Describing Distributions on Histograms

Goals

- Compare and contrast (orally) bar graphs and histograms, recognizing that descriptions of shape, center, and spread don't pertain to bar graphs.
- Describe (orally and in writing) the overall shape and features of a distribution represented on a histogram, including peaks, clusters, gaps, and symmetry.
- Identify histograms that display distributions with specific features.

Learning Targets

- I can describe the shape and features of a histogram and explain what they mean in the context of the data.
- I can distinguish histograms and bar graphs.

Lesson Narrative

In this lesson, students explore various shapes and features of a distribution displayed in a histogram. They use the structure to look for symmetry, peaks, clusters, gaps, and any unusual values in histograms. Students also begin to consider how features such as symmetry might affect how we characterize a data set. This work is informal, but helps to prepare students to better understand measures of center and spread later in the unit. In this lesson, students also distinguish between the uses and construction of bar graphs and histograms.

Student Learning Goal

Let's describe distributions displayed in histograms.

Access for Students with Diverse Abilities

• Action and Expression (Activity 2)

Access for Multilingual Learners

- MLR2: Collect and Display (Activity 2)
- MLR8: Discussion Supports (Activity 1)

Instructional Routines

- · Card Sort
- Which Three Go Together?

Required Materials

Materials to Gather

• Math Community Chart: Activity 1

Materials To Copy

- Sorting Histograms Cards (1 copy for every 4 students): Activity 1
- Getting to School Handout (1 copy for every 2 students): Activity 2

Required Preparation

Lesson:

The "Getting to School" activity requires students to use data previously collected on their travel methods and times. Organize the data into the tables in the blackline master ahead of time or allow time for students to do it themselves. Either make a copy for every 2 students, or display the completed tables for all to see during the activity.

Lesson Timeline

5_{min}

Warm-up

20 min

Activity 1

10 min

Activity 2

10 min

Lesson Synthesis

Assessment

5 min

Cool-down

Warm-up

Which Three Go Together: Histograms



Activity Narrative

This Warm-up prompts students to compare four histograms. It gives students a reason to use language precisely. It gives the teacher an opportunity to hear how students use terminology and talk about characteristics of the items in comparison to one another.

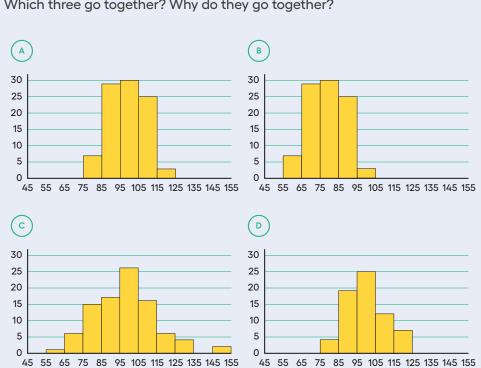
Launch



Arrange students in groups of 2-4. Display the histograms for all to see. Give students 1 minute of quiet think time and ask them to indicate when they have noticed three histograms that go together and can explain why. Next, tell students to share their response with their group, and then together to find as many sets of three as they can.

Student Task Statement

Which three go together? Why do they go together?



Sample responses:

- A, B, and C go together because they all have about 100 values in their data sets.
- A, B, and D go together because the spread from the least possible value to the greatest is 50.
- A, C, and D go together because their center is around 100.
- B, C, and D go together because they are shaped like a hill or a bell-shape.

Instructional Routines

Which Three Go Together?

ilclass.com/r/10690736





Student Workbook

Instructional Routines

Card Sort

ilclass.com/r/10783726

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



Activity Synthesis

Invite each group to share one reason why a particular set of three go together. Record and display the responses for all to see. After each response, ask the class if they agree or disagree. Because there is no single correct answer to the question of which three go together, attend to students' explanations and ensure that the reasons given are correct.

During the discussion, prompt students to explain the meaning of any terminology that they use, such as "center," "spread," or "distribution," and to clarify their reasoning as needed. Consider asking:

○ "How do you know ... ?"

"What do you mean by ...?"

"Can you say that in another way?"

Activity 1

Card Sort: Sorting Histograms



Activity Narrative

Students sort different histograms during this activity. A sorting task gives students opportunities to analyze representations, statements, and structures closely and make connections.

As students work, encourage them to refine their descriptions of the distributions using more precise language and mathematical terms.

Students may be familiar with the geometric meaning of symmetry. It is used in a similar way here to describe the shape of a distribution.

Launch

Math Community

Display the Math Community Chart for all to see. Give students a brief quiet think time to read the norms, or invite a student to read them out loud. Tell students that during this activity they are going to practice looking for their classmates putting the norms into action. At the end of the activity, students can share what norms they saw and how those norms supported the mathematical community during the activity.

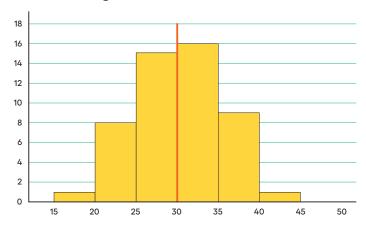
Arrange students in groups of 2 and distribute pre-cut cards. Allow students to familiarize themselves with the representations on the cards:

- 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/ started sorting their cards.
- Discuss as many different types of categories as time allows.

Attend to the language that students use to describe their categories and distributions, giving them opportunities to describe their distributions more precisely. Highlight the use of terms like "symmetric," "gaps," or "clusters." After a brief discussion, invite students to complete the remaining questions.

If the idea of symmetry is not brought up, ask students if they notice any symmetry in any of the distributions. If further discussion on the topic is helpful, display the image of the histogram here for all to see. Explain to students that a diagram of a distribution—a dot plot or a histogram—is described as symmetrical if you can draw a line on the diagram and the parts on one side of the line mirror the parts on the other side. Many distributions are not perfectly symmetrical, but are close to or approximately symmetrical.

The histogram here shows an approximately symmetrical distribution. When a line is drawn at the center (such as the line at 30) the two sides are roughly mirror images. If you were to fold the histogram at the line, the two sides would be close to matching.



Student Task Statement

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

- **1.** Sort the histograms into 2 groups based on whether they are approximately symmetrical or not. Be prepared to explain how you know where each system belongs.
 - Sample response: Approximately symmetrical: Histograms F, I, and J. Students may consider Histograms C, E and K as approximately symmetrical as well. Not approximately symmetrical: Histograms other than the ones previously listed.
- 2. Histograms are also described by how many major peaks they have. Histogram A is an example of a distribution with a single peak that is not symmetrical. Sort the histograms into new categories based on whether they have a single peak or not.
 - Sample response: Histograms with one peak: A, B, D, E, G, H, J, L. Histograms C and I are fairly uniform without an obvious peak. Histogram F may have one peak or it may have 3 peaks. Histogram K seems to have multiple peaks.
- **3.** Some histograms have a gap, a space between two bars where there are no data points. Sort the histograms into new categories based on whether they have one or more gaps or have no gaps.
 - Sample response: With gaps: Histograms A, D, E, G, K and L.
- **4.** Some histograms have a few data points that are unusually far from the center like in Histogram A. Sort the histograms into groups that have this feature.

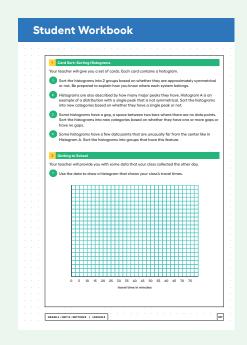
Sample response: With values far from the center: Histograms A, E, G and L. Some may consider including Histogram D, but because there are quite a few data points in the upper group, this really looks more like a distribution that just has a gap.

Access for Multilingual Learners (Activity 1, Student Task)

MLR8: Discussion Supports.

Think aloud and use gestures to emphasize the meaning of symmetry used for distributions. For example, show both hands side-by-side with palms outward, then bring one hand over to match the other with palms together.

Advances: Listening, Representing



Activity Synthesis

Students will have had a chance to discuss the different features of a distribution in small groups. Use the whole-class discussion to prompt students to think about what the features might mean, and whether or how they affect the way we characterize a distribution. Remind students that we have been using the center of a distribution to talk about what is typical in a group. Discuss some of these questions:

- "Look at the histograms that you think show symmetry. When a distribution is approximately symmetrical, where might its center be?"
 For a symmetrical or approximately symmetrical distribution the center is usually at the line of symmetry.
- "Is it easy to find the peaks on a distribution? Are there any histograms that your group debated about?"
 - Sometimes it is easy to find a peak because there is a much larger bar in the histogram compared to the others. We debated about some of them like Histogram F because we might consider there being 3 peaks or maybe just I.
- "Look at the histograms that show gaps. How might a gap (such as that in Histogram K) affect our description of what is typical in a group?"
 - When there are gaps, there might be 2 groups within the data that are different in an important way. It may be important to separate them to discuss what is typical.
- "Look at the histograms that have values that are far away from other values. Do unusual values (such as those in Histogram G) affect our description of center and spread? If those unusual values weren't there, would our description of center and spread change?"

They might pull the center closer to these values. The spread appears much greater with these unusual values because they indicate that there is a lot of variability.

It is okay if there is some disagreement about how the histograms are categorized. Depending on the situation, it may be important to consider things like peaks more or less precisely. For example, if the data in Histogram F are very precise like from a chemistry experiment, it may be important to investigate the 3 possible peaks more closely. If the data are less precise like how people are rating a TV show, then it may be okay to say there is only 1 peak or that the data are fairly spread out and there are no peaks.

Expect students' answers to be very informal. The goal of the discussion is to raise students' awareness that the shape and features of distributions may affect how we characterize the data. This experience provides a conceptual foundation that would help students make sense of measures of center (mean and median) and measures of spread (mean absolute deviation, interquartile range, and range) later.

Conclude the discussion by inviting 2–3 students to share a norm that they identified in action. Provide this sentence frame to help students organize their thoughts in a clear, precise way:

 "I noticed our norm ' 	 ' in action today.	and it really	helped	me/my
group because				

Activity 2

Getting to School



Activity Narrative

In this activity, students draw a bar graph and histogram, and then they describe the distributions shown on each display. Although the two visual displays may appear similar at first glance, there are important distinctions between the representations. Students notice differences in how we might characterize distributions in bar graphs and those in histograms, including how we describe typical values or categories. Along the way, students consolidate their understanding about categorical and numerical data.

Monitor for groups that draw a bar graph with the bars in a different order.

Launch

Students will need the data on their travel methods and times, collected at the beginning of the unit. Distribute or display the data collected for these questions from the survey given earlier in the unit. Alternatively, complete the tables in the blackline master ahead of time.

Arrange students in groups of 2. Give one copy of the blackline master to each group of students. Display the data from the prior survey or the completed frequency tables for all to see, or give a copy to each group of 2 students. Give students 5–6 minutes to complete the activity. Ask one partner to create a bar graph to represent the data on the class's travel methods and the other to create a histogram to represent the data on travel times, and then answer the questions together.

Student Task Statement

Your teacher will provide you with some data that your class collected the other day.

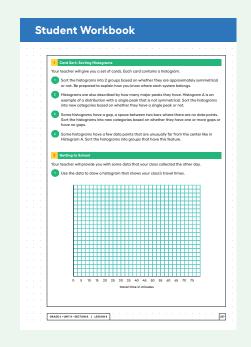
Answers and graphs vary based on class data.

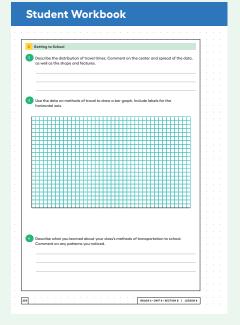
- 1. Use the data to draw a histogram that shows your class's travel times.

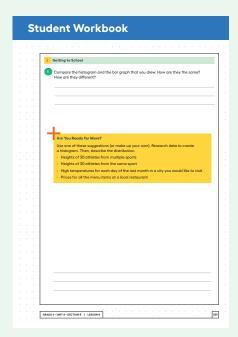
 The histogram should match class data.
- **2.** Describe the distribution of travel times. Comment on the center and spread of the data, as well as the shape and features.
 - Students should describe the distribution in terms of center, spread, shape, or other previously discussed features of distribution.
- **3.** Use the data on methods of travel to draw a bar graph. Include labels for the horizontal axis.

The bar graph should match class data.

- **4.** Describe what you learned about your class's methods of transportation to school. Comment on any patterns you noticed.
 - Students may identify categories that are most or least common. They may also describe the distribution in terms of a fraction or percentage of data values, or point out both common and unusual characteristics.







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

Action and Expression: Internalize Executive Functions.

To support organization, provide students with a graphic organizer such as a Venn diagram to compare bar graphs and histograms.

Supports accessibility for: Language, Organization

Access for Multilingual Learners (Activity 2, Synthesis)

MLR2: Collect and Display.

Collect the language that students use to compare histograms and bar graphs. Display words and phrases such as "numerical data," "categorical data," "order," and "bars" in a chart, such as a Venn diagram. During the synthesis, invite students to suggest ways to update the display:

"What are some other words or phrases we should include?" Invite students to borrow language from the display as needed.

Advances: Conversing, Reading

5. Compare the histogram and the bar graph that you drew. How are they the same? How are they different?

Sample responses:

- Bar graphs and histograms both use heights of bars to show frequency.

 The more frequently a value appears, the taller the bar that represents it.
- The graphs are different in that histograms show numerical data and they group data values into bins (so we cannot tell how many of a particular value there are). Bar graphs show categorical data, so the horizontal axis shows different categories or labels, rather than a number line. The frequency tells us exactly how many times a particular category appears in the data set.

Are You Ready for More?

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

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

Answers vary.

Activity Synthesis

The purpose of the discussion is for students to recognize the differences between histograms and bar graphs.

Invite previously selected groups to share an accurate histogram and bar graphs with bars in different orders. Then, solicit several observations about how the two graphical displays compare. Ask questions such as:

- "How are the bar graphs and histograms alike? How are they different?"
 They both use bars to represent the frequency of data. Bar graphs represent categorical data and histograms represent numerical data.
- Can we use a bar graph to display the data on travel times? Why or why not?"

No, because travel times are numerical. We could convert the times to categories by classifying different intervals as "short" or "long" and then make a bar graph from that, but this would require an extra step and a discussion about what times are considered in each category.

Can we use a histogram to display the data on methods of travel? Why or why not?"

No, histograms can be used only with numerical data.

Then, invite students to share their descriptions of the distributions shown on each type of display. Ask questions such as:

"How are your descriptions of the distribution for travel methods different from those for travel times?"

For travel methods, the description can only comment on individual bars in comparison to other ones. For travel times, we can talk about the shape, center, and spread of the distribution.

"Can you talk about the shape of a distribution shown on a bar graph?
Can you talk about its center and spread? Why or why not?"

No, because the bars can be rearranged in a bar graph. Which bar is in the center is not important, and the spread doesn't make sense as a numerical value. We could talk about variability by mentioning how many different bars there are and their relative heights, but it does not make sense to give that as a number.

Students should recognize that only the distribution of numerical data can be described in terms of shape, center, or spread. We cannot analyze these features for a distribution of categorical data on a bar graph because a bar graph does not use a number line. This means that the bars can be drawn anywhere, in any order, and with any kind of spacing, so shape, center, and spread would have no meaning.

Lesson Synthesis

In this lesson, students studied the shapes and features of distributions that are represented by histograms.

"What does it mean for a histogram to have symmetry?"

The left and right side of the distribution are almost the same.

"What is a 'peak' in a distribution? Is it always in the middle, or can it be to one side?"

A peak is a part of the distribution where the frequency is greater than the nearby values. Peaks do not need to be in the middle of the distribution.

"Can a distribution have more than one peak?"

Yes, the distribution could have multiple peaks. One example is when a distribution has a U-shape where there might be a peak on either end.

"What does it mean for a histogram to show a cluster (or more than one cluster)?"

It means that the data appear in smaller groups within the overall distribution.

"What does it mean for a histogram to show a gap?"

A gap is an interval with 0 frequency between intervals with some positive frequency. Often, gaps separate clusters.

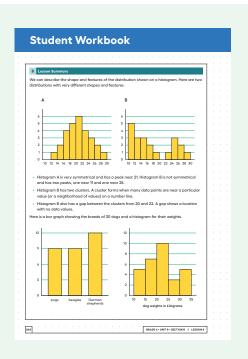
We also contrast bar graphs and histograms.

"When do we use a histogram and when do we use a bar graph?"

A histogram represents numerical data and a bar graph represents categorical data.

"What are the major differences between how a histogram is drawn and how a bar graph is drawn?"

A histogram is drawn on top of a portion of a number line and a bar graph is labeled with the categories. This means that the bars in a bar graph can be rearranged to any order, but histogram bars must be located over the appropriate interval.

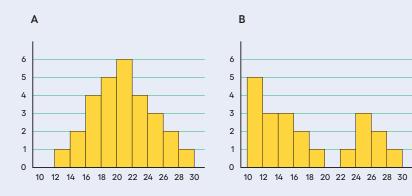


Student Workbook



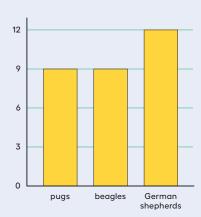
Lesson Summary

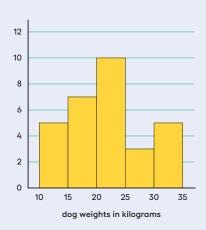
We can describe the shape and features of the distribution shown on a histogram. Here are two distributions with very different shapes and features.



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

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





Bar graphs and histograms may look alike, but they have different uses.

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

Cool-down

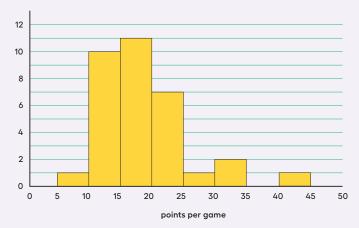
128

Point Spread

5 min

Student Task Statement

Here is a histogram that shows the number of points scored by a college basketball player during the 2008 season. Describe the shape and features of the data. Mention the center and spread as well as any symmetry, gaps, peaks, or other features that you notice.



Sample response: The distribution is not symmetrical because there is a peak on the left. The histogram shows a gap between 35 and 40, so there is no game in which the player scored 35, 36, 37, 38 or 39 points. There was one game that was unusually high scoring, between 40 and 44 points. The peak is between 15 and 20 points. The center is around 20 points, and the data is spread out pretty far above this center.

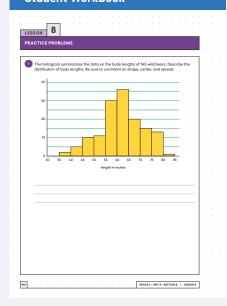
Responding To Student Thinking

Press Pause

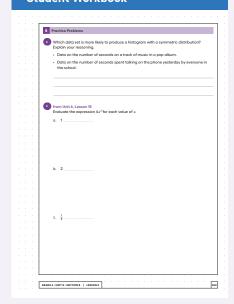
If students struggle with using words such as "cluster," "outlier," "symmetrical," "gap," or "center," use this cool down as the first draft for students to use the language routine *Stronger and Clearer Each Time*. End with several strong revised examples of how students describe the data display.

8

Practice Problems Student Workbook



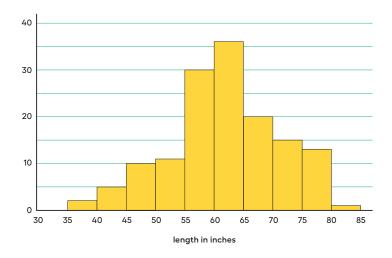
Student Workbook



Problem 1

The histogram summarizes the data on the body lengths of 143 wild bears. Describe the distribution of body lengths. Be sure to comment on shape, center, and spread.

5 Problems



Sample response: The distribution of body lengths is approximately symmetrical. A typical body length for the bears in the group studied is about 60 inches. There is a lot of variability in the body lengths of the bears, with the shortest length being somewhere between 35 and 40 inches and the longest length being somewhere between 80 and 85 inches.

Problem 2

Which data set is more likely to produce a histogram with a symmetric distribution? Explain your reasoning.

- Data on the number of seconds on a track of music in a pop album.
- Data on the number of seconds spent talking on the phone yesterday by everyone in the school.

Data on the number of seconds on a track of music in a pop album. Sample explanation: Most pop songs are around the same amount of time, but most people in the school will not talk much on the phone while a few people will talk a lot so there will be a peak near zero and a few very long times for some people.

Problem 3

from Unit 6, Lesson 15

Evaluate the expression $4x^3$ for each value of x.

a. 1

4

b. 2

32

c. $\frac{1}{2}$ $\frac{1}{2}$ (or equivalent)

Problem 4

Decide if each data set might produce one or more gaps when represented by a histogram. For each data set that you think might produce gaps, briefly describe or give an example of how the values in the data set might do so.

a. The ages of students in a sixth-grade class.

No

b. The ages of people in an elementary school.

Yes

Sample reasoning: The data set might show a lot of observations that are between 5 and 12 years old (the students' ages) and older than 20 years (the ages of staff and teachers), but no observations between 12 and 20.

c. The ages of people eating in a family restaurant.

Yes

Sample reasoning: The data set might show a lot of observations of children, then their parents. Observations between 12 and 25 are less likely.

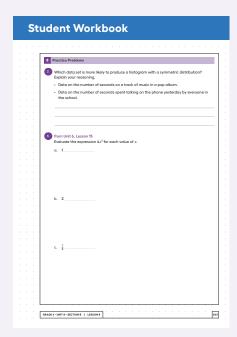
d. The ages of people who watch football.

No

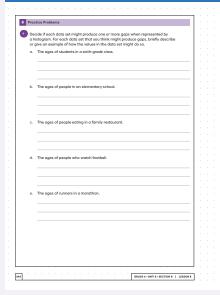
e. The ages of runners in a marathon.

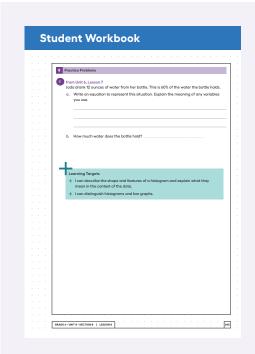
Yes

Sample reasoning: Most of the runners might be adults between 24 and 60 years of age, but there might be a few runners who are in their late teens or older runners in their 80's.



Student Workbook





Problem 5

from Unit 6, Lesson 7

Jada drank 12 ounces of water from her bottle. This is 60% of the water the bottle holds.

a. Write an equation to represent this situation. Explain the meaning of any variables you use.

Sample responses: $I2 = \frac{60}{100} b$ or I2 = 0.6b, where b is the number of ounces of water the bottle holds

b. How much water does the bottle hold?

$$b = 20$$

LESSON 8 • PRACTICE PROBLEMS