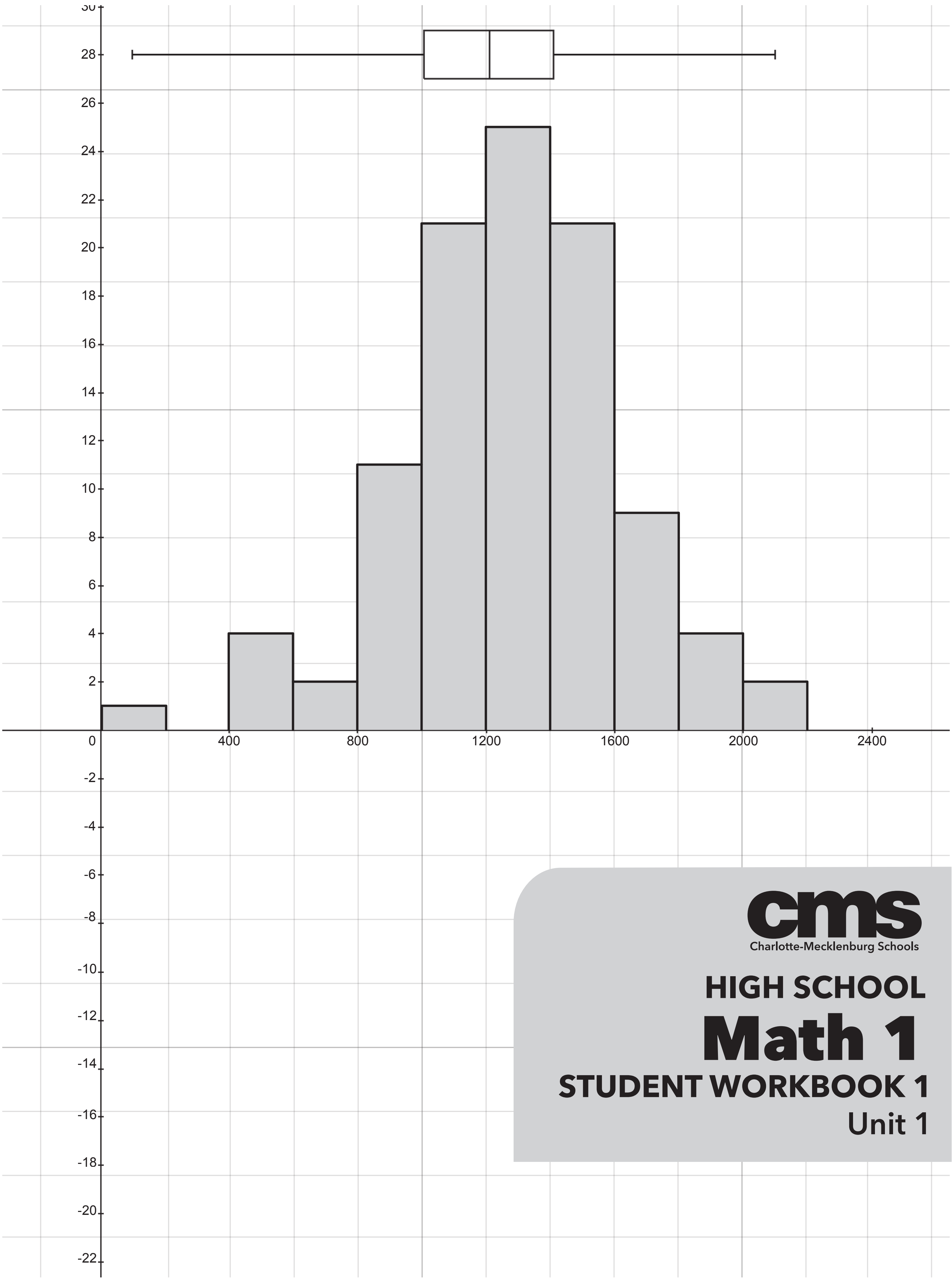


Charlotte-Mecklenburg Schools

HIGH SCHOOL
Math 1
STUDENT WORKBOOK 1
Unit 1

Table of Contents

Lesson 1: Welcome to Math!

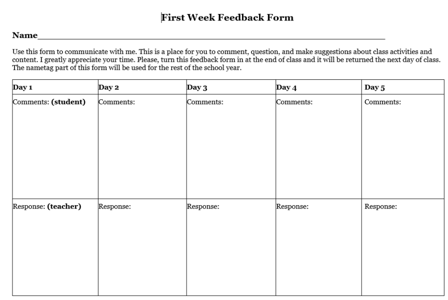
Learning Targets

- I can articulate the big ideas we will explore in this course.
- I understand the reasoning for and will strive to meet the expectations communicated by my teacher.
- I know my classmates and can recognize the value I will add to this classroom community.

Warm-up: Name Tent

Welcome to Math 1!

1. As you enter the classroom, locate your assigned seat.
2. On your desk, you will find a printout titled “First Week Feedback Form.” Please write your name on the Name line and then fold the paper in half so that the printed side is inside the fold to make a tent.
3. On the outside of the tent, please write your name on both sides (large enough for everyone to be able to read as needed).
4. Introduce yourself to the classmates sitting closest to you.
5. If you have time, you may decorate the outside of your name tent.



First Week Feedback Form

Name______________________________

Use this form to communicate with me. This is a place for you to comment, question, and make suggestions about class activities and content. I greatly appreciate your time. Please, turn this feedback form in at the end of class and it will be returned the next day of class. The nametag part of this form will be used for the rest of the school year.

Day 1	**Day 2**	**Day 3**	**Day 4**	**Day 5**
Comments: **(student)**	Comments:	Comments:	Comments:	Comments:
Response: **(teacher)**	Response:	Response:	Response:	Response:

Activity 1: What Is Math 1?

What am I going to learn and do in Math 1?

1. Briefly skim the Math 1 course overview provided by your teacher.
2. Use the graphic organizer below to record at least one thing you notice and one thing you wonder about Math 1. You can certainly list more than one!

I Noticed...	I Wondered...

3. Be ready to share what you noticed and wondered with your class.

Activity 2: What Does It Mean To Be a Mathematician in This Class?

In your small group, reflect on the video, “Math Is for Everybody” by Francis Su, by discussing the following questions with your group:

- Do you believe that some people are “math people” and others are not?

- What makes someone a mathematician?

- In what ways can someone work to get better at a sport, playing a musical instrument, or some other hobby?

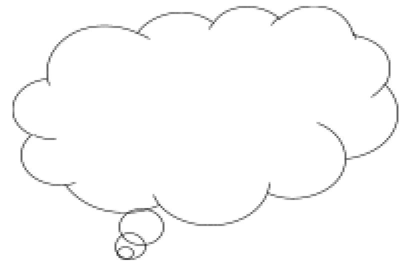
- How can a student who believes math is for everybody approach this course?

Activity 3: What Does My Teacher Expect From Me?

Briefly review the course syllabus. Feel free to highlight and annotate your copy.

Lesson Debrief

What are you most excited about in this class? Write one or two sentences OR draw a picture representing your response.



Lesson 2: Building a Math Community

Learning Targets

- I know my classmates and can recognize the value I will add to this classroom community.
- I understand that making mistakes and engaging in a productive struggle are necessary steps in growing as a mathematician.

Warm-up: Check Your Readiness (Part One)

Check Your Readiness questions allow your teacher to learn about the prior knowledge students have related to topics in an upcoming unit. In other words, they help your teacher determine whether or not an upcoming lesson will create a productive struggle for all learners and if not, help your teacher find ways to support each student.

It is really important that your responses to these questions accurately represent what you know. **Answer what you can to the best of your ability. If you are stuck on a question, it is okay to name what you do know or understand and what questions you have.**

Activity 1: Math Empathy Interview Protocol

Goals: To gain a deeper understanding of another student's experience in math

Norms for the interview:

- Interviewee:
  - Share only what you feel comfortable sharing
- Interviewer:
  - Seek to understand, not confirm (for example, do not agree or disagree with statements given)
  - Ask once, clearly
  - Probe: tell me more… what was that like for you

During an empathy interview, the interviewer asks the questions and the interviewee answers them. The interviewer can only ask probing questions, not share their own experiences.

Instead of writing down responses to each of the questions below, listen intently for big ideas so that you can capture three things that stood out to you.

After 6 minutes, the interviewer and interviewee switch roles.

Empathy Interview Questions:

1. Tell me about a time where you felt successful in math.

    - Why do you think you were successful?

    - What did you do?

    - What did others do (your classmates, teacher, friends)?

2. Tell me about a time where math was hard.

    - How did it feel?

    - What did you do?

    - Why was it hard?

- What do you wish would have happened?

- What would have helped?

3. What comes to mind when I say, “There is no such thing as a math person”?

4. Describe or draw a picture on a whiteboard or scrap piece of paper of what happens in your brain when you see a math problem you don’t understand.

5. What advice would you give another student who thinks they aren’t good at math?

6. Draw a picture on a whiteboard or scrap piece of paper of what you think about when you hear “growth mindset in math.”

The mathematician I interviewed was:	Three things that stood out to me from the interview were:
	1. 2. 3.

Lesson Debrief

Lesson 2 Summary and Glossary

All students make mistakes and experience struggle at some point in learning mathematics. There are different ways to respond to that struggle, and some ways are more productive than others.

Making mistakes and engaging in productive struggle are valuable parts of learning mathematics, and both provide a strong foundation to build a safe classroom community.

Lesson 3: Great Group Work

Learning Targets

- I know my classmates and can recognize the value I will add to this classroom community.
- I can articulate what great group work looks like.

Warm-up: Check Your Readiness (Part Two)

Check Your Readiness questions allow your teacher to learn about the prior knowledge students have related to topics in an upcoming unit. In other words, they help your teacher determine whether or not an upcoming lesson will create a productive struggle for all learners and if not, help your teacher find ways to support each student.

It is really important that your responses to these questions accurately represent what you know. **Answer what you can to the best of your ability. If you are stuck on a question, it is okay to name what you do know or understand and what questions you have.**

Activity 1: 100 Numbers to Get Students Talking[1]

Preparation:

- Arrange desks to accommodate a group of four students.
- Clear tables of everything.
- Each person in the group needs a different colored highlighter.
- Each group needs one copy of the 1–100 number sheet and must keep it UPSIDE DOWN until the teacher says, “Go.”

Directions:

Your group will have 3 minutes to find as many numbers from 1–100 in order. Group members will take turns going in a circle highlighting each number as they find it.

- Student 1 highlights 1.
- Student 2 highlights 2.
- Student 3 highlights 3.
- Student 4 highlights 4.
- When it is student 1’s turn again, highlight the next number.
- Keep the cycle going, finding as many numbers as possible in the provided time.

All group members can assist in finding the numbers, but each student has to highlight the number assigned.

[1] Adapted from Sara VanDerWerf

Lesson Debrief

Great group work in math looks like …

Lesson 4: Getting to Know You

Learning Targets

- I can tell statistical questions from non-statistical questions and can explain the difference.
- I can tell the difference between numerical and categorical data.

Bridge

Here are some survey questions. After you answer the survey questions, record answers from one other classmate.

1. How many hours of sleep do you typically get on a school night? Answer to the nearest hour. ________
2. How do you travel to school on most days? Choose one.
   - Walk
   - Bike
   - Scooter or skateboard
   - Car
   - School bus
   - Public transport
   - Other
3. What is the difference between the answers to question 1 and question 2?

____

Warm-up: Types of Data

Which one doesn't belong? Explain your reasoning.

a. How many potato chips are in this bag of chips?	b. What is the typical number of chips in a bag of chips?
c. What type of chips are these?	d. What type of chips do students in this class prefer?

Activity 1: Representing Data About You and Your Classmates

Your teacher will assign your group a set of three questions.

1. As a group, write another question (#4) that will require data collected from the class to answer.

2. For each of the four questions, write a survey question that will help your group collect data from the class that can be analyzed to answer the questions.
   - Survey question for #1:
   - Survey question for #2:
   - Survey question for #3:
   - Survey question for #4:

3. Assign one survey question to each member of your group. Ask your one survey question to 15 classmates. Record classmates' responses below.

Classmate's name		Survey question response
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		

4. After collecting the data, return to your group and together summarize the data for each question in a sentence or two.

    - Summary of data for survey question #1:

    - Summary of data for survey question #2:

    - Summary of data for survey question #3:

    - Summary of data for survey question #4:

5. With your group, decide what the responses for survey question number 1 have in common. Then do the same for survey questions number 2 and number 3.

6. Does the question you wrote (#4) fit best with the questions from numbers 1, 2, or 3? Explain your reasoning.

Are You Ready For More?

1. Find a news article that uses numerical data to discuss a statistical question.

2. Find a news article that uses categorical data to discuss a statistical question.

Lesson Debrief

Lesson 4 Summary and Glossary

Statistics is about using data to solve problems or make decisions. The question that you ask determines the type of data that you collect and whether or not there is variability in the data collected.
There are two types of data:

Numerical data: Data where the values are numbers, measurements, or quantities; also called measurement or quantitative data. For example, the weights of 10 different dogs are numerical data.

Categorical data: Data where the values are categories. For example, the breeds of 10 different dogs are categorical data.

In earlier grades, you learned that there is variability in a data set if not all of the values in the data set are the same. In this lesson, you learned about two specific types of questions.

Statistical question: A question that can only be answered by collecting data and there is an expectation for the data to have variability.

Here are some examples of statistical questions:

- "What is the average class size at this school?" would produce numerical data with some variability.
- "What are the favorite colors of students in this class?" would produce categorical data with some variability.

Non-statistical question: A question that can be answered by a specific measurement or procedure where no variability is anticipated.

Here are some examples of non-statistical questions:

- "How many students are on the roster for this class?" would produce numerical data that do not vary. There is only one value in the data set, so there is no variability.
- "What shape is this table?" would produce categorical data that do not vary. There is only one value in the data set, so there is no variability.

Unit 1 Lesson 4 Practice Problems

1. Write a survey question for which you would expect to collect numerical data.

____

2. Write a survey question for which you would expect to collect categorical data.

____

3. Select **all** the statistical questions.

   a. What is the typical amount of rainfall for the month of June in the Galapagos Islands?
   b. How much did it rain yesterday at the Mexico City International Airport?
   c. Why do you like to listen to music?
   d. How many songs do students in the class usually listen to each day?
   e. How many songs did you listen to today?
   f. What is the capital of Canada?
   g. How long does it typically take for 2nd graders to walk a lap around the track?

4. Read each question. Think about the data you might collect to answer it and whether you expect to see variability in the data. Complete each blank with “Yes” or “No.” If you decide a question is statistical, determine if its answers will give numerical data or categorical data. If you decide a question is not statistical, then skip the question about the type of data.[1]

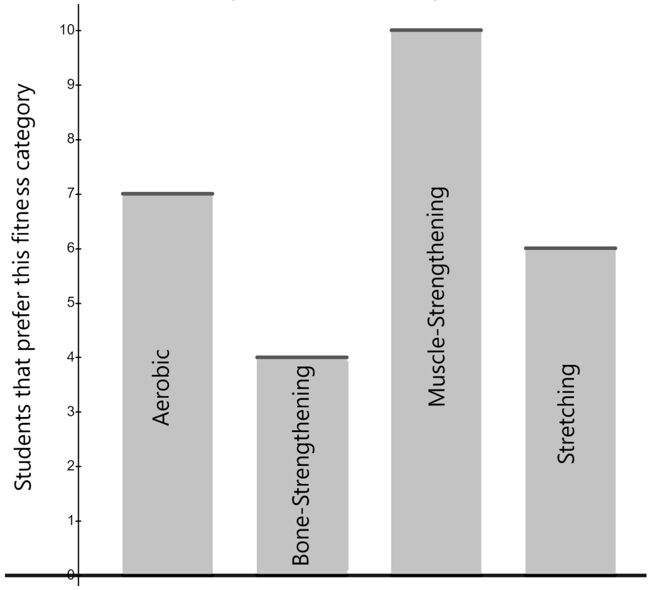
   a. How many cups of water do my classmates drink each day?
      - Is variability expected in the data? ________
      - Is the question statistical? ________
      - Will the answers give numerical data or categorical data? __________________

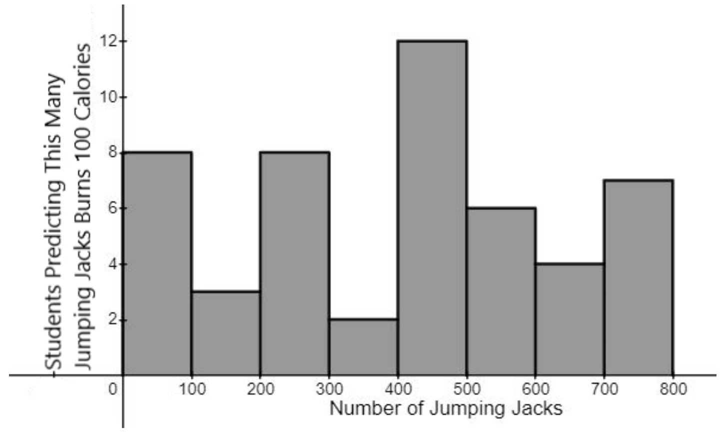
   b. Where in town does our math teacher live?
      - Is variability expected in the data? ________
      - Is the question statistical? ________
      - Will the answers give numerical data or categorical data? __________________

   c. How many minutes does it take students in my class to get ready for school in the morning?
      - Is variability expected in the data?________
      - Is the question statistical? ________
      - Will the answers give numerical data or categorical data? __________________

(Addressing NC.6.SP.1)

5. Examine the bar graph and histogram below.





a. What differences do you notice in each representation?

b. What question could have been asked to give the answers that are represented in the bar graph?

c. What question could have been asked to give the answers that are represented in the histogram?

d. How do the questions you came up with differ from one another?

(Addressing NC.6.SP.4)

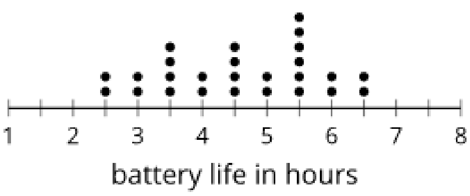
Lesson 5: Data Representations

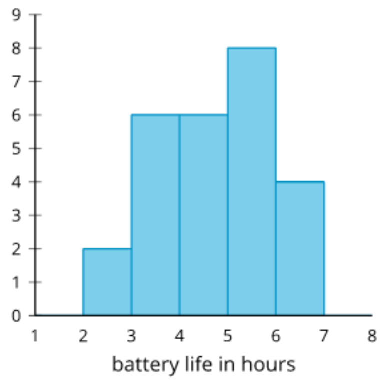
Learning Targets

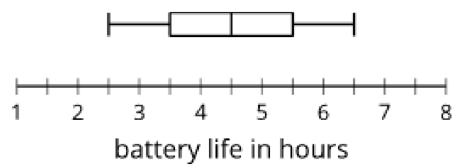
- I can find the five-number summary for data.
- I can use a histogram or box plot to represent data.

Warm-up: Battery Life

The dot plot, histogram, and box plot summarize the hours of battery life for 26 cell phones constantly streaming video. What do you notice? What do you wonder?







What do you notice?

What do you wonder?

Activity 1: Broad Jump Distances: Histogram

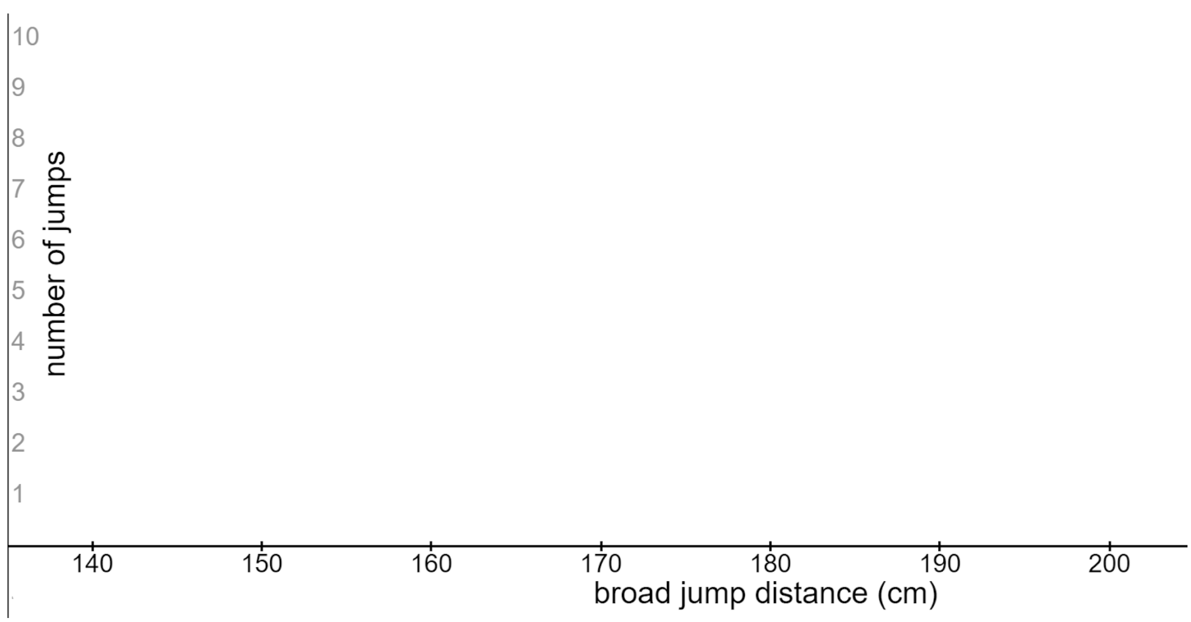
A histogram can be used to represent the distribution of numerical data.

1. The data represent the distances for all broad jumps (standing long jump), measured in centimeters (cm), at a JV track meet. Use the information to complete the frequency table.

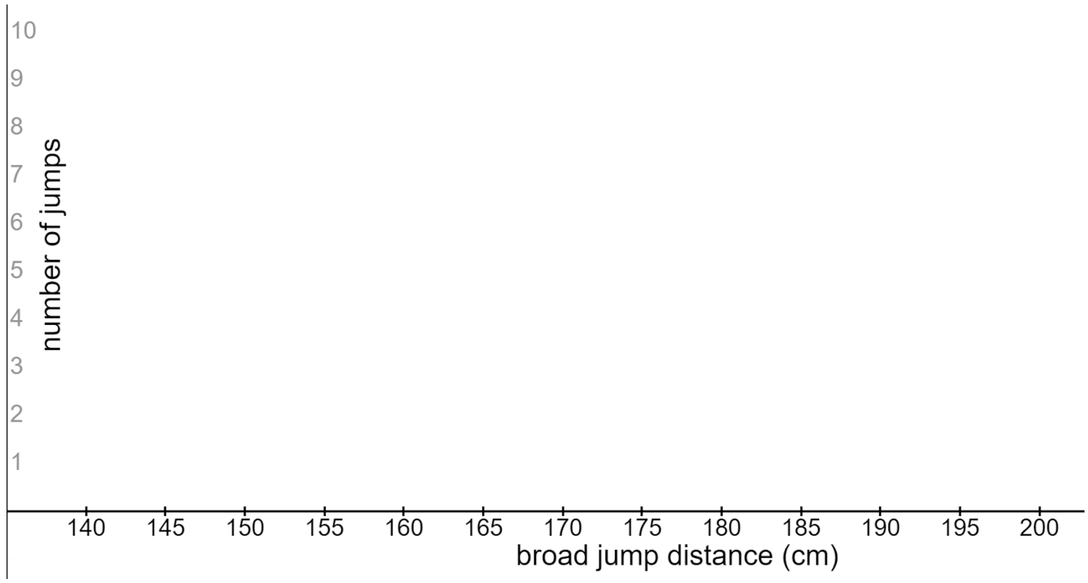
147	152	153	155	157	160	161	162
163	165	165	165	165	168	170	172
172	175	175	175	176	177	178	180
181	182	185	188	189	190		

Jump distance (cm)	Frequency
140–150	
150–160	
160–170	
170–180	
180–190	
190–200	

2. Use the set of axes and the information in your table to create a histogram.



3. The histogram you created has intervals of width 10 (like 140–150 and 150–160). Use the set of axes and data to create another histogram with an interval of width 5.



How does this histogram differ from the other one?

Are You Ready For More?

It often takes some playing around with the interval widths to figure out which gives the best sense of the shape of the distribution.

1. What might be a problem with using interval widths that are too large?

2. What might be a problem with using interval widths that are too small?

3. What other considerations might go into choosing the width of an interval?

Activity 2: Broad Jump Distances: Box Plot

A box plot can also be used to represent the distribution of numerical data.

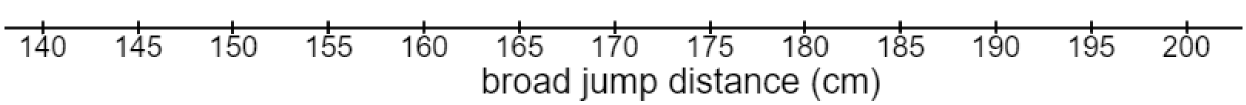
Minimum	Q1	Median	Q3	Maximum

1. Using the same data as the previous activity for long jump distances, find the median and add it to the table. What does the median represent for these data?

2. Find the median of the least 15 values. Because this number splits the data into the first and second quarters, it is called the first quartile. Add this value to the table under Q1. What does this value mean in this situation?

3. Find the value (the third quartile) that splits the data into the third and fourth quarters and add it to the table under Q3. Add the minimum and maximum values to the table.

4. Use the five-number summary to create a box plot that represents the length of the broad jumps in cm at the JV track meet.



Lesson Debrief

Lesson 5 Summary and Glossary

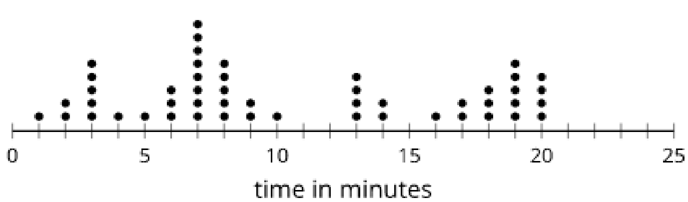
The table shows a list of the number of minutes people could intensely focus on a task before needing a break. Fifty (50) people of different ages are represented.

19	7	1	16	20	2	7	19
9	13	3	9	18	13	20	8
3	14	13	2	8	5	17	7
18	17	8	8	7	6	2	20
7	7	10	7	6	19	3	18
8	19	7	13	20	14	6	3
19	4						

In a situation like this, it is helpful to represent the **distribution** graphically to better notice any patterns or other interesting features in the data.

Distribution: How many of each value or each category there are in a data set that is numerical or categorical.

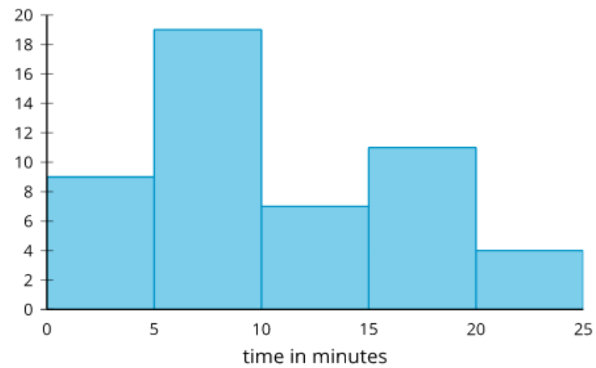
A **dot plot** shows all of the individual values in the data set:



Dot plots are useful when the data set is not too large. In this view, it is easy to see quite a few people lost focus at around 3, 7, 13, and 19 minutes, and nobody lost focus at 11, 12, or 15 minutes. If the data set is very large (more than 100 values, for example) or if there are many different values that are not exactly the same, it may be hard to see all of the dots on a dot plot.

Dot plot: A way of displaying data that shows all the values in the data set. It is useful when the data set is not too large.

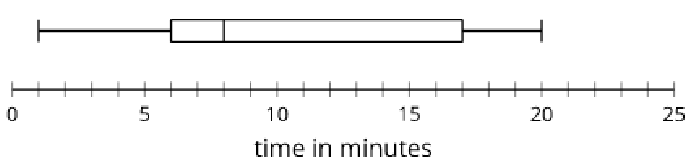
A **histogram** is useful for larger data sets and for data sets that are not restricted to whole numbers:



In this view, it is easy to see that most people lose focus between 5 and 10 minutes or between 15 and 20 minutes, while only 4 of the 50 people got distracted between 20 and 25 minutes. In a histogram, each interval includes the number at the lower end of the interval but not the upper end. For example, the tallest bar displays values that are greater than or equal to 5 minutes but less than 10 minutes.

Histogram: A graphical display of data using bars of different heights. In a histogram, each bar groups numbers into intervals. In each interval the lower boundary value is included in the interval, and the upper boundary value is not.

Box plots are created using the **five-number summary**:



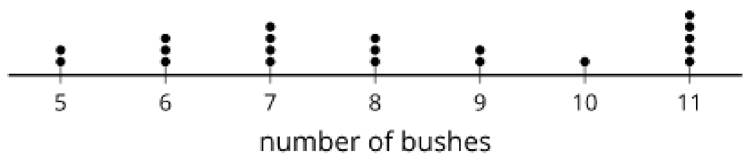
The vertical lines are drawn at the minimum (1 minute), Q1 (6 minutes), median (8 minutes), Q3 (17 minutes), and maximum (20 minutes). These 5 values split the data into four sections each representing approximately one-fourth of the data. Like the histogram, the box plot does not show individual data values, but other features such as quartiles, range, and median are seen more easily.

Five-number summary: These five values: the minimum value, the first quartile, the median, the third quartile, and the maximum value.

Box plot: A graphical display of data that shows the five-number summary.

Unit 1 Lesson 5 Practice Problems

1. The dot plot displays the number of bushes in the yards for houses in a neighborhood. What is the median?



2. The data set represents the shoe sizes of 19 students in a fifth-grade physical education class.

   4, 5, 5, 5, 6, 6, 6, 6, 7, 7, 7, 7, 7.5, 7.5, 8, 8, 8.5, 8.5, 9

   Create a box plot to represent the distribution of the data.

3. The data set represents the number of pages in the last book read by each of 20 students over the summer.

   163, 170, 171, 173, 175, 205, 220, 220, 220, 253, 267, 281, 305, 305, 305, 355, 371, 388, 402, 431

   Create a histogram to represent the distribution of the data.

4. Each set of data was collected from surveys to answer statistical questions. Select **all** of the data sets that represent numerical data.

    a. {1, 1.2, 1.4, 1.4, 1.5, 1.6, 1.8, 1.9, 2, 2, 2.1, 2.5}

    b. {Red, Red, Yellow, Yellow, Blue, Blue, Blue}

    c. {45, 60, 60, 70, 75, 80, 85, 90, 90, 100, 100, 100}

    d. {-7, -5, -3, -1, -1, -1, 0}

    e. {98.2, 98.4, 98.4, 98.6, 98.6, 98.6, 98.6, 98.7, 98.8, 98.8}

    f. {Yes, Yes, Yes, Yes, Maybe, Maybe, No, No, No}

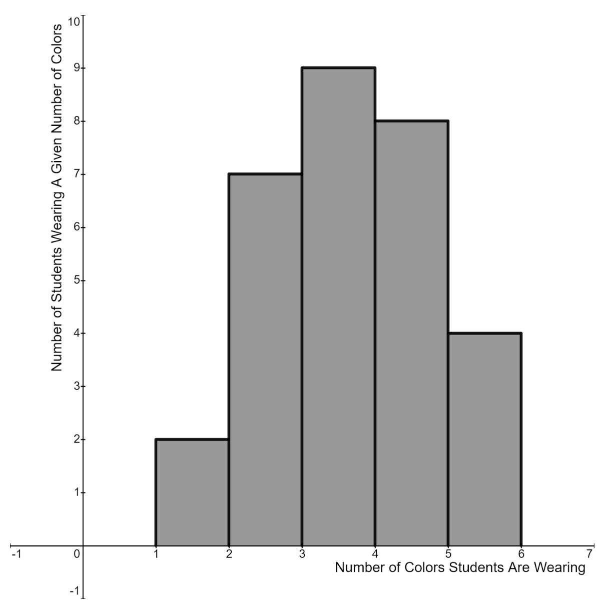
    g. {A, A, A, B, B, B, C, C, C}

(From Unit 1, Lesson 4)

5. Is “What is the typical distance a moped can be driven on a single tank of gas?” a statistical question? Explain your reasoning.

(From Unit 1, Lesson 4)

6. Here is a data display representing the number of colors students are wearing.



    a. How many students are wearing 1–2 colors?

    b. How many students are wearing 3–4 colors?

    c. How many students were surveyed?

    d. How might these representations look different if 200 students were surveyed?

(Addressing NC.6.SP.4)

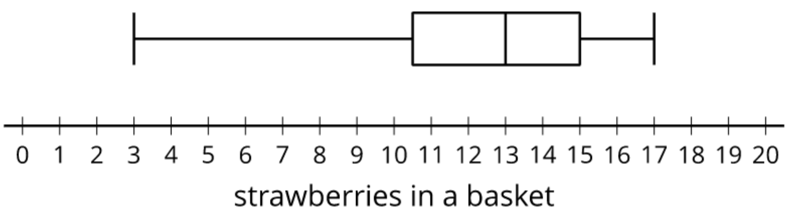
Lesson 6: A Gallery of Data

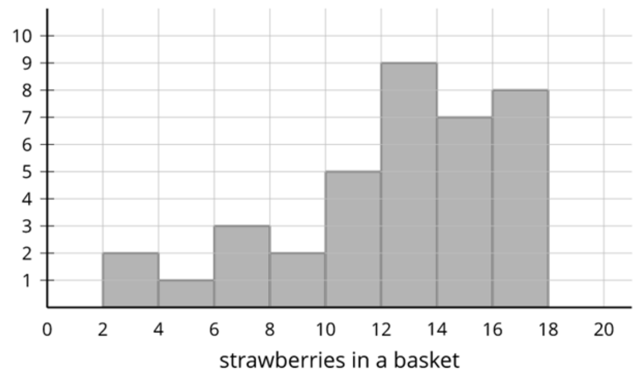
Learning Target

- I can graphically represent the data I collected and critique the representations of others.

Bridge

There are several baskets on a table, and each basket contains a certain number of strawberries. Here are two data displays showing the number of strawberries in each basket.



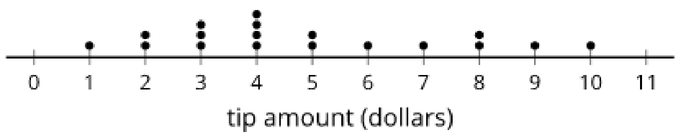


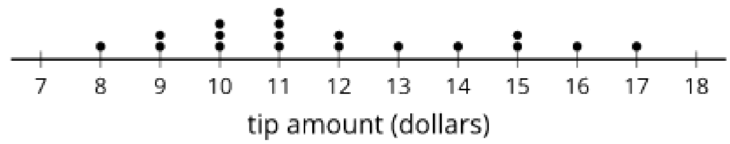
Kiran makes these claims. For each claim, decide whether you agree or disagree. Explain your reasoning using at least one of the data displays.

Claims	Agree/Disagree	Explain your reasoning
a. 3 is the least amount of strawberries in any basket.		
b. The range of the number of strawberries in baskets can be found using either of the two data displays.		
c. The total number of baskets can be found using either the box plot or histogram.		

Warm-up: Dot Plots

The dot plots represent the distribution of the amount of tips, in dollars, left at two different restaurants on the same night. What do you notice? What do you wonder?





What do you notice?

What do you wonder?

Activity 1: Data Displays

Your teacher will assign your group a statistical question. As a group:

1. Create a histogram and box plot to display the distribution of the data.

2. Write three comments that interpret the data.

As you visit each display during the gallery walk, write a sentence or two summarizing the information in the display.

Lesson Debrief

Lesson 6 Summary and Glossary

We can represent a distribution of data in several different forms, including lists, dot plots, histograms, and box plots. A list displays all of the values in a data set and can be organized in different ways.

This list shows the pH for 30 different water samples.

5.9	7.6	7.5	8.2	7.6	8.6	8.1	7.9
6.1	6.3	6.9	7.1	8.4	6.5	7.2	6.8
7.3	8.1	5.8	7.5	7.1	8.4	8.0	7.2
7.4	6.5	6.8	7.0	7.4	7.6		

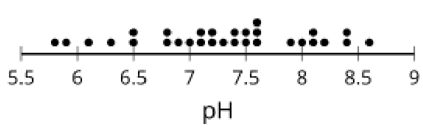
Here is the same list organized in order from least to greatest.

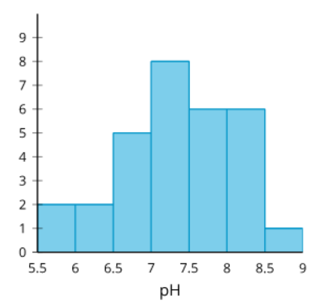
5.8	5.9	6.1	6.3	6.5	6.5	6.8	6.8
6.9	7.0	7.1	7.1	7.2	7.2	7.3	7.4
7.4	7.5	7.5	7.6	7.6	7.6	7.9	8.0
8.1	8.1	8.2	8.4	8.4	8.6		

With the list organized, you can more easily:

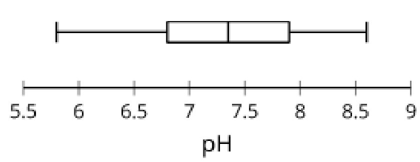
- interpret the data
- calculate the values of the five-number summary
- estimate or calculate the mean
- create a dot plot, histogram or box plot

Here is a dot plot and histogram representing the distribution of the data in the list.





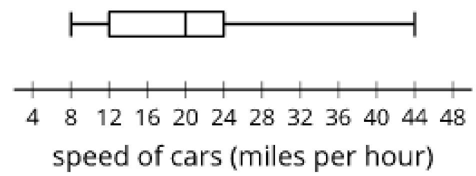
To make a histogram, first decide on the intervals you will use to group data together. Count the number of values from the data set in each interval and draw a bar over that interval at a height that matches the count. In the pH histogram, there are five water samples that have a pH between 6.5 and 7 (including 6.5, but not 7).
Here is a box plot representing the distribution of the same data as the dot plot and histogram.



To create a box plot, find the five-number summary: the minimum, first quartile, median, third quartile, and maximum values for the data set. Draw a vertical mark at each of these values, then connect the pieces as in the example. For the pH box plot, we can see that the minimum is about 5.8, the median is about 7.4, and the third quartile is around 7.9.

Unit 1 Lesson 6 Practice Problems

1. The box plot represents the distribution of speeds, in miles per hour, of 100 cars as they passed through a busy intersection.



   a. What is the smallest value in the data set? Interpret this value in the situation.

   b. What is the largest value in the data set? Interpret this value in the situation.

   c. What is the median? Interpret this value in the situation.

   d. What is the first quartile (Q1)? Interpret this value in the situation.

   e. What is the third quartile (Q3)? Interpret this value in the situation.

2. The data set represents the number of eggs produced by a small group of chickens each day for ten days: 7, 7, 7, 7, 7, 8, 8, 8, 8, 9.

   Select **all** the values that could represent the typical number of eggs produced in a day.

   a. 7.5 eggs

   b. 7.6 eggs

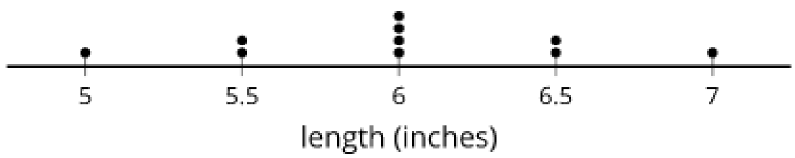
   c. 7.7 eggs

   d. 8 eggs

   e. 9 eggs

3. Choose one of the more interesting questions you or a classmate asked and collect data from a larger group, such as more students from the school. Create a data display and compare results from the data collected in class.

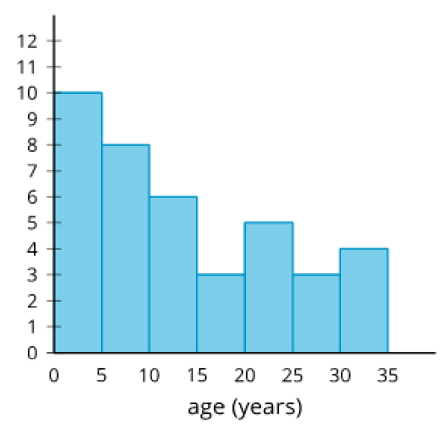
4. The dot plot displays the lengths of pencils (in inches) used by students in a class. What is the mean?



(From Unit 1, Lesson 5)

5. The histogram represents ages of 40 people at a store that sells children's clothes. Which interval contains the median?

   a. The interval from 0 to 5 years.
   b. The interval from 5 to 10 years.
   c. The interval from 10 to 15 years.
   d. The interval from 15 to 20 years.



(From Unit 1, Lesson 5)

6. The data set represents the responses, in degrees Fahrenheit, collected to answer the question, “How hot is the sidewalk during the school day?”

   92, 95, 95, 95, 98, 100, 100, 100, 103, 105, 105, 111, 112, 115, 115, 116, 117, 117, 118, 119, 119, 119, 119, 119, 119

   a. Create a box plot to represent the distribution of the data.

   b. Create a histogram to represent the distribution of the data.

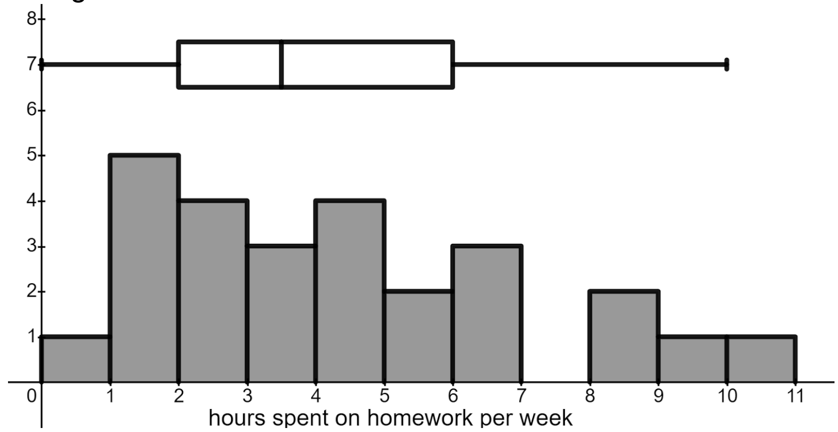
   c. Which display gives you a better overall understanding of the data? Explain your reasoning.

   (From Unit 1, Lesson 5)

7. Is “What is the area of the floor in this classroom?” a statistical question? Explain your reasoning.

   (From Unit 1, Lesson 4)

8. Here are two data displays that show the number of hours per week the same group of 26 NC Math 1 students reported spending on homework.[1]



Kai makes these claims. For each claim, decide whether you agree or disagree. Explain your reasoning using at least one of the data displays.

Claims	**Agree/Disagree**	**Explain your reasoning**
10 is the most amount of hours spent in a week on homework.		
The range of the number of hours spent on homework per week can be found using either of the two data displays.		
The total number of students spending less than 3 hours per week on homework can be found using either the box plot or histogram.		

(Addressing NC.6.SP.4; NC.6.SP.5a)

[1] Adapted from IM 6–8 Math https://curriculum.illustrativemathematics.org/MS/index.html, which was originally developed by Open Up Resources and authored by Illustrative Mathematics, and is copyright 2017–2019 by Open Up Resources. It is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0). OUR's 6–8 Math Curriculum is available at https://openupresources.org/math-curriculum/. Adaptations and updates to IM 6–8 Math are copyright 2019 by Illustrative Mathematics, and are licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

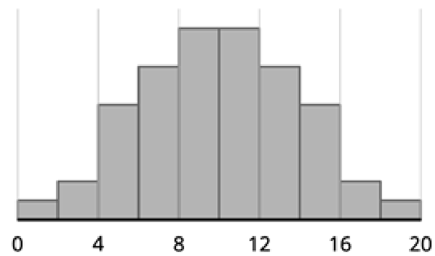
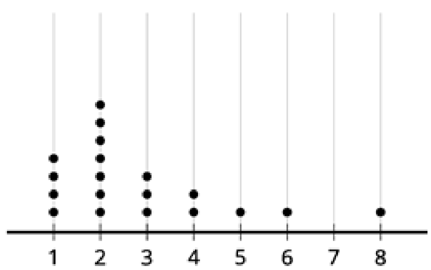
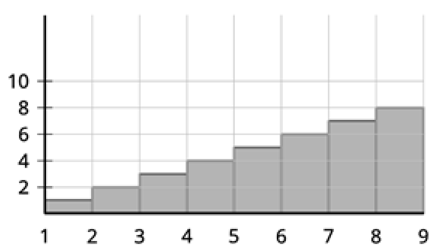
Lesson 7: The Shape of Distributions

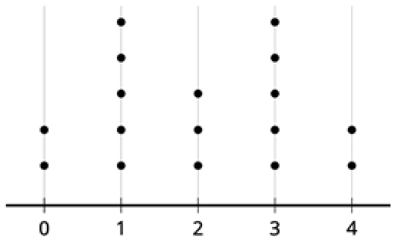
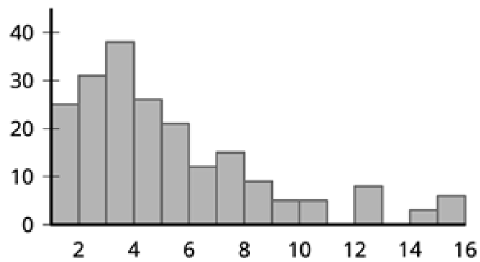
Learning Target

- I can describe the shape of a distribution using the terms "symmetric," "skewed," "bell-shaped," "uniform," and "bimodal."

Bridge

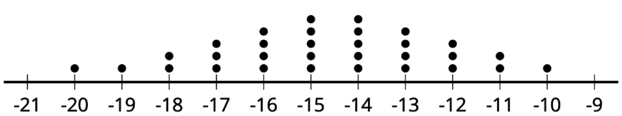
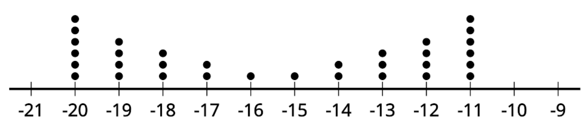
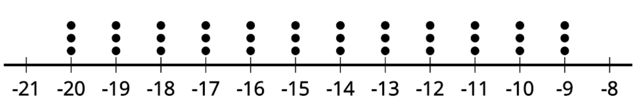
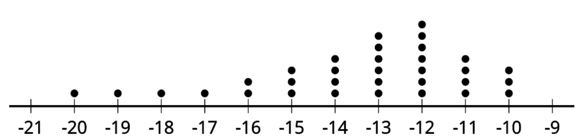
For each picture and description (a-e), state whether you agree or disagree, and explain your reasoning.

Claims	Agree/ Disagree	Explain your reasoning
a. Bell-shaped since there is a central peak for symmetric data values that are less frequent on the ends. 		
b. Symmetric because if the distribution was cut in half, both sides would be the same shape. 		
c. Uniform because there seems to be the same amount of data points across the entire distribution. 		

Claims	Agree/ Disagree	Explain your reasoning
d. Symmetric because if the distribution was cut in half, both sides would be the same shape. 		
e. Skewed left since most of the data values are on the left side of the distribution. 		

Warm-up: Distribution Shape

Which one doesn't belong? Explain your reasoning.

a. 	b. 
c. 	d. 

Activity 1: Matching Distributions

Take turns with your partner matching two different data displays that represent the distribution of the same set of data.

1. For each set that you find, explain to your partner how you know it's a match.
2. For each set that your partner finds, listen carefully to their explanation. If you disagree, discuss your thinking and work to reach an agreement.
3. When finished with all 10 matches, write a description of the shape of each distribution.

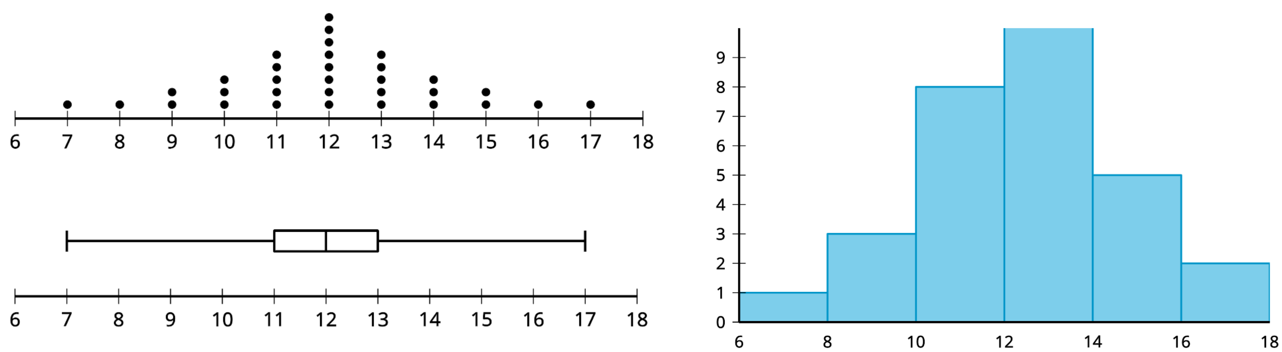
Display	Description of the shape of the distribution
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Lesson Debrief

Lesson 7 Summary and Glossary

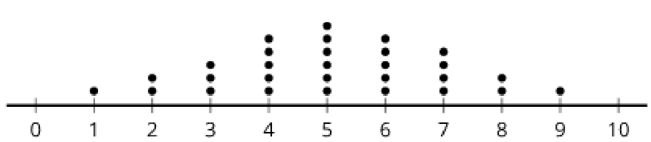
We can describe the shape of distributions as *symmetric*, *bell-shaped, skewed*, *uniform*, or *bimodal*.

Here is a dot plot, histogram, and box plot representing the distribution of the same data set. This data set has a **symmetric distribution**, which means there is a vertical line of symmetry in the center of the data display. In the dot plot shown, the distribution is symmetric about the data value 12. The histogram does not look perfectly symmetric because data is grouped together.



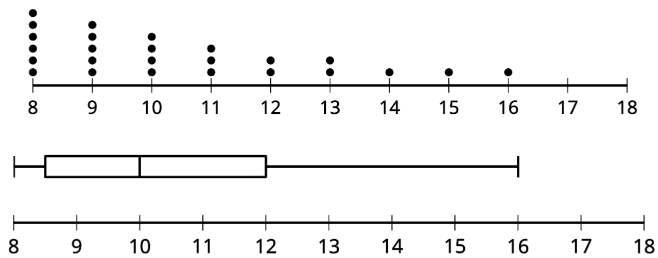
Symmetric distribution: A distribution that has a vertical line of symmetry in the center of the data display.

A **bell-shaped distribution** has a dot plot that takes the form of a bell with most of the data clustered near the center and fewer points farther from the center. Bell-shaped distributions are always symmetric or close to it.

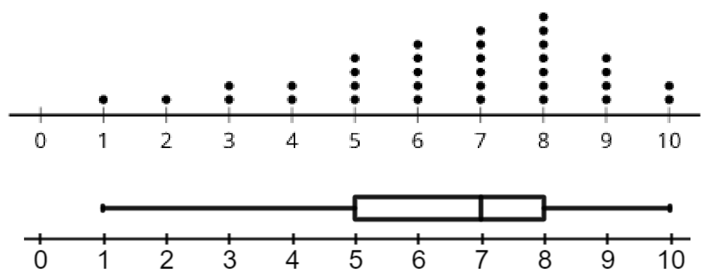


Bell-shaped distribution: A distribution that has a dot plot that takes the form of a bell with most of the data clustered near the center and fewer points farther from the center.

Here is a dot plot and box plot representing a **skewed distribution**. The data are **skewed right** because most of the data points are near the 8 to 10 interval, but there are many points to the right.

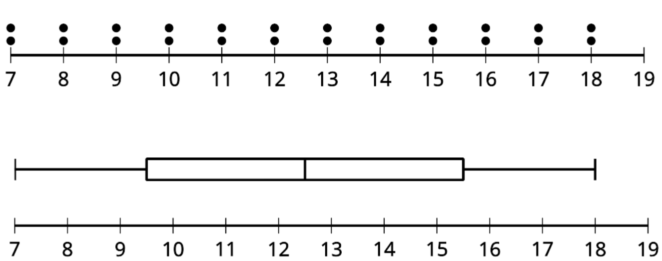


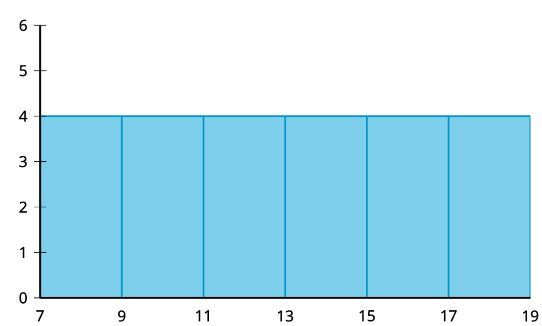
This dot plot and box plot show data that are **skewed left.** The data values on the left, such as 1, 2, and 3, are further from the bulk of the data than the data values on the right.



Skewed distribution: A distribution where one side of the distribution has more values farther from the bulk of the data than the other side. If the extreme values are on the right, the distribution is skewed right. If the extreme values are on the left, the distribution is skewed left.

A **uniform distribution** has the data values evenly distributed throughout the range of the data. This causes the distribution to look like a rectangle.

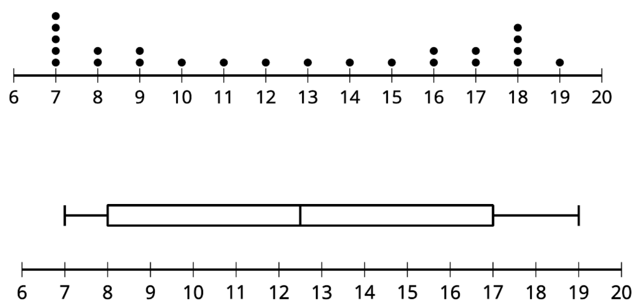


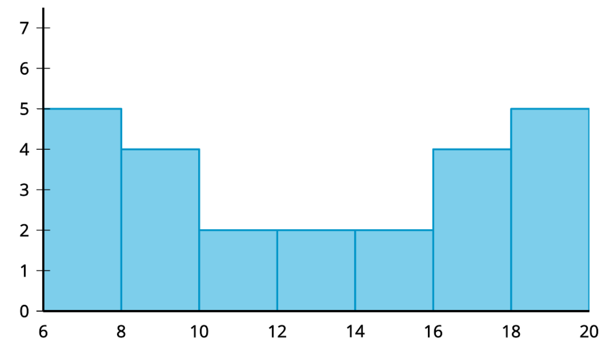


The box plot does not provide enough information to describe the shape of the distribution as uniform, though the even length of each quarter does suggest that the distribution may be approximately symmetric.

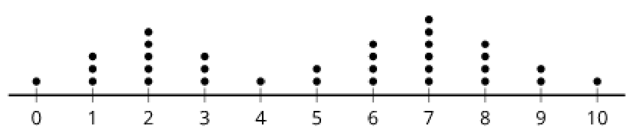
Uniform distribution: A distribution that has the values evenly distributed throughout the range of the data.

The following distribution is **bimodal**. The data cluster around two distinct peaks. Notice that you can see this in the dot plot and histogram, but not the box plot.





Bimodal distributions are not always symmetric, as in the dot plot shown below.

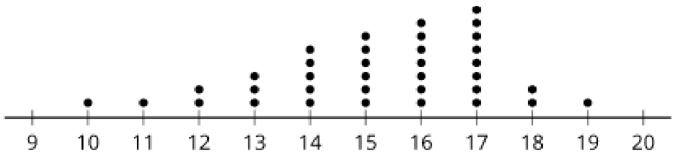


Bimodal distribution: A distribution that has two very common data values seen in a dot plot or histogram as distinct peaks.

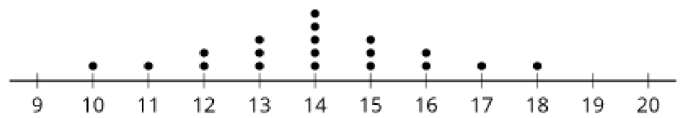
Unit 1 Lesson 7 Practice Problems

1. Which of the dot plots shows a symmetric distribution?

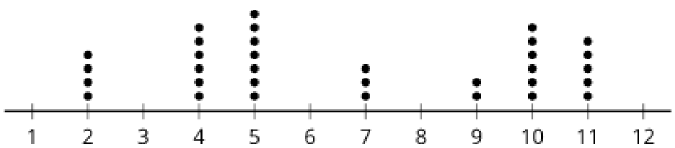
a.



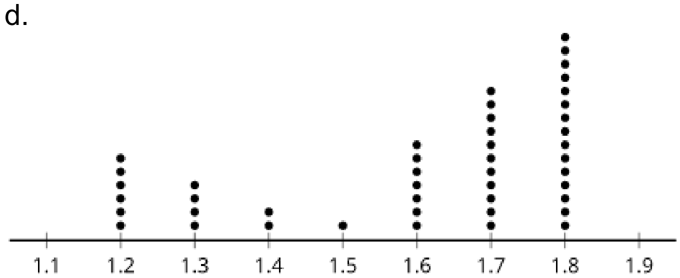
b.



c.

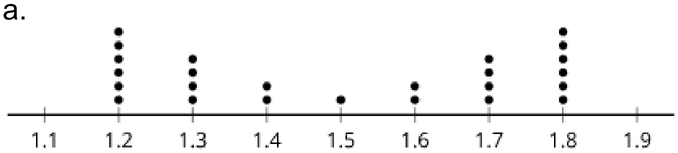


d.

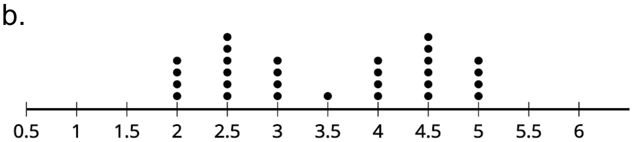


2. Which of the dot plots shows a skewed right distribution?

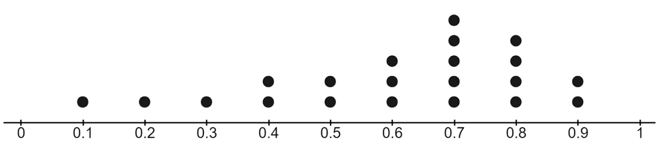
a.



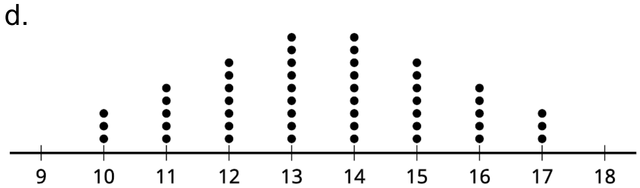
b.



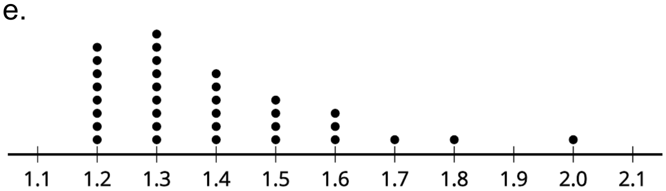
c.



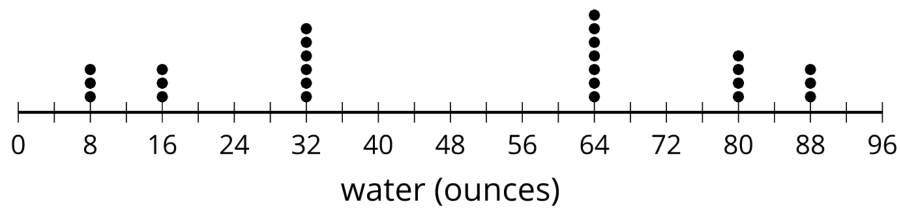
d.



e.



3. The data represent the number of ounces of water that 26 students drank before donating blood: 8, 8, 8, 16, 16, 16, 32, 32, 32, 32, 32, 32, 64, 64, 64, 64, 64, 64, 64, 80, 80, 80, 80, 88, 88, 88. Here is a dot plot that represents the data:



a. Create a box plot for the data.

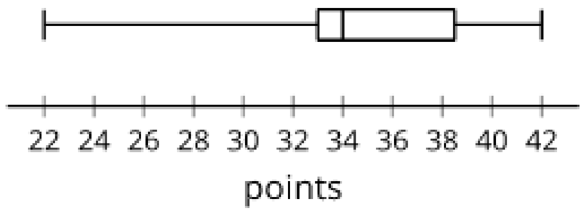
b. What information about the data is provided by the box plot that is not provided by the dot plot?

c. What information about the data is provided by the dot plot that is not provided by the box plot?

d. It was recommended that students drink 48 or more ounces of water. How could you use a histogram to easily display the number of students who drank the recommended amount?

(From Unit 1, Lesson 6)

4. The box plot represents the distribution of the number of points scored by a cross country team at 12 meets.



   a. If possible, find the mean. If not possible, explain why not.

   b. If possible, find the median. If not possible, explain why not.

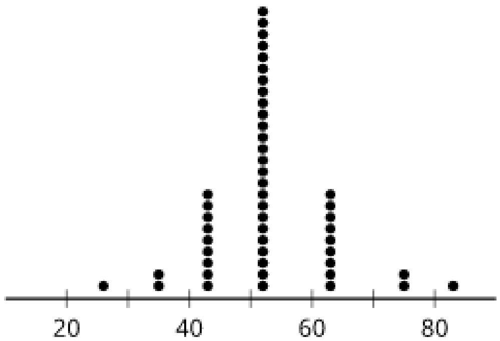
   c. Explain why it is impossible to tell if the cross country team ever scored 30 points at a meet.

(From Unit 1, Lesson 5)

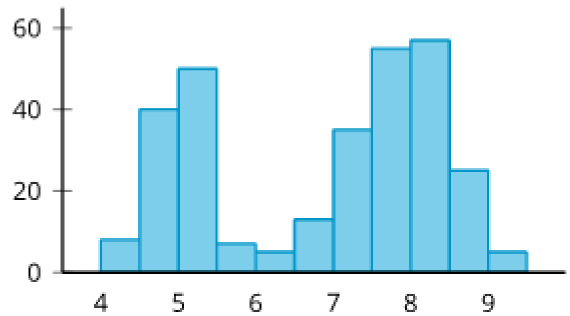
5. Describe the approximate shape of each distribution using the terms:

- approximately
- symmetric
- bell-shaped
- skewed left
- skewed right
- uniform
- bimodal

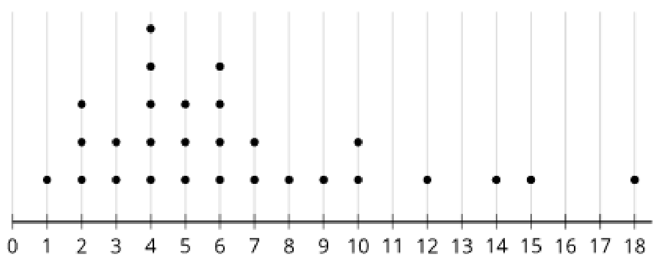
a.



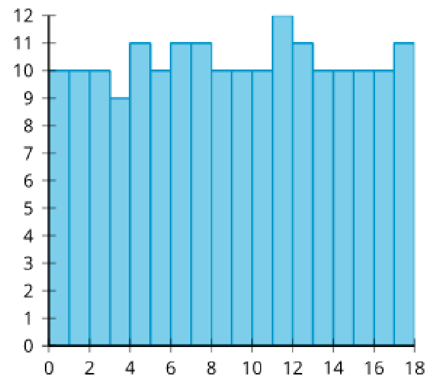
b.



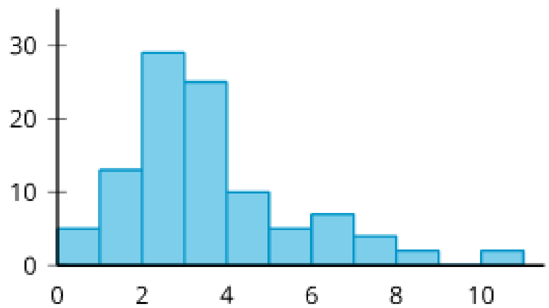
c.



d.



e.



(Addressing NC.6.SP.2)

Lesson 8: Calculating Measures of Center and Variability

Learning Target

- I can calculate interquartile range, mean, and median for a set of data.

Bridge

The first column of the table below is a list of measures of center and measures of variability. The second column describes the steps you take to calculate the measures. Match each measure with the way it is computed.

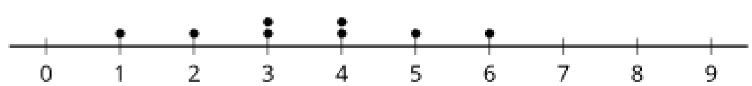
Measures	Calculations	Match
1. Median	a. Add up all of the values in a data set, then divide by the number of values in the set.	1. ________
2. Interquartile range	b. The difference between the first and the third quartiles.	2. ________
3. Mean	c. List the values in the data set in order, then find the middle value. If there are two "middle values," find the mean of those two values.	3. ________

Warm-up: Calculating Centers

Decide if each statement is true or false. Explain your reasoning.

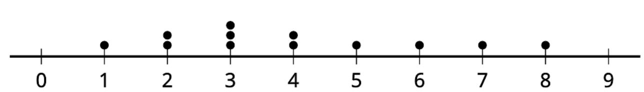
1. The mean can be found by adding all the numbers in a data set and dividing by the number of values in the data set.

2. The mean of the data in the dot plot is 4.



3. The median of the data set 4, 5, 9, 1, 10 is 9.

4. The median of the data in the dot plot is 3.5.



Activity 1: Heartbeats

The heart rates of eight high school students are listed in beats per minute:

72	75	81	76	76	77	79	78

1. What is the interquartile range?

2. How many values in the data set are:
   a. less than Q1?

   b. between Q1 and the median?

   c. between the median and Q3?

   d. greater than Q3?

Are You Ready For More?

1. A pod of dolphins contains 800 dolphins of various ages and lengths. The median length of dolphins in this pod is 5.8 feet. What information does this tell you about the length of dolphins in this pod?

2. The same vocabulary test with 50 questions is given to 600 students from fifth to tenth grades, and the number of correct responses is collected for each student in this group. The interquartile range is 40 correct responses. What information does this tell you about the number of correct responses for students taking this test? What would a box plot look like for the results of the vocabulary test?

Activity 2: Pennies on a Stick

Suppose there are six pennies taped onto a meter stick so that the mean position is the 50-centimeter mark. What information does this tell you about the position of the pennies along the meter stick?

1. Find possible locations for the six pennies.

2. Find a different set of possible locations for the six pennies.

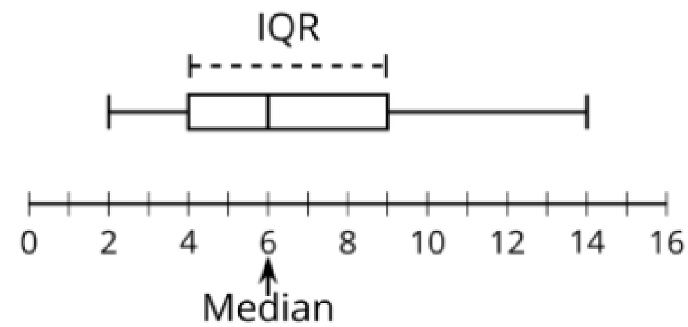
Lesson Debrief

Lesson 8 Summary and Glossary

In this lesson, we focused on the **interquartile range**, or **IQR**.

To calculate the IQR, subtract the value of the first quartile from the value of the third quartile. Recall that the first and third quartile are included in the five number summary.

The IQR is a measure of variability. Measures of variability tell you how much the values in a data set tend to differ from one another. A greater measure of variability means that the data are more spread out while a smaller measure of variability means that the data are more consistent and close to the measure of center.

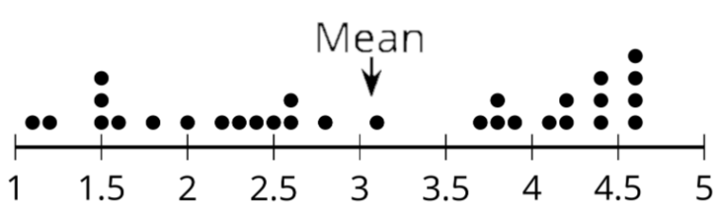


Interquartile Range (IQR): A measure of variability calculated by subtracting the first quartile from the third quartile.

Also in this lesson, we revisited two terms from previous learning:

Median: A measure of center that divides the data so that half of the values are greater than the median and half of the values are less than the median. Medians are easiest to see in a box plot.

Mean: Also called the average, it is the value you get by adding up all the values in the set and dividing by the number of values in the set. The mean also represents the balance point of a distribution.



Unit 1 Lesson 8 Practice Problems

1. The data set represents the number of errors on a typing test.

5	6	8	8	9	9	10	10	10	12

   a. What is the median? Interpret this value in the situation.

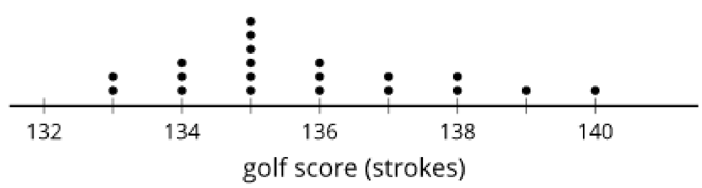
   b. What is the IQR?

2. The data set represents the heights, in centimeters, of ten model bridges made for an engineering competition.

13	14	14	16	16	16	16	18	18	19

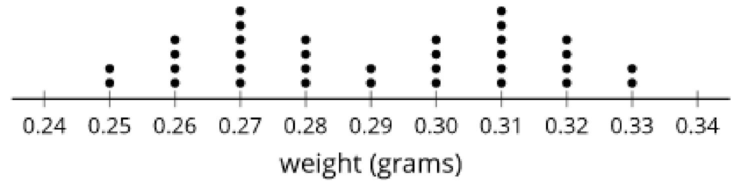
   What is the mean?

3. Describe the shape of the distribution shown in the dot plot. The dot plot displays the golf scores from a golf tournament.



(From Unit 1, Lesson 7)

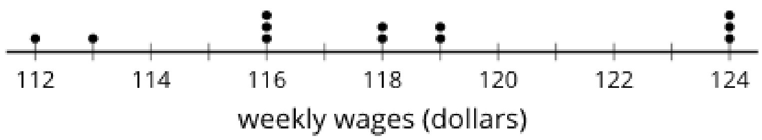
4. The dot plot shows the weight, in grams, of several different rocks. Select **all** the terms that describe the shape of the distribution.



   a. bell-shaped
   b. bimodal
   c. skewed
   d. symmetric
   e. uniform

(From Unit 1, Lesson 7)

5. The dot plot represents the distribution of wages earned during a one-week period by 12 college students.



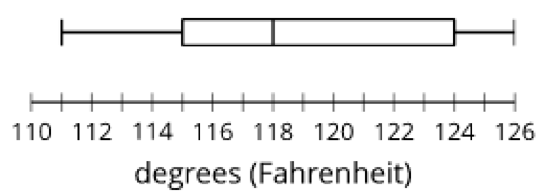
   a. What is the mean? Interpret this value based on the situation.

   b. What is the median? Interpret this value based on the situation.

   c. Would a box plot of the same data have allowed you to find both the mean and the median?

(From Unit 1, Lesson 6)

6. The box plot displays the temperature of saunas in degrees Fahrenheit. What is the median?



(From Unit 1, Lesson 5)

7. Here are the shoe sizes of some grade 9 and grade 12 students.

Grade 9 shoe sizes:

6, 8, 6.5, 7.5, 7, 6.5, 9, 6, 8.5, 7.5, 8, 10, 11, 8, 9

Grade 12 shoe sizes:

10, 9, 10.5, 8.5, 10, 9, 9.5, 8, 8, 11, 9, 9.5, 11, 10.5, 8.5

a. Create a box plot or histogram to represent both sets of data.

b. Complete the table.

	Mean	**Median**	**IQR**
Grade 9 shoe sizes			
Grade 12 shoe sizes			

c. Does one grade's shoe sizes have more variation than the other? Explain how you know.

(Addressing NC.6.SP.3.a; NC.7.SP.3.a)

Lessons 9 & 10: Checkpoint

Learning Targets

- I can create graphic representations of data and calculate statistics using technology.
- I can apply what I have learned to analyze and create graphical representations of data sets to answer questions and make informed decisions.

Station B: Technology

Access www.desmos.com/calculator on an internet browser.

To create a list in Desmos, first name the list and set it equal to the set of numbers. For example: a data set of the numbers 4, 3, 2, 4, 1 would be entered as L=[4, 3, 2, 4, 1]. You may use any letter you'd like to name your list but you must include the equal sign and the square brackets with numbers separated by commas.

1. Create a list of data using the numbers from the set of numbers below.

   Data Set: [7, 8, 4, 13, 5, 15, 14, 8, 12, 2, 8, 13, 12, 13, 6, 1, 9, 4, 9, 15, 21, 7, 9, 2, 5, 1, 9, 8, 9, 11, 4, 10, 6, 3, 13, 9, 12, 4, 6, 8, 7, 11]

2. On a new line, type "boxplot(L)" to create a box plot of the data set. Remember, if you named your data set with a different letter, that letter should be in parentheses instead of L.

   - How many values are in the data set? Explain your reasoning.

   - On a new line, type "mean(L)"
     - What is the mean of the data set?

   - On a new line, type "stats(L)"
     - What is the minimum of the data set?
     - What is the first quartile of the data set?
     - What is the median of the data set?
     - What is the third quartile of the data set?
     - What is the maximum of the data set?

- Use the Q1 and Q3 values to calculate the interquartile range (IQR). What is the IQR?

- Next, add the number 32 to your data set.
  - What happens to the mean?
  - What happens to the median?
  - Discuss with your partner or reflect individually: Which center of measure (mean or median) was most affected by the addition of the 32? Why?

3. Open up www.desmos.com/calculator on a new tab.
   - From the new blank desmos screen, create a data set using the same set of numbers, without the number 32.
   - Create a histogram of the data by typing “histogram(L)” on a new line.
   - Change the Bin Alignment to Left.
   - Describe the shape of the data.

   - Click on Bin Width. Notice that the notation changes to histogram(L,1). Replace 1 with 2 so it reads histogram(L,2).
   - How did the display change as a result of adjusting the Bin Width?

   - Discuss with your partner or reflect individually: Why might changing the Bin Width be useful when representing a data set?

4. Open up www.desmos.com/calculator on a new tab to answer the following questions about the data set below.

   The data below represent the amount of corn, in bushels per acre, harvested from different locations.

   133, 133, 134, 134, 134, 135, 135, 135, 135, 135, 135, 136, 136, 136, 137, 137, 138, 138, 139, 140

   - Use technology to create a histogram and a box plot.

   - What is the shape of the distribution?

   - Compare the information displayed by the histogram and box plot.

Station D: Where Did The Distribution Come From?

With a partner:

1. Choose a survey card. This is a question that resulted in one of the matched sets of data displays.
2. Discuss the card and make an educated guess about which set of data displays represents the data from that survey question.
3. Explain your choice. Use the questions below as a guide for your explanation:
    - How did you use the shape of the data to come up with your question?
    - Would you always expect your question to result in a [symmetric, skewed, bell-shaped, bimodal, uniform] distribution?
4. There is one survey card for each set of data displays.
5. Repeat the process, as time permits.

Survey question	**Data display**	**Explanation**
a.		
b.		
c.		
d.		
e.		
f.		
g.		
h.		
i.		
j.		

Station G: Long Jump[1]

The track and field coach has to select a girl for the long jump at the regional championship. Three girls are in contention. The results of a school jump-off, in meters, are given in the accompanying table.

Elena	Jada	Priya
3.25	3.55	3.67
3.95	3.88	3.78
4.28	3.61	3.92
2.95	3.97	3.62
3.66	3.75	3.85
3.81	3.59	3.73

1. Who do you think should go to the championship? Write a letter to the coach with your decision and an explanation of your reasoning.

2. What are the long-jump records for your school, district, and/or state? When did they occur? Be sure to include the units. What do you notice and wonder about what you discovered?

[1] Adapted from AchievetheCore.org

Unit 1 Lessons 9 & 10 Practice Problems

1. *Technology required.* The data represent the average customer ratings for several items sold online.

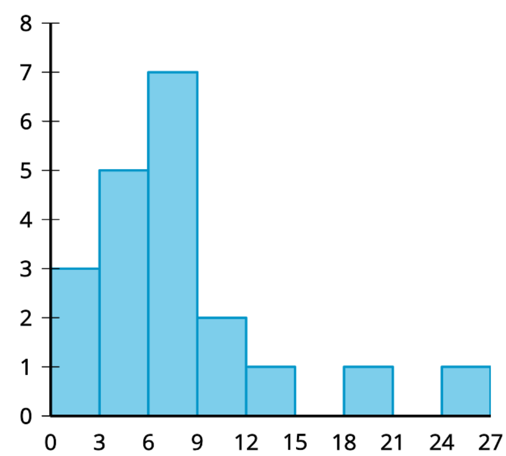
0.5	1	1.2	1.3	2.1	2.1	2.1	2.3
2.5	2.6	3.5	3.6	3.7	4	4.1	4.1
4.2	4.2	4.5	4.7	4.8			

a. Use technology to create a histogram for the data with intervals 0–1, 1–2, and so on.

b. Describe the shape of the distribution.

c. Which interval has the highest frequency?

2.



a. Describe the shape of the distribution.

b. How many values are represented by the histogram?

c. Write a statistical question that could have produced the data set summarized in the histogram.

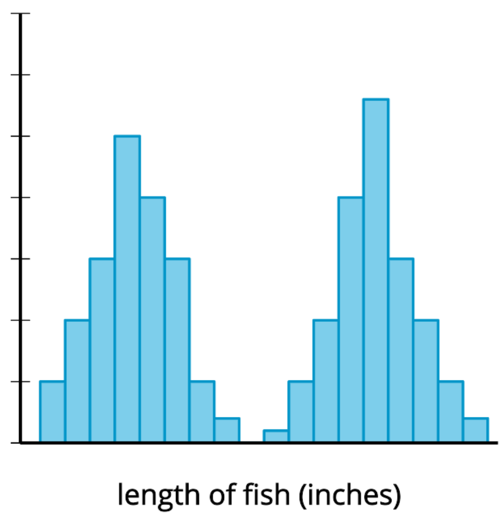
(From Unit 1, Lesson 4)

3. The dot plot represents the distribution of satisfaction ratings for a landscaping company on a scale of 1 to 10. Twenty-five customers were surveyed. On average, what was the satisfaction rating of the landscaping company?

(From Unit 1, Lesson 3)

4. This distribution shows the length in inches of fish caught and released from a nearby lake.

    a. Describe the shape of the distribution.



    b. Make an educated guess about what could cause the distribution to have this shape.

(From Unit 1, Lesson 7)

Lesson 11: The Effect of Extremes

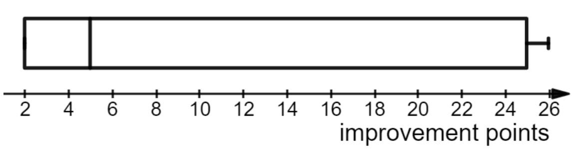
Learning Targets

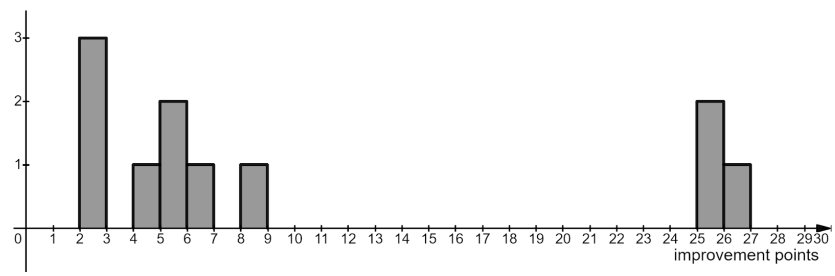
- I can describe how some data points can affect the mean and median.
- I can use the shape of a distribution to compare the mean and median.

Bridge

A gym scores their clients on a scale of 1–100 and claims that they can help their clients improve on their exercises. After two months with a gym membership, eleven gym members are assessed to see how much they have improved on their exercises.

Here are two different data displays of the same data that represent how much the eleven gym members have improved.





1. What is the mean improvement among these members?

2. What is the median?

3. Which measure of center is a better representation of the members' improvement? Explain your reasoning.

Activity 1: Separated by Skew

1. Use technology to create a histogram that represents the distribution of the data below.

   6 7 8 8 9 9 9 10 10 10 10 11 11 11 12 12 13 14

   Describe the shape of the distribution.

2. Find the mean and median of the data.

3. Find the mean and median of the data with two additional values included as described.

	Mean	Median
a. Add two values to the original data set that are greater than 14.		
b. Add two values to the original data set that are less than 6.		
c. Add one value that is greater than 14 and one value that is less than 6 to the original data set.		
d. Add the two values, 50 and 100, to the original data set.		

4. Share your work with your group. What do you notice is happening with the mean and median based on the additional values?

5. Change the values so that the distribution fits the description given to you by your teacher, then find the mean and median.

6. Find another group that created a distribution with a different description. Explain your work and listen to their explanation, then compare your measures of center.

Are You Ready For More?

The mean and the median are by far the most common measures of center for numerical data. There are other measures of center, though, that are sometimes used. For each measure of center, list some possible advantages and disadvantages. Be sure to consider how it is affected by extremes.

1. *Interquartile mean*: The mean of only those points between the first quartile and the third quartile.

2. *Midhinge*: The mean of the first quartile and the third quartile.

3. *Midrange*: The mean of the minimum and maximum value.

4. *Trimean:* The mean of the first quartile, the median, the median again, and the third quartile. So we are averaging four numbers as the median is counted twice.

Activity 2: Plots Matching Measures

Use technology to create a histogram with at least 10 values for each of the conditions listed. Each histogram must have at least three values that are different.

1. A distribution that has both mean and median of 10

2. A distribution that has both mean and median of -15

3. A distribution that has a median of 2.5 and a mean greater than the median

4. A distribution that has a median of 5 and a median greater than the mean

Lesson Debrief

Lesson 11 Summary and Glossary

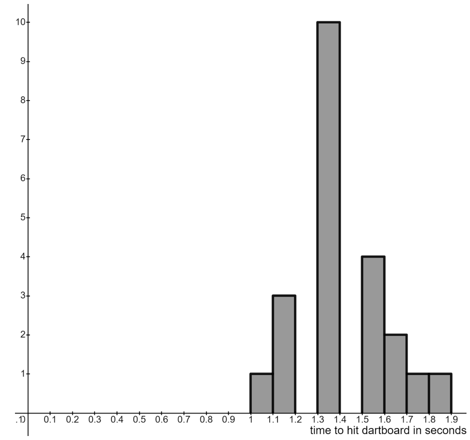
Have you heard the word **statistic** before? You have been using different statistics in all the data that you have examined.

Statistic: A quantity that is calculated from sample data, such as mean or median.

How do you decide when to use the mean or median to describe the center of a data set?

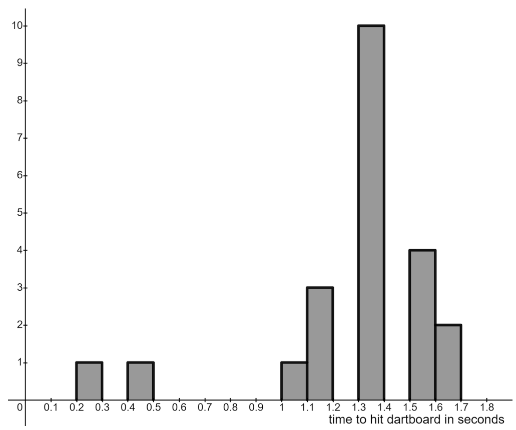
- The mean gives equal importance to each value. The mean usually represents the typical values well when the data sets have a symmetric distribution. On the other hand, the mean can be greatly affected by changes to even a single value.
- The median tells you the middle value in the data set, so changes to a single value usually do not affect the median much. Because of this, the median is more appropriate for data sets that are not very symmetric.

Here is a histogram showing the amount of time a dart takes to hit a target in seconds. The data produces a symmetric distribution.



Since the distribution is symmetric, the mean should be used to describe a typical value of the data.

Here is a histogram using the same data, but with two of the values changed, resulting in a skewed distribution.



Since the distribution is skewed, the median should be used to describe a typical value of the data. The extreme values at the lower end make the mean lower than most values in the data set. The median is more resistant to those extreme values, so it is the preferred measure of center.

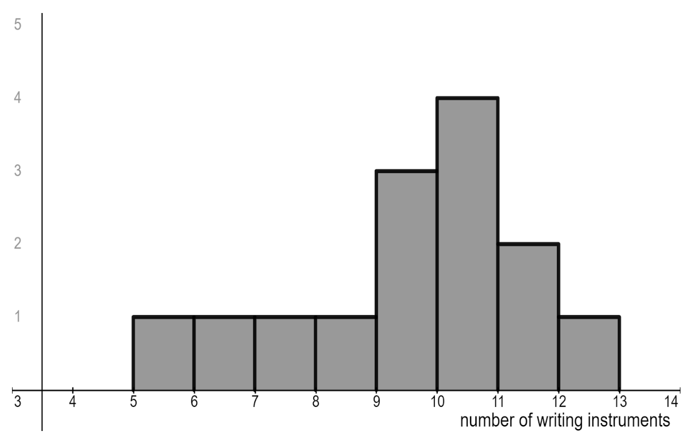
The distribution above was skewed left, making the mean lower than the median. Likewise, if a distribution is skewed right, the mean will almost always be higher than the median.

Unit 1 Lesson 11 Practice Problems

1. Select **all** the distribution shapes for which it is most often appropriate to use the mean.
    a. bell-shaped
    b. bimodal
    c. skewed
    d. symmetric
    e. uniform

2. For which distribution shape is it usually appropriate to use the median when summarizing the data?
    a. bell-shaped
    b. skewed
    c. symmetric
    d. uniform

3. The number of writing instruments in some teachers' desks is displayed in the histogram. Which is greater, the mean or the median? Explain your reasoning using the shape of the distribution.



4. A student has these scores on their assignments. The teacher is considering dropping the lowest score. What effect does eliminating the lowest value, 0, from the data set have on the mean and median?

    0, 40, 60, 70, 75, 80, 85, 95, 95, 100

(From Unit 1, Lessons 9 & 10)

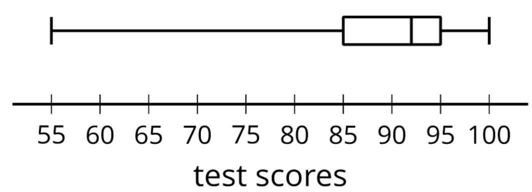
5. The data set 2, 2, 4, 4, 5, 5, 6, 7, 9, 15 represents the number of books purchased by the first 10 customers in a bookstore.

   a. What is the five-number summary?

   b. When the maximum, 15, is removed from the data set, what is the five-number summary?

(From Unit 1, Lessons 9 & 10)

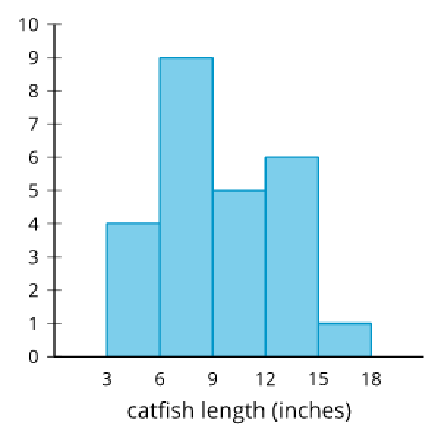
6. The box plot summarizes the test scores for 100 students: Which term best describes the shape of the distribution?



   a. bell-shaped
   b. skewed right
   c. skewed left
   d. symmetric

(From Unit 1, Lesson 7)

7. The histogram represents the distribution of lengths, in inches, of 25 catfish caught in a lake.



   a. If possible, find the mean. If not possible, explain why not.

   b. If possible, find the median. If not possible, explain why not.

   c. Were any of the fish caught 12 inches long?

   d. Were any of the fish caught 19 inches long?

(From Unit 1, Lesson 5)

8. For five days, Tyler has recorded how long his walks to school take in minutes. The mean for his data is 11 minutes.[1]

    a. Without calculating, predict if each of the data sets shown could be Tyler's. Explain your reasoning.

        - data set A (11, 8, 7, 9, 8)
        - data set B (12, 7, 13, 9, 14)
        - data set C (11, 20, 6, 9, 10)
        - data set D (8, 10, 9, 11, 11)

    b. Determine which data set is Tyler's. Explain how you know.

(Addressing NC.6.SP.3.a)

Lesson 12: Standard Deviation

Learning Targets

- I can describe standard deviation as a measure of variability.
- I can use technology to compute standard deviation.

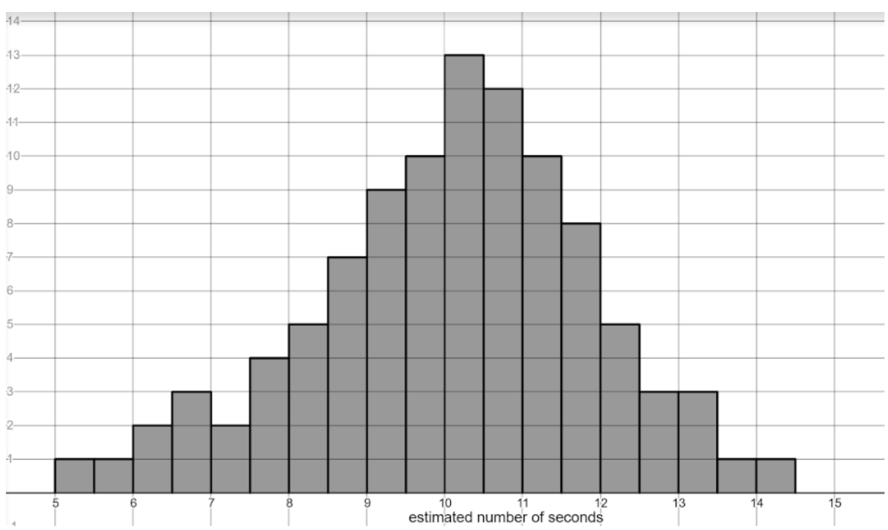
Warm-up: Estimating Time

The passing of time is constant, but our estimation of how much time has passed can vary. How close are we at estimating 10 seconds?

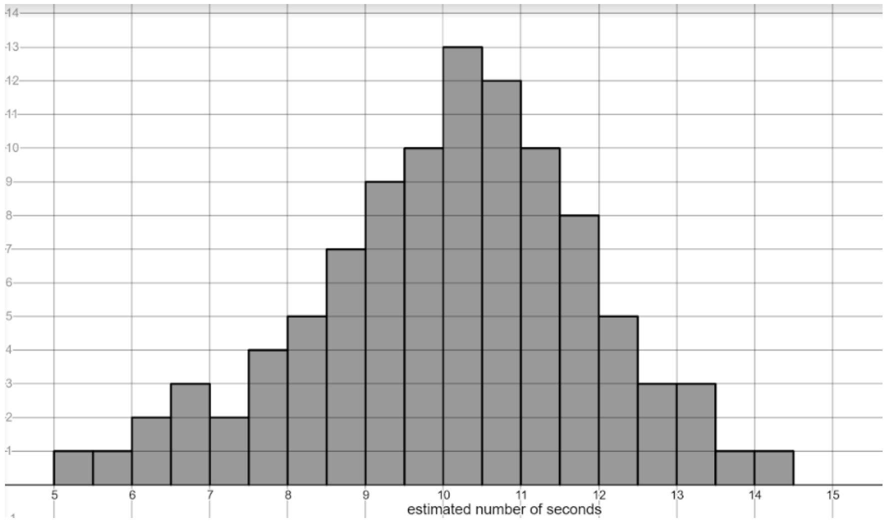
Open the Estimating Time app. When ready, press the button to start the timer. When you believe 10 seconds has passed press the button again to stop the timer. Record the amount of time that actually passed.

1. How close was your estimate to being 10 seconds?

2. This histogram shows the estimates of 100 students.



- Describe the shape.



- Estimate the center of the data set and mark it on the horizontal axis.

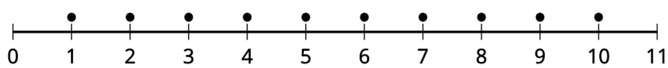
- Mark on the histogram:
  - the estimates that are the best
  - the estimates that are close
  - the estimates that are not so close

3. Consider the estimates that are close: some will be too high and some will be too low. Approximately how far from the mean are these values?

Activity 1: Investigating Standard Deviation

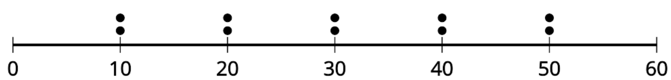
1. Use technology to find the mean and the standard deviation for the data in the dot plots.

Partner 1 Dot plots:



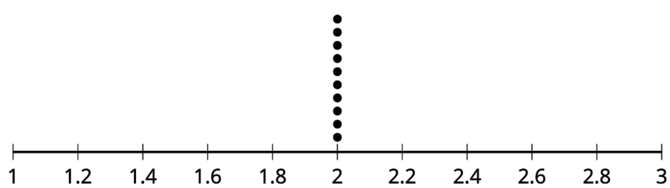
Mean: __________

Standard deviation: ___________



Mean: __________

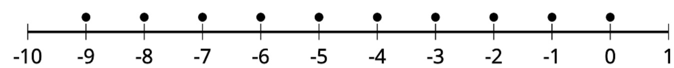
Standard deviation: ___________



Mean: __________

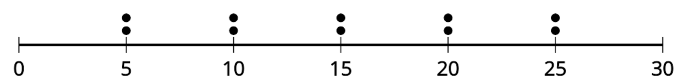
Standard deviation: ___________

Partner 2 Dot plots:



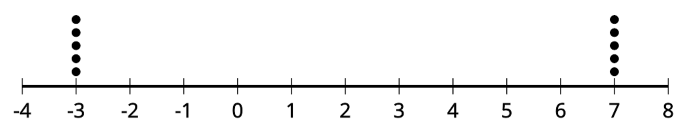
Mean: __________

Standard deviation: ___________



Mean: __________

Standard deviation: ___________



Mean: __________

Standard deviation: ___________

2. What do you notice about the mean and standard deviation you and your partner found for the three dot plots?

Are You Ready For More?

Invent some data that fits the conditions. Be prepared to share your data set and reasoning for choice of values.

1.
    a. 10 numbers with a standard deviation equal to the standard deviation of your first dot plot with a mean of 6.

    b. 10 numbers with a standard deviation equal to the standard deviation of your first dot plot with a mean of 12.

2.
    a. 10 numbers with a standard deviation three times greater than the data in the first row.

    b. 10 numbers with a standard deviation four times greater than the data in the first row.

Activity 2: Investigating Variability

Begin with this data set:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20

1. Use technology to find the mean, standard deviation, and median. Calculate the interquartile range (IQR).

2. How do the standard deviation and mean change when you remove the greatest value from the data set? How do they change if you add a value to the data set that is twice the greatest value?

3. What do you predict will happen to the standard deviation and mean when you remove the least value from the data set? Check to see if your prediction was correct.

4. What happens to the standard deviation and mean when you add a value to the data set equal to the mean? Add a second value equal to the mean. What happens?

5. Add, change, and remove values from the data set to answer the question: What appears to change more easily, the standard deviation or the interquartile range? Explain your reasoning.

Are You Ready For More?

How is the standard deviation calculated?

1. Using the original data set, calculate the deviation of each point from the mean by subtracting the mean from each data point.

2. If we just tried to take a mean of those deviations, what would we get?

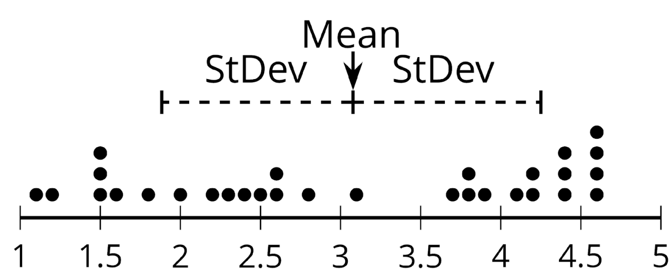
3. There are two common ways to turn negative values into more useful positive values: take the absolute value or square the value. To find the standard deviation, we first square each of the deviations and then find the mean of those numbers. Finally, we take the square root of that mean. Compute the standard deviation of the original data set.

Lesson Debrief

Lesson 12 Summary and Glossary

We can describe the variability of a distribution using the **standard deviation**. The standard deviation is a measure of variability that tells you how spread out the data points are from the mean. The standard deviation is large when the data are more spread out, and the standard deviation is small when the data are close together.

Since standard deviation is a measure of variability, two data sets with the same shape will have the same standard deviation. For instance, the data sets {1, 3, 5} and {4, 6, 8} have the same standard deviation because the distances from the mean for both data sets are 2, 0, and 2. Unlike IQR, the standard deviation changes a lot when you add or remove certain values from a data set.



Standard deviation: A measure of variability based on how spread out the data points are from the mean.

Unit 1 Lesson 12 Practice Problems

1. The shoe size for all the pairs of shoes in a person's closet are recorded.

7	7	7	7	7	7	7	7	7	7

   a. What is the mean?

   b. What is the standard deviation?

2. Here is a data set:

1	2	3	3	4	4	4	4	5	5	6	7

   a. What happens to the mean and standard deviation of the data set when the 7 is changed to a 70?

   b. For the data set with the value of 70, why would the median be a better choice for the measure of center than the mean?

3. Here is a data set with a mean 57 of and a standard deviation of 6.9:

44	54	57	57	60	61	66

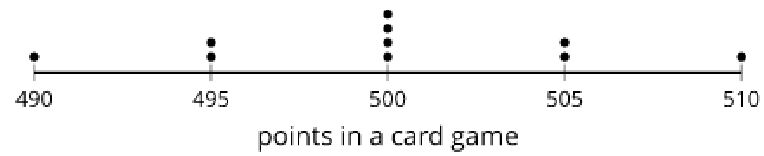
   a. Each value in this new data set increases each original value by three. Without calculating, find the new mean and standard deviation of this data set:

47	57	60	60	63	64	69

   b. Each value in this new data set is twice as big as the values in the original data set. Without calculating, find the new mean and standard deviation of this data set:

88	108	114	114	120	122	132

4. Which of these best estimates the standard deviation of points in a card game?

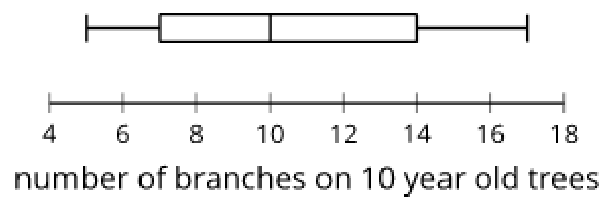


   a. 5 points
   b. 20 points
   c. 50 points
   d. 500 points

5. Select **all** the distribution shapes for which the mean and median *must be* about the same.
   a. bell-shaped
   b. bimodal
   c. skewed
   d. symmetric
   e. uniform

(From Unit 1, Lesson 11)

6. What is the IQR?



   a. 5 branches
   b. 7 branches
   c. 10 branches
   d. 12 branches

(From Unit 1, Lesson 13)

7. The data represent the number of cans collected by different classes for a service project.

12	14	22	14	18	23	42	13	9	19	22	14

   a. Find the mean.
   b. Find the median.
   c. Eliminate the greatest value, 42, from the data set. Explain how the measures of center change.

(From Unit 1, Lesson 8)

Lesson 13: Comparing and Contrasting Data Distributions

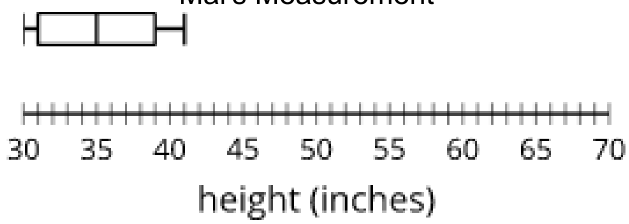
Learning Target

- I can arrange data sets in order of variability given graphic representations.

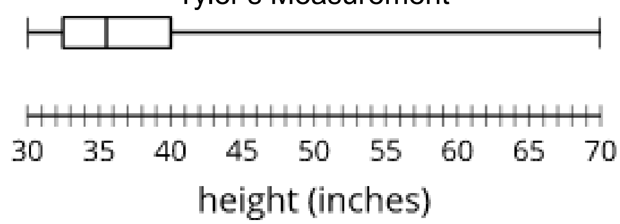
Bridge

Mai and Tyler both visit the same preschool classroom and measure the heights of people in the room in inches. The summary of their results are shown in the box plot and statistics.

Mai's Measurement



Tyler's Measurement



What do you notice?

What do you wonder?

Warm-up: Math Talk: Mean

Evaluate the mean of each data set mentally.

- 27, 30, 33
- 61, 71, 81, 91, 101
- 0, 100, 100, 100, 100
- 0, 5, 6, 7, 12

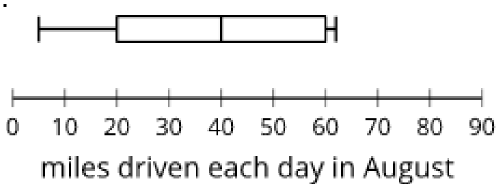
Activity 1: Describing Data Distributions

1. Your teacher will give you a set of cards. Take turns with your partner to match a data display with a written statement.
    a. For each match that you find, explain to your partner how you know it’s a match.
    b. For each match that your partner finds, listen carefully to their explanation. If you disagree, discuss your thinking and work to reach an agreement.
2. After matching, determine if the mean or median is more appropriate for describing the center of the data set based on the shape of the distribution. Discuss your reasoning with your partner. If it is not given, calculate (if possible) or estimate the appropriate measure of center. Be prepared to explain your reasoning.

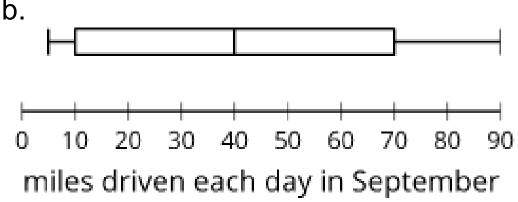
Activity 2: Visual Variability and Statistics

Each box plot summarizes the number of miles driven each day for 30 days in each month. The box plots represent, in order, the months of August, September, October, November, and December.

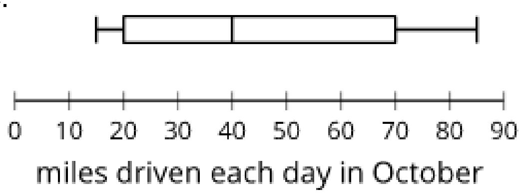
a.



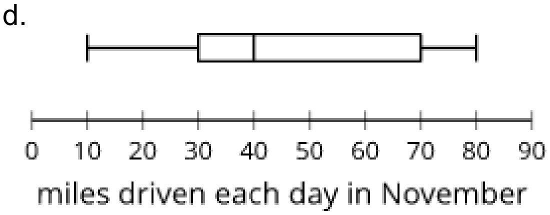
b.



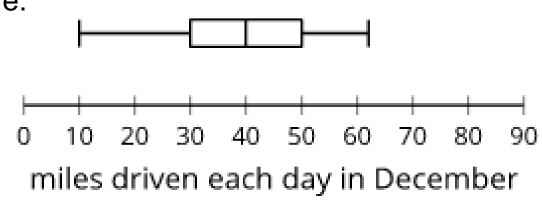
c.



d.



e.

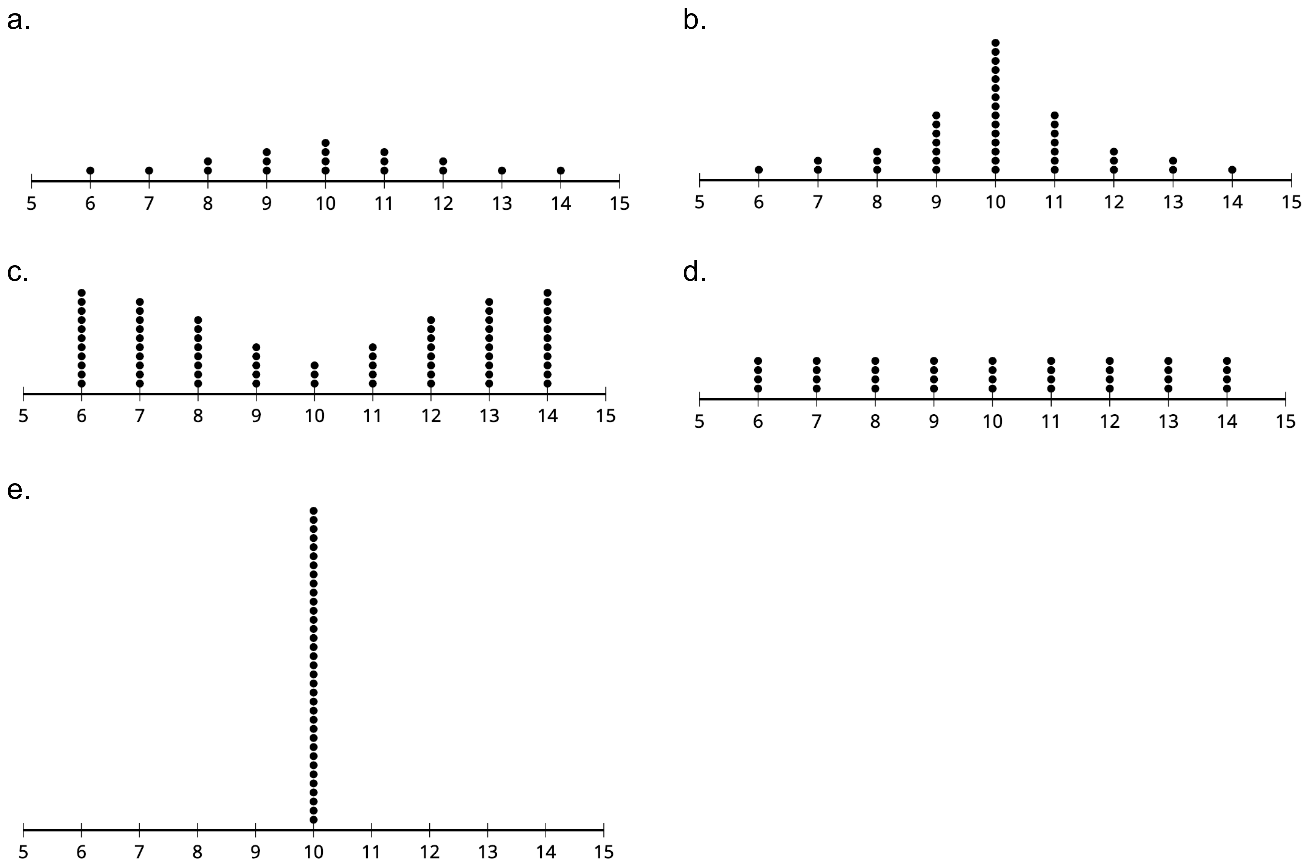


1. The five box plots have the same median. Explain why the median is more appropriate for describing the center of the data set than the mean for these distributions.

2. Arrange the box plots in order of least variability to greatest variability. Check with another group to see if they agree.

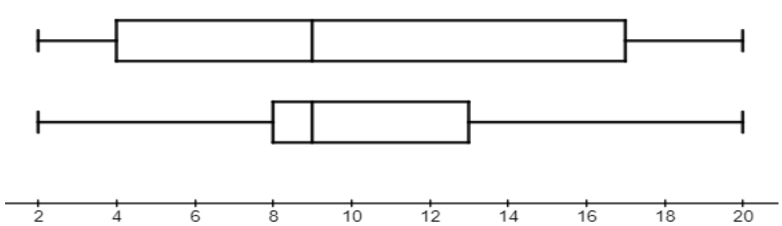
3. The five dot plots have the same mean. Explain why the mean is more appropriate for describing the center of the data set than the median.

4. Arrange the dot plots in order of least variability to greatest variability. Check with another group to see if they agree.

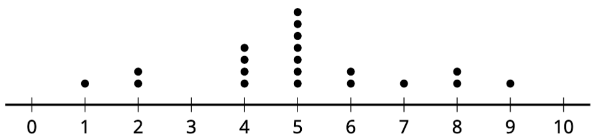


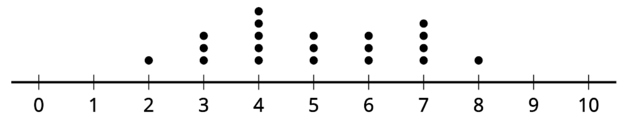
Are You Ready For More?

1. These two box plots have the same median and the same maximum and minimum values. How could we compare the variability of the two distributions?



2. These two dot plots have the same mean and the same standard deviation. How could we compare the variability of the two distributions?

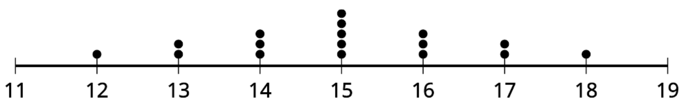


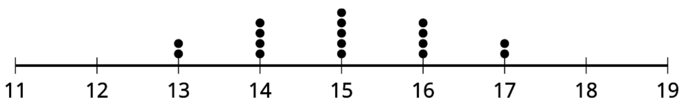


Lesson Debrief

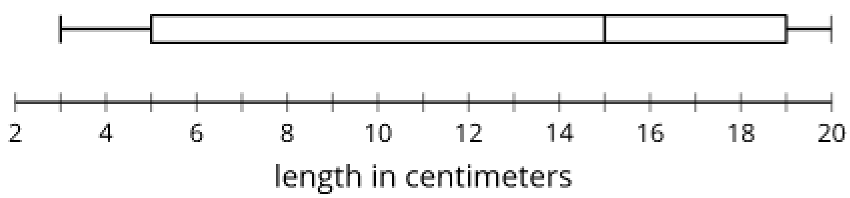
Lesson 13 Summary and Glossary

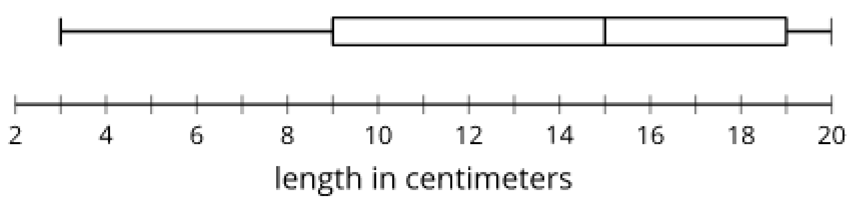
Here are two dot plots, each with a mean of 15 centimeters, displaying the length of sea scallop shells in centimeters.





Notice that both dot plots show a symmetric distribution, so the mean and the standard deviation are appropriate choices for describing center and variability. The data in the first dot plot appear to be more spread apart than the data in the second dot plot, so you can say that the first data set appears to have greater variability than the second data set. This is confirmed by the standard deviation. The standard deviation of the first data set is 1.53cm and the standard deviation of the second data set is approximately 1.19cm. The greater the standard deviation of the data, the greater the variability of the data.
These two box plots represent the distributions of the lengths in centimeters of a different group of sea scallop shells, each with a median of 15 centimeters.





Notice that neither of the box plots have a symmetric distribution. The median and the IQR are appropriate choices for describing center and variability for these data sets. The data points in the middle half of the first box plot appear to be more spread apart, or show greater variability, than data points in the middle half of the second box plot. The IQR of the first distribution is 14 cm and 10 cm for the second data set. The IQR measures the difference between the median of the second half of the data, Q3, and the median of the first half, Q1, of the data, so it is not impacted by the minimum or the maximum value in the data set. It is a measure of the spread of the middle 50% of the data.

The standard deviation is calculated using every value in the data set, while the IQR is calculated using only the values for Q1 and Q3.

Unit 1 Lesson 13 Practice Problems

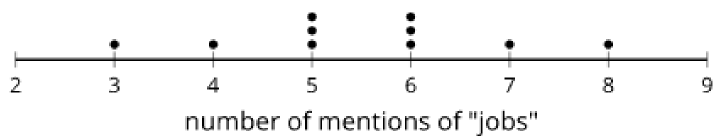
1. In science class, Clare and Lin estimate the mass of eight different objects that actually weigh 2,000 grams each. Some summary statistics for each set of eight measurements:

Clare
- mean: 2,000 grams
- standard deviation: 275 grams
- median: 2,000 grams
- IQR: 500 grams

Lin
- mean: 2,000 grams
- standard deviation: 225 grams
- median: 1,950 grams
- IQR: 350 grams

Which student was better at estimating the mass of the objects? Explain your reasoning.

2. A reporter counts the number of times a politician talks about jobs in their campaign speeches. What is the standard deviation of the data represented in the dot plot?



   a. 1.4 mentions

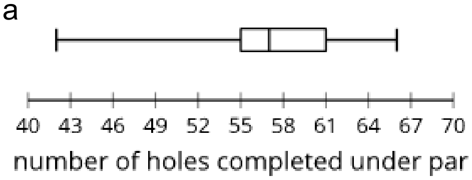
   b. 2.1 mentions

   c. 2.5 mentions

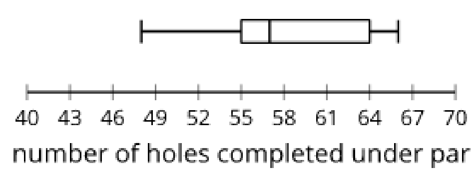
   d. 5.5 mentions

3. Four amateur miniature golfers attempt to finish 100 holes of golf under par several times. “Under par” means it took them fewer strokes than what was expected for that hole. For each round of 100, the number of holes they successfully complete under par is recorded. Due to the presence of extreme values, box plots were determined to be the best representation for the data. List the four box plots in order of variability from least to greatest.

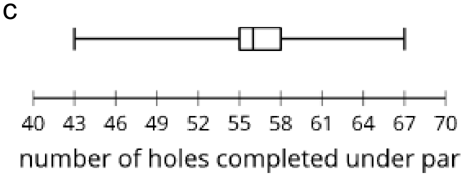
player a



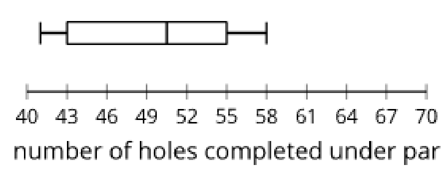
player b



player c



player d



4. Select **all** the distribution shapes for which the median *could be* much less than the mean.

    a. symmetric

    b. bell-shaped

    c. skewed left

    d. skewed right

    e. bimodal

(From Unit 1, Lesson 11)

5.

    a. What is the five-number summary for the data 0, 2, 2, 4, 5, 5, 5, 5, 7, 11?

    b. When the minimum, 0, is removed from the data set, what is the five-number summary?

(From Unit 1, Lesson 6)

6. What effect does eliminating the highest value, 180, from this data set have on the mean and median?

    25, 50, 50, 60, 70, 85, 85, 90, 90, 180

(From Unit 1, Lesson 8)

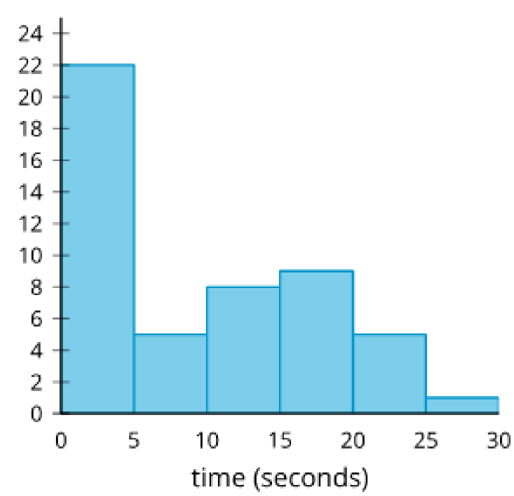
7. The histogram represents the distribution of the number of seconds it took for each of 50 students to find the answer to a trivia question using the internet. Which interval contains the median?

    a. 0 to 5 seconds

    b. 5 to 10 seconds

    c. 10 to 15 seconds

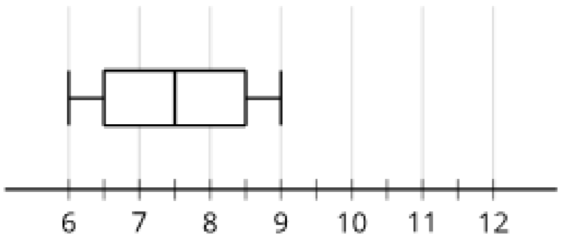
    d. 15 to 20 seconds



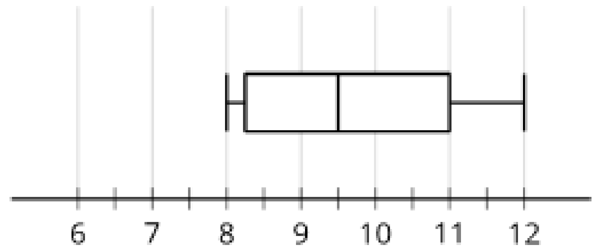
(From Unit 1, Lesson 6)

8. Here are the shoe sizes from two cohorts in the military.

Cohort A



Cohort B



a. Which cohort has more variability? Explain your reasoning.

b. Does at least one person from cohort A have a bigger shoe size than someone from cohort B? Explain your reasoning.

c. Compare the measures of center.

(Addressing NC.6.SP.3b)

Lesson 14: Interpreting Data Sets

Learning Targets

- I can compare two data sets using the measures of center and variability.
- I can tell how an outlier will impact mean, median, IQR, or standard deviation.

Bridge

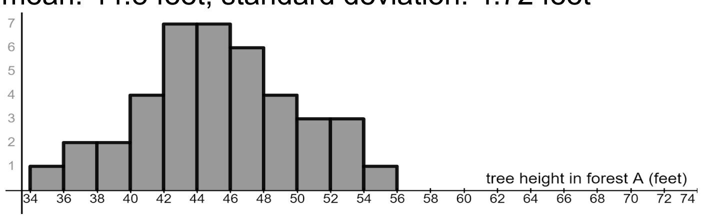
A health inspector for the county is trying to decide which restaurant to go to for their final inspection of the week. Here are each restaurant's health inspection ratings over the past year. Based on the restaurants' ratings, which restaurant should the inspector go to? Use the median and interquartile range to support your choice.

- Restaurant A: 88, 87, 89, 90, 87, 85, 88, 91, 86, 86, 88, 89
- Restaurant B: 90, 65, 89, 50, 94, 93, 95, 95, 75, 70, 88, 89

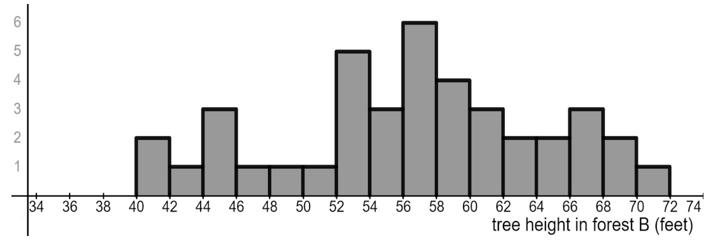
Activity 1: Interpreting Differences Between Distributions

1. The heights of the 40 trees in each of two forests are collected.

Forest A
mean: 44.8 feet, standard deviation: 4.72 feet



Forest B
mean: 56.03 feet, standard deviation: 7.87 feet



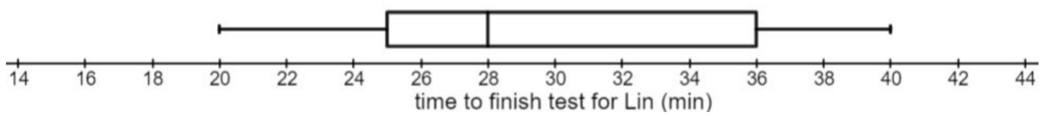
   a. Interpret the measure of center in terms of the situation.

   b. Interpret the measure of variability in terms of the situation.

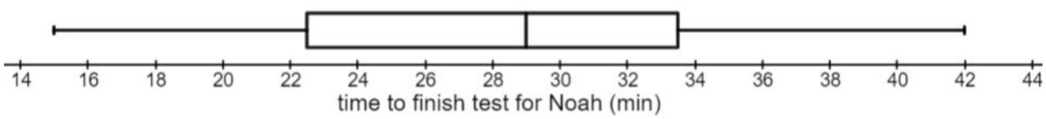
   c. Use these two measures to compare the two data sets.

2. The number of minutes it takes Lin and Noah to finish their tests in German class is collected for the year.

Lin median: 28 minutes, IQR: 11 minutes



Noah median: 29 minutes, IQR: 11 minutes



a. Interpret the measure of center in terms of the situation.

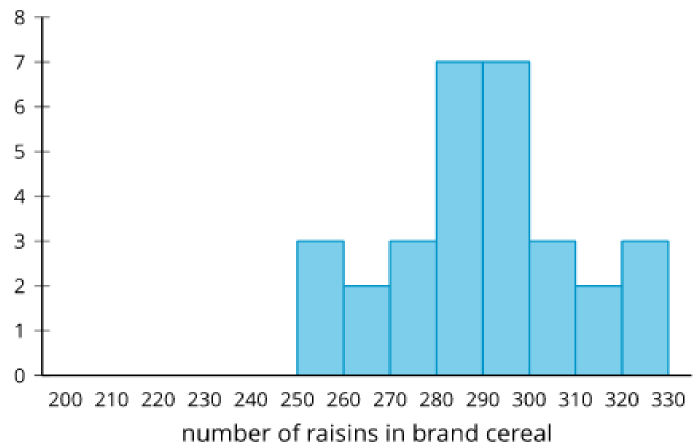
b. Interpret the measure of variability in terms of the situation.

c. Compare the two data sets.

3. The number of raisins in a cereal with a name brand and the generic version of the same cereal are collected for several boxes.

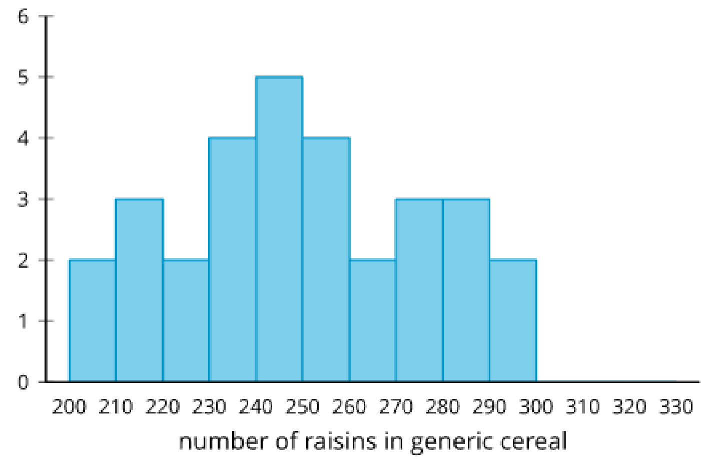
Brand cereal

mean: 289.1 raisins, standard deviation: 19.8 raisins



Generic cereal

mean: 249.2 raisins, standard deviation: 26.4 raisins



a. Interpret the measure of center in terms of the situation.

b. Interpret the measure of variability in terms of the situation.

c. Compare the two data sets.

Are You Ready For More?

Standard deviation gives a natural scale as to how far above or below the mean of a data point is. This is incredibly useful for comparing points from two different distributions.

For example, they say you cannot compare apples and oranges, but here is a way. The average weight of a granny smith apple is 128 grams with a standard deviation of about 10 grams. The average weight of a navel orange is 140 grams with a standard deviation of about 14 grams. If we have a 148-gram granny smith apple and a 161-gram navel orange, we might wonder which is larger for its species even though they are both about 20 grams above their respective mean. We could say that the apple, which is 2 standard deviations above its mean, is larger for its species than the orange, which is only 1.5 standard deviations above its mean.

1. The tallest tree in forest A is 55 feet. How many standard deviations above the mean height is it?

2. The tallest tree in forest B is 70 feet. How many standard deviations above the mean height is it?

3. Which tree is taller in its forest?

Activity 2: Investigating Outliers

This data set is the average amount of money, in thousands of dollars, per year, spent on each person in the country (per capita spending) for health care in 43 countries in 2019. This data can also be found here: http://bit.ly/L14dataset. Use this data for questions 1–3.

1.154	1.213	1.337	1.907	1.973	1.996	2.008	2.159
2.222	2.230	2.354	2.384	2.579	2.638	2.790	2.932
3.224	3.379	3.384	3.428	3.616	3.649	4.204	4.224
4.262	4.578	4.653	4.811	4.823	5.187	5.276	5.376
5.418	5.428	5.558	5.568	5.765	5.782	5.851	6.646
6.647	7.732	11.072					

Source: OECD (2021), Health spending (indicator)

1. One value in the set is an outlier. Which one is it? How do you know?

2. Use technology to create a histogram and a box plot of the data.
   a. How does the outlier appear in the histogram?

   b. How does the outlier appear in the box plot?

3. Use technology to find the mean, standard deviation, median, and IQR.

The maximum value in this data set represents the spending for the United States. Although outliers should not be removed without considering their cause, it is important to see how influential outliers can be for various statistics.

4. Remove the value for the United States from the data set.
   a. Use technology to calculate the new mean, standard deviation, median, and IQR.

   b. How do the mean, standard deviation, median, and IQR of the data set with the outlier removed compare to those of the original data set?

5. A statistician is curious why the United States is an outlier in relation to these other countries. What are some questions he could research to learn more?

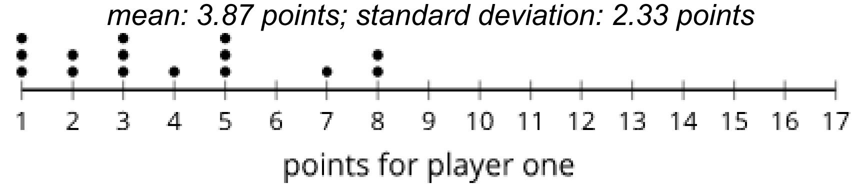
Lesson Debrief

Lesson 14 Summary and Glossary

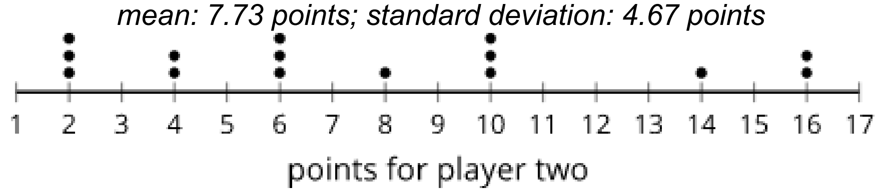
When comparing data sets it is useful to interpret the differences in the measure of center and measure of variability. Both the measure of center and the measure of variability have the same unit as the data set it describes. In the example below, the units are points.

The first dot plot shows the number of points that player one on a basketball team made during each of 15 games. The second dot plot shows the number of points scored by player two during the same 15 games.

mean: 3.87 points; standard deviation: 2.33 points

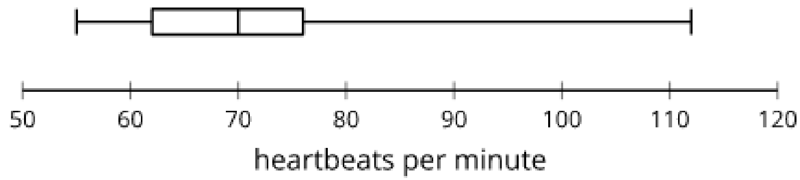


mean: 7.73 points; standard deviation: 4.67 points



On average, player two scored 7.73 points in a game, about four more points than player one's average. Player one is more consistent in scoring her typical number of points. Player two has greater variability in points than player one because the data is more spread out. This is shown in the standard deviation for the second distribution being greater than the standard deviation for the first distribution.

Outlier: A data value that is unusual in that it differs quite a bit from the other values in the data set.

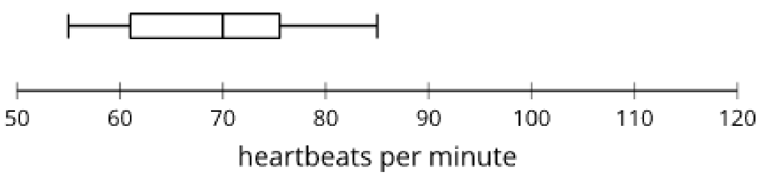


The box plot displays the resting heart rate, in beats per minute (bpm), of 50 athletes taken five minutes after a workout.

mean: 69.78 bpm
standard deviation: 10.71 bpm
minimum: 55 bpm
Q1: 62 bpm
median: 70 bpm
Q3: 76 bpm
maximum: 112 bpm

Suppose someone tells us that the maximum value of 112 bpm is an outlier. Once the outlier is removed, the box plot and summary statistics are:

mean: 68.92 bpm
standard deviation: 8.9 bpm
minimum: 55 bpm
Q1: 61 bpm
median: 70 bpm
Q3: 75.5 bpm
maximum: 85 bpm



The mean decreased by 0.86 bpm, and the median remained the same. The standard deviation decreased by 1.81 bpm. Based on the standard deviation, the data set with the outlier removed shows much less variability than the original data set containing the outlier. Since the mean and standard deviation use all of the numerical values, removing one very large data point can affect these statistics in important ways.

The median remained the same after the removal of the outlier and the IQR increased slightly. These measures of center and variability are much more resistant to change than the median and standard deviation. The median and IQR measure the middle of the data based on the number of values rather than the actual numerical values themselves, so the loss of a single value will not often have a great effect on these statistics.

Unit 1 Lesson 14 Practice Problems

1. Three drivers competed in the same fifteen drag races. The mean and standard deviation for the race times of each of the drivers are given.

   - Driver A had a mean race time of 4.01 seconds and a standard deviation of 0.05 seconds.
   - Driver B had a mean race time of 3.96 seconds and a standard deviation of 0.12 seconds.
   - Driver C had a mean race time of 3.99 seconds and a standard deviation of 0.19 seconds.

   a. Which driver had the fastest typical race time?

   b. Which driver's race times were the most variable?

   c. Which driver do you predict will win the next drag race? Support your prediction using the mean and standard deviation.

2. The number of letters received in the mail over the past week is recorded.

2	3	5	5	5	15

   a. Which value appears to be an outlier?

   b. Explain what would happen to the mean and standard deviation if the 15 were removed from the data set.

3. A group of pennies made in 2018 are weighed. The mean is approximately 2.5 grams with a standard deviation of 0.02 grams.

   Interpret the mean and standard deviation in terms of the context.

4. The height of everyone in a fourth grade classroom, including the teacher, was collected. They had a mean height of 50 inches and a standard deviation of 4.4 inches.

   If the height of the teacher is removed, what effect might that have on the mean and standard deviation? Explain your reasoning.

5. The mean exam score for the first group of twenty examinees applying for a security job is 35.3 with a standard deviation of 3.6. The mean exam score for the second group of twenty examinees is 34.1 with a standard deviation of 0.5. Both distributions are close to symmetric in shape.

   a. Use the mean and standard deviation to compare the scores of the two groups.

   b. The minimum score required to get an in-person interview is 33. Which group do you think has more people get in-person interviews?

6. Andre records how long it takes him (in minutes) to hike a mountain each day for six days.

50	52	58	55	59	50

a. Use technology to calculate the mean number of minutes it takes Andre to hike a mountain.

b. What do you think will happen to the mean time for the week if Andre decides to take pictures of the trees, waterfalls, and wildlife throughout the hike for the seventh day?

c. Calculate the mean number of minutes including the time it took Andre on the seventh day: 50, 52, 58, 55, 59, 50, 130.

d. If Andre didn't stop to take pictures on the seventh day, he thinks he could have finished the trail in 60 minutes. Calculate the mean hiking time using Andre's estimate for the seventh day: 50, 52, 58, 55, 59, 50, 60.

7. The number of different species of plants in some gardens is recorded.

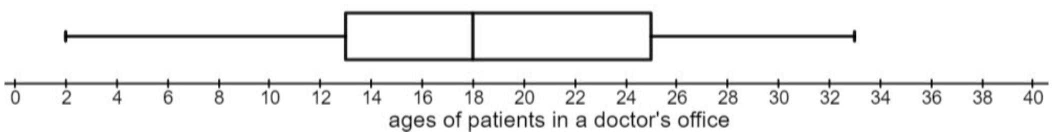
1	2	3	4	4	5	5	6	7	8

a. What is the mean?

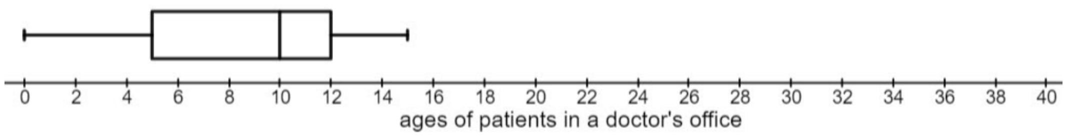
b. What is the standard deviation?

(From Unit 1, Lesson 13)

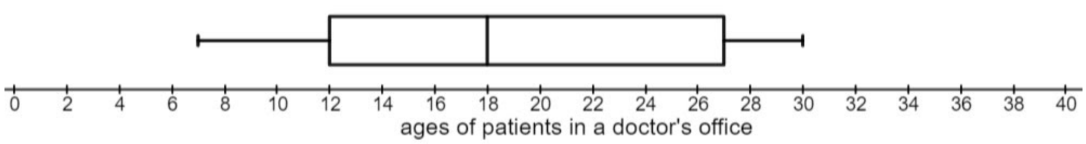
8. List the data distributions from the smallest IQR to the largest IQR and write a comparison between the smallest and largest.



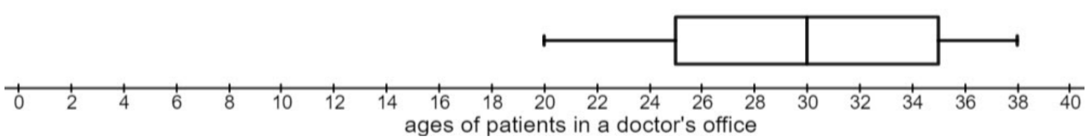
a.



b.



c.



d.

(From Unit 1, Lesson 13)

9. At the end of last year, teachers were rated by their students on a 0–10 scale. Two of the teachers' ratings are given. Whose class would you register for? Explain your reasoning.

   a. Teacher A: 9, 8, 10, 10, 7, 1, 8, 1, 2, 8

   b. Teacher B: 9, 8, 8, 7, 9, 7, 7, 9, 7, 8

(Addressing NC.7.SP.3.a)

Lesson 15: Comparing Data Sets

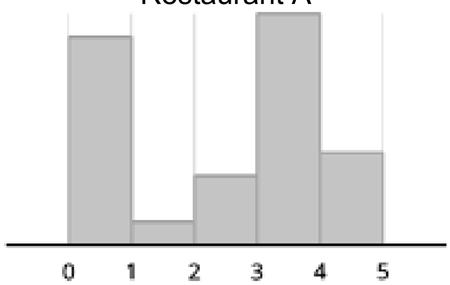
Learning Target

- I can compare and contrast situations using measures of center and measures of variability.

Bridge

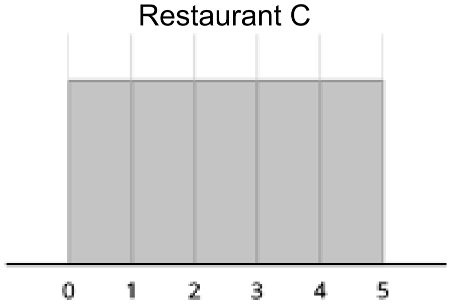
Each histogram represents the number of star ratings for a different restaurant. The ratings range from 0–4, with 0 representing a very poor experience and 4 representing an excellent experience.

Restaurant A

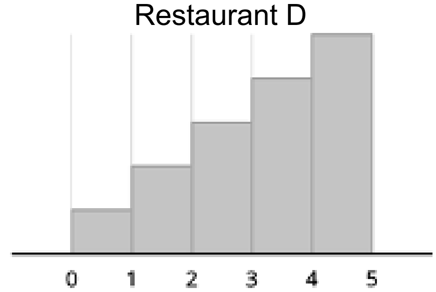


Restaurant B

Restaurant C



Restaurant D

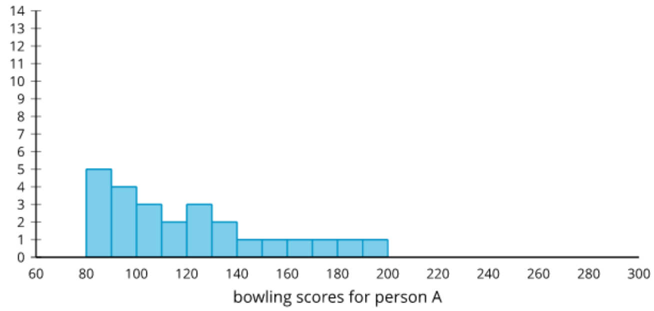


Questions	Explain your reasoning
1. Which restaurant do reviewers like the most?	
2. Which restaurant do reviewers like the least?	
3. Which restaurant received mostly mixed reviews?	
4. Which restaurant would you choose to try?	

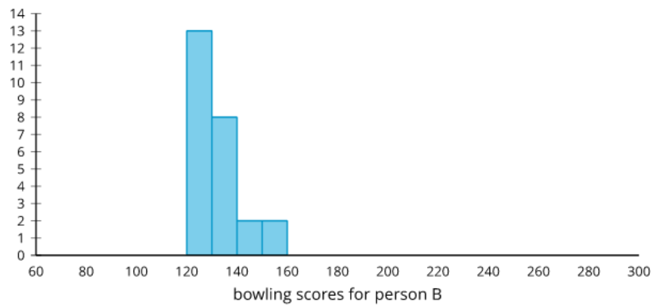
Warm-up: Bowling Partners

Each histogram shows the bowling scores for the last 25 games played by each person. Choose two of these people to join your bowling team. Explain your reasoning.

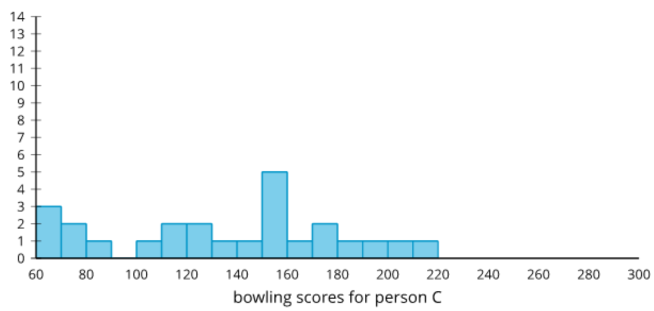
Person A
- mean: 118.96
- median: 111
- standard deviation: 32.96
- interquartile range: 44



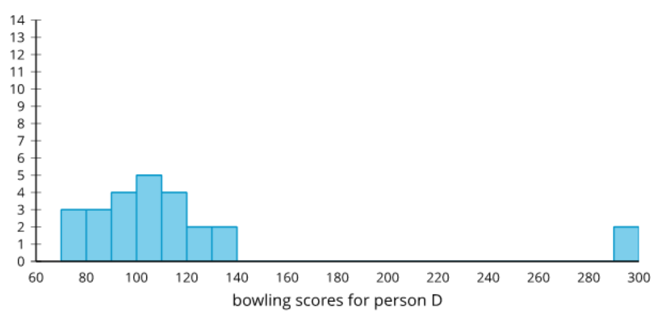
Person B
- mean: 131.08
- median: 129
- standard deviation: 8.64
- interquartile range: 8



Person C
- mean: 133.92
- median: 145
- standard deviation: 45.04
- interquartile range: 74



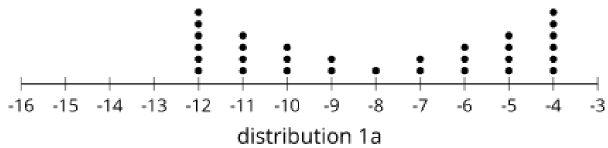
Person D
- mean: 116.56
- median: 103
- standard deviation: 56.22
- interquartile range: 31.5



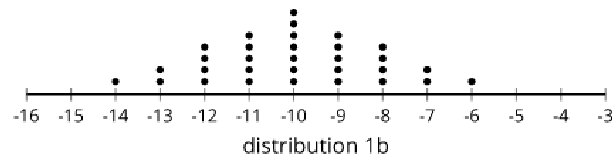
Activity 1: Comparing Measures

For each group of data sets:

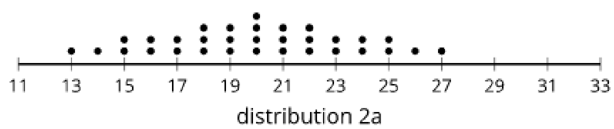
- Determine the best measure of center and measure of variability to use based on the shape of the distribution.
- Determine which set has the greatest measure of center.
- Determine which set has the greatest measure of variability.
- Be prepared to explain your reasoning.



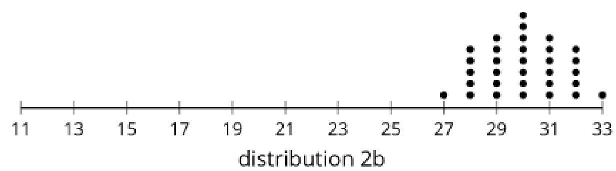
1a



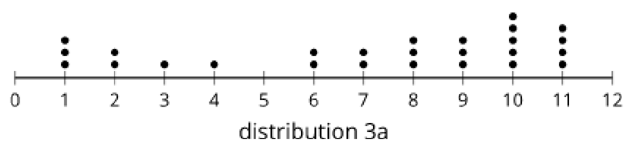
1b



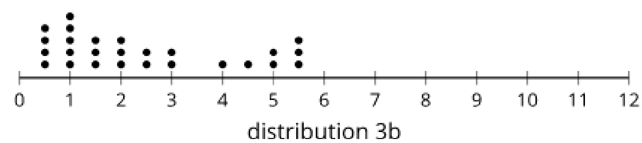
2a



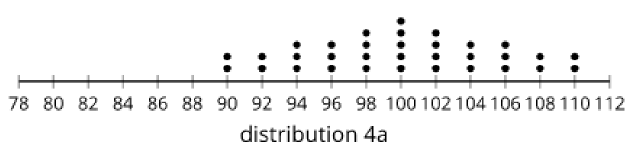
2b



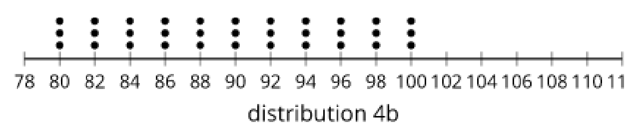
3a



3b



4a

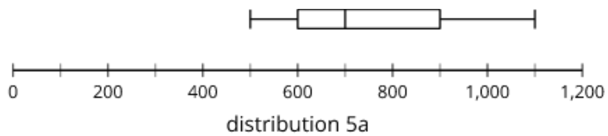


4b

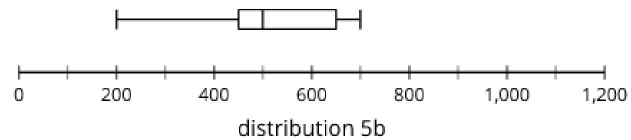
(continued) For each group of data sets:

- Determine the best measure of center and measure of variability to use based on the shape of the distribution.
- Determine which set has the greatest measure of center.
- Determine which set has the greatest measure of variability.
- Be prepared to explain your reasoning.

5a



5b



6a

A political podcast has mostly reviews that either love the podcast or hate it.

6b

A cooking podcast has reviews that neither hate nor love the podcast.

7a

Stress testing concrete from site A has all 12 samples break at 450 pounds per square inch (psi).

7b

Stress testing concrete from site B has samples break every 10 psi starting at 450 psi until the last core is broken at 560 psi.

7c

Stress testing concrete from site C has 6 samples break at 430 psi and the other 6 break at 460 psi.

Activity 2: Comparing Marathon Times

1. Do some research on marathons to answer the following questions:
   a. How long is a marathon?

   b. Name a place that is approximately that same distance from where you live or go to school.

   c. How long does it typically take someone to finish a marathon?

      i. What factors play into how long it takes someone?

      ii. Name some assumptions you made to determine that “typical” time.

2. Here are the marathon finishing times for two different age groups, in minutes. What are some possible mathematical questions that could be asked about these data sets?
   - Ages 30–39 {232, 238, 240, 243, 243, 245, 251, 252, 258, 259, 260, 262, 263, 265, 265, 265, 270, 278, 279, 283, 284, 291, 291, 293, 293, 300, 305, 310, 312, 319, 323, 324, 376, 377, 379}

   - Ages 40–49 {250, 260, 262, 262, 263, 265, 271, 272,272, 283, 284, 298, 298, 299, 299, 300, 302, 304, 310, 319, 320, 321, 322, 324, 338, 338, 342, 343, 348, 348, 362, 363, 379, 380, 382, 385, 442, 448, 450}

3. Use technology to create a data display for each age group of runners and then describe the distribution for each.

4. Which age group tends to take longer to run the marathon? Explain your reasoning.

5. Which age group has more variability in its finishing times? Explain your reasoning.

Are You Ready For More?

1. How do you think finish times for a 20–29 age range will compare to these two distributions?

2. Find some actual marathon times for this group and create a data display using technology to help compare.

Lesson Debrief

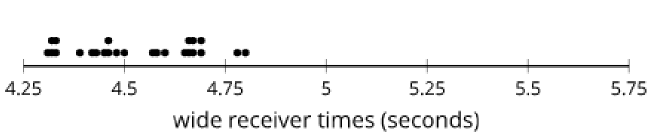
Lesson 15 Summary and Glossary

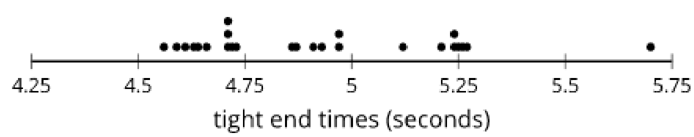
To compare data sets, it is helpful to look at the measures of center and measures of variability. The shape of the distribution can help choose the most useful measure of center and measure of variability.

When distributions are symmetric or approximately symmetric, the mean is the preferred measure of center and should be paired with the standard deviation as the preferred measure of variability. When distributions are skewed or when outliers are present, the median is usually a better measure of center and should be paired with the interquartile range (IQR) as the preferred measure of variability.

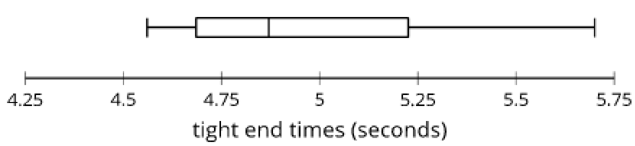
Once the appropriate measure of center and measure of variability are selected, these measures can be compared for data sets with similar shapes.

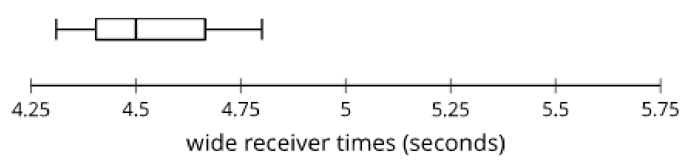
For example, let's compare the number of seconds it takes football players to complete a 40-yard dash at two different positions. First, we can look at a dot plot of the data to see that the tight end times do not seem symmetric, so we should probably find the median and IQR for both sets of data to compare information.





The median and IQR could be computed from the values, but can also be determined from a box plot.



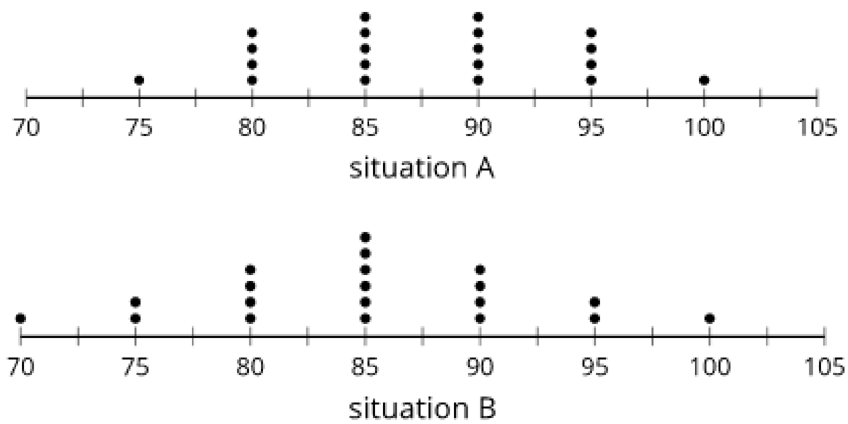


This shows that the tight end times have a greater median (about 4.9 seconds) compared to the median of wide receiver times (about 4.5 seconds). The IQR is also greater for the tight end times (about 0.5 seconds) compared to the IQR for the wide receiver times (about 0.25 seconds).

This means that the tight ends tend to be slower in the 40-yard dash when compared to the wide receivers. The tight ends also have greater variability in their times. Together, this can be taken to mean that, in general, a typical wide receiver is faster than a typical tight end, and the wide receivers tend to have more similar times to one another than the tight ends do to one another.

Unit 1 Lesson 15 Practice Problems

1. Twenty students participated in a psychology experiment that measured their heart rates in two different situations.



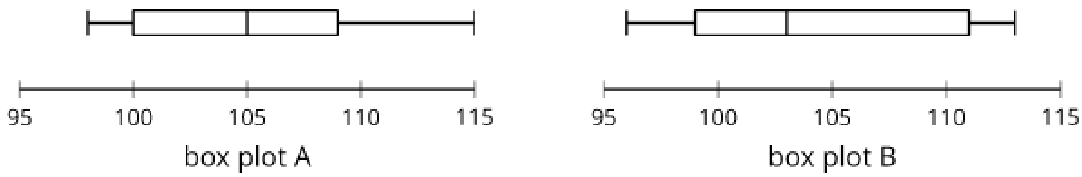
   a. What are the appropriate measures of center and variability to use with the data? Explain your reasoning.

   b. Which situation shows a greater typical heart rate?

2.
   a. Invent two situations that you think would result in distributions with similar measures of variability. Explain your reasoning.

   b. Invent two situations that you think would result in distributions with different measures of variability. Explain your reasoning.

3. Here are box plots for two data sets:



   a. Which data set has a greater median?

   b. Which data set has a greater IQR?

4. The data set and some summary statistics are listed.

   11.5, 12.3, 13.5, 15.6, 16.7, 17.2, 18.4, 19, 19.5, 21.5
   mean: 16.52
   median: 16.95
   standard deviation: 3.11
   IQR: 5.5

   a. How does adding 5 to each of the values in the data set impact the shape of the distribution?

   b. How does adding 5 to each of the values in the data set impact the measures of center?

   c. How does adding 5 to each of the values in the data set impact the measures of variability?

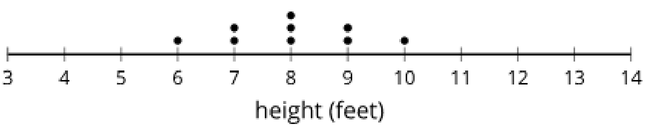
(From Unit 1, Lesson 6)

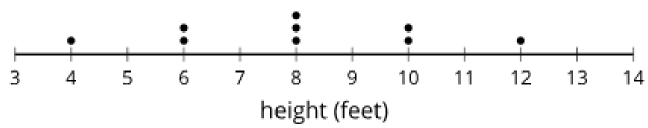
5. The depth of two lakes is measured at multiple spots. For the first lake, the mean depth is about 45 feet with a standard deviation of 8 feet. For the second lake, the mean depth is about 60 feet with a standard deviation of 27 feet.

   Noah says the second lake is generally deeper than the first lake. Do you agree with Noah?

   (From Unit 1, Lesson 14)

6. The dot plots display the height, rounded to the nearest foot, of maple trees from two different tree farms.





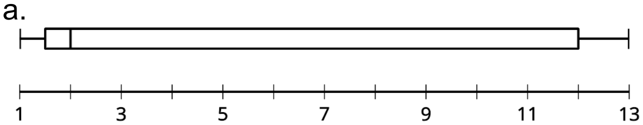
   a. Compare the mean and standard deviation of the two data sets.

   b. What does the standard deviation tell you about the trees at these farms?

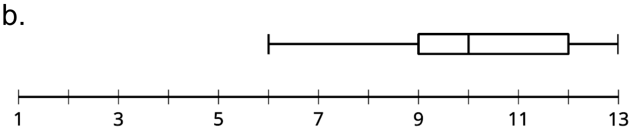
   (From Unit 1, Lesson 13)

7. Which box plot has an IQR of 10?

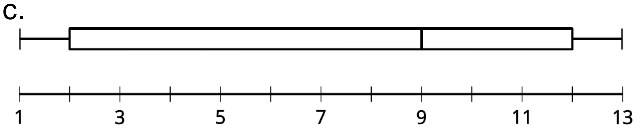
a.



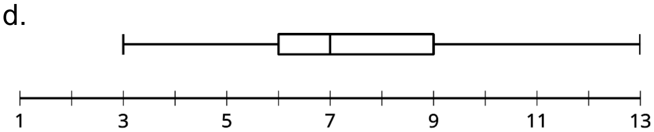
b.



c.



d.



(From Unit 1, Lesson 12)

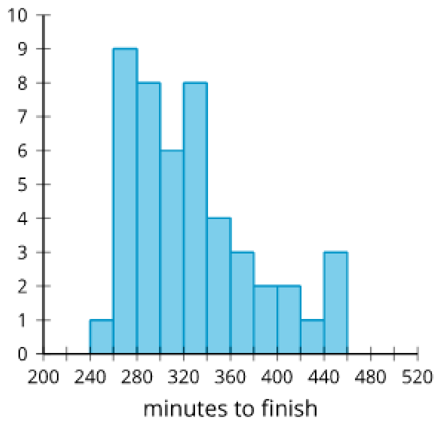
8. What effect does eliminating the lowest value, -6, from the data set have on the mean and median?

   -6, 3, 3, 3, 3, 5, 6, 6, 8, 10

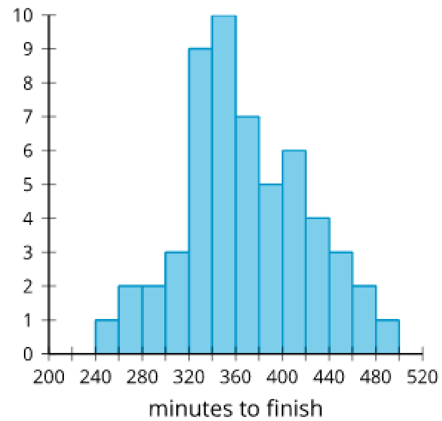
   (From Unit 1, Lesson 8)

9. These distributions represent marathon times for different groups.

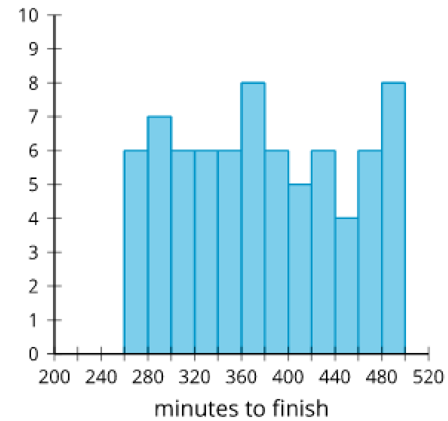
Group 1



Group 2



Group 3



   a. Which display is most likely to represent the marathon times for people aged 20–30? Explain your reasoning.

   b. Which display is most likely to represent the marathon times for every tenth person to cross the finish line? Explain your reasoning.

   c. Which display is most likely to represent the marathon times for people aged 40–50? Explain your reasoning.

   (Addressing NC.6.SP.2)