Unit 8, Lesson 2

Statistical Questions

Goals

- Justify (orally) whether a question is "statistical" based on whether variability is expected in the data that could be collected.
- Match survey questions to data sets representing possible responses and justify (in writing) why they match.

Learning Targets

- I can tell statistical questions from non-statistical questions and explain the difference.
- I can tell when data has variability.

Lesson Narrative

In this lesson, students continue to analyze questions and the kinds of responses they can expect from those questions. They begin to recognize **variability** in data and learn about **statistical questions** and how they differ from non-statistical questions. In order to define variability, students categorize data sets and name the categories to make use of structure by seeking mathematically important similarities between the objects (MP7).

Standards

Building On	**5.MD.B.2**
Addressing	**6.SP.A, 6.SP.A.1, 6.SP.B, 6.SP.B.5.b**

Instructional Routines

- Card Sort
- MLR2: Collect and Display
- MLR8: Discussion Supports
- Take Turns

Required Materials

Materials to Gather

- Dot stickers: Activity 1
- Rulers: Activity 1

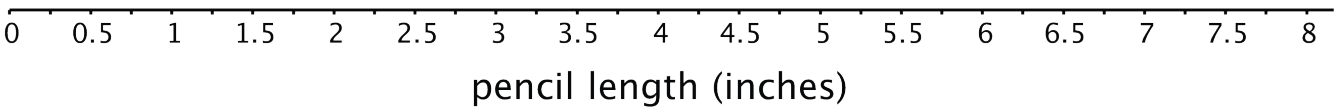
Materials to Copy

- Sifting for Statistical Questions Cards (1 copy for every 1 students): Activity 4

Required Preparation

Lesson:

Prior to the *Warm-up*, prepare a large class dot plot for the class to use with a horizontal axis labeled "pencil length (inches)" and intervals of $\frac{1}{4}$ from 0 to 8.



One dot sticker for each student to include in the class dot plot.

Prior to the "What Makes a Statistical Question" activity, prepare a two-column table to record students' observations, and display it for all to see.

Print and cut up cards from the "Sifting for Statistical Questions" blackline master. Prepare one set of cards for every two students.

Student Facing Learning Goals

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

2.1 Pencils on A Plot

5 min

Warm-up

Activity Narrative

The purpose of this *Warm-up* is for students to review how to represent measurements on a dot plot and how to interpret the data.

Standards

Building On	**5.MD.B.2**
Addressing	**6.SP.B**

Launch

Arrange students in groups of 2. Distribute rulers marked in inches to each group, and ensure that each student has a pencil.

Display the large class dot plot prepared before class for all to see and access. Tell students to measure the length of their pencil to the nearest $\frac{1}{4}$ inch and record their measurement as a dot on the class dot plot. Give each student a dot sticker as a way to record their measurement.

When the class data is recorded, give students 1 minute of quiet work time. Then, ask partners to briefly share their responses and follow with a whole-class discussion.

Student Task Statement

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.

Student Response

Answers vary. There may or may not be one most common pencil length, depending on the distribution of the data.

Building on Student Thinking

Some students may struggle with subtracting the shortest pencil length from the longest. Ask if they could use the horizontal axis to find the difference (for example, by adding up from the shorter length to the longer one).

Activity Synthesis

The purpose of the discussion is for students to recognize the usefulness of the dot plot structure.

Ask a student to share their responses for each of the questions. Record and display their reasoning for all to see. After the student shares, ask the class if they agree or disagree and why. Some discussion may arise about the interpretation of the most common pencil size. It is ok to allow some ambiguity at this time.

To involve more students in the conversation, consider asking some of the following questions:

- Who can restate ___'s reasoning in a different way?
- Did anyone have the same response but would explain it differently?
- Did anyone find the difference between the shortest and longest lengths in a different way?
- Does anyone want to add on to _____'s reasoning?

2.2 What's in the Data?

15 min

Activity Narrative

In this activity, students reason abstractly and quantitatively (MP2) about numerical data sets to match them with questions that are likely to produce the data. They see that some survey questions lead to responses that are expected to vary when posed to different people, but others produce responses that are likely to be the same. Students categorize data sets based on whether more than one different value is present and make use of this structure (MP7) to define variability.

As students match questions with data sets, look out for different plausible explanations for their choices. Their matches are reasonable if they can explain why the given data could be responses to a question. Identify one student to share the response to each question. Also notice students who offer different but equally reasonable explanations for the same data set, and invite them to share later.

Standards

Addressing **6.SP.A, 6.SP.B.5.b**

Launch

Tell students that they will be looking at numerical data sets and thinking about what question could produce the responses in each data set. Emphasize that they need to be able to support their matching decisions with reasonable explanations.

If necessary, guide students to understand how to read the table by asking:

- "What values are included in Data Set A?" (0, 1, 1, 3, 0, 0, 0, 2, 1, 1)

- "To which data set does the number 8 belong?" (Data Set C)
- "How many people answered '1' to the question that produced the data for Data Set A?" (4)

Keep students in groups of 2. Give them 5 minutes of quiet work time and 1–2 minutes to share their responses with their partner.

Student Task Statement

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?
   - Question 2: How many books did you read in the last year?
   - Question 3: What grade are you in?
   - Question 4: How many dogs and cats do you have?
   - Question 5: How many inches are in 1 foot?
2. How are survey Questions 3 and 5 different from the other questions?

Student Response

1.
   - Question 1: C. Sample reasoning: The numbers in the other data sets are unlikely or impossible (they are too large or too small, are greater than 10, or all the same). When we flip a coin the chance of getting a head is one-half, and the numbers in the data are around 5, which is half of 10.
   - Question 2: E. Sample reasoning: The numbers make sense for the number of books read in a year. It is unlikely that they all read the same number of books or that some read 0 books.
   - Question 3: D. Sample reasoning: The students being interviewed are sixth-grade students.
   - Question 4: A. Sample reasoning: Values of 0–3 make sense for the number of pets people might have.
   - Question 5: B. Sample reasoning: The number of inches in a foot is always 12, regardless of who answers the question.
2. Sample response: Everyone surveyed gave the same answer for these 2 questions.

Building on Student Thinking

Some students may have trouble matching questions and data sets because they do not attend carefully to the range of possible solutions. For example, they may not notice that a data set with 11 as a data value cannot be a response to the first question about flipping a coin 10 times. Ask them to study the questions and data values more closely, and to look for values that seem unlikely or impossible for a given context.

Activity Synthesis

The purpose of this discussion is for students to define **variability** and recognize when it is present.

Select previously identified students to share their choices and explanations. Briefly poll the class after each explanation to see if others made the same choice for the same reason. If not, invite students with different explanations to share.

Discuss how the question about grade level and the one about number of inches in a foot are different from the others. If not mentioned by students, highlight the idea of variability. Explain that we use the term variability to describe data sets in which not every data value is the same. Data sets B and D are unlike the other sets because they show no variability.

2.3 What Makes a Statistical Question?

15 min

Activity Narrative

In this activity, students use both the experience of reasoning about questions and the idea of variability to define **statistical questions** as questions that can be answered by using data in which variability is expected.

For example, the question, "What is the favorite subject of students in my class?" is a statistical question because we need data about favorite subjects and we can expect students to have different preferences. The question, "What is the counselor's favorite subject?", is not a statistical question because it can be answered by collecting a single data value. Even if multiple responses were collected, the responses are not expected to show variability.

As students analyze and discuss examples and non-examples of statistical questions, listen for groups who distinguish the two in terms of the data needed to answer the questions. For example, some questions may require collecting data that will probably show some variability while other questions may have only a single response. Invite them to share later.

Access for English Language Learners

This activity uses the Collect and Display math language routine to advance conversing and reading as students clarify, build on, or make connections to mathematical language.

Standards

Addressing **6.SP.A.1**

Instructional Routines

- MLR2: Collect and Display

Launch

Arrange students in groups of 3–4. Give students 1–2 minutes of quiet time to study the examples and non-examples of

statistical questions and then 4–5 minutes to discuss with their group how the two sets are different and generate a rough definition of statistical questions.

While students discuss the first question in their groups, use *Collect and Display* to create a shared reference that captures students' developing mathematical language. Collect the language that students use to describe statistical questions and non-statistical questions. Display words and phrases such as "variability," "single answer," and "different values."

Pause the class after the first question. Direct students' attention to the reference created using *Collect and Display*. Ask students to share how they distinguish between statistical and non-statistical questions. Invite students to borrow language from the display as needed, and update the reference to include additional phrases as they respond.

If not mentioned by students, explain that we use the term variability to describe data sets in which not every data value is the same. In contrast, finding out the color of the principal's car, whether Elena has a cell phone, and Diego's reading preference does not require data, or any data collected are not expected to vary.

Give students an additional 3–5 minutes to complete the remaining questions before the whole-class discussion.

Access for Students with Disabilities

Representation: Develop Language and Symbols. Maintain a visible display to record new vocabulary. Invite students to suggest details (words or pictures) that will help them remember the meaning of the phrase. *Supports accessibility for: Language, Memory*

Student Task Statement

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:
   a. How are the three statistical questions alike? What do they have in common?
   b. How are the three non-statistical questions alike? What do they have in common?
   c. How can you find answers to the statistical questions? How about answers to non-statistical questions?
   d. 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? _____

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? _____

Student Response

1. Sample explanations:
   - The statistical questions are all about a group (a group of cars, or a group of students). Their answers would require some work to figure out (counting, comparing, calculating, etc.)
   - The non-statistical questions are all about an individual. Their answers seem to be pretty straightforward and can be easily found.
   - To answer statistical questions some research and data collection would be needed (for example, counting the number of cars of each color; asking each student in the school if they have a cell phone or if they prefer fiction or nonfiction). To answer the non-statistical questions one could simply ask the principal, Elena, or Diego the relevant question.
   - A statistical question requires gathering and studying data to answer.
2. The answers to both parts should match.
   a. Yes
   b. No
   c. Yes
   d. No
   e. No

Building on Student Thinking

Students might think that if the response to a question requires counting or some kind of analysis then the question is statistical. Though statistical questions do require analysis, help students see that the starting point for distinguishing a statistical question is to see whether the data used to answer it have variability, which would then determine if analysis is called for.

Activity Synthesis

Invite students to share their responses to the second set of questions. Be sure that students understand that a

question is statistical if we need data to answer it and that the data are expected to have variability.

Determining whether a question is statistical or not may depend on the situation in which it is being asked. Here are two examples to ask students:

- "Who is the tallest student in the class?"
- "How long is the longest river in the United States?"

While the tallest student may be obvious in some classrooms, it is helpful to remember that this is not true in all classrooms. The students in a class are often close (but not identical) in height, and finding out who is tallest requires collecting data about different heights then analyzing it to determine who is the tallest. So the question, "Who is the tallest person in the class?", is generally a statistical question, but there may be specific situations in which it is not.

Likewise, while the longest river in the country can be easily researched, it is helpful to remember that this was not always the case. The answer may be considered a fact now, but the question was once a statistical question—at some point, lengths of rivers were collected and compared in order to answer it.

To tell statistical questions from non-statistical ones, it is useful to look closely at the context of the questions and what it takes to answer them.

2.4 Card Sort: Sifting for Statistical Questions

15 min

Optional

Activity Narrative

This optional activity provides additional practice in determining what it means for a question to be statistical in nature.

Students develop a deeper understanding of statistical questions by studying a wider range of examples and non-examples. Students sort a variety of questions and explain why they are or are not statistical. During the card sort, they explain their reasoning, critique the reasoning of others, and attempt to persuade one another (MP3). Students also begin writing statistical questions and think about the data that might be used to answer the questions.

As students work, encourage them to refine their descriptions of statistical questions using more precise language and mathematical terms (MP6).

Access for English Language Learners

This activity uses the *Collect and Display* math language routine to advance conversing and reading as students clarify, build on, or make connections to mathematical language.

Standards

Addressing **6.SP.A.1**

Instructional Routines

- Card Sort
- MLR2: Collect and Display
- MLR8: Discussion Supports
- Take Turns

Launch

Arrange students in groups of 2 and distribute pre-cut cards. Allow students to familiarize themselves with the representations on the cards:

- Give students 1 minute to sort the cards into categories of their choosing.
- Pause the class after students have sorted the cards.
- Select groups to share their categories and how they sorted their cards or started sorting their cards.
- Discuss as many different types of categories as time allows.

Use *Collect and Display* to direct attention to words collected and displayed from an earlier activity. Ask students to suggest ways to update the display: "Are there any new words or phrases that you would like to add?" "Is there any language you would like to revise or remove?" Encourage students to use the display as a reference.

After a brief discussion, invite students to complete the remaining questions.

The word "typical" appears for the first time in this activity (in one of the questions to be sorted: "What is a typical number of students per class in your school?"). The term is connected to the idea of center and spread later in the unit, but is used informally here. If needed, explain that we can think of "typical" as meaning what is common or what can be expected in a given group.

Access for Students with Disabilities

Representation: Develop Language and Symbols. Maintain a display of important terms and vocabulary. Invite students to suggest language or diagrams to include that will support their understanding of statistical questions. Terms may include "typical value" and examples and non-examples of statistical questions.
Supports accessibility for: Conceptual Processing, Language

Access for English Language Learners

MLR8 Discussion Supports. Students should take turns finding a match and explaining their reasoning to their partner. Display the following sentence frame for all to see: "I noticed _____ , so I matched . . . ." Encourage students to challenge each other when they disagree.
Advances: Conversing

Student Task Statement

Your teacher will give you a set of cards. Each card contains a question.

Sort the cards into two groups based on whether it is a statistical question or not. Be prepared to explain how you know where each question belongs.

Student Response

Statistical questions: A, C, D, E, F, I, J, K

Not statistical questions: B, G, H, L

Building on Student Thinking

Students may think that the number of possible answers to a question is what defines a statistical question. For

example, they may say that the question, “How many cups of water do my classmates drink each day?”, is not a statistical question because there is one answer. Ask students how they would arrive at the answer.

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 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.

Extension Student Response

I agree with Han. While it is true that only one person lives the farthest from school, that information is not likely to be a known fact, so to find out who lives the farthest requires analyzing data that can be expected to vary.

Activity Synthesis

Most of the discussions happen in small groups. Bring the class together to discuss any remaining disagreements or questions.

Direct students' attention to the reference created using *Collect and Display*. Ask students to share how they decided which cards went into each group. Invite students to borrow language from the display as needed, and update the reference to include additional phrases as they respond. (For example, “We decided that the height of the door is not statistical because it requires only a single measurement with no variability.”)

Ask the class:

- “Do you all agree with the list of the statistical questions?”
- “If not, which one(s) are harder to distinguish? Why?”
- “Now that you have seen more examples of statistical questions, what new insights do you have about them?”

Ensure that students understand the difference between statistical questions and survey questions. A survey question is what is used to collect data. A statistical question is one that is answered using collected data.

For example, the question, “Are most residents of this building older or younger than 30?” is a statistical question, because answering it requires collecting and analyzing the ages of the residents. It is not a question you would ask people directly, though, so it is not a survey question.

The related survey question is “How old are you?” because it can be used to gather data about the ages of people in a group being studied. On its own, though, it is not a statistical question because there is a single response without any variability when it is asked.

Lesson Synthesis

In addition to looking at numerical and categorical data more closely, students explored whether data show **variability**,

and the kinds of questions the data sets could help answer.

Here are some questions for discussion:

- "What does it mean to say that data have variability?" (Not all the data values are the same.)
- "When might we expect data to have variability?" (When the question we are trying to answer is about a feature or a characteristic of a group that has different members, with different features or characteristics.)
- "When might we expect data to have no variability?" (When the question we are trying to answer is about an individual, or about a feature that all group members have in common.)
- "Give some examples of data that would show variability and some that would not show variability."
- "What is a **statistical question**?" (A question that can be answered using data that are expected to have variability.)
- "Give some examples of statistical and non-statistical questions."
  - "What kinds of data are needed to answer them?"
  - "How might you collect the data?"
  - "What units of measurement are involved?"

2.5 Questions about Temperature

5 min

Cool-down

Standards

Addressing **6.SP.A.1**

Student Task Statement

Here are two questions:

Question A: Over the past 10 years, what is the warmest temperature recorded, in degrees Fahrenheit, for the month of December in Miami, Florida?

Question B: At what temperature does water freeze in Miami, Florida?

1. Decide if each question is statistical or non-statistical. Explain your reasoning.
2. If you decide that a question is statistical, describe how you would find the answer. What data would you collect?

Student Response

1.
   - Question A is a statistical question. Sample reasoning: The temperature in Miami in December changes from day to day and from year to year.
   - Question B is not a statistical question. Water freezes at sea level at 32 degrees Fahrenheit. This is a known fact.
2. To answer Question A (about the warmest temperature), find the temperature records for the past ten years and look for the highest value in degrees Fahrenheit.

Responding to Student Thinking

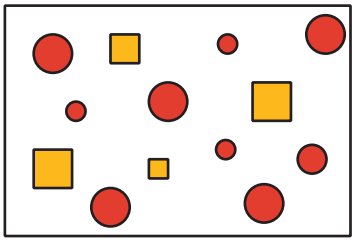
Points to Emphasize
If students struggle with identifying whether questions are statistical or non-statistical, use the questions and data sets in upcoming lessons as opportunities to ask whether the questions posed were statistical or non-statistical.

Lesson 2 Summary

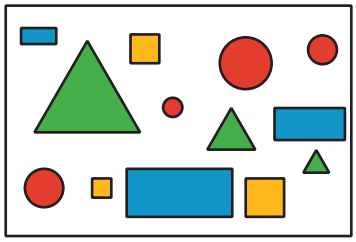
We often collect data to answer questions about something. The data we collect may show **variability**, which means the data values are not all the same.

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 be answered only 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. It would only require one person to count the books to get the answer, so there is no variability. 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.

Glossary

- statistical question
- variability

Lesson 2 Practice Problems

1 Student Task Statement

Sixth-grade students are asked, “What grade are you in?” Explain why this is not a statistical question.

Solution

This is not a statistical question, because all of the students are in the sixth grade, so there wouldn't be any variability in the data collected.

2 Student Task Statement

Lin and her friends go out for a snack after school. These questions come up during their trip. Select **all** the questions that are statistical questions.

A. How far are we from school?

B. What is the most popular snack sold at this shop this week?

C. What does a group of 4 people typically spend on snacks at this shop?

D. Do kids usually prefer to get salty or sweet snacks?

E. How many snacks are there to choose from?

Solution

B, C, D

3 Student Task Statement

Here is a list of questions about the students and teachers at a school. Select **all** the questions that are statistical questions.

A. What is the most popular lunch choice?

B. What school do these students attend?

C. How many math teachers are in the school?

D. What is a common age for the teachers at the school?

E. About how many hours of sleep do students generally get on a school night?

F. How do students usually travel from home to school?

Solution

A, D, E, F

4 Student Task Statement

Here is a list of statistical questions. What data would you collect and analyze to answer each question? For numerical data, include the unit of measurement that you would use.

a. What is a typical height of female athletes on a team in the most recent international sporting event?

b. Are most adults in the school football fans?

c. How long do drivers generally need to wait at a red light in Washington, DC?

Solution

Sample responses:

a. Collect the heights, in inches, of female athletes on the team.

b. Collect responses from the teachers about whether or not they are football fans.

c. Determine the number of seconds between the moment a car stops at a red light and the moment it starts moving again for multiple cars at multiple intersections in Washington, DC.

5 from Unit 7, Lesson 13

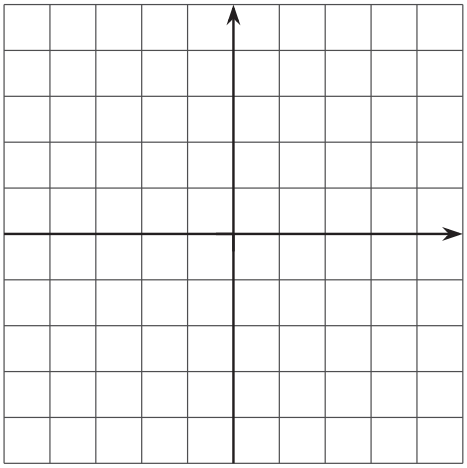
Student Task Statement

Describe the scale that you would use on the coordinate plane to plot each set of points. What value would you assign to each unit of the grid?

a. $(1, -6), (-7, -8), (-3, 7), (0, 9)$

b. $(-20, -30), (-40, 10), (20, -10), (5, -20)$

c. $(\frac{-1}{3}, -1), (\frac{2}{3}, -1\frac{1}{3}), (\frac{-4}{3}, \frac{2}{3}), (\frac{1}{6}, 0)$



Solution

Sample responses:

a. Axes labeled by multiples of 2
b. Axes labeled by multiples of 10
c. Axes labeled by multiples of $\frac{1}{3}$

6 from Unit 7, Lesson 9

Student Task Statement

Noah's water bottle contains more than 1 quart of water but less than $1\frac{1}{2}$ quarts. Let w be the amount of water in Noah's bottle, in quarts. Select **all** the true statements.

A. w could be $\frac{3}{4}$.
B. w could be 1.
C. $w > 1$
D. w could be $\frac{4}{3}$.
E. w could be $\frac{5}{4}$.
F. w could be $\frac{5}{3}$.
G. $w > 1.5$

Solution

C, D, E

7 from Unit 7, Lesson 7

Student Task Statement

Order these numbers from least to greatest:

$|-17|$ $|-18|$ -18 $|19|$ 20

Solution

-18 $|-17|$ $|-18|$ $|19|$ 20