Lesson 1: Got Data?

Goals

- Ask survey questions (orally) and record responses (in writing). Include units of measurement when reporting numerical data (orally and in writing).
- Comprehend and use the terms “numerical” and “categorical” to describe data sets (orally and in writing).
- Interpret various representations of data sets and determine whether it is reasonable that a verbal description represents a given numerical data set.

Learning Targets

- I can collect the correct data to answer a question and use the correct units.
- I can explain the difference between categorical and numerical data.

Lesson Narrative

Students begin the unit by interacting closely with data. They collect data about themselves by measuring and answering survey questions, studying the different types of responses collected, and identifying the appropriate variables and units being measured.

Students learn about **categorical** and **numerical** data. They determine whether a particular survey question will produce one type of data or the other. They also get reacquainted with dot plots (often called line plots in earlier grades) as a way to represent data and make sense of what the data points mean in context (MP2).

Alignments

Building On

- 2.MD.D.9: Generate measurement data by measuring lengths of several objects to the nearest whole unit, or by making repeated measurements of the same object. Show the measurements by making a line plot, where the horizontal scale is marked off in whole-number units.
- 4.MD.A.1: Know relative sizes of measurement units within one system of units including km, m, cm; kg, g; lb, oz.; l, ml; hr, min, sec. Within a single system of measurement, express measurements in a larger unit in terms of a smaller unit. Record measurement equivalents in a two-column table. For example, know that 1 ft is 12 times as long as 1 in. Express the length of a 4 ft snake as 48 in. Generate a conversion table for feet and inches listing the number pairs $(1, 12)$, $(2, 24)$, $(3, 36)$, ...
- 5.MD.B: Represent and interpret data.

Addressing

- 6.SP.B: Summarize and describe distributions.

Building Towards

- 6.SP.B.4: Display numerical data in plots on a number line, including dot plots, histograms, and box plots.

Instructional Routines

- Group Presentations
- MLR2: Collect and Display
- Notice and Wonder

Required Materials

Blank paper
Measuring tapes
Rulers
Sticky notes

Required Preparation

For the activity Surveying the Class:

Choose 4–5 survey questions and measurement activities in advance. Be sure to include questions and activities that would produce both categorical and numerical data. The questions about how and how long it takes students travel to school (the first two prompts) and students' heights in centimeters (the third prompt) will be used in a later lesson, so be sure to include these questions. Provide sticky notes to each student on which they can record their responses. Set up one position in the room for each selected question where students may put their sticky notes to have a visual display of responses.

To collect measurements, prepare measuring stations equipped with the necessary tools (e.g., rulers, measuring tape, etc.), instructions on how to measure, and a way to record the measurements. Students can then rotate through the stations.

Student Learning Goals

Let's explore different kinds of data.

1.1 Dots of Data

Warm Up: 10 minutes
The purpose of this warm-up is to review students' prior knowledge about Representation of numerical data. Students may be familiar with line plot from previous grades but unfamiliar with the term **dot plot**, which is what will be used in this unit and beyond. Students learn that both terms are commonly used for the same type of diagram.

Students examine a dot plot of data and consider which contexts may make sense for the data shown. Then students invent their own context for the data and interpret what additional data in that set would mean in their context.

This is an opportunity to see how students make sense of a data representation in the context of situations (MP2), which is central to the work of the unit.

Building On

- 5.MD.B

Building Towards

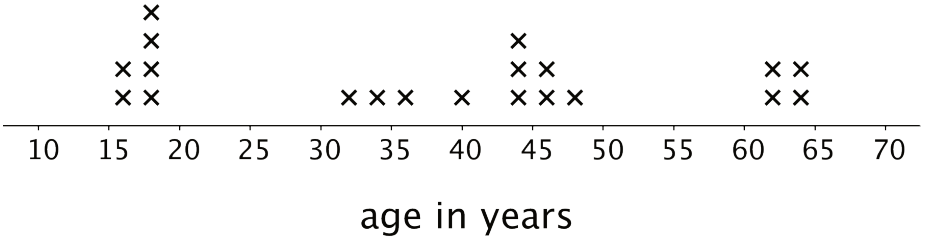
- 6.SP.B.4

Instructional Routines

- Notice and Wonder

Launch

Display the image for all to see.



Ask students what they notice and wonder about the image. Students should notice:

- The numbers represent age based on the label.
- Each X represents one person. For example, it looks like there is 1 40-year old represented but 3 44-year olds.
- There are a total of 20 people represented.

Students may wonder:

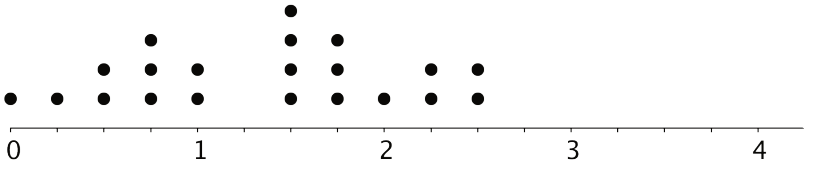
- Why are these people important?
- Why is there such a large age range?
- Why is there nobody in their 20s represented?

If not mentioned by students, remind them that the display is a line plot. Tell students that in the task they will see and use the same type of representation, but it is called a dot plot, and that dots are used instead of Xs.

Give students 2 minutes of quiet work time to complete the task, followed by a whole-class discussion.

Student Task Statement

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?

Student Response

1. Reasoning varies. Sample responses:
    a. No. We cannot have partial children.
    b. Yes. The distances are reasonable for the context and can be fractional.
    c. Yes. The hours are reasonable for the context and can be fractional.
    d. No. The unit of measurement (pounds) does not work with the context. Elephants weigh much more than several pounds.
    e. Yes. A teacher could grade a homework assignment on a scale of 1 to 4 and assign partial points.

2. Answers vary. Sample responses:
    a. Time spent napping, in hours.
    b. A dot represents the length of nap (in hours) on one day.
    c. A day without a nap.

Activity Synthesis

The purpose of the discussion is for students to understand how to read a dot plot including what dots represent in context.

Poll students on their responses to each scenario presented in the first question and ask a few students to explain their reasoning for each. After each student shares their explanation, ask the class if they agree or disagree and why.

Invite a few students to share another label they think could be used with the set of data and what each dot would represent. Based on each response, ask the class the following the questions:

- "What would a data point of 0 mean in the context mentioned?"
- "What would a data point of $3\frac{1}{4}$ mean in the context mentioned?"

1.2 Surveying the Class

20 minutes

Students begin their statistical explorations with data collection. They answer several survey questions, take some measurements, or do both. This data gathering activity serves several purposes: to give students firsthand experience in gathering and organizing data, to prompt students to notice different types of data that could be collected, and to provide the class with authentic data that can later be analyzed as students gain and expand the skills to do so.

The task statements show a range of survey questions and ideas for collecting measurements. To optimize the data gathering, have several survey questions and measuring activities identified ahead of time and consider how to best collect the responses or measurements.

Building On

- 2.MD.D.9

Addressing

- 6.SP.B

Instructional Routines

- Group Presentations
- MLR2: Collect and Display

Launch

Explain to students that they will gather data to learn more about the students in the class. Tell them which data sets will be collected and give instructions on the collection process. If students are to do a gallery walk or rotate through measuring stations, consider arranging them into groups of 3–4 to facilitate the rotation.

Access for Students with Disabilities

Action and Expression: Internalize Executive Functions. Chunk this task into more manageable parts to support students who benefit from support with organizational skills in problem solving. For example, demonstrate the data collection process and provide graphic organizers for collecting data from the survey questions.
Supports accessibility for: Organization; Attention

Anticipated Misconceptions

When taking measurements, students might not remember to attend to the right units, to start from 0, or to hold the measuring tool so that measurements can be precisely taken. Remind students about these issues as needed.

Student Task Statement

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.
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

Student Response

Data collected vary.

Activity Synthesis

Tell students to put their responses to the selected questions in the appropriate place in the room. Allow them to do a gallery walk of the data sets discussing with their group things they notice about each question. After students have had a chance to view the data, draw their attention to the twelve survey questions in the task statements. Ask them to think about the types of responses they produce. Students are likely to notice that responses to some questions are numbers and others

are not. Explain that responses that are measurements or quantities are called **numerical data**. For example, the first question about the travel time to school produces numerical data because the responses are quantities, measured in minutes.

Point out that responses to other questions are not quantities but can be sorted into categories. Explain that these types of responses are called **categorical data**. For example, the second question about ways of traveling to school produces categorical data because the responses can be sorted into categories.

Explain that numerical and categorical data will continue to be investigated in upcoming lessons.

The data from the first three questions regarding travel time, travel method, and height will be used in future lessons. Collect this data in an organized way to redistribute and use with students later.

Access for English Language Learners

Representing, Listening: MLR2 Collect and Display. During the gallery walk, listen for words and phrases students use to describe data that is "numerical" (measurements or quantitative values) and data that is "categorical" (where the values represent categories). Display students' words and phrases grouped into two lists prior to introducing the terms and definitions. This will help students to connect their informal use of language to vocabulary.
Design Principle(s): Optimize output (for describing); Support sense-making

1.3 Numerical and Categorical Data

10 minutes
In the previous activity, students responded to survey questions and collected data. They learned that data can be categorical or numerical. In this activity, students practice distinguishing **categorical** and **numerical** data, using the same survey questions and additional ones. They think about the kind of responses these questions would yield. For numerical responses, they consider the units of measurement. For categorical responses, they identify the characteristic being studied.

Note that some data may have numbers for their values but are categorical rather than numerical. Area codes and zip codes are examples of such categorical data. They are not quantities or measurements, but rather labels that happen to be numbers. We can meaningfully compare quantities or measurements (e.g., 6 minutes is greater than 4 minutes, or 7 years is less than 12 years), but we cannot do the same with the numbers in area codes or zip codes. It would not make sense to say, for example, that the zip code is 19104 is less than 63105.

As students work and discuss, notice students who could succinctly articulate the variables being investigated in both numerical and categorical questions. Also notice any disagreements partners might have about whether a question results in numerical or categorical data. The sixth survey question (rating the importance of environmental issues) and last survey question (hours spent on

out-of-school activities) are likely to generate conversations as they may appear less straightforward than the others.

Building On

- 5.MD.B

Addressing

- 6.SP.B

Launch

Arrange students in groups of 2. Tell students they will need the list of survey questions from the previous activity. Give them 3–4 minutes of quiet work time to complete the first three questions. Ask them to briefly discuss their responses with their partner before completing the last question.

Access for Students with Disabilities

Action and Expression: Internalize Executive Functions. Provide students with a Venn diagram with which to compare the similarities and differences between numerical and categorical data. *Supports accessibility for: Language; Organization*

Anticipated Misconceptions

Students may mistake numbers such as area codes, zip codes, or the numbers we use to represent months (e.g., 1 for January) as numerical data. Be sure to discuss this common confusion if arises. See the Activity Narrative section for ideas for addressing it.

Student Task Statement

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?

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.

Student Response

1. Answers vary. Questions 3, 4, 5, 6, 8, 9, 12 produce numerical data. Sample response:
   a. Question 3. Value being measured: height of student without shoes on. Unit of measurement: centimeters.

   b. Question 12. Values being measured: amount of time spent on different activities. Unit of measurement: minutes.

2. Answers vary. Questions 7, 10, 11 produce categorical data. Sample response:
   a. Question 7. Characteristic studied: whether student is the only child.

   b. Question 11. Characteristic studied: first activity after school.

3. a. Numerical

   b. Numerical

   c. Categorical

   d. Categorical

   e. Categorical

   f. Categorical

   g. Numerical

4. Answers vary. Sample response: Numerical: The number of people in a student's household. Categorical: The name of the street where a student lives.

Are You Ready for More?

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

This list 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 list 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.

Student Response

1. Answers vary. Sample response: Alike: They both collected the data of the same students and in the same order. Different: Priya recorded the names of the month (e.g., January, April, etc.), while Han recorded the numbers showing the order in which the months appear in the year (e.g., 5 for May, 7 for July, etc.).
2. Birth month is categorical. Sample explanation: The numbers that Han wrote are labels or names for months, which cannot be measured or compared (i.e., we don't say that month 5 is less than month 7). They appear in a certain order, but they are not quantities.

Activity Synthesis

The purpose of the discussion is for students to be able to identify data collected as either numerical or categorical.

Select a few previously identified students to share their responses to the first two questions. After each student shares, ask the rest of the class if they agree or disagree and discuss any

disagreements. Then, poll the class on their responses to the third set of questions. If not mentioned by students, explain how some categorical data are comprised of numbers, as noted in the Activity Narrative section.

Lesson Synthesis

In this lesson, we collected and explored different types of data. We noticed that certain survey questions produce responses that are quantities or measurements; we call these responses **numerical data**. Other questions produce responses that are not measurements or quantities but can be sorted into categories; we call these **categorical data**.

- "What are some examples of categorical data?"
- "What are some examples of numerical data?"
- "What is a **dot plot**?" (A dot plot is a representation of numerical data.)
- "How does it represent data?" (It represents each data value with a point above a number line.)

Consider creating a permanent display of the vocabulary throughout the unit including numerical data, categorical data, and dot plot for this lesson.

1.4 What's the Question?

Cool Down: 5 minutes

Building On

- 4.MD.A.1
- 5.MD.B

Addressing

- 6.SP.B

Student Task Statement

1. Would each survey question produce categorical data or numerical data?
    a. What is your favorite vegetable?
    b. Have you been to the capital city of your state?
    c. How old is the youngest person in your family?
    d. In which zip code do you live?
    e. What is the first letter of your name?
    f. How many hours do you spend outdoors each day?
2. Andre collected data measured in centimeters.

8.5 10.5 7.8 9.5 8.1 9.0 10.2 9.6 11.2 10.9 12.7 9.8

What could he be investigating? Select **all** that apply.

a. The weight of a dozen eggs.

b. The length of leaves from a tree.

c. The height of cups and mugs in a cupboard.

d. The length of songs on a music CD.

e. The length of colored pencils in a box.

Student Response

1. a. Categorical

   b. Categorical

   c. Numerical

   d. Categorical

   e. Categorical

   f. Numerical

2. b, c, e

Student Lesson Summary

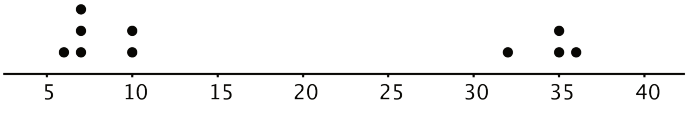
The table contains data about 10 dogs.

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.

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.

Glossary

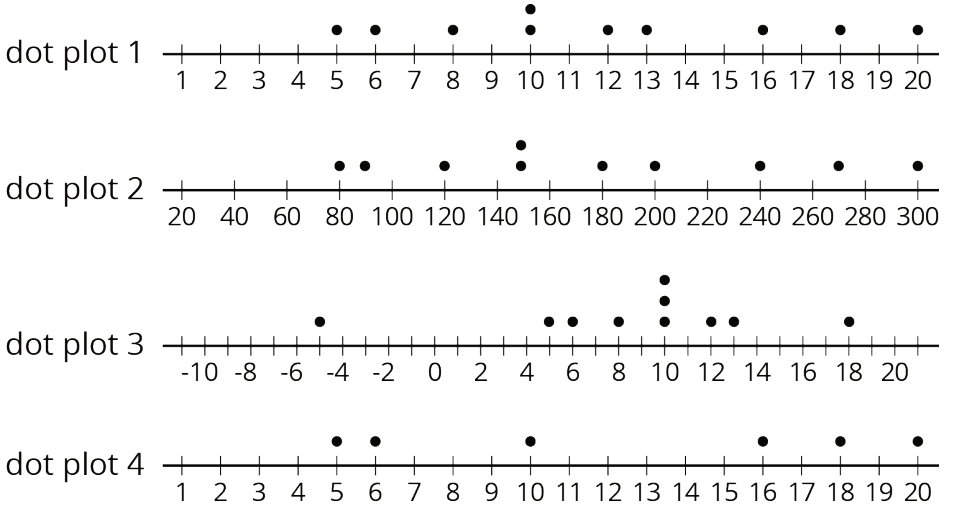
- categorical data
- dot plot
- numerical data

Lesson 1 Practice Problems

Problem 1

Statement

Tyler asked 10 students at his school how much time in minutes it takes them to get from home to school. Determine if each of these dot plots could represent the data Tyler collected. Explain your reasoning for each dot plot.



Solution

Answers vary. Sample responses:

- dot plot 1: This could be a dot plot of the time it takes to get to school in minutes for 10 students. The times range from 5 minutes to 20 minutes, which seems reasonable.
- dot plot 2: This couldn't be a dot plot of the time it takes to get to school in minutes because the values seem too big. The shortest time would be 80 minutes, which is more than an hour. The longest time would be 300 minutes, which is 5 hours. These don't seem like reasonable times that would be responses to the question that Tyler asked.
- dot plot 3: This couldn't be a dot plot of the time it takes to get to school in minutes because there are some negative values represented in the dot plot. The time it takes to get to school can't have a negative value.

- dot plot 4: This couldn't be a dot plot of the time it takes to get to school in minutes for 10 students because there are not 10 data values represented on the dot plot.

Problem 2

Statement

Here is a list of questions. For each question, decide if the responses will produce numerical data or categorical data and give two possible responses.

a. What is your favorite breakfast food?

b. How did you get to school this morning?

c. How many different teachers do you have?

d. What is the last thing you ate or drank?

e. How many minutes did it take you to get ready this morning—from waking up to leaving for school?

Solution

a. Categorical. Sample responses: cereal, toast

b. Categorical. Sample responses: walked, took the bus

c. Numerical. Sample responses: 3, 5

d. Categorical. Sample responses: water, apple

e. Numerical. Sample responses: 30, 45

Problem 3

Statement

a. Write two questions that you could ask the students in your class that would result in categorical data. For each question, explain how you know that responses to it would produce categorical data.

b. Write two questions that you could ask the students in your class that would result in numerical data. For each question, explain how you know that responses to it would produce numerical data.

Solution

Answers vary. Sample responses:

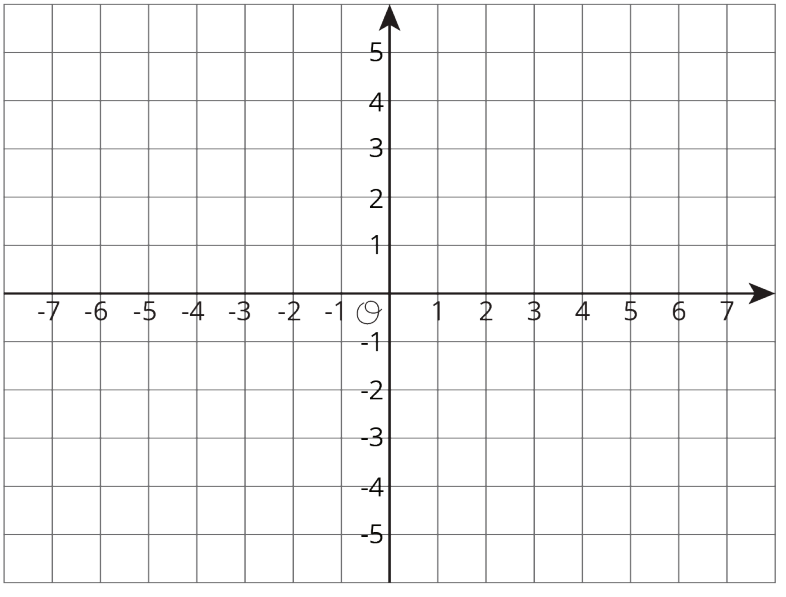
1. What is your favorite ice cream flavor? What town do you live in? These questions result in categories.

2. How many push-ups can you do in 1 minute? What is the distance from your home to the school? These questions result in quantities or measurements.

Problem 4

Statement

Triangle DEF has vertices $D = (-4, -4)$, $E = (-2, -4)$, and $F = (-3, -1)$.



a. Plot the triangle in the coordinate plane and label the vertices.

b. Name the coordinates of 3 points that are inside the triangle.

c. What is the area of the triangle? Show your reasoning.

Solution

a. Three points and the edges connecting them are plotted.

b. Answers vary. Sample response: $(-3, -2)$, $(-3, -3)$, $(-3, -2.5)$

c. 3 square units; the base is 2 units, the height is 3 units, and $3 \cdot 2 \div 2 = 3$.

(From Unit 7, Lesson 15.)