said, “I think 3 kilograms is too low to describe a typical weight. Half of the dots are for backpacks that are heavier than 3 kilograms, so I would use a larger value.”

   Do you agree with either of them? Explain your reasoning.

NAME DATE PERIOD

5.2: On the Phone

Twenty-five sixth-grade students were asked to estimate how many hours a week they spend talking on the phone. This dot plot represents their reported number of hours of phone usage per week.



1. a. How many of the students reported not talking on the phone during the week? Explain how you know.

   b. What percentage of the students reported not talking on the phone?

2. a. What is the largest number of hours a student spent talking on the phone per week?

   b. What percentage of the group reported talking on the phone for this amount of time?

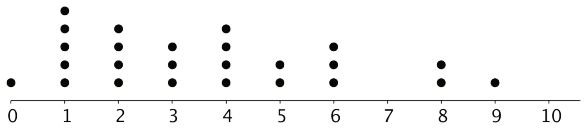
3. a. How many hours would you say that these students typically spend talking on the phone?

   b. How many minutes per day would that be?

NAME DATE PERIOD

4. a. How would you describe the **spread** of the data? Would you consider these students' amounts of time on the phone to be alike or different? Explain your reasoning.

b. Here is the dot plot from an earlier activity. It shows the number of hours per week the same group of 25 sixth-grade students reported spending on homework.



hours spent on homework per week

Overall, are these students more alike in the amount of time they spend talking on the phone or in the amount of time they spend on homework? Explain your reasoning.

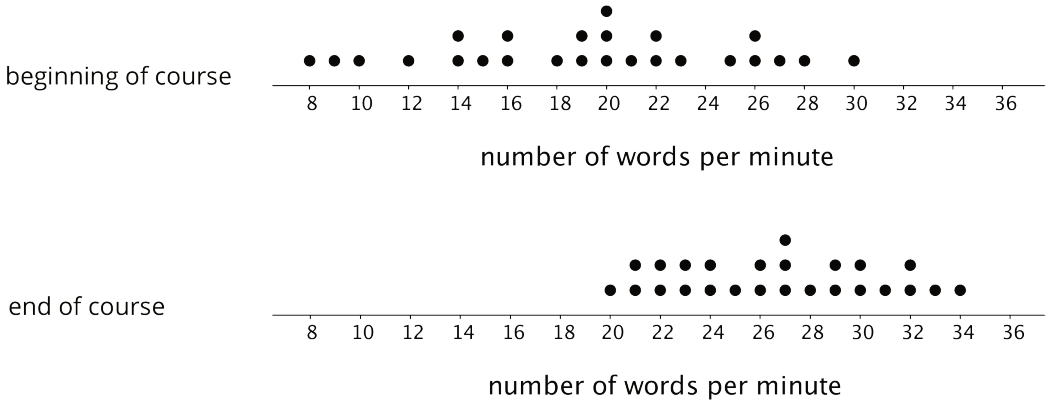
5. Suppose someone claimed that these sixth-grade students spend too much time on the phone. Do you agree? Use your analysis of the dot plot to support your answer.

NAME DATE PERIOD

5.3: Click-Clack

1. A keyboarding teacher wondered: "Do typing speeds of students improve after taking a keyboarding course?" Explain why her question is a statistical question.

2. The teacher recorded the number of words that her students could type per minute at the beginning of a course and again at the end. The two dot plots show the two data sets.



Based on the dot plots, do you agree with each of the following statements about this group of students? Be prepared to explain your reasoning.

a. Overall, the students' typing speed did not improve. They typed at the same speed at the end of the course as they did at the beginning.

b. 20 words per minute is a good estimate for how fast, in general, the students typed at the beginning of the course.

c. 20 words per minute is a good description of the **center** of the data set at the end of the course.

d. There was more variability in the typing speeds at the beginning of the course than at the end, so the students' typing speeds were more alike at the end.

NAME DATE PERIOD

3. Overall, how fast would you say that the students typed after completing the course? What would you consider the center of the end-of-course data?

Are you ready for more?

Use one of these suggestions (or make up your own). Research to create a dot plot with at least 10 values. Then, describe the center and spread of the distribution.

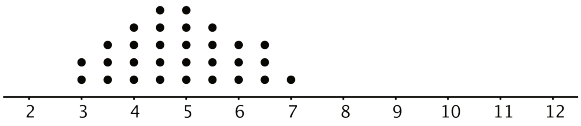
- Points scored by your favorite sports team in its last 10 games
- Length of your 10 favorite movies (in minutes)
- Ages of your favorite 10 celebrities

NAME DATE PERIOD

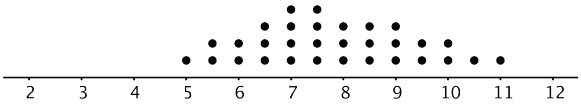
Lesson 5 Summary

One way to describe what is typical or characteristic for a data set is by looking at the **center** and **spread** of its distribution.

Let's compare the distribution of cat weights and dog weights shown on these dot plots.



cat weights in kilograms



dog weights in kilograms

The collection of points for the cat data is further to the left on the number line than the dog data. Based on the dot plots, we may describe the center of the distribution for cat weights to be between 4 and 5 kilograms and the center for dog weights to be between 7 and 8 kilograms.

We often say that values at or near the center of a distribution are typical for that group. This means that a weight of 4–5 kilograms is typical for a cat in the data set, and weight of 7–8 kilograms is typical for a dog.

We also see that the dog weights are more spread out than the cat weights. The difference between the heaviest and lightest cats is only 4 kilograms, but the difference between the heaviest and lightest dogs is 6 kilograms.

A distribution with greater spread tells us that the data have greater variability. In this case, we could say that the cats are more similar in their weights than the dogs.

In future lessons, we will discuss how to measure the center and spread of a distribution.

Lesson 5 Glossary Terms

- spread
- center

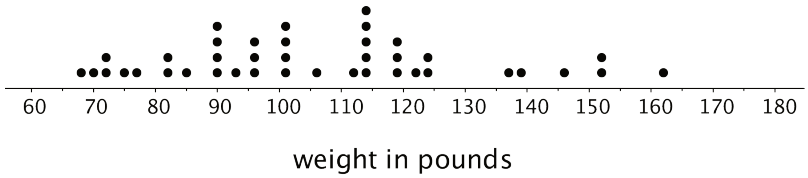
NAME DATE PERIOD

Unit 8, Lesson 6: Histograms

Let's explore how histograms represent data sets.

6.1: Dog Show (Part 1)

Here is a dot plot showing the weights, in pounds, of 40 dogs at a dog show.



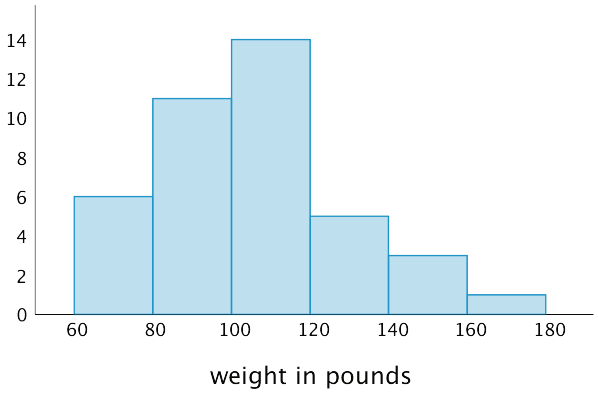
1. Write two statistical questions that can be answered using the dot plot.

2. What would you consider a typical weight for a dog at this dog show? Explain your reasoning.

NAME DATE PERIOD

6.2: Dog Show (Part 2)

Here is a **histogram** that shows some dog weights in pounds.



Each bar includes the left-end value but not the right-end value. For example, the first bar includes dogs that weigh 60 pounds and 68 pounds but not 80 pounds.

1. Use the histogram to answer the following questions.

   a. How many dogs weigh at least 100 pounds?

   b. How many dogs weigh exactly 70 pounds?

   c. How many dogs weigh at least 120 and less than 160 pounds?

   d. How much does the heaviest dog at the show weigh?

   e. What would you consider a typical weight for a dog at this dog show? Explain your reasoning.

2. Discuss with a partner:

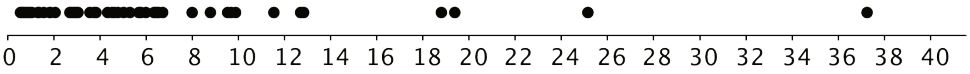
   - If you used the dot plot to answer the same five questions you just answered, how would your answers be different?

   - How are the histogram and the dot plot alike? How are they different?

NAME DATE PERIOD

6.3: Population of States

Every ten years, the United States conducts a census, which is an effort to count the entire population. The dot plot shows the population data from the 2010 census for each of the fifty states and the District of Columbia (DC).



population of states in millions

1. Here are some statistical questions about the population of the fifty states and DC. How difficult would it be to answer the questions using the *dot plot*?

   In the middle column, rate each question with an E (easy to answer), H (hard to answer), or I (impossible to answer). Be prepared to explain your reasoning.

statistical question	using the dot plot	using the histogram
a. How many states have populations greater than 15 million?		
b. Which states have populations greater than 15 million?		
c. How many states have populations less than 5 million?		
d. What is a typical state population?		
e. Are there more states with fewer than 5 million people, or more states with between 5 and 10 million people?		
f. How would you describe the distribution of state populations?		

NAME DATE PERIOD

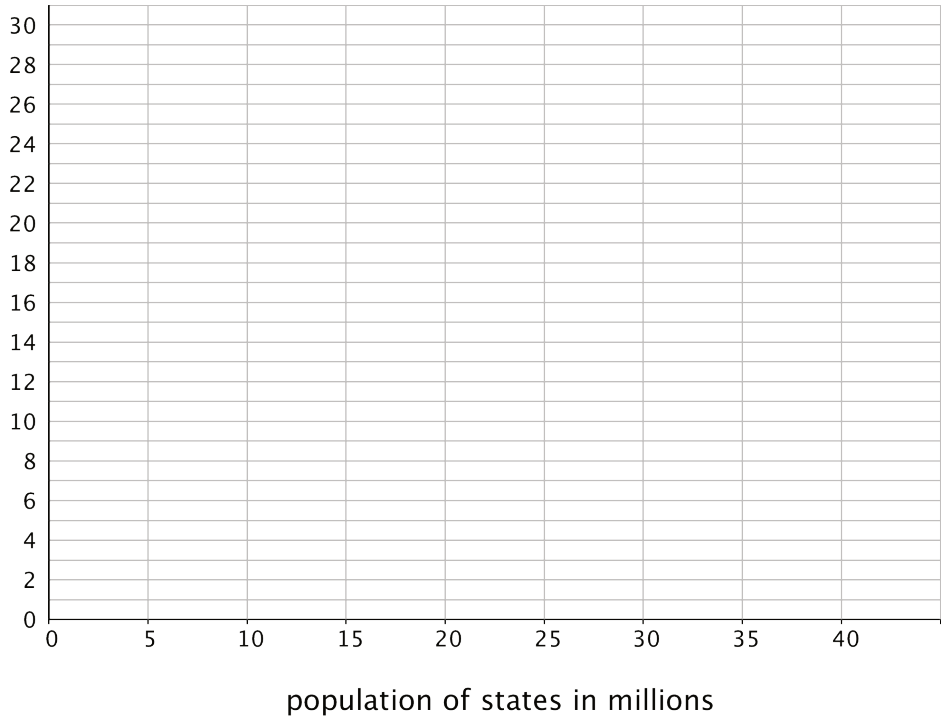
2. Here are the population data for all states and the District of Columbia from the 2010 census. Use the information to complete the table.

Alabama	4.78	Illinois	12.83	Montana	0.99	Rhode Island	1.05
Alaska	0.71	Indiana	6.48	Nebraska	1.83	South Carolina	4.63
Arizona	6.39	Iowa	3.05	Nevada	2.70	South Dakota	0.81
Arkansas	2.92	Kansas	2.85	New Hampshire	1.32	Tennessee	6.35
California	37.25	Kentucky	4.34	New Jersey	8.79	Texas	25.15
Colorado	5.03	Louisiana	4.53	New Mexico	2.06	Utah	2.76
Connecticut	3.57	Maine	1.33	New York	19.38	Vermont	0.63
Delaware	0.90	Maryland	5.77	North Carolina	9.54	Virginia	8.00
District of Columbia	0.60	Massachusetts	6.55	North Dakota	0.67	Washington	6.72
Florida	18.80	Michigan	9.88	Ohio	11.54	West Virginia	1.85
Georgia	9.69	Minnesota	5.30	Oklahoma	3.75	Wisconsin	5.69
Hawaii	1.36	Mississippi	2.97	Oregon	3.83	Wyoming	0.56
Idaho	1.57	Missouri	5.99	Pennsylvania	12.70		

population (millions)	frequency
0–5	
5–10	
10–15	
15–20	
20–25	
25–30	
30–35	
35–40	

NAME DATE PERIOD

3. Use the grid and the information in your table to create a histogram.



4. Return to the statistical questions at the beginning of the activity. Which ones are now easier to answer?

   In the last column of the table, rate each question with an E (easy), H (hard), and I (impossible) based on how difficult it is to answer them. Be prepared to explain your reasoning.

Are you ready for more?

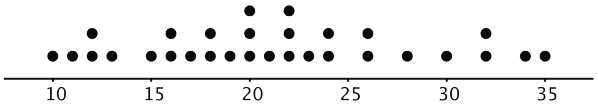
Think of two more statistical questions that can be answered using the data about populations of states. Then, decide whether each question can be answered using the dot plot, the histogram, or both.

NAME DATE PERIOD

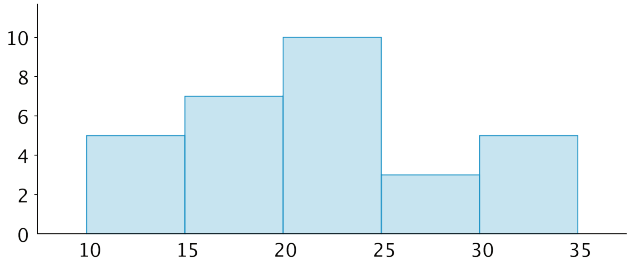
Lesson 6 Summary

In addition to using dot plots, we can also represent distributions of numerical data using **histograms**.

Here is a dot plot that shows the weights, in kilograms, of 30 dogs, followed by a histogram that shows the same distribution.



dog weights in kilograms



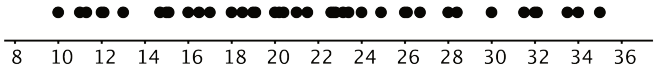
dog weights in kilograms

In a histogram, data values are placed in groups or "bins" of a certain size, and each group is represented with a bar. The height of the bar tells us the frequency for that group.

For example, the height of the tallest bar is 10, and the bar represents weights from 20 to less than 25 kilograms, so there are 10 dogs whose weights fall in that group. Similarly, there are 3 dogs that weigh anywhere from 25 to less than 30 kilograms.

Notice that the histogram and the dot plot have a similar shape. The dot plot has the advantage of showing all of the data values, but the histogram is easier to draw and to interpret when there are a lot of values or when the values are all different.

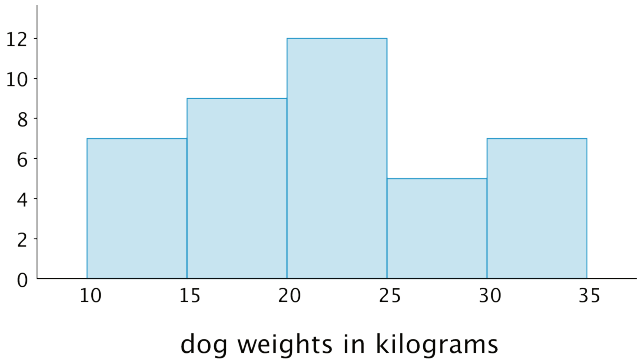
Here is a dot plot showing the weight distribution of 40 dogs. The weights were measured to the nearest 0.1 kilogram instead of the nearest kilogram.



dog weights in kilograms

NAME DATE PERIOD

Here is a histogram showing the same distribution.



In this case, it is difficult to make sense of the distribution from the dot plot because the dots are so close together and all in one line. The histogram of the same data set does a much better job showing the distribution of weights, even though we can't see the individual data values.

Lesson 6 Glossary Terms

- histogram

NAME DATE PERIOD

Unit 8, Lesson 7: Using Histograms to Answer Statistical Questions

Let's draw histograms and use them to answer questions.

7.1: Which One Doesn't Belong: Questions

Here are four questions about the population of Alaska. Which question does not belong? Be prepared to explain your reasoning.

A. In general, at what age do Alaska residents retire?

B. At what age can Alaskans vote?

C. What is the age difference between the youngest and oldest Alaska residents with a full-time job?

D. Which age group is the largest part of the population: 18 years or younger, 19–25 years, 25–34 years, 35–44 years, 45–54 years, 55–64 years, or 65 years or older?

7.2: Measuring Earthworms

An earthworm farmer set up several containers of a certain species of earthworms so that he could learn about their lengths. The lengths of the earthworms provide information about their ages. The farmer measured the lengths of 25 earthworms in one of the containers. Each length was measured in millimeters.

1. Using a ruler, draw a line segment for each length:
    - 20 millimeters
    - 40 millimeters
    - 60 millimeters
    - 80 millimeters
    - 100 millimeters

NAME DATE PERIOD

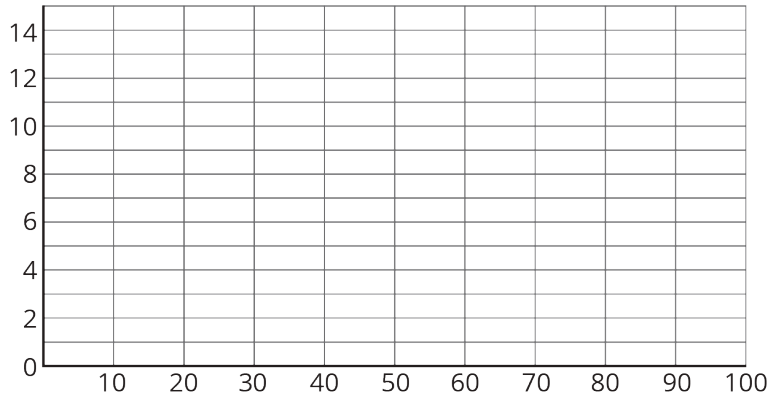
a. Here are the lengths, in millimeters, of the 25 earthworms.

6	11	18	19	20	23	23	25	25	26
27	27	28	29	32	33	41	42	48	52
54	59	60	77	93					

Complete the table for the lengths of the 25 earthworms.

length	frequency
0 millimeters to less than 20 millimeters	
20 millimeters to less than 40 millimeters	
40 millimeters to less than 60 millimeters	
60 millimeters to less than 80 millimeters	
80 millimeters to less than 100 millimeters	

2. Use the grid and the information in the table to draw a histogram for the worm length data. Be sure to label the axes of your histogram.



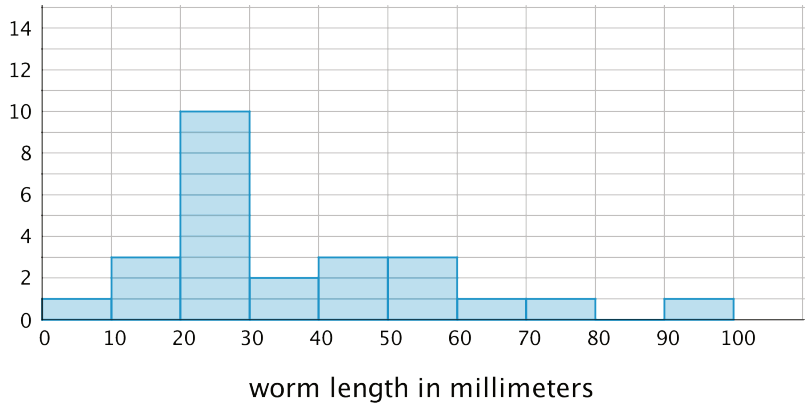
3. Based on the histogram, what is a typical length for these 25 earthworms? Explain how you know.

NAME DATE PERIOD

4. Write 1–2 sentences to describe the spread of the data. Do most of the worms have a length that is close to your estimate of a typical length, or are they very different in length?

Are you ready for more?

Here is another histogram for the earthworm measurement data. In this histogram, the measurements are in different groupings.



a. Based on this histogram, what is your estimate of a typical length for the 25 earthworms?

b. Compare this histogram with the one you drew. How are the distributions of data summarized in the two histograms the same? How are they different?

c. Compare your estimates of a typical earthworm length for the two histograms. Did you reach different conclusions about a typical earthworm length from the two

NAME DATE PERIOD

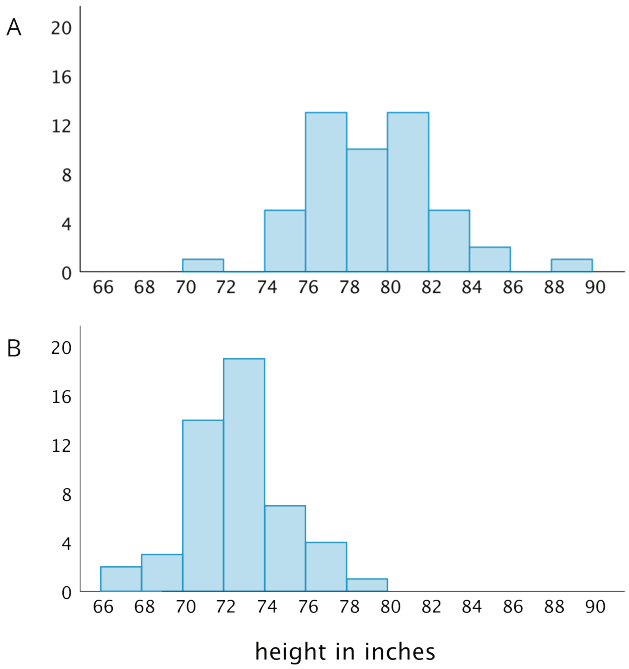
histograms?

7.3: Tall and Taller Players

Professional basketball players tend to be taller than professional baseball players.

Here are two histograms that show height distributions of 50 male professional baseball players and 50 male professional basketball players.

1. Decide which histogram shows the heights of baseball players and which shows the heights of basketball players. Be prepared to explain your reasoning.



2. Write 2–3 sentences that describe the distribution of the heights of the basketball players. Comment on the center and spread of the data.

NAME DATE PERIOD

3. Write 2–3 sentences that describe the distribution of the heights of the baseball players. Comment on the center and spread of the data.

Lesson 7 Summary

Here are the weights, in kilograms, of 30 dogs.

10	11	12	12	13	15	16	16	17	18
18	19	20	20	20	21	22	22	22	23
24	24	26	26	28	30	32	32	34	34

Before we draw a histogram, let's consider a couple of questions.

- What are the smallest and largest values in our data set? This gives us an idea of the distance on the number line that our histogram will cover. In this case, the minimum is 10 and the maximum is 34, so our number line needs to extend from 10 to 35 at the very least.

  (Remember the convention we use to mark off the number line for a histogram: we include the left boundary of a bar but exclude the right boundary. If 34 is the right boundary of the last bar, it won't be included in that bar, so the number line needs to go a little greater than the maximum value.)

- What group size or bin size seems reasonable here? We could organize the weights into bins of 2 kilograms (10, 12, 14, . . .), 5 kilograms, (10, 15, 20, 25, . . .), 10 kilograms (10, 20, 30, . . .), or any other size. The smaller the bins, the more bars we will have, and vice versa.

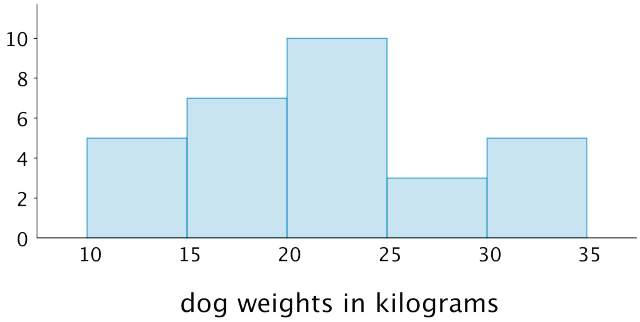
NAME DATE PERIOD

Let's use bins of 5 kilograms for the dog weights. The boundaries of our bins will be: 10, 15, 20, 25, 30, 35. We stop at 35 because it is greater than the maximum.

Next, we find the frequency for the values in each group. It is helpful to organize the values in a table.

weights in kilograms	frequency
10 to less than 15	5
15 to less than 20	7
20 to less than 25	10
25 to less than 30	3
30 to less than 35	5

Now we can draw the histogram.



The histogram allows us to learn more about the dog weight distribution and describe its center and spread.

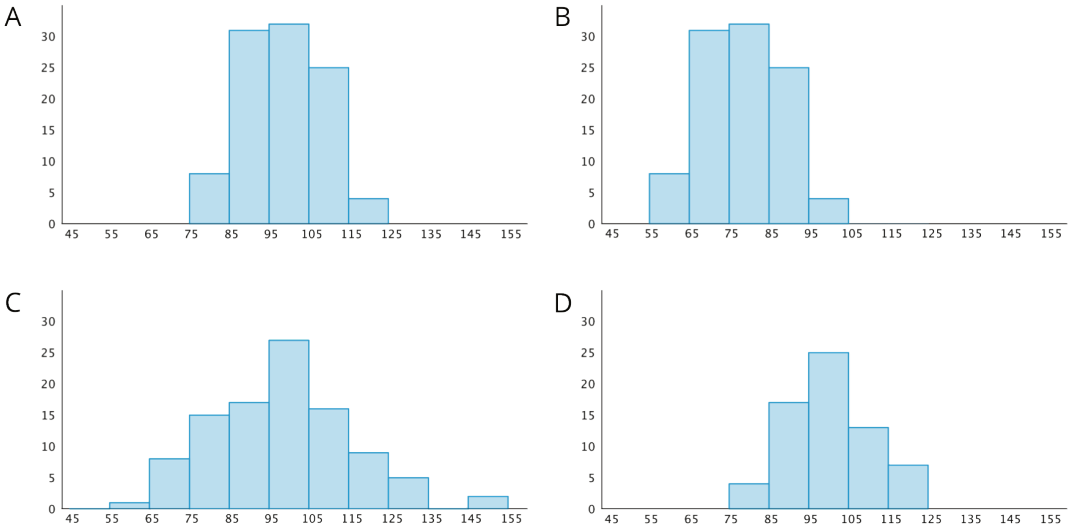
NAME DATE PERIOD

Unit 8, Lesson 8: Describing Distributions on Histograms

Let's describe distributions displayed in histograms.

8.1: Which One Doesn't Belong: Histograms

Which histogram does not belong? Be prepared to explain your reasoning.



8.2: Sorting Histograms

1. Your teacher will give your group a set of histogram cards. Sort them into two piles—one for histograms that are approximately symmetrical, and another for those that are not.

2. Discuss your sorting decisions with another group. Do both groups agree which cards should go in each pile? If not, discuss the reasons behind the differences and see if you can reach agreement. Record your final decisions.

    - Histograms that are approximately symmetrical:

    - Histograms that are not approximately symmetrical:

3. Histograms are also described by how many major peaks they have. Histogram A is an example of a distribution with a single peak that is not symmetrical.

    Which other histograms have this feature?

NAME DATE PERIOD

4. Some histograms have a gap, a space between two bars where there are no data points. For example, if some students in a class have 7 or more siblings, but the rest of the students have 0, 1, or 2 siblings, the histogram for this data set would show gaps between the bars because no students have 3, 4, 5, or 6 siblings.

   Which histograms do you think show one or more gaps?

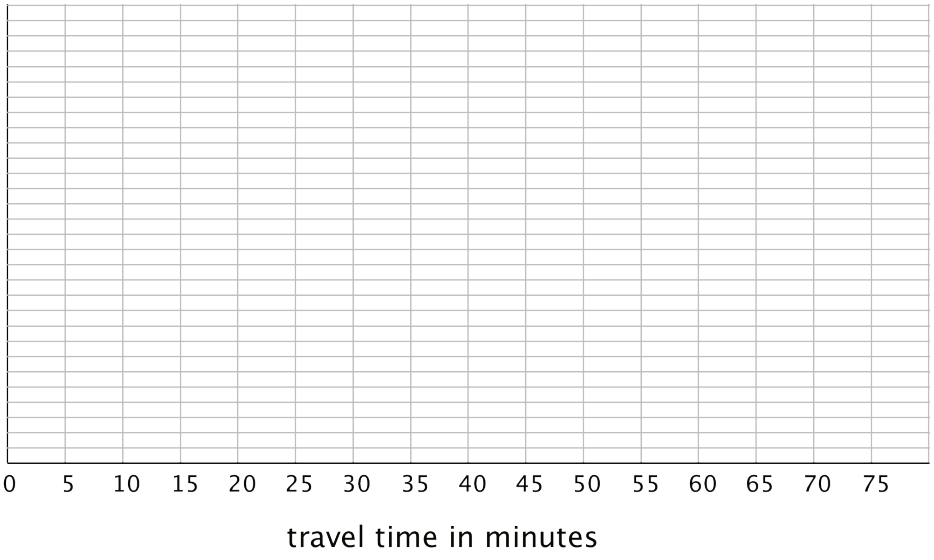
5. Sometimes there are a few data points in a data set that are far from the center. Histogram A is an example of a distribution with this feature.

   Would you describe any of the other histograms as having this feature? If so, which ones?

8.3: Getting to School

Your teacher will provide the data that your class collected on how students travel to school and their travel times.

1. Use the data to draw a histogram that shows your class's travel times.



2. Write a couple of sentences to describe the distribution of travel times. Comment on the center and spread of the data, as well as the shape and features.

NAME DATE PERIOD

3. Use the data on methods of travel to draw a graph. Include labels for the horizontal axis.

4. Write 2–3 sentences to describe what you learned about your class's methods of transportation to school. Comment on any patterns you noticed.

5. Compare the histogram and the bar graph that you drew. Discuss your thinking with your partner:

    - How are they alike?
    - How are they different?

Are you ready for more?

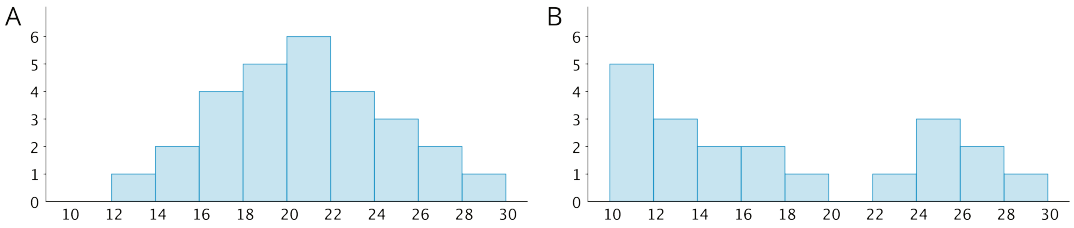
Use one of these suggestions (or make up your own). Research data to create a histogram. Then, describe the distribution.

- Heights of 30 athletes from multiple sports
- Heights of 30 athletes from the same sport
- High temperatures for each day of the last month in a city you would like to visit
- Prices for all the menu items at a local restaurant

NAME DATE PERIOD

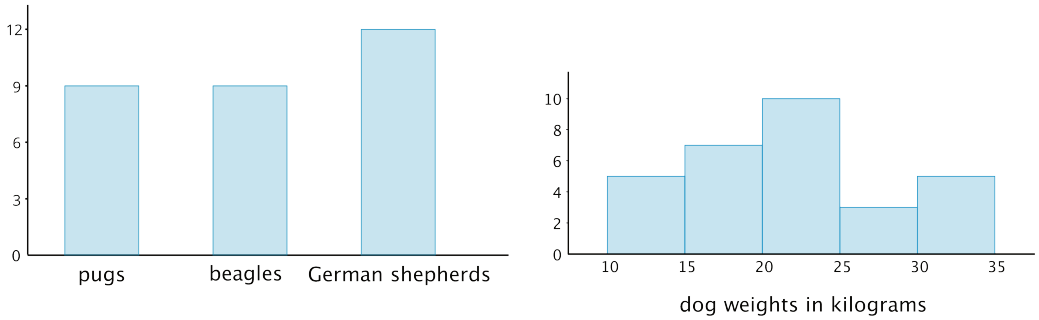
Lesson 8 Summary

We can describe the shape and features of the distribution shown on a histogram. Here are two distributions with very different shapes and features.



- Histogram A is very symmetrical and has a peak near 21. Histogram B is not symmetrical and has two peaks, one near 11 and one near 25.
- Histogram B has two clusters. A cluster forms when many data points are near a particular value (or a neighborhood of values) on a number line.
- Histogram B also has a gap between 20 and 22. A gap shows a location with no data values.

Here is a bar graph showing the breeds of 30 dogs and a histogram for their weights.



Bar graphs and histograms may seem alike, but they are very different.

- Bar graphs represent categorical data. Histograms represent numerical data.
- Bar graphs have spaces between the bars. Histograms show a space between bars *only* when no data values fall between the bars.
- Bars in a bar graph can be in any order. Histograms must be in numerical order.
- In a bar graph, the number of bars depends on the number of categories. In a histogram, we choose how many bars to use.