

Using Data to Solve Problems (Optional)

Goals

- Recognize that different graphical displays offer different insights into a distribution. Choose an appropriate graphical display to represent a data set, and justify the choice (orally and in writing).
- Recognize that different measures of center and variability offer different insights into a data set. Choose an appropriate measure of center and variability to describe a data set, and justify the choice (orally and in writing).

Learning Targets

- I can decide whether the mean and MAD, or the median and IQR, would be more appropriate for describing the center and spread of a data set.
- I can draw an appropriate graphical representation for a set of data.
- I can explain what the mean and MAD, or the median and IQR, tell us in the context of a situation, and I can use them to answer questions.

Access for Students with Diverse Abilities

- Representation (Activity 2)

Access for Multilingual Learners

- MLR8: Discussion Supports (Activity 2)

Instructional Routines

- Notice and Wonder

Required Materials

Materials to Gather

- Straightedges: Activity 2
- Tools for creating a visual display: Activity 3, Activity 4
- Internet-enabled device: Activity 4

Required Preparation

Lesson:

Provide access to straightedges for students to use when drawing box plots. For the main activity, preview the background information about the yellow perch fish. Prepare tools for creating a visual display, one set for every 3–4 students.

Lesson Narrative

This optional lesson is a good opportunity for students to use the information that they have learned in the unit and to apply it to different situations, but it may be shortened to fit time constraints. After the Warm-up, reinforcing interpretations of box plots in context, a few options are available for activities depending on the needs of the class.

In the first pair of activities, students have the option of calculating measures of center and variability by hand or by using a digital applet to do the calculation, and they then focus on interpreting these measures in context. They also have the option of creating visual representations of the distribution or viewing them in a digital applet, and they then decide which representation best shows the distribution.

If students are proficient with calculating statistics and understanding visual representations of data, the next activity gives students a chance to interpret real-world data and draw a conclusion about an environmental situation. Students analyze the summary of data and use it to construct an argument about the state of fish in a lake.

Lesson Timeline

5
min

Warm-up

15
min

Activity 1

15
min

Activity 2

30
min

Activity 3

20
min

Activity 4

10
min

Lesson Synthesis

5
min

Cool-down

Assessment

Using Data to Solve Problems (Optional)

Lesson Narrative (continued)

A third option is digital only and gives students a chance to work with a large data set. This can highlight the power of summarizing a distribution with a measure of center and a measure of variability to tell the story, rather than looking at the raw data or a visualization that shows all of the information.

Student Learning Goal

Let's compare data sets using visual displays.

Warm-up

Wild Bears

5 min

Activity Narrative

This *Warm-up* allows students to review two important ideas of this unit: interpreting data in a box plot and writing statistical questions based on a data set. Students write statistical questions based on given box plots, and then they trade questions to answer questions written by another student.

Launch

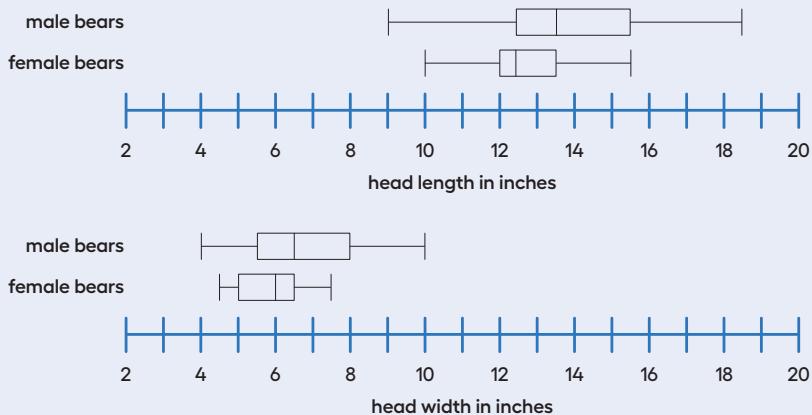


Arrange students in groups of 2. Tell students that, for the first question, one partner should write two questions about the head lengths and the other partner should write two questions about the head widths. For the second question, they should exchange and review each other's questions. If their partner's question does not seem to be a statistical question, suggest a revision so that it becomes a statistical question, and then answer the question. Remind students to consider units of measurement.

Give students 2 minutes of quiet work time for the first question and 2 minutes for collaboration afterward.

Student Task Statement

In one study on wild bears, researchers measured the head lengths and head widths, in inches, of 143 wild bears. The box plots summarize the data from the study.



1. Write four statistical questions that could be answered using the box plots: two questions about the head length and two questions about the head width.

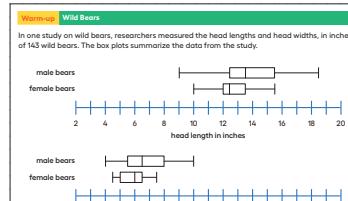
Sample responses:

- What is a typical head length, in inches, for male bears in the data set?
- Do female bears generally have longer heads than male bears?
- Which data set shows more variability in head widths: male bears or female bears?
- What is the widest head width, in inches, for male bears?
- How do the ranges of head widths compare? Which group—male bears or female bears—has a larger range, and by how much?

Student Workbook

Using Data to Solve Problems

Let's compare data sets using visual displays.



1. Write four statistical questions that could be answered using the box plots: two questions about the head length and two questions about the head width.
2. Trade questions with your partner.
 - Decide if each question is a statistical question.
 - Use the box plots to answer each question.

GRADE 6 • UNIT 8 • SECTION E | LESSON 18

Instructional Routines

Notice and Wonder
ilclass.com/r/10694948

Please log in to the site
 before using the QR
 code or URL.

**2. Trade questions with your partner.**

a. Decide if each question is a statistical question.

b. Use the box plots to answer each question.

Sample responses:

- A typical head length for male bears is about 13.5 inches.
- No, female bears generally have shorter heads than male bears.
- The data for male bears show more variability in head width.
- The widest head width for male bears is 10 inches.
- Male bears have a larger range of head widths. It is nearly twice as large as female bears' range of head widths. The range for male bears is about 6 inches, and the range for female bears is a little over 3 inches.

Activity Synthesis

Ask several students to share their questions about the head width and head length. Record and display their responses for all to see. After each student shares, ask the class if they agree or disagree that it is a statistical question. If they agree, ask how they would find the answer, or ask for the answer itself. If they disagree, ask how they could rewrite the question so that it is a statistical question.

Activity 1: Optional**Homework Analysis (Part 1)**

15
min

Activity Narrative**There is a digital version of this activity.**

In this activity, students compare and contrast different measures of center and variability for data sets that have gaps and are not symmetrical. They interpret mean, MAD, median, and IQR in the context of a situation. Unlike many of the data sets that students have seen so far, this one shows values that could roughly divide into three parts: the days when there is little or no homework, the days when there is a moderate number of homework problems, and the days when the assignment is relatively large. Because of this distribution, finding a typical number of homework problems (or whether it would be helpful to identify a typical number) is not obvious, prompting students to interpret measures of center and spread more carefully.

As students work and discuss, identify at least one student or group that decides that the mean and MAD are appropriate measures of center and spread and can explain their reasoning, and another that decides to go with the median and IQR and could support their choice. Invite them to share during a whole-class discussion.

In the digital version of the activity, students use an applet to calculate the IQR and MAD. The applet allows students to focus on finding the less calculation-heavy statistics. Use the digital version if students need more work on interpreting the statistics in context rather than on calculating their values.

Launch 

Keep students in groups of 2.

Display the dot plot.



Give students a moment of quiet time to look at the data on homework problems and identify at least one thing that they notice and one thing that they wonder about. Give them another brief moment to share their observation and question with their partner. Then, ask a few students to share their responses with the class.

Students are likely to notice that the data values are quite spread out, that there are some days with no homework and others with quite a few problems, that there is not an obvious cluster, or that the number of problems could be roughly grouped into three kinds (a little, moderate, and a lot). They are likely to wonder why the numbers are so spread out and varied.

Briefly discuss the following questions to encourage students to think about the data contextually:

- ❑ “Why might the homework assignment data show this distribution? What are some possible explanations?”

When only one problem was assigned, the problem might be particularly challenging or might require considerable work or collaboration. Another possibility: There might be an upcoming assessment, so the homework load was reduced. When many problems were assigned, the problems might be quick exercises with short answers, or the assignment might be review materials for an entire chapter.

- ❑ “How might we describe ‘a typical number of homework problems’ in this case?”

We could try using median, but because the numbers are so spread out, it’s hard to say what is typical.

- ❑ “Which do you predict would be higher: the mean or the median number of problems? Why?”

I predict the mean will be greater than the median because there are a couple of days with a lot more problems than the rest.

Give students 8–10 minutes to complete the task, either independently or collaboratively. Ask students to think quietly about the last question before discussing their response with their partner.

Student Workbook

1 Homework Analysis (Part 1)

Over a two-week period, Mai records the number of homework problems she has each school day.

2	15	20	0	5	25	1	0	10	12
---	----	----	---	---	----	---	---	----	----

1 Calculate these values. Show your reasoning.

- The mean number of math homework problems _____
- The mean absolute deviation (MAD) _____

2 Interpret the mean and MAD. What do they tell you about the number of homework problems Mai had over these two weeks?

Student Workbook

1 Homework Analysis (Part 1)

3 Find or calculate the following values and show your reasoning.

- The median, quartiles, maximum, and minimum of Mai's data _____
- The interquartile range (IQR) _____

4 Which pair of measures of center and variability—mean and MAD, or median and IQR—do you think summarizes the distribution of Mai's math homework assignments better? Explain your reasoning.

Student Task Statement

Over a two-week period, Mai records the number of homework problems she has each school day.

2 15 20 0 5 25 1 0 10 12

- 1. Calculate these values. Show your reasoning.**

- a. The mean number of math homework problems

The mean is 9 homework problems per day.

$$2 + 15 + 20 + 0 + 5 + 25 + 1 + 0 + 10 + 12 = \frac{90}{10} = 9.$$

- b. The mean absolute deviation (MAD)

The MAD is 7.4 homework problems per day. This is computed using the absolute deviations from the mean: 7, 6, 11, 9, 4, 16, 8, 9, 1, 3. The average of these deviations is $\frac{7+6+11+9+4+16+8+9+1+3}{10} = \frac{74}{10}$.

- 2. Interpret the mean and MAD. What do they tell you about the number of homework problems Mai had over these two weeks?**

Sample response: The mean tells us that a typical number of homework problems given in a day is 9. Because the MAD is 7.4, which is almost as large as the mean, the distribution is widely spread out. There is a lot of variation in the number of homework problems given each day.

- 3. Find or calculate the following values and show your reasoning.**

- a. The median, quartiles, maximum, and minimum of Mai's data

The data listed in order are: 0, 0, 1, 2, 5, 10, 12, 15, 20, 25. This means that Q1 is 1, Q2, or the median, is 7.5 (the average of the 5th and 6th data points), and Q3 is 15. The minimum is 0, and the maximum is 25.

- b. The interquartile range (IQR)

The IQR is 14, which is the difference between Q3 and Q1.

- 4. Which pair of measures of center and variability—mean and MAD, or median and IQR—do you think summarizes the distribution of Mai's math homework assignments better? Explain your reasoning.**

Sample response: Both the MAD and IQR show the large variability in the data, and the mean and median are pretty close in value. I think the median and IQR better summarize the center and spread of the data. There are more days on which not much homework was given, and the median is lower than the mean.

Activity Synthesis

Briefly discuss students' interpretations of the measures they just calculated:

What do the mean of 9 and MAD of 7.4 tell us? How can we interpret them in this context?

The mean tells us that she could've done 9 problems per day and completed the same total number of problems. The MAD being close to the value of the mean means that there is a large spread to the number of problems due each day.

- “What do the median of 7.5 and IQR of 14 tell us?”

The median tells us that half of the days had fewer than 7.5 problems assigned and half had more. The IQR means that the middle half of the distribution includes values that are up to 14 problems apart.

Then, select two or more previously identified students to share their responses about which measures of center and spread are appropriate for summarizing the data set. After each person shares, briefly poll the class to see if others reasoned the same way about the measures. Summarize by asking:

- “Now that you have two pairs of measures of center and spread, how would you respond if someone asked you for a typical number of homework problems that Mai did over these 2 weeks. Is the question easier to answer now?”

We have some values to say, but neither does a great job describing Mai’s homework. The most important feature about this distribution is the wide spread.

Students should walk away with increased awareness that, in some cases, measures of center and spread do not always paint a full picture of what the actual data set entails, and that the measures should be interpreted with care.

Activity 2: Optional

Homework Analysis (Part 2)

15
min

Activity Narrative

There is a digital version of this activity.

In this activity, students show a data set using three different kinds of graphical representations—a dot plot, a box plot, and histograms using different bin sizes—and decide which are more useful or more appropriate for communicating the distribution.

As students work and discuss, identify those who draw clear graphical displays, those who notice that the different displays offer different insights about the data distribution, and those who advocate for using different representations to display Jada’s data. Ask them to share with the class later.

In the digital version of the activity, students use an applet to create the graphical representations. The applet allows students to focus on interpreting the information. Use the digital version if students need more work on interpreting the statistics in context rather than on drawing the representations.

Launch



Arrange students in groups of 3–4. Provide access to straightedges.

Explain to students that they will now represent Jada’s homework data graphically and think about which representations might appropriately communicate the distribution of her data. Give students 4–5 quiet minutes to draw a dot plot and a box plot (the first two questions), and then another 4–5 minutes to collaborate on drawing histograms with different bin sizes. Ask each student in a group to be in charge of one histogram with a particular bin size. After all representations are drawn, students should analyze them and discuss the last question in their group.

Student Workbook

Homework Analysis (Part 2)

Jada wants to know whether a dot plot, a histogram, or a box plot would best show the distribution of her homework data.

0 0 4 4 5 5 7 8 10 10 12 14 16 18 20 22 24 26

1 Use the axis to make a dot plot to represent the data, and indicate the mean of 6.6 with a triangle. The MAD is 4.32.

2 Draw a box plot that represents Jada's homework data.

GRADE 6 • UNIT 8 • SECTION E | LESSON 18

23

Student Workbook

Homework Analysis (Part 2)

3 Work with your group to draw three histograms to represent Jada's homework data. The width of the bars in each histogram should represent a different number of homework problems.

a. The width of one bar represents 10 problems.

b. The width of one bar represents 5 problems.

GRADE 6 • UNIT 8 • SECTION E | LESSON 18

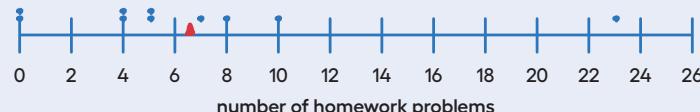
23a

Student Task Statement

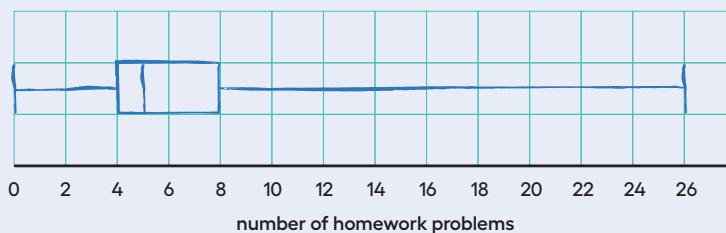
Jada wants to know whether a dot plot, a histogram, or a box plot would best show the distribution of her homework data.

0 0 4 4 5 5 7 8 10 12 14 16 18 20 22 24 26

1. Use the axis to make a dot plot to represent the data, and indicate the mean of 6.6 with a triangle. The MAD is 4.32.

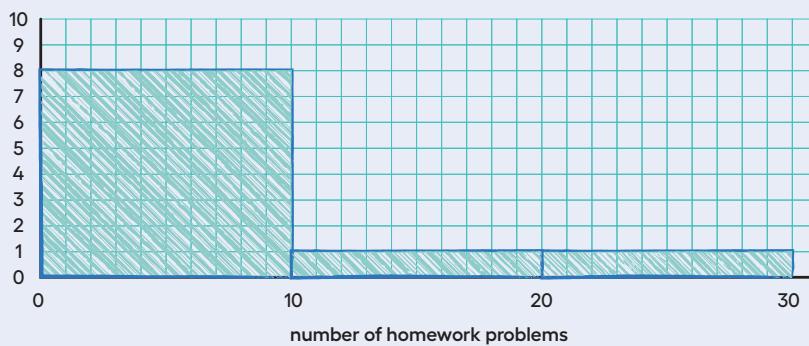


2. Draw a box plot that represents Jada's homework data.

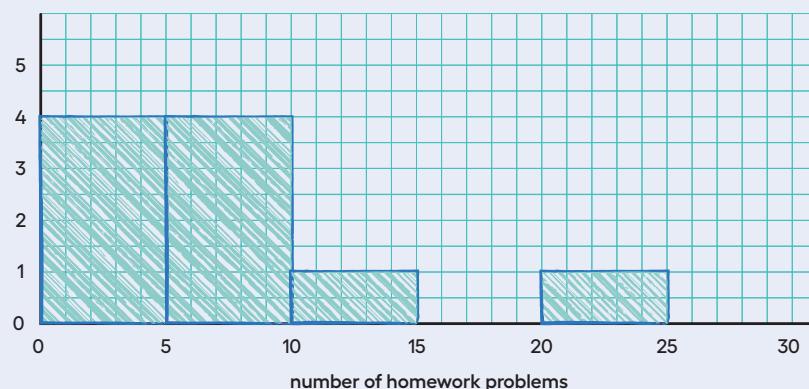


3. Work with your group to draw three histograms to represent Jada's homework data. The width of the bars in each histogram should represent a different number of homework problems.

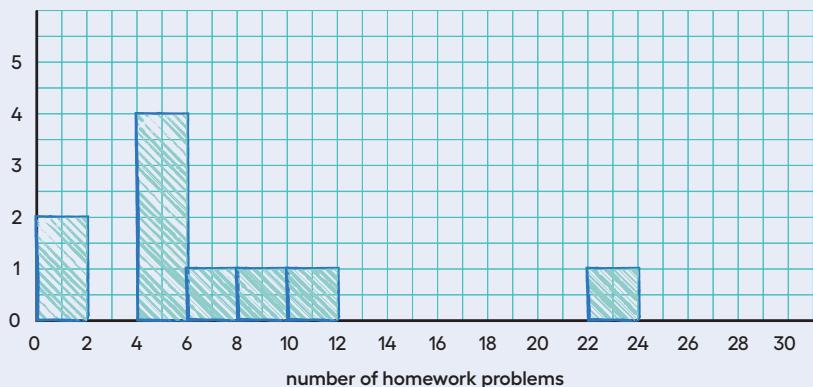
- a. The width of one bar represents 10 problems.



- b. The width of one bar represents 5 problems.



- c. The width of one bar represents 2 problems.

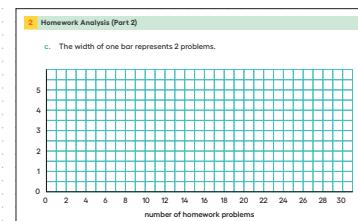


4. Which of the five representations should Jada use to summarize her data? Explain your reasoning.

Sample responses:

- o Jada should use the dot plot. Because it shows all the data points, we could use it to find the mean and MAD, or the median and IQR, to summarize the data. With only 10 points, it is easiest to understand the data with all of it visible.
- o Jada should use the box plot. It shows that for about three-quarters of the days Jada had 8 or fewer problems, so we know that days with more than 8 problems are not typical.
- o Jada should use the histogram with a bin size of 5. The histogram with a bin size of 2 is almost the same as the dot plot and doesn't help show an overall distribution very well. The histogram with a bin size of 10 puts almost all of the data in the first bar and loses too much information. The histogram with bin size of 5 is a good middle ground where we can see the shape of the distribution, but it is not too detailed.

Student Workbook



GRADE 6 • UNIT 8 • SECTION E | LESSON 18

Access for Students with Diverse Abilities (Activity 2, Synthesis)

Representation: Internalize Comprehension.

Use color coding and annotations to highlight connections between representations in a problem. For example, color code differences between how dot plots, box plots, and histograms represent the data.

Supports accessibility for: Visual-Spatial Processing

Activity Synthesis

Invite previously identified students to share their dot plot, box plot, and histograms. Display their drawings for all to see. Then, select several students or groups to share their response to their choice of which representation is best and their explanation. If not already mentioned by students, discuss the different insights that each display offers, or different challenges that each poses. For instance, consider asking these questions about each data display:

“What information can we get from this display?”

The dot plot shows all of the values and is good for estimating the center and spread. The box plot clearly shows the five-number summary. A histogram can show the shape of the distribution, but loses some information about the exact data values.

“Does it give us a meaningful snapshot of the distribution?”

Each display could be useful for different purposes. In this case, students may defend their choice using arguments similar to those listed in the response section.

“What characteristics of a different data set would make this representation more useful?”

Dot plots are good when there are not many different values. Box plots are good for when there is a lot of data or when the median and IQR are useful. The histograms are useful when there is a lot of data and the shape of the distribution is important.

Help students see that, in this case, none of the representations here are ill-suited to represent the data set, but a couple of them allow us to describe the distribution of the data set more easily because of how they summarize the data values in some ways.

Activity 3: Optional

Will the Yellow Perch Survive?

30
min

Activity Narrative

In this culminating activity, students use what they have learned in the unit to answer statistical questions about a species of fish in the Great Lakes region. They use a histogram to represent the given data distribution, decide on appropriate measures of center and variability, and use their analyses to draw conclusions about a certain fish population.

Launch



Tell students that they will now look at an example in which data analysis could be used to help conservation efforts. Provide students with this background information.

The yellow perch is a freshwater fish that is a popular food for people in the Great Lakes region (Minnesota, Wisconsin, Michigan, Illinois, Indiana, Ohio, Pennsylvania, and New York). In past research, samples of yellow perch taken from the Great Lakes seemed to be mostly male and mostly old. People worried that yellow perch might not survive, and efforts were made to limit commercial and individual fishing in order to try to increase the number of younger fish. An important part of these efforts is to periodically check the typical age of the fish in the Great Lakes.

The Wisconsin Department of Natural Resources and the Great Lakes Water Institute collected data from samples of yellow perch in Lake Michigan. Students at Rufus King High School in Milwaukee, Wisconsin, participated in the research. They evaluated the data and presented their findings in a student-conducted press conference. Explain to students that, in this task, they will investigate some of the same questions that these students addressed in their research.

Arrange students in groups of 3–4. Provide access to straightedges. Give students 7–8 minutes of quiet work time for the first three questions, and then 10–12 minutes to discuss their responses, complete the remainder of the task, and prepare a brief presentation on their responses to the last set of questions.

Give each group access to tools for creating a visual display. Ask them to support their conclusions with specific pieces of evidence, such as their histogram, their analysis of the distribution, measures of center and spread, and other features of interest.

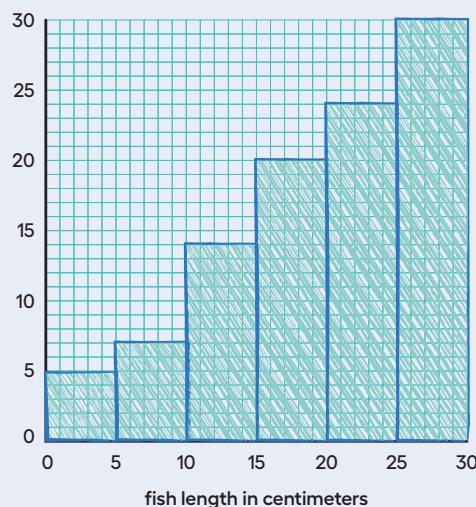
Student Task Statement

Scientists studying the yellow perch, a species of fish, believe that the length of a fish is related to its age. This means that the longer the fish, the older it is. Adult yellow perch vary in size, but they are usually between 10 and 25 centimeters.

Scientists at the Great Lakes Water Institute caught, measured, and released yellow perch at several locations in Lake Michigan. This summary is based on a sample of yellow perch from one of these locations.

length of fish in centimeters	number of fish
0 to less than 5	5
5 to less than 10	7
10 to less than 15	14
15 to less than 20	20
20 to less than 25	24
25 to less than 30	30

1. Use the data to make a histogram that shows the lengths of the captured yellow perch. Each bar should contain the lengths shown in each row in the table.



2. How many fish were measured? How do you know?

100 fish were measured.

The numbers of fish in all length groups add up to 100.

**Access for Multilingual Learners
(Activity 2, Student Task)****MLR8: Discussion Supports.**

At the appropriate time, give students 2–3 minutes to make sure that everyone in their group can explain what age group they would use to describe the typical fish and whether the researchers should be worried. Invite groups to rehearse what they will say when they share with the whole class.

Advances: Speaking, Conversing, Representing

Student Workbook

Will the Yellow Perch Survive?
Scientists studying the yellow perch, a species of fish, believe that the length of a fish is related to its age. This means that the longer the fish, the older it is. Adult yellow perch vary in size, but they are usually between 10 and 25 centimeters.
Scientists at the Great Lakes Water Institute caught, measured, and released yellow perch at several locations in Lake Michigan. This summary is based on a sample of yellow perch from one of these locations.

length of fish in centimeters	number of fish
0 to less than 5	5
5 to less than 10	7
10 to less than 15	14
15 to less than 20	20
20 to less than 25	24
25 to less than 30	30

Use the data to make a histogram that shows the lengths of the captured yellow perch. Each bar should contain the lengths shown in each row in the table.

Student Workbook

1 Will the Yellow Perch Survive?

- How many fish were measured? How do you know?

- Use the histogram to answer these questions.
 - How would you describe the shape of the distribution?

 - Estimate the median length for this sample. Describe how you made this estimate.

 - Predict whether the mean length of this sample is greater than, less than, or nearly equal to the median length for this sample of fish? Explain your prediction.

 - Would you use the mean or the median to describe a typical length of the fish being studied? Explain your reasoning.

GRADE 6 • UNIT 8 • SECTION E | LESSON 18

Student Workbook

1 Will the Yellow Perch Survive?

- Based on your work so far:
 - Would you describe a typical age for the yellow perch in this sample as: young, adult, or old? Explain your reasoning.

 - Some researchers are concerned about the survival of the yellow perch. Do you think the lengths (or the ages) of the fish in this sample are something to worry about? Explain your reasoning.

4 Butterfly Measurements

Navigate to this activity in the digital version of the materials.

The applet contains 127 measurements of wingspans and body length for butterflies captured in an area.

Select one of the measurements and use the applet to help create a display that summarizes the information, including any values calculated. Make sure that your display contains:

- At least one display of the distribution discussed in this unit, as well as a description of why you chose that method to display the distribution.
- A value for a measure of center, along with an explanation of why you chose that measure.
- A few sentences describing the distribution, including any additional features you notice

GRADE 6 • UNIT 8 • SECTION E | LESSON 18

3. Use the histogram to answer these questions.

a. How would you describe the shape of the distribution?

Sample response: The distribution is not symmetrical and has a peak in the range 25–30 cm.

b. Estimate the median length for this sample. Describe how you made this estimate.

Sample response: I estimate the median to be 22–23 cm. I look in the table for where the 50th and 51st value would be and see that it is in the 20 to less than 25 cm group.

c. Predict whether the mean length of this sample is greater than, less than, or nearly equal to the median length for this sample of fish? Explain your prediction.

Sample response: I predict the mean to be less than the median. Because the data has a peak on the right, the values that are to the left of the median (or less than the median) pull the mean down.

d. Would you use the mean or the median to describe a typical length of the fish being studied? Explain your reasoning.

Sample response: I would use the median, because it would better describe where the center of the data is (the data is not symmetric).

4. Based on your work so far:

a. Would you describe a typical age for the yellow perch in this sample as: young, adult, or old? Explain your reasoning.

Sample response: I would describe a typical fish as old because its length is on the higher end of the range of adult sizes.

b. Some researchers are concerned about the survival of the yellow perch. Do you think the lengths (or the ages) of the fish in this sample are something to worry about? Explain your reasoning.

Sample response: The lengths seem to be something to worry about. If the yellow perch in the Great Lakes tend to be old, and there are not many young fish around, the species might not survive.

Activity Synthesis

To allow all groups a chance to present, consider putting 2–3 groups together and asking them to present their work to each other. Groups that are not the first to present should focus on sharing new insights that have not been mentioned by the preceding groups. Invite students who are not presenting to attend carefully to the reasoning of the presenting group and to ask clarifying questions.

If time permits, highlight some conclusions that students draw about whether the fish in the sample are young, adult age, or old, and whether researchers should be worried.

Tell students that several years after the students at Rufus King High School participated in the research, newer samples of yellow perch showed more favorable length-age distributions: More of the the fish are smaller or younger.

Activity 4: Optional**Butterfly Measurements****20
min****Activity Narrative**

There is a digital version of this activity.

This optional activity provides students the opportunity to work with a large data set that would not be reasonable to work with by hand.

Throughout much of the unit, students have worked with smaller data sets so that they could check their understanding by hand and get a sense of the data as a whole. Sometimes, though, this makes the calculation of statistics seem unnecessary because many of them could be clearly seen. The power of these statistics comes out with larger data sets when the values are less obvious.

Due to the nature of the task, only a digital option is available.

Launch

Distribute internet-enabled devices, and give students instructions to navigate to this lesson in the digital version of the materials.

Student Task Statement

Navigate to this activity in the digital version of the materials.

The applet contains 127 measurements of wingspans and body length for butterflies captured in an area.

Select one of the measurements and use the applet to help create a display that summarizes the information, including any values calculated.

Make sure that your display contains:

- At least one display of the distribution discussed in this unit, as well as a description of why you chose that method to display the distribution
- A value for a measure of center, along with an explanation of why you chose that measure
- A value for a measure of variability, along with an explanation of why you chose that measure
- A few sentences describing the distribution, including any additional features you notice

Student Workbook**Will the Yellow Perch Survive?**

- Based on your work so far:
- Would you describe a typical age for the yellow perch in this sample as: young, adult, or old? Explain your reasoning.

- Some researchers are concerned about the survival of the yellow perch. Do you think the lengths (or the ages) of the fish in this sample are something to worry about? Explain your reasoning.

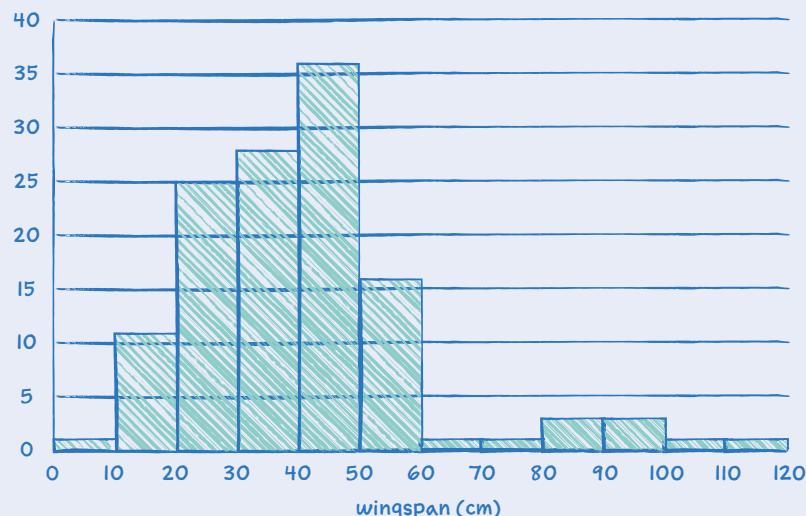
Butterfly Measurements

Navigate to this activity in the digital version of the materials. The applet contains 127 measurements of wingspans and body length for butterflies captured in an area.

Select one of the measurements and use the applet to help create a display that summarizes the information, including any values calculated. Make sure that your display contains:

- At least one display of the distribution discussed in this unit, as well as a description of why you chose that method to display the distribution
- A value for a measure of center, along with an explanation of why you chose that measure
- A value for a measure of variability, along with an explanation of why you chose that measure
- A few sentences describing the distribution, including any additional features you notice

Sample response:



This histogram with a bin size of 10 summarizes the data well by showing the shape of the distribution without getting lost in the details.

Most of the butterflies have a wingspan of about 38.9 centimeters based on the median. The IQR is about 21.8 cm, so there is a lot of variability in these data. The median and IQR are used here because the distribution is not approximately symmetric. It looks like the typical values are a little less than 40, but there are a few large butterflies that pull up the mean.

Activity Synthesis

Invite previously selected groups to share their visual displays. Encourage students to ask questions about the choice of data display and measures of center and variability that went into creating the display. Here are questions for discussion, if not already mentioned by students:

- ❑ “Which measure of center did you use to describe the distribution and why? What does it mean in this situation?”
We selected the median because the distribution didn’t appear to be almost symmetric. It means that half of the butterflies have a wingspan of 38.9 cm or longer and half have wingspans that are shorter.
- ❑ “Which measure of variability did you use to describe the distribution, and what does it mean in this situation?”
Because we used the median for the measure of center, we used the IQR to measure the variability. In this situation it means that the middle half of the butterfly wingspans are within 21.8 cm of one another.

Lesson Synthesis

In this lesson we practice finding measures of center and variability (mean, MAD, median, and IQR) and making sense of them in the context of the given situation. We notice that they give us different insights into the distribution of a data set.

“What do the mean and MAD tell us?”

The mean tells us the fair share or balance point of the distribution, and the MAD tells us the average distance a value is from the mean.

“How do we interpret this statement: ‘Noah’s mean number of homework problems per day is 10 and the MAD is 6.’?”

If we were to distribute Noah’s assignments so that the number of problems he has each day is the same, he would have 10 per day. The MAD of 6 tells us that there is some variability in the number of problems assigned, so not all days have exactly 10 problems assigned. The average distance between the number of problems assigned and the mean of 10 is 6.

“What do the median and IQR tell us?”

The median tells us the value for which half the data set is equal to or greater and half the data set is equal to or less, and the IQR tells us the range for the middle half of the data set.

“How do we interpret this statement: ‘Lin’s median number of homework problems per day is 10, and the IQR is 6.’?”

One half of Lin’s assignments involve 10 or fewer problems, and the other half involve 10 or more problems. The IQR tells us that the middle half of the data are all within 6 problems of one another.

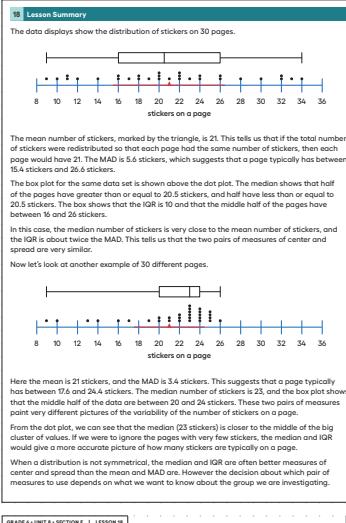
We also looked at different ways to graphically represent a numerical distribution.

“What are the ways we can represent a data set?”

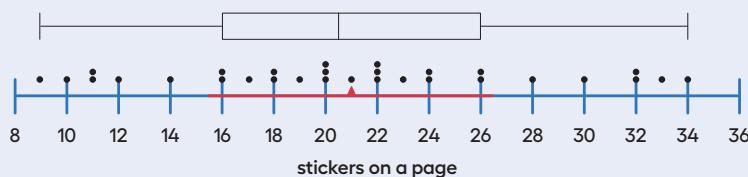
Dot plot, histogram, box plot.

“Which representations are helpful for summarizing a distribution?”

It varies depending on the distribution we’re studying and what information we want to know.

Student Workbook**Lesson Summary**

The data displays show the distribution of stickers on 30 pages.

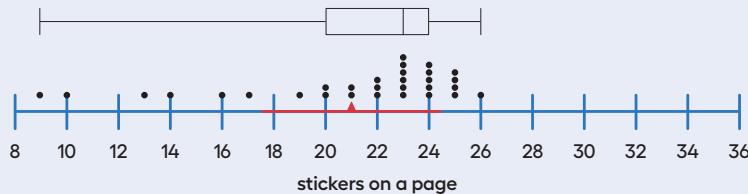


The mean number of stickers, marked by the triangle, is 21. This tells us that if the total number of stickers were redistributed so that each page had the same number of stickers, then each page would have 21. The MAD is 5.6 stickers, which suggests that a page typically has between 15.4 stickers and 26.6 stickers.

The box plot for the same data set is shown above the dot plot. The median shows that half of the pages have greater than or equal to 20.5 stickers, and half have less than or equal to 20.5 stickers. This tells us that the IQR is 10 and that the middle half of the pages have between 16 and 26 stickers.

In this case, the median number of stickers is very close to the mean number of stickers, and the IQR is about twice the MAD. This tells us that the two pairs of measures of center and spread are very similar.

Now let's look at another example of 30 different pages.



Here the mean is 21 stickers, and the MAD is 3.4 stickers. This suggests that a page typically has between 17.6 and 24.4 stickers. The median number of stickers is 23, and the box plot shows that the middle half of the data are between 20 and 24 stickers. These two pairs of measures paint very different pictures of the variability of the number of stickers on a page.

From the dot plot, we can see that the median (23 stickers) is closer to the middle of the big cluster of values. If we were to ignore the pages with very few stickers, the median and IQR would give a more accurate picture of how many stickers are typically on a page.

When a distribution is not symmetrical, the median and IQR are often better measures of center and spread than the mean and MAD are. However the decision about which pair of measures to use depends on what we want to know about the group we are investigating.

Cool-down**Time Spent on Chores**5
min**Student Task Statement**

Lin surveys her classmates on the number of hours they spend doing chores each week. She represents her data with a dot plot and a histogram.



- Lin thinks that she can find the median, the minimum, and the maximum of the data set using both the dot plot and the histogram. Do you agree? Explain your reasoning.

Sample response: Disagree. The dot plot makes it possible to find the median, the minimum, and the maximum fairly easily because it shows each data value individually. The histogram makes it possible to estimate these values, but it is impossible to tell the exact values because the data points are grouped together.

- Should Lin use the mean and MAD, or the median and IQR to summarize her data? Explain your reasoning.

Sample response: Lin should use the median and IQR because the data is not approximately symmetrical and has values far from the center. There are a few larger values that are not similar to most of the other values