Lesson 6: Interpreting Histograms

Goals

- Compare and contrast (orally) dot plots and histograms in terms of how useful they are for answering different statistical questions.
- Create a histogram to represent a data set.
- Interpret a histogram to answer (in writing) statistical questions about a data set.

Learning Targets

- I can recognize when a histogram is an appropriate graphical display of a data set.
- I can use a histogram to get information about the distribution of data and explain what it means in a real-world situation.

Lesson Narrative

In this lesson students are introduced to **histograms**. They learn that, like a dot plot, a histogram can be used to show the distribution of a numerical data set, but unlike a dot plot, a histogram shows the frequencies of groups of values, rather than individual values. Students analyze the structures of dot plots and histograms displaying the same data sets and determine what information is easier to to understand from each type of display (MP7). Students read and interpret histograms in context (MP2) to prepare them to create a histogram.

Alignments

Addressing

- 6.SP.A.1: Recognize a statistical question as one that anticipates variability in the data related to the question and accounts for it in the answers. For example, “How old am I?” is not a statistical question, but “How old are the students in my school?” is a statistical question because one anticipates variability in students' ages.
- 6.SP.A.3: Recognize that a measure of center for a numerical data set summarizes all of its values with a single number, while a measure of variation describes how its values vary with a single number.
- 6.SP.B.4: Display numerical data in plots on a number line, including dot plots, histograms, and box plots.
- 6.SP.B.5.b: Describing the nature of the attribute under investigation, including how it was measured and its units of measurement.

Instructional Routines

- MLR2: Collect and Display
- MLR8: Discussion Supports

Required Materials

Straightedges
A rigid edge that can be used for drawing line segments. Sometimes a ruler is okay to use as a straightedge, but sometimes it is preferable to use an unruled straightedge, like a blank index card.

Student Learning Goals

Let's explore how histograms represent data sets.

6.1 Dog Show (Part 1)

Warm Up: 5 minutes
The purpose of this warm-up is to connect the analytical work students have done with dot plots in previous lessons with statistical questions. This activity reminds students that we gather, display, and analyze data in order to answer statistical questions. This work will be helpful as students contrast dot plots and histograms in subsequent activities.

Addressing

- 6.SP.A.1
- 6.SP.A.3

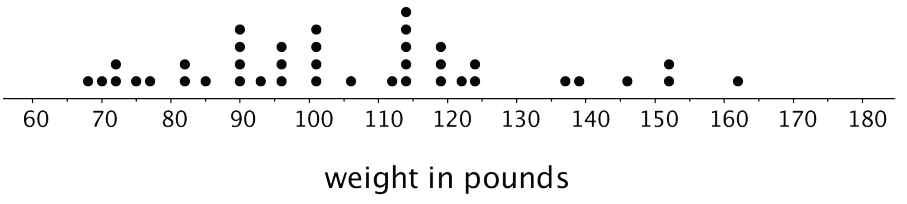
Launch

Arrange students in groups of 2. Give students 1 minute of quiet work time, followed by 2 minutes to share their responses with a partner. Ask students to decide, during partner discussion, if each question proposed by their partner is a statistical question that can be answered using the dot plot. Follow with a whole-class discussion.

If students have trouble getting started, consider giving a sample question that can be answered using the data on the dot plot (e.g., “How many dogs weigh more than 100 pounds?”)

Student Task Statement

Here is a dot plot showing the weights, in pounds, of 40 dogs at a dog show.



1. Write two statistical questions that can be answered using the dot plot.
2. What would you consider a typical weight for a dog at this dog show? Explain your reasoning.

Student Response

1. Answers vary. Sample questions:
    - How many dogs weigh exactly 70 pounds?
    - How many dogs weigh more 80 pounds but less than 150 pounds?
    - How much does the heaviest dog at the dog show weigh?
    - How many times as heavy as the lightest dog is the heaviest dog?
    - How alike or different are the weights of the dog at the show?
2. Answers vary. Sample responses:
    - About 114 pounds, because the largest percentage of the dots are at 114, and it seems to be about where the center of the data is.
    - About 100 pounds, because about half of the dogs are 100 pounds or lighter, and half are heavier than 100 pounds.

Activity Synthesis

Ask students to share questions they agreed were statistical questions that could be answered using the dot plot. Record and display their responses for all to see. If there is time, consider asking students how they would find the answer to some of the statistical questions.

Ask students to share a typical weight for a dog at this dog show and why they think it is typical. Record and display their responses for all to see. After each student shares, ask the class if they agree or disagree.

6.2 Dog Show (Part 2)

10 minutes

This activity introduces students to **histograms**. By now, students have developed a good sense of dot plots as a tool for representing distributions. They use this understanding to make sense of a different form of data representation. The data set shown on the first histogram is the same one from the preceding warm-up, so students are familiar with its distribution. This allows them to focus on making sense of the features of the new representation and comparing them to the corresponding dot plot.

Note that in all histograms in this unit, the left-end boundary of each bin or bar is included and the right-end boundary is excluded. For example, the number 5 would not be included in the 0–5 bin, but would be included in the 5–10 bin.

Addressing

- 6.SP.B.4
- 6.SP.B.5.b

Instructional Routines

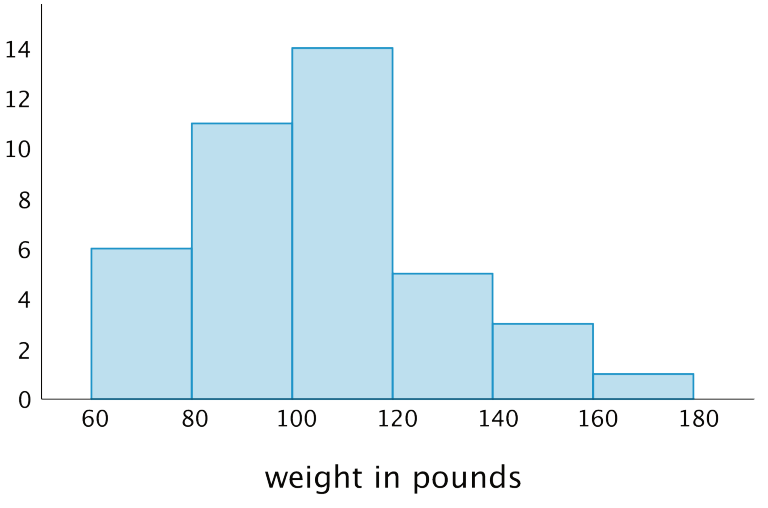
- MLR2: Collect and Display

Launch

Explain to students that they will now explore histograms, another way to represent numerical data. Give students 3–4 minutes of quiet work time, and then 2–3 minutes to share their responses with a partner. Follow with a whole-class discussion.

Student Task Statement

Here is a **histogram** that shows some dog weights in pounds.



Each bar includes the left-end value but not the right-end value. For example, the first bar includes dogs that weigh 60 pounds and 68 pounds but not 80 pounds.

1. Use the histogram to answer the following questions.
    a. How many dogs weigh at least 100 pounds?
    b. How many dogs weigh exactly 70 pounds?
    c. How many dogs weigh at least 120 and less than 160 pounds?
    d. How much does the heaviest dog at the show weigh?
    e. What would you consider a typical weight for a dog at this dog show? Explain your reasoning.
2. Discuss with a partner:
    - If you used the dot plot to answer the same five questions you just answered, how would your answers be different?
    - How are the histogram and the dot plot alike? How are they different?

Student Response

1. a. 23 dogs.

   b. Unknown. It could be anywhere between 0 and 6.

   c. 8 dogs.

   d. The exact weight cannot be determined, but it weighs at least 160 pounds but less than 180 pounds.

   e. Answers vary. Sample response: Around 100 pounds. The largest percentage (35%) of the weights fall in the third bar (at least 100 pounds and less than 120 pounds), and it is approximately the middle of the data.

2. Answers vary. Sample responses:
   - They are alike in that they are both built on number lines, show the same total number of data values, and show how the values are spread out. They are different in that the dot plot shows individual data points and the histogram groups the data points together.
   - With the dot plot we can see the values of individual points and tell how many there are. With the histogram, we can't tell how many data points have a specific values; we only know how many points fall into a specific range.

Activity Synthesis

Ask a few students to briefly share their responses to the first set of questions to make sure students are able to read and interpret the graph correctly.

Focus the whole-class discussion on the last question. Select a few students or groups to share their observations about whether or how their answers to the statistical questions would change if they were to use a dot plot to answer them, and about how histograms and dot plots compare. If not already mentioned by students, highlight that, in a histogram:

- Data values are grouped into "bins" and represented as vertical bars.
- The height of a bar reflects the combined frequency of the values in that bin.
- A histogram uses a number line.

At this point students do not yet need to see the merits or limits of each graphical display; this work will be done in upcoming lessons. Students should recognize, however, how the structures of the two displays are different (MP7) and start to see that the structural differences affect the insights we are able to glean from the displays.

Access for Students with Disabilities

Representation: Develop Language and Symbols. Create a display of important terms and vocabulary. Include the following terms and maintain the display for reference throughout the unit: histogram.
Supports accessibility for: Memory; Language

Access for English Language Learners

Representing, Listening, Speaking: MLR2 Collect and Display. Record and display the language students use as they discuss how dot plots and histograms are alike and different. Highlight language related to individual specific data values and groups of data values, center, and spread. Keep the display visible for reference in upcoming lessons. This will help students connect unique and shared characteristics of histograms and dot plots.
Design Principle(s): Support sense-making; Maximize meta-awareness

6.3 Population of States

20 minutes
In this activity, students continue to develop their understanding of histograms. They begin to notice that a dot plot may not be best for representing a data set with a lot of variability (or where few values are repeated) or when a data set has a large number of values. Histograms may help us visualize a distribution more clearly in these situations. Students organize a data set into "bins" and draw a histogram to display the distribution.

As students work and discuss, listen for explanations for why certain questions might be easy, hard, or impossible to answer using each graphical display.

Addressing

- 6.SP.B.4

Instructional Routines

- MLR8: Discussion Supports

Launch

Give students a brief overview of census and population data, as some students may not be familiar with them. Refer to the dot plot of the population data and discuss questions such as:

- "How many total dots are there?" (51)

- "What's the population of the state with the largest population? Do you know what state that is?" (Between 37 and 38 million. It's California.)
- "Look at the leftmost dot. What state might it represent? Approximately what is its population?" (The leftmost dot represents Wyoming, with a population of around half a million.)
- "Do you know the approximate population of our state? Where do you think we are in the dot plot?"

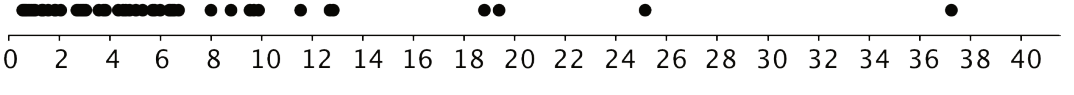
Explain to students that they will now draw a histogram to represent the population data. Remind them that histograms organize data values into "bins" or groups. In this case, the bins sizes are already decided for them. Then, arrange students in groups of 3–4. Provide access to straightedges. Give students 10–12 minutes to complete the activity. Encourage them to discuss their work within their group as needed.

Access for Students with Disabilities

Action and Expression: Internalize Executive Functions. To support development of organizational skills, check in with students within the first 2-3 minutes of work time. Check to make sure students have developed a way to keep track of counting the 2010 census data.
Supports accessibility for: Organization; Attention

Student Task Statement

Every ten years, the United States conducts a census, which is an effort to count the entire population. The dot plot shows the population data from the 2010 census for each of the fifty states and the District of Columbia (DC).



population of states in millions

1. Here are some statistical questions about the population of the fifty states and DC. How difficult would it be to answer the questions using the *dot plot*?

   In the middle column, rate each question with an E (easy to answer), H (hard to answer), or I (impossible to answer). Be prepared to explain your reasoning.

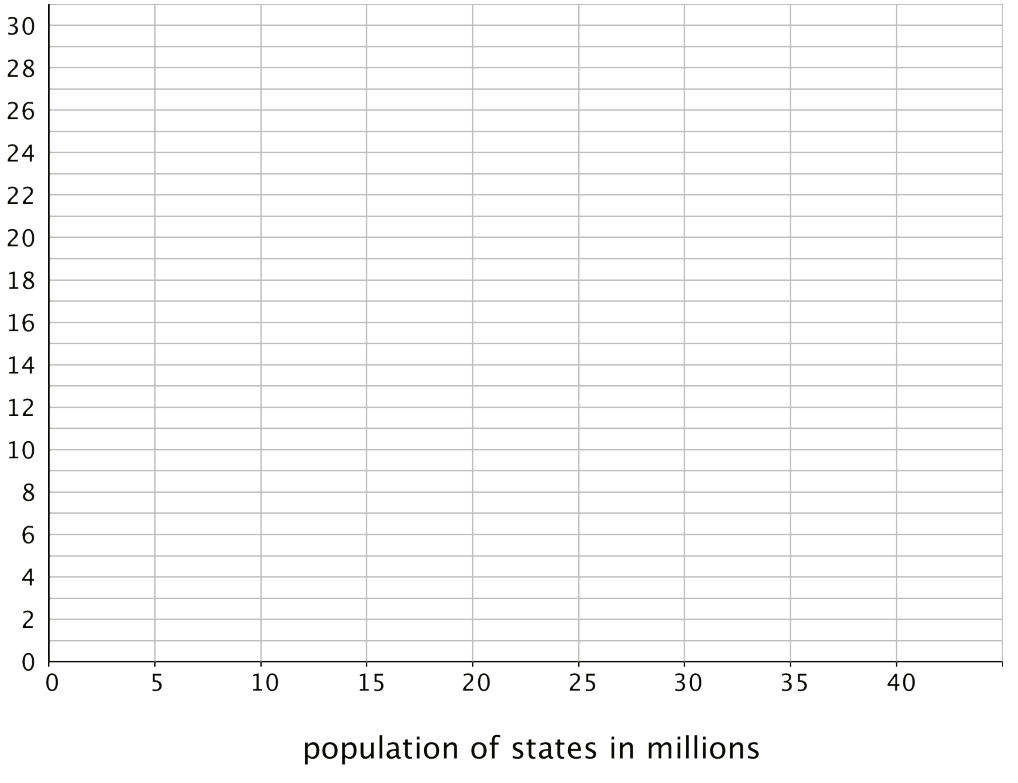
statistical question	using the dot plot	using the histogram
a. How many states have populations greater than 15 million?		
b. Which states have populations greater than 15 million?		
c. How many states have populations less than 5 million?		
d. What is a typical state population?		
e. Are there more states with fewer than 5 million people or more states with between 5 and 10 million people?		
f. How would you describe the distribution of state populations?		

2. Here are the population data for all states and the District of Columbia from the 2010 census. Use the information to complete the table.

Alabama	4.78	Illinois	12.83	Montana	0.99	Rhode Island	1.05
Alaska	0.71	Indiana	6.48	Nebraska	1.83	South Carolina	4.63
Arizona	6.39	Iowa	3.05	Nevada	2.70	South Dakota	0.81
Arkansas	2.92	Kansas	2.85	New Hampshire	1.32	Tennessee	6.35
California	37.25	Kentucky	4.34	New Jersey	8.79	Texas	25.15
Colorado	5.03	Louisiana	4.53	New Mexico	2.06	Utah	2.76
Connecticut	3.57	Maine	1.33	New York	19.38	Vermont	0.63
Delaware	0.90	Maryland	5.77	North Carolina	9.54	Virginia	8.00
District of Columbia	0.60	Massachusetts	6.55	North Dakota	0.67	Washington	6.72
Florida	18.80	Michigan	9.88	Ohio	11.54	West Virginia	1.85
Georgia	9.69	Minnesota	5.30	Oklahoma	3.75	Wisconsin	5.69
Hawaii	1.36	Mississippi	2.97	Oregon	3.83	Wyoming	0.56
Idaho	1.57	Missouri	5.99	Pennsylvania	12.70		

population (millions)	frequency
0–5	
5–10	
10–15	
15–20	
20–25	
25–30	
30–35	
35–40	

3. Use the grid and the information in your table to create a histogram.



4. Return to the statistical questions at the beginning of the activity. Which ones are now easier to answer?

   In the last column of the table, rate each question with an E (easy), H (hard), and I (impossible) based on how difficult it is to answer them. Be prepared to explain your reasoning.

Student Response

1. a. Easy (E). Unless some dots are lying directly on top of one another, there are four states with a population greater than 15 million.

   b. Impossible (I). Since the dots are not labeled, it is impossible to tell which states have a population greater than 15 million.

   c. Impossible (I). Since the dots are so close together below 5 million, it's impossible to count how many there are.

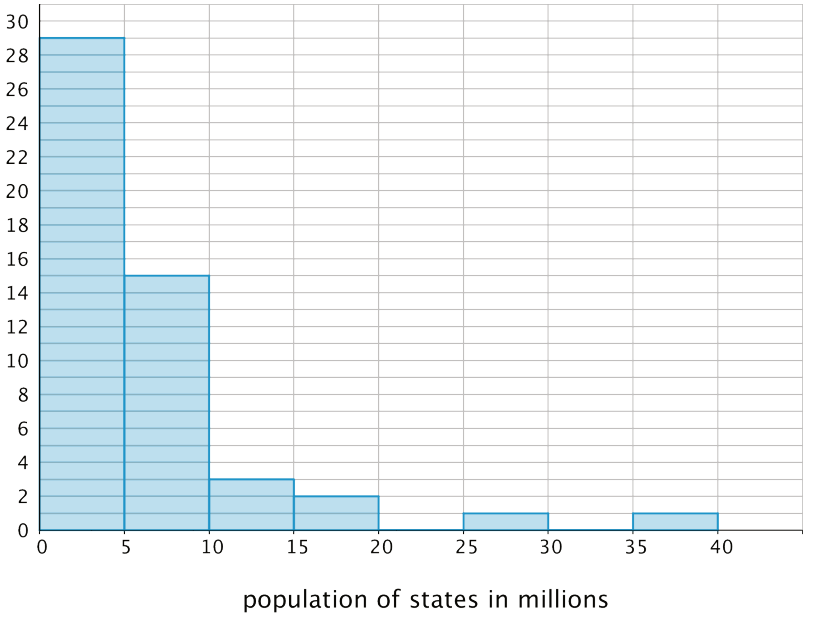
   d. Hard (H). Since so many dots are indistinguishable, it's hard to determine a typical state population.

   e. Hard (H). It appears that there are more dots for populations that are less than 5 million than for those between 5 and 10 million, but we can't be sure because dots might be right on top of each other.

   f. Hard (H). Since the dots overlap a lot, it is difficult to give a good estimate for the center and spread.

2.

population in millions	frequency
0–5	29
5–10	15
10–15	3
15–20	2
20–25	0
25–30	1
30–35	0
35–40	1

3. 

4. Revisiting the questions from the first problem:
   a. Still easy (E).
   b. Still impossible (I), based on the histogram alone.
   c. Now it is easy (E) to tell how many states had a population below 5 million. It was previously impossible (I).
   d. From the histogram, we can estimate that a typical state population has fewer than 10 million, but it is hard (H) to be more precise than that at this point. It was previously hard (H).
   e. Using the histogram it is easy (E) to tell how many states have fewer than 5 million people and how many have between 5 and 10 million (there are more states in the smaller population category). It was previously hard (H).
   f. It is easier (E) to describe the data distribution more precisely because the histogram shows the population sizes in intervals of 5 million people.

Are You Ready for More?

Think of two more statistical questions that can be answered using the data about populations of states. Then, decide whether each question can be answered using the dot plot, the histogram, or both.

Student Response

Answers vary.

Activity Synthesis

Much of the discussion about how to construct histograms should have happened in small groups. Address unresolved questions about drawing histograms if they are relatively simple. Otherwise, consider waiting until students have more opportunities to draw histograms in upcoming lessons.

Focus the discussion on comparing the effectiveness of dot plots and histograms to help us answer statistical questions.

Select a few students or groups to share how their ratings of "easy," "hard," and "impossible," changed when they transitioned from using dot plots to using histograms to answer statistical questions about populations of states. Then, discuss and compare the two displays more generally. Solicit as many ideas and observations as possible to these questions:

- "What are some benefits of histograms?"
- "When might histograms be preferable to dot plots?"
- "When might dot plots be preferable to histograms?"

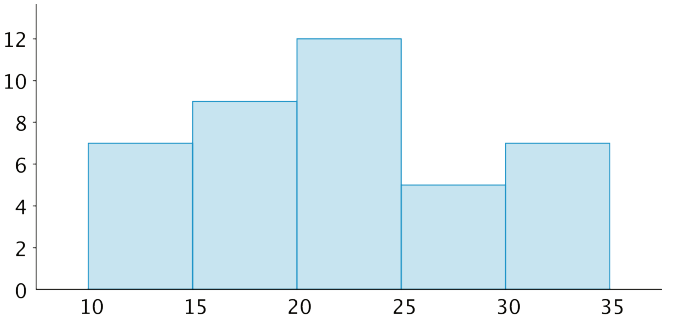
Students should walk away from the activity recognizing that in cases when a data set has many numerical values, especially if the values do not repeat, a histogram can give us a better visualization of the distribution. In such a case, creating a dot plot can be very difficult to do including finding a scale that can meaningfully display the data while a histogram will be easier to create and display the information in a way that is easier to understand at a glance.

Access for English Language Learners

Representing, Writing, Conversing: MLR8 Discussion Supports. Give students sentence frames during the discussion, such as: "Histograms are easy (or hard or impossible) to use when _____, because . . .". This will help students make decisions about the type of graph to use to display different types of data sets.
Design Principle(s): Optimize output (for generalization)

Lesson Synthesis

In this lesson, we learn about a different way to represent the distribution of numerical data—using a **histogram**. This histogram, for instance, represents the distribution for the weights of some dogs.



- "What could the smallest dog weigh? The largest?" (10 kilograms up to almost 40 kilograms)
- "What does the bar between 25 and 30 tell you?" (5 dogs weigh between 25 and just under 30 kilograms)
- "What can you say about the dogs who weigh between 10 and 20 kg?" (There are 16 total dogs in this range including 7 dogs between 10 and 15 kg and 9 between 15 and 20 kg)
- "In general, what information does a histogram allow us to see? How is it different from a dot plot?" (A bigger picture of the distribution is shown in the histogram, but some of the detail is lost when compared to a dot plot. For example, this histogram does not show the weight of any individual dogs.)
- "When might it be more useful to use a histogram than a dot plot?" (When the data is very spread out, when there are not very many data points with the same value, or when an overall idea of the distribution is more important than a detailed view.)

6.4 Rain in Miami

Cool Down: 5 minutes

Addressing

- 6.SP.B.4

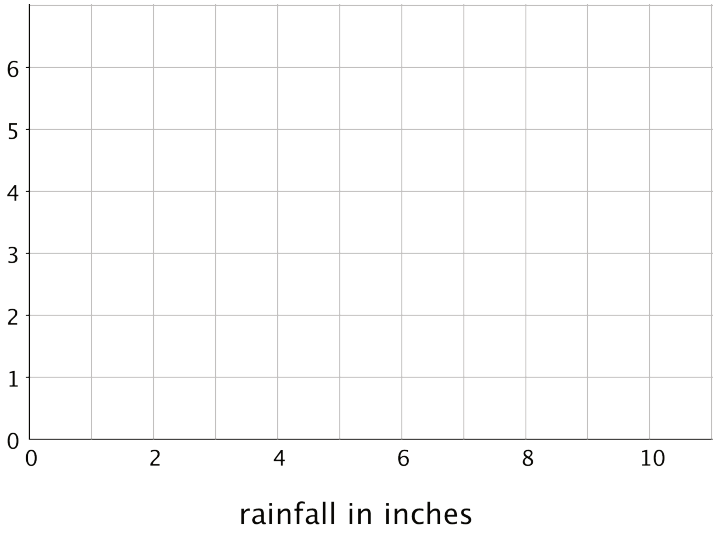
Student Task Statement

Here is the average amount of rainfall, in inches, for each month in Miami, Florida.

month	rainfall (inches)	month	rainfall (inches)
January	1.61	July	6.5
February	2.24	August	8.9
March	2.99	September	9.84
April	3.14	October	6.34
May	5.35	November	3.27
June	9.69	December	2.05

1. Complete the frequency table and use it to make a histogram.

rainfall (inches)	frequency
0–2	
2–4	
4–6	
6–8	
8–10	

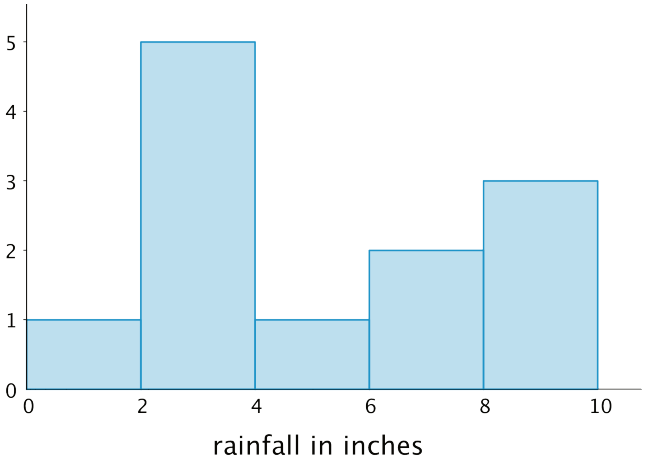


2. What is a typical amount of rainfall in one month in Miami?

Student Response

1.

rainfall (inches)	frequency
0–2	1
2–4	5
4–6	1
6–8	2
8–10	3

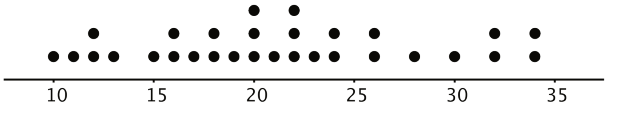


2. Answers vary. Sample response: It is difficult to say what typical amount of rainfall is. A typical month could be between 2 and 4 inches, this happens in 5 months. But for another 5 months, the rainfall is more than 6 inches, between 6 and 10 inches.

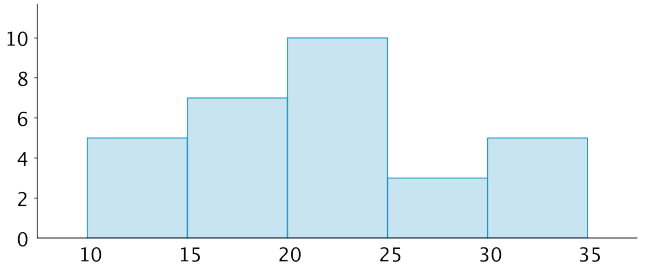
Student Lesson Summary

In addition to using dot plots, we can also represent distributions of numerical data using **histograms**.

Here is a dot plot that shows the weights, in kilograms, of 30 dogs, followed by a histogram that shows the same distribution.



dog weights in kilograms



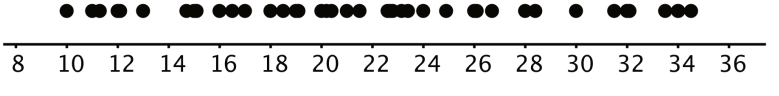
dog weights in kilograms

In a histogram, data values are placed in groups or "bins" of a certain size, and each group is represented with a bar. The height of the bar tells us the frequency for that group.

For example, the height of the tallest bar is 10, and the bar represents weights from 20 to less than 25 kilograms, so there are 10 dogs whose weights fall in that group. Similarly, there are 3 dogs that weigh anywhere from 25 to less than 30 kilograms.

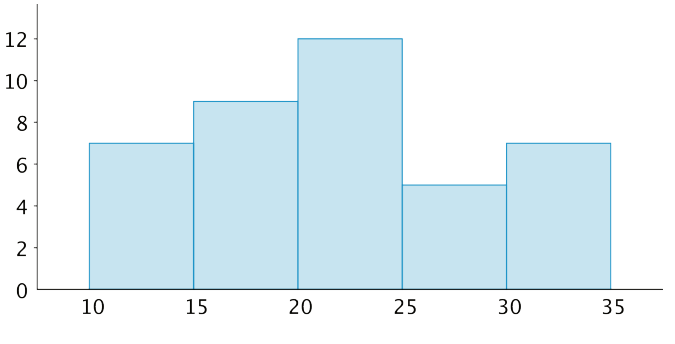
Notice that the histogram and the dot plot have a similar shape. The dot plot has the advantage of showing all of the data values, but the histogram is easier to draw and to interpret when there are a lot of values or when the values are all different.

Here is a dot plot showing the weight distribution of 40 dogs. The weights were measured to the nearest 0.1 kilogram instead of the nearest kilogram.



dog weights in kilograms

Here is a histogram showing the same distribution.



dog weights in kilograms

In this case, it is difficult to make sense of the distribution from the dot plot because the dots are so close together and all in one line. The histogram of the same data set does a much better job showing the distribution of weights, even though we can't see the individual data values.

Glossary

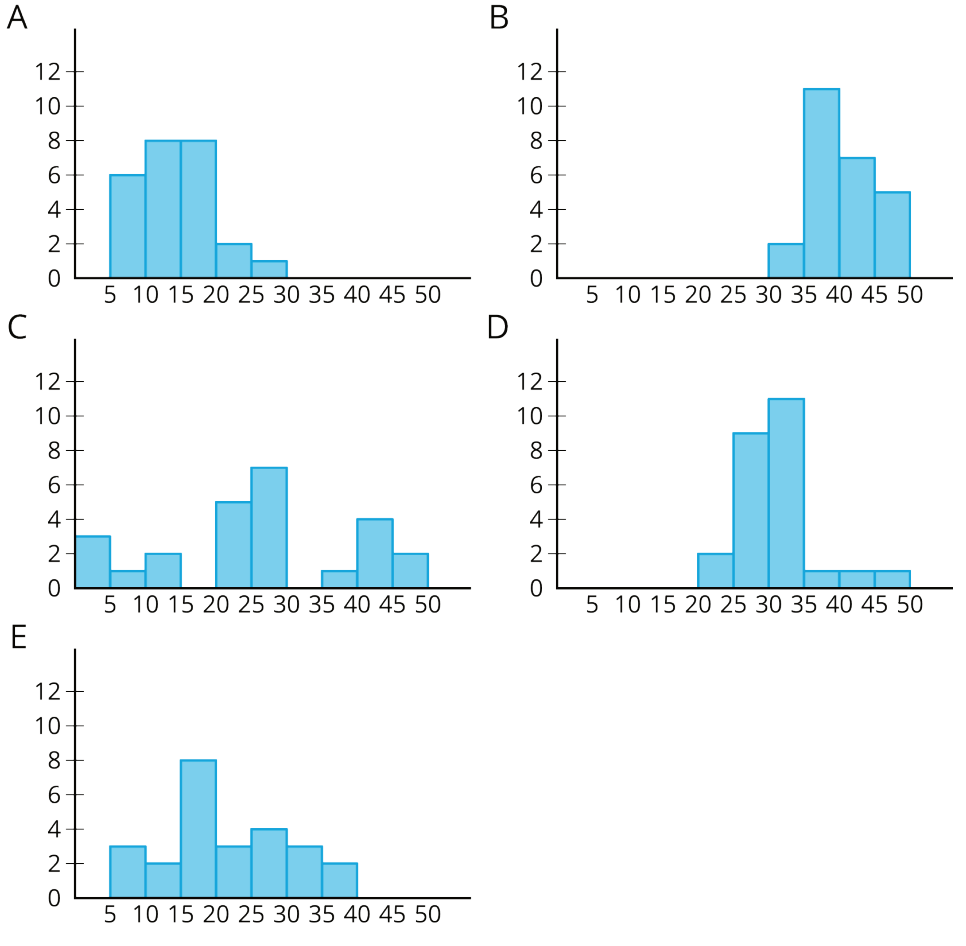
- histogram

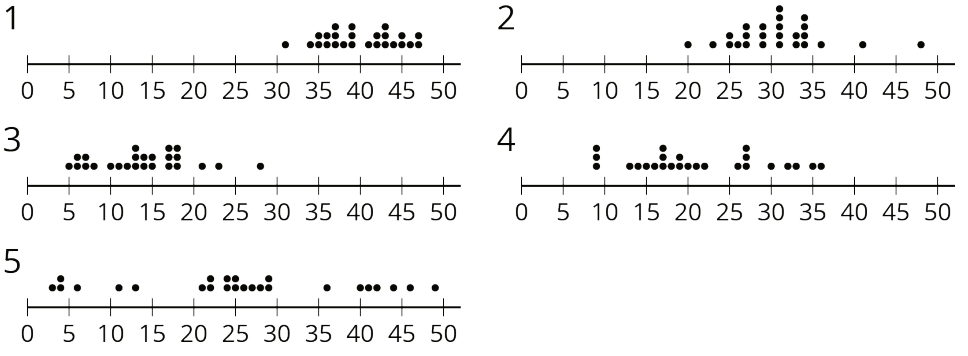
Lesson 6 Practice Problems

Problem 1

Statement

Match histograms A through E to dot plots 1 through 5 so that each match represents the same data set.





Solution

1. B

2. D

3. A

4. E

5. C

Problem 2

Statement

$(-2, 3)$ is one vertex of a square on a coordinate plane. Name three points that could be the other vertices.

Solution

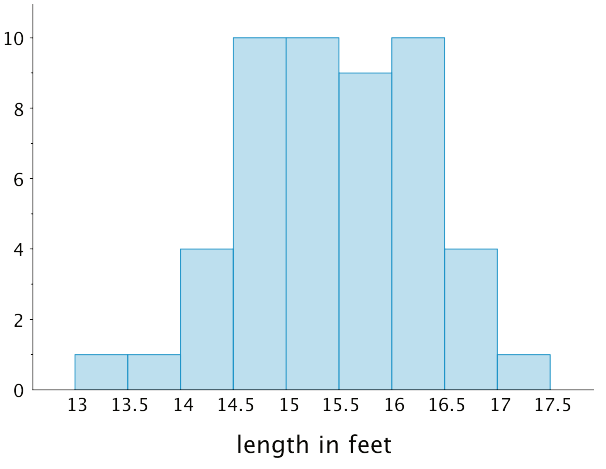
Answers vary. Sample response: $(2, 3)$, $(2, -1)$, $(-2, -1)$

(From Unit 7, Lesson 12.)

Problem 3

Statement

Here is a histogram that summarizes the lengths, in feet, of a group of adult female sharks. Select **all** the statements that are true, according to the histogram.



A. A total of 9 sharks were measured.

B. A total of 50 sharks were measured.

C. The longest shark that was measured was 10 feet long.

D. Most of the sharks that were measured were over 16 feet long.

E. Two of the sharks that were measured were less than 14 feet long.

Solution

["B", "E"]

Problem 4

Statement

This table shows the times, in minutes, it took 40 sixth-grade students to run 1 mile.

time (minutes)	frequency
4 to less than 6	1
6 to less than 8	5
8 to less than 10	13
10 to less than 12	12
12 to less than 14	7
14 to less than 16	2

Draw a histogram for the information in the table.

Solution