

What's Your Question?

1

Understanding the Statistical Process

WARM UP

Ms. White asked the 25 sixth graders in her class, "How many pets do typical 6th graders in our class have?" Ms. White summarized the responses in the table.

Number of Students	Number of Pets
8	0
10	1
4	2
2	3
1	38

1. About how many pets does each sixth grader in Ms. White's class own? How did you make your decision?

LEARNING GOALS

- Recognize and design statistical questions and anticipate variability in data related to the question.
- Differentiate between surveys, observational studies, and experiments.
- Describe the four stages of the statistical process.
- Discuss the different types of data that can be collected, displayed, and analyzed.
- Analyze and interpret bar graphs and circle graphs.

KEY TERMS

- variability
- data
- statistical question
- statistical process
- categorical data
- quantitative data
- population
- sample
- survey
- observational study
- experiment
- bar graph
- circle graph
- frequency
- mode

You have been solving mathematical problems throughout this course. Now, you are going to study statistical problems. How are mathematics and statistics similar and different?

Getting Started

In statistics, **variability** means that the value of the attribute being studied can change from one person or thing to another.

Data are categories, numbers, or observations gathered in response to a statistical question.

Statistical or Not, That Is the Question

Have you ever wondered, “How much money do professional athletes make?” Or, “How long are the books assigned to sixth graders?” If so, you have asked a statistical question. If you have sought out the answer to your question, you have engaged in the statistical process.

Cut out the questions provided at the end of the lesson. Read each question and sort them into as many groups as you would like. There must be more than one group and there must be at least two questions per group.

1. Record your groups and the questions in each group.

In this module, you will begin your formal study of statistics and the statistical process. Statistics is a problem-solving process, also called an investigative process, because the heart of statistics is about determining a possible answer to a question that has *variability*.

Statistical problem solving begins with a *statistical question*. A **statistical question** is a question that anticipates an answer based on *data* that vary.

2. Which questions from your sort are statistical questions? Explain how you would expect the answers to those questions to vary.

Keep your questions. You will need them later in the lesson.



Formulating Statistical Questions



The **statistical process** has four components:

- Formulating a statistical question.
- Collecting appropriate data.
- Analyzing the data graphically and numerically.
- Interpreting the results of the analysis.

This lesson provides an overview of the statistical process, but you will continue to use the process throughout your study of statistics. Statistics is about posing interesting questions that you want to answer about varying attributes.

Analyze the questions posed by Bianca and Rajan.

Bianca



"What clubs am I in?"
"How many students are in the Chess Club?"

Rajan



"What clubs do my classmates belong to?"
"How many members do the clubs at my school have?"

1. Explain why Bianca's questions are not statistical questions but Rajan's are.

2. What kinds of answers do you expect from Rajan's questions?

Categorical data, or *qualitative data*, are data for which each piece of data fits into exactly one of several different groups or categories.

Quantitative data, or *numerical data*, are data for which each piece of data can be placed on a numerical scale and compared.

Answering a statistical question requires collecting variable data. You will learn about two types of data: *categorical data* and *quantitative data*.

3. Would the answers to Rajan’s questions be categorical or quantitative?

4. Gather the statistical questions from the *Statistical or Not* activity. Which questions have categorical answers and which have quantitative answers?

5. For each question, determine if it is a statistical question. If it is not, rewrite it as a statistical question. Then, state if the data would be categorical or quantitative.

a. How many text messages did you send and receive yesterday?

b. What are the most popular school mascots?

c. How much time did you spend watching TV or playing video games last weekend?

Just as graphs can be described as discrete or continuous, quantitative data can be described as discrete or continuous.



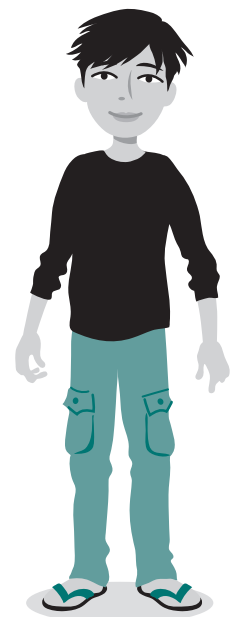
d. How many hours do 6th graders sleep each night?

Some data, like area codes, are numbers but are not quantitative variables. This data serves as a label, or category.

e. What is your favorite sport?

6. Write at least 2 additional statistical questions that you would be interested in answering. State if the data would be categorical or quantitative.

“Don't worry about the answer to the question, if there even is one. What would you like to know?”





For this activity, let's consider the topic of school lunches.

1. Write three statistical questions that you can ask about school lunches.

a.

b.

c.

The second component of the statistical process is to collect the data to answer the statistical question.

A statistical question can be answered by collecting data from an entire *population* or, more commonly, from a *sample* of the population. A **population** is an entire set of items from which data are collected. A **sample** is a selection from a population.

For example, to answer the question "How tall are 6th graders?" using the population of all 6th graders, you would need to determine the heights of every 6th grader in the world. However, you could choose to answer the question by collecting data from a sample of 6th graders—the 6th graders at your school.

Three common methods of data collection are *surveys*, *observational studies*, and *experiments*. In a **survey**, people are asked one or more questions. Similarly, in an **observational study**, the researcher (you!) collects data by observing the variable of interest. In an **experiment**, the researcher imposes a condition and observes the results.

You could conduct an experiment to investigate if 6th graders perform better on an assessment if they read a textbook or watch a video about the material. You would randomly assign half the students to read the text and half the students to watch the video. All students would be given the same assessment. You would compare the scores of the students in the two groups.

2. For each statistical question you wrote in Question 1, identify the population and sample of interest.

a.

b.

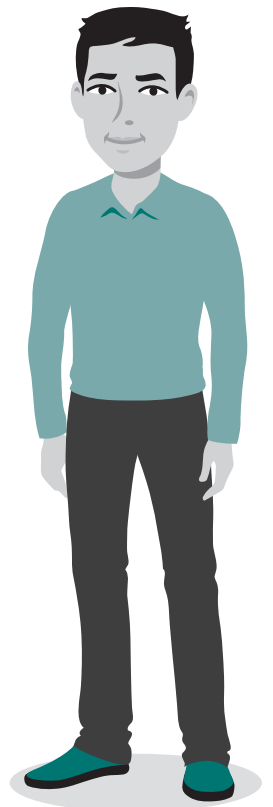
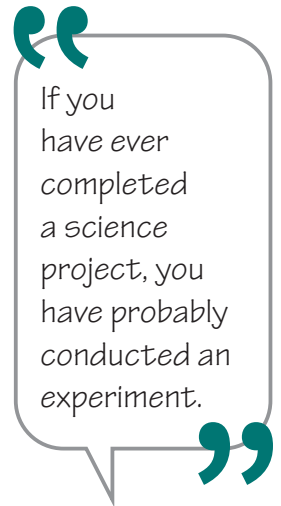
c.

3. Do you think a survey, observational study, or experiment would be the best way to collect the data to answer your statistical questions? Explain your reasoning.

a.

b.

c.



ACTIVITY
1.3

Analyzing and Interpreting Data



Formulating a statistical question is Part 1 of the statistical process.

Suppose you are interested in characteristics of sixth graders at your school.

- 1. Formulate three categorical statistical questions and survey your class to obtain a sample.**

Collecting data is Part 2 of the statistical process.

Analyzing the collected data is Part 3 of the statistical process.

In the statistical process, after you collect the data, it is time to analyze and interpret the results. Analysis includes selecting the most appropriate graphical display and numerical summaries for your question and your method of data collection.

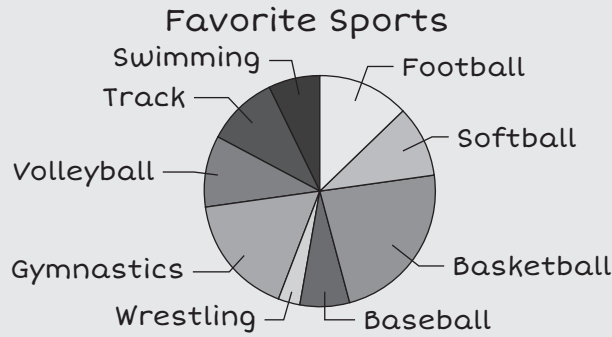
You already have experience displaying and summarizing categorical data using *bar graphs* and *circle graphs*.

A **bar graph** displays categorical data using either horizontal or vertical bars on a graph. The height or length of each bar indicates the value for that category.

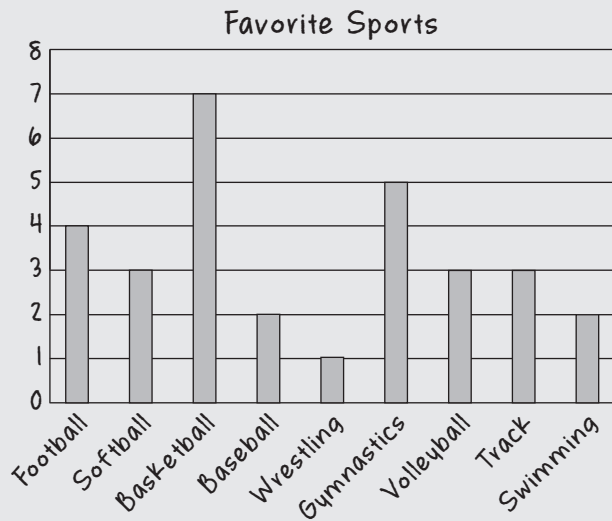
A **circle graph**, often called a pie chart, displays categorical data using sectors, or “wedges,” of a circle. It shows how parts of the whole relate to the whole and how parts of the whole relate to the other parts. The area of each sector corresponds to the percentage of the part in relation to the whole.

Nicole and Neal were interested in the favorite sports of 6th graders. They surveyed their class of 30 students. Then, they displayed their class’s data in different ways. Analyze each graph.

Nicole



Neal



2. How are the graphs similar? How are they different?

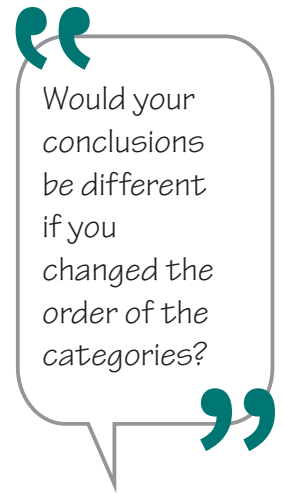
In order to create the graphs, Nicole and Neal determined the *frequency* of each response and recorded the frequencies in a frequency table. A **frequency** is the number of times an item or number occurs in a data set. Once the frequency is known, you can determine the *mode*. The **mode** is the value or values that occur most frequently in a data set.

Sport	Frequency (f)
Football	4
Softball	3
Basketball	7
Baseball	2
Wrestling	1
Gymnastics	5
Volleyball	3
Track	3
Swimming	2

Interpreting the data, or drawing conclusions, is Part 4 of the statistical process.

3. What conclusions can you make about the most popular sport in Nicole and Neal's class? Use the table and graphs to explain your reasoning.

4. Compile your class's responses to the 3 survey questions you asked in Question 1. Record the frequency of each response in a table.



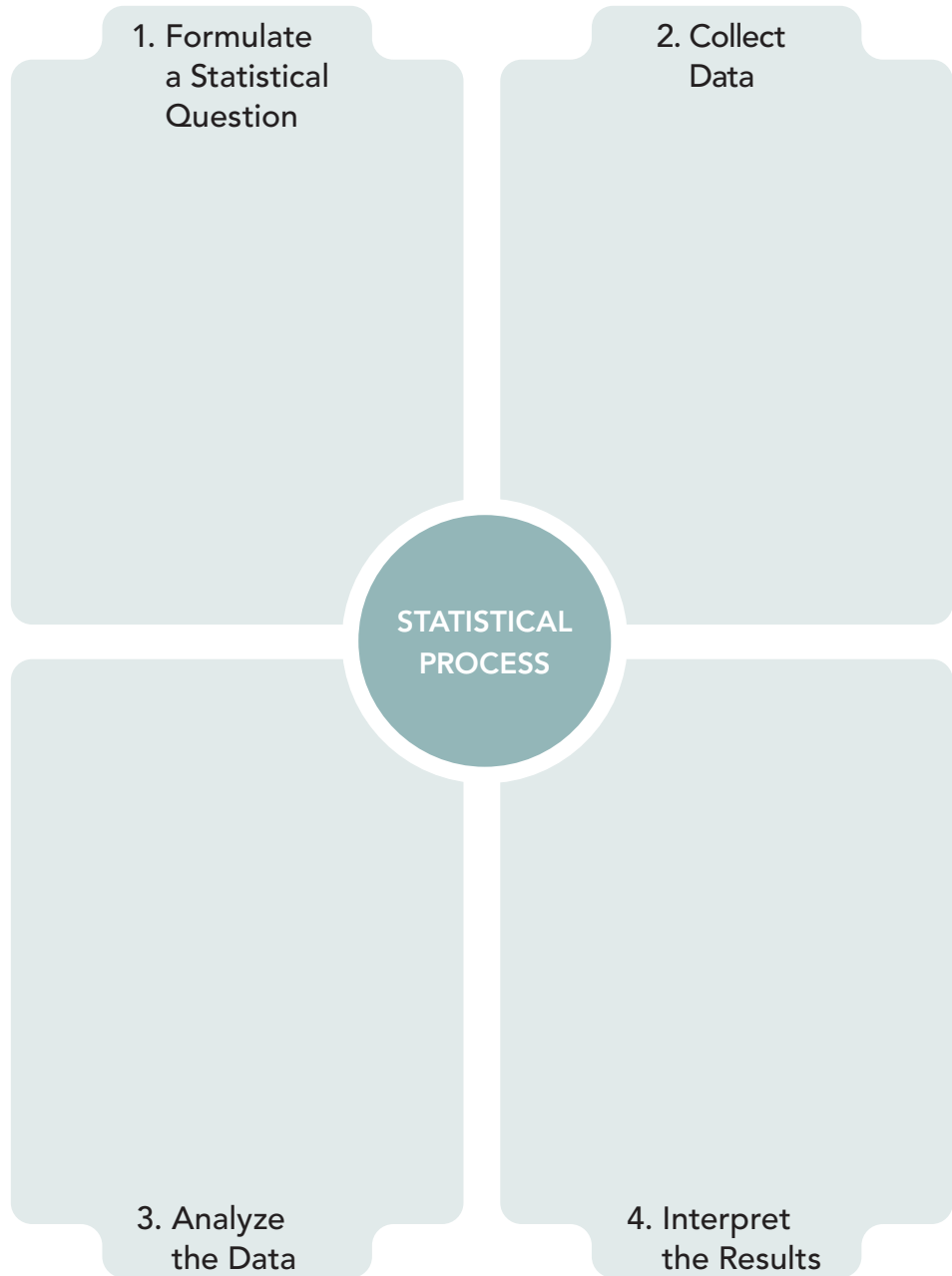
5. Create a graphical display for your assigned survey question. What conclusions can you make about your class based on your graph?



TALK the TALK

(Graphically) Organizing the Process

Complete the graphic organizer for the statistical process. In each section, summarize what you know about the component and provide examples.



A

What is your favorite sport?

B

How many TVs are in your house?

C

What is the most popular favorite color in the school?

D

How far do I travel to school?

E

What grade did I earn on my last math test?

F

How many siblings do I have?

G

How many siblings do 6th graders have?

H

How tall are 6th graders?

I

How many TVs are in a U.S. household?

J

What is your favorite color?

K

How many clubs are at my school?

L

How tall am I?

M

What kinds of sports do 6th graders prefer?

N

How far do the students at my school travel to get to school each day?

O

What grades were earned on the last math test?



Assignment

Write

Match each definition to its corresponding term.

1. an entire set of items from which data can be selected
 2. the information that is collected from an experiment, study, or survey
 3. a question that anticipates variability
 4. imposing a condition to test a specific result
 5. a method for collecting information by asking one or more questions
 6. a method for collecting information by observing a phenomenon in action
 7. a subset of a population
 8. the value of an attribute, or quality, being studied can change from one person or thing to another
 9. data for which each piece of data fits into exactly one of several different groups or categories
 10. data which can be placed on a numerical scale and compared, and can consist of discrete or continuous variables
 11. a graph that shows how parts of the whole relate to the whole and how parts relate to other parts
 12. a way of displaying categorical data by using either horizontal or vertical bars so that the height or length of the bars indicates the value for that category
 13. the number of times an item or number occurs in a data set
 14. the observation or value that occurs the most
- a. data
 - b. experiment
 - c. bar graph
 - d. variability
 - e. statistical question
 - f. categorical data
 - g. circle graph
 - h. survey
 - i. observational study
 - j. population
 - k. sample
 - l. frequency
 - m. mode
 - n. quantitative data

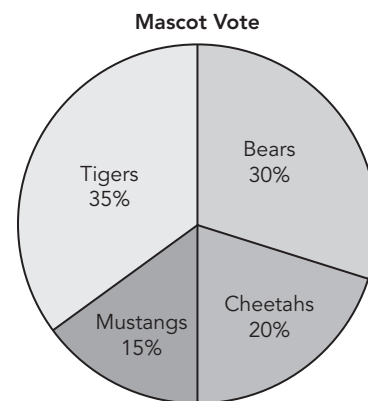
Remember

There are four components to the statistical process:

- Formulate a statistical question.
- Collect data.
- Analyze the data using graphical displays and numerical summaries.
- Interpret the results in terms of the original statistical question and context.

Practice

- Determine whether each given question is a statistical question. If not, rewrite it to make it a statistical question.
 - How many people in your class like to play video games?
 - Is pizza your favorite food?
 - What time do you go to bed on school nights?
- Determine whether a survey, observational study, or experiment would be the best way to answer each given statistical question.
 - "How many of the students in your class ate breakfast this morning?"
 - "Which students in your school can run a 40-meter sprint the fastest?"
 - "How many students in your class can type at least 30 words per minute?"
 - "How many students in your class ride the bus to school each day?"
- Determine whether each set of given data are categorical or quantitative. If the data are quantitative, determine whether they are discrete or continuous.
 - Each student in your math class records their height.
 - The members of the Horse Club list the types of horses they have.
 - The members of the Horse Club list the numbers of horses they each have.
- Tamara claims that Sweet Grove apple juice tastes better than Juicy Bushels apple juice. Isaac claims that there is no difference between the 2 types of apple juice. Tamara and Isaac would like to find the answer to the following question: Do more 6th graders prefer Sweet Grove apple juice or Juicy Bushels apple juice?
 - Is this a statistical question? Explain your reasoning.
 - Explain how this question can be answered with an experiment.
- The circle graph shows the results of the vote for the new school mascot.
 - If 400 students voted, how many students voted for the Cheetahs?
 - Create a bar graph to display the information, in terms of frequency, of each mascot.
 - What conclusions about the question of what mascot should be adopted can you make based on the graphs?



Stretch

In 1945, George Polya published a book about mathematical problem solving. He outlined a four-step process for problem solving:

1. Understand the Problem
2. Devise a Plan
3. Carry out the Plan
4. Look Back

Research the four steps and explain how the four-component statistical problem-solving process is similar to and different from Polya's four steps for mathematical problem solving.

Review

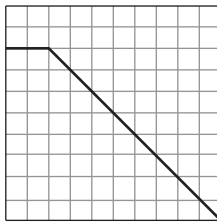
1. Choose the graph that best represents each scenario. Explain your reasoning.

a. Carla fills a mug with tea. Every few minutes Carla takes a drink from the mug.

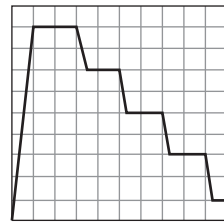
A.



B.

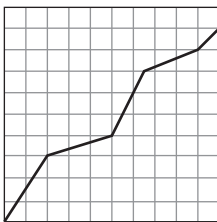


C.

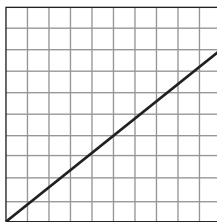


b. When Jamal rides his bike up a hill, his speed decreased. When he rides down a hill, his speed increased.

A.



B.



C.



2. Use absolute value equations to justify each answer.

- a. Determine the distance between the horizontal lines that contain points $A(7, 5)$ and $B(-4, -8)$.
- b. Determine the distance between the vertical lines that contain points $A(7, 5)$ and $B(-4, -8)$.

3. Insert a $>$, $<$, or $=$ symbol to make each number sentence true.

a. $-9\frac{1}{8}$ _____ -9.4

b. 0.006 _____ 1%

