Quartiles and Interquartile Range

Goals

- Calculate the range and interquartile range (IQR) of a data set and interpret (orally and in writing) what they tell us about the situation.
- Comprehend that "interquartile range (IQR)" is another measure of variability that describes the span of the middle half of the data.
- Identify and interpret (in writing) the numbers in the five-number summary for a data set, that is, the minimum, first quartile (Q1), median (Q2), third quartile (Q3), and maximum.

Learning Targets

- I can use the IQR to describe the spread of data.
- I know what quartiles and the interquartile range (IQR) measure and what they tell us about the data.
- When given a list of data values or a dot plot, I can find the quartiles and interquartile range (IQR) for the data.

Lesson Narrative

In this lesson, students learn that they can further decompose a data set into 4 equal groups and use the **quartiles** to describe a distribution. They learn that the three quartiles along with the maximum and minimum values of the data set make up a five-number summary.

Students also explore the **range** and **interquartile range** (**IQR**) of a distribution as two ways to measure its spread. Students reason abstractly and quantitatively as they find and interpret the IQR as describing the distribution of the middle half of the data.

Student Learning Goal

Let's look at other measures for describing distributions.

Lesson Timeline

5 min Warm-up

min
Activity 1

15

15 min

Activity 2

10 min

Lesson Synthesis

Access for Students with Diverse Abilities

• Engagement (Activity 1)

Access for Multilingual Learners

 MLR8: Discussion Supports (Activity 2)

Instructional Routines

• Notice and Wonder

Assessment

5_{min}

Cool-down

Inspire Math

Sami Reindeer Herders video



Go Online

Before the lesson, show this video to review the real-world connection.

ilclass.com/l/614246

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



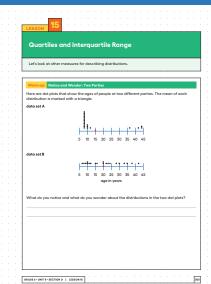
Instructional Routines

Notice and Wonder ilclass.com/r/10694948

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



Student Workbook



Warm-up

Notice and Wonder: Two Parties



Activity Narrative

The purpose of this *Warm-up* is to elicit the idea that there are multiple ways to describe variability, which will be useful when students learn about range and interquartile range in a later activity. While students may notice and wonder many things about these dot plots, variability and how to measure it are the important discussion points.

When students articulate what they notice and wonder, they have an opportunity to attend to precision in the language that they use to describe what they see. They might first propose less formal or imprecise language, and then restate their observation with more precise language in order to communicate more clearly.

Launch

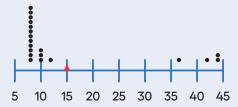


Arrange students in groups of 2. Display the dot plots for all to see. Ask students to think of at least one thing that they notice and at least one thing they wonder. Give students 1 minute of quiet think time, and then 1 minute to discuss with their partner the things that they notice and wonder about.

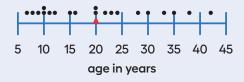
Student Task Statement

Here are dot plots that show the ages of people at two different parties. The mean of each distribution is marked with a triangle.

data set A



data set B



What do you notice and what do you wonder about the distributions in the two dot plots?

Students may notice:

- The mean of the two data sets are different. The mean for the second data set is 5 years higher than that for the first.
- The range in values of the two data sets are about the same.
- Most points in the first data set are clustered around 8 and 10 with only a few that are much higher.
- The mean for the first data set is located where there are no points.
- The points in the second data set are not clustered anywhere. They are distributed along the dot plot, between 5 and 42 years.
- The MAD values are close for the two data sets.

Students may wonder:

- · Why the data distributions look so different.
- · How the MAD values are the same or close.
- If there could be other distributions that look very different from these two but also have the same MAD.

Activity Synthesis

Ask students to share the things they noticed and wondered. Record and display their responses for all to see, without editing or commentary. If possible, record the relevant reasoning on or near the dot plots. Next, ask students,

"Is there anything on this list that you are wondering about now?"

Encourage students to respectfully disagree, ask for clarification, or point out contradicting information.

If the idea that the MAD does not describe the variability of these two sets well does not come up during the conversation, ask students to discuss that idea.

Two key ideas to uncover here are:

- The MAD is a way to summarize variation from the mean, but the single number does not always tell us how the data are distributed.
- The same MAD could result from very different distributions.

Activity 1

The Five-Number Summary



Activity Narrative

This activity introduces students to the *five-number summary* and the process of identifying the five numbers. Students learn how to partition the data into four sets: using the median to decompose the data into upper and lower halves, and then finding the middle of each half to further decompose it into quarters. They learn that each value that decomposes the data into four parts is called a **quartile**, and the three quartiles are the first quartile (Q1), second quartile (Q2, or the median), and third quartile (Q3). Together with the minimum and maximum values of the data set, the quartiles provide a five-number summary that can be used to describe a data set without listing or showing each data value.

Access for Students with Diverse Abilities (Activity 1, Student Task)

Engagement: Develop Effort and Persistence.

Chunk this task into more manageable parts. Have students complete the first question, then the next 2, and finally the last one. Check in with students to provide feedback and encouragement after each chunk. Ensure that students understand how the data are being divided.

Supports accessibility for: Attention, Social-Emotional Functioning

Student Workbook



Students reason abstractly and quantitatively as they identify and interpret the guartiles in the context of the situation given.

Launch



Arrange students in groups of 2. Give groups 8–10 minutes to complete the activity. Follow with a whole-class discussion.

Ask students if they have attended different kinds of parties with different age ranges. If a few different possibilities don't come up, share some examples:

- A birthday party for a sixth grader may have most people there around the same age.
- A family get-together may have a wide range of ages fairly evenly distributed.
- A dinner party with mostly your parents' friends may have mostly older people.

Remind students that they previously summarized variability by finding the MAD, which involves calculating the distance of each data point from the mean and then finding the average of those distances. Explain that we will now explore another way to describe variability and summarize the distribution of data. Instead of measuring how far away data points are from the mean, we will decompose a data set into four equal parts and use the markers that partition the data into quarters to summarize the spread of data.

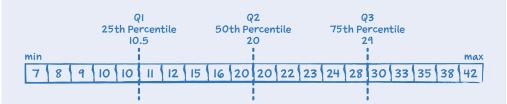
If necessary, remind students how to find the median, especially when there are an even number of values in the data set.

Student Task Statement

Here are the ages of the people at one party, listed from least to greatest.

7	8	9	10	10	11	12	15	16	20
20	22	23	24	28	30	33	35	38	42

- 1. Split the data into 4 equal parts using 3 values called quartiles.
 - **a.** Find the median of the data set and label it Q2 for second quartile. This splits the data into an upper half and a lower half.
 - **b.** Find the middle value of the lower half of the data, without including the median. Label this value Q1 for first quartile.
 - **c.** Find the middle value of the upper half of the data, without including the median. Label this value Q3 for third quartile.



2. Label the least value in the set "minimum" and the greatest value "maximum."

See diagram.

3. The values you have identified make up the *five-number summary* for the data set. Record them here.

minimum: 7 years Q1: 10.5 years Q2: 20 years Q3: 29 years maximum: 42 years

- **4.** The median of this data set is 20. This tells us that half of the people at the party were 20 years old or younger, and the other half were 20 or older. What do each of these other values tell us about the ages of the people at the party?
 - a. The third quartile

Q3 tells us that a quarter of the party goers are over the age of 29 years old, and the rest are younger.

b. The minimum

The minimum tells us that the youngest person at the party is 7 years old.

c. The maximum

The maximum tells us that the oldest person at the party is 42 years old.

Are You Ready for More?

There was another party where 21 people attended. Here is the five-number summary of their ages.

minimum: 5 Q1: 6 Q2: 27 Q3: 32 maximum: 60

1. Do you think this party had more children or fewer children than the earlier one? Explain your reasoning.

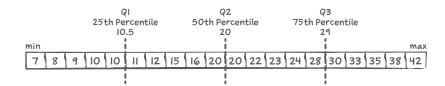
Sample response: There are about the same, or possibly more, kids. Because the first quartile (QI) is 6 years old, there are at least 6 and as many as IO children at this party.

2. Were there more children or adults at this party? Explain your reasoning.

Sample response: There are more adults at this party. The median age is 27 years old, an adult. Besides this adult, half of the other guests are adults aged 27 or older.

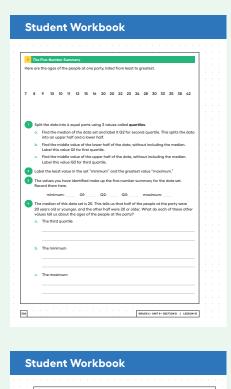
Activity Synthesis

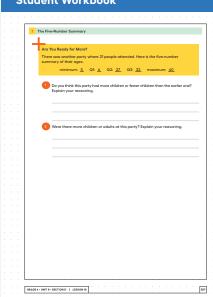
Ask a student to display the data set that they have decomposed and labeled, or display the diagram for all to see.



Focus the conversation on students' interpretation of the five numbers. As students discuss their solutions, color code or annotate the five-number summary on the data set and diagram. Discuss:

"In this context, what do the minimum and maximum values tell us?"
the ages of the youngest and oldest partygoers





"In this context, what does Q1 (10.5) tell us?"

That a quarter of the partygoers are 10.5 years old or younger

"What does Q3 (29) tell us?"

That three quarters of the partygoers are 29 years old or younger

(*) "How do the five numbers help us to see the distribution of the data?"

It divides the values in the data into sections containing one-fourth of the values each. This gives us an idea about the distribution of the data by looking at how varied each section is.

Activity 2

Range and Interquartile Range



Activity Narrative

In this lesson, students extend their understanding of variability to finding the range and interquartile range (IQR) of a data set. The range is the difference between the maximum and minimum values in the data, and the IQR is the difference between the third and first quartiles. While the range tells us how spread out (or close together) the overall data values are, the IQR tells us how spread out (or close together) the middle half of the data values are.

Students identify the range and IQR of a data set and analyze distributions with different IQRs. They reason abstractly and quantitatively as they use the IQR to describe the variability of data.

Launch



Tell students that they will write the five-number summary of a distribution shown on a dot plot. Give students a moment of quiet time to look at the dot plot in the first question and think about how they might identify the quartiles. Then, ask students to share their ideas. Students might suggest these strategies.

- List the values of all the data points, put them in order, and then count off the values to find the median and the other two quartiles.
- Count the points by 3's (because the data set is to be decomposed into 4 equal parts and 12 ÷ 4 = 3) and mark the end of the first set with Q1, the end of the second set with Q2, etc.
- Divide the points into two halves (by counting 6 points from the left or from the right), and then divide each half into two halves.

It is not necessary that all of these ideas are brought up at this point, but if no students mentioned the first approach (listing all values), mention it. The concrete process of writing out all the values, in order, is likely to be accessible to most students.

Arrange students in groups of 2. Give students 3–4 minutes of quiet work time for the first question, and 5–7 minutes to discuss their work with their partner and to complete the rest of the activity. Follow with a whole-class discussion.

Student Task Statement

1. Here is a dot plot that shows the lengths of Elena's bus rides to school, over 12 days.



Write the five-number summary for this data set. Show your reasoning.

The five-number summary is minimum: 6; QI: 7.5; Q2: 8.5; Q3: 9.5; maximum: I2. This is found by ordering the travel times in a list. The QI is the average of 7 and 8, the 3rd and 4th values. The Q2 is the average of 8 and 9, the 6th and 7th values. The Q3 is the average of 9 and I0, the 9th and I0th values.

2. The **range** is one way to describe the spread of values in a data set. It is the difference between the maximum and minimum. What is the range of Elena's travel times?

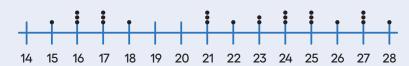
The range of Elena's data set is 6. The smallest data point is 6, and the largest data point is 12; the difference between these is 6.

- **3.** Another way to describe the spread of values in a data set is the **interquartile range (IQR)**. It is the difference between the third quartile (Q3) and the first quartile (Q1).
 - a. What is the interquartile range (IQR) of Elena's travel times?

The 1st and 3rd quartiles are 7.5 and 9.5. The difference between these values, or the IQR, is 2.

- b. What fraction of the data values are between Q1 and Q3?
 - $\frac{1}{2}$ of the data set is between the upper and lower quartiles.
- 4. Here are 2 more dot plots.

data set A



data set B



Without doing any calculations, predict:

a. Which data set has the smaller range?

Data set A has the smaller range. The minimum and maximum are closer together in Data set A than in Data set B.

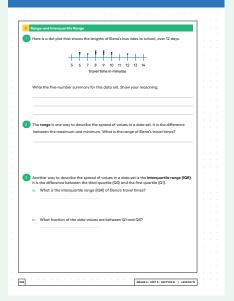
b. Which data set has the smaller IQR?

Data set B has the smaller IQR. The middle part of the distribution in Data set B is closer together than in Data set A, so the distance between QI and Q3 should be smaller.

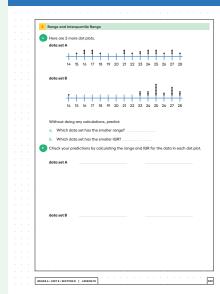
Building on Student Thinking

When finding the IQR of the dot plots in the last question, students might neglect to divide the data set into four parts. Or they might instead divide the distance between the maximum and minimum into four parts (rather than dividing the data points into four parts). Remind students about the conversation at the start of the task about listing all the values or counting off the data points in order to find the quartiles.

Student Workbook



Student Workbook



Access for Multilingual Learners (Activity 2, Synthesis)

MLR8: Discussion Supports.

During group work, invite students to take turns sharing their responses. Ask students to restate what they heard using precise mathematical language and their own words. Display the sentence frame, "I heard you say ..." Original speakers can agree or clarify for their partner.

Advances: Listening, Speaking

5. Check your predictions by calculating the range and IQR for the data in each dot plot.

The range for Data set A is 13 because the maximum value is 28 and the minimum value is 15. The range for Data set B is 14 because the maximum value is 28 and the minimum value is 14. The IQR of Data set A is 8 because the QI is 17, and the Q3 is 25. The IQR of Data set B is 3 because the QI is 24, and the Q3 is 27.

Activity Synthesis

Ensure that students know how to find the range and IQR, and then focus the discussion on interpreting these two measures and how they provide information about a distribution.

Select a couple of students to share the range and IQR of Elena's data. Ask:

- "What does a range of 6 minutes tell us about Elena's travel times?"
 Elena's travel times vary by 6 minutes at most, or that the difference between the shortest commute and the longest one is 6 minutes.
- "What does an IQR of 2 minutes tell us about her travel time?"
 The middle half of Elena's travel times vary by 2 minutes. This means that most of the time, Elena's trips are within 2 minutes of each other.

Then, select a few other students to explain their response to the last questions. Discuss:

"Without calculating, how did you determine which data set had the smaller range?"

The dot plot that has the narrower overall spread would have the smaller range because the distance between the greatest and least values is smaller.

"How did you determine which one had the smaller IQR?"

The distribution that has a middle half of the points that seem more clustered together would have the smaller IQR.

"In general, what does a larger range tell us?"

A wider spread in the data overall, or more variability in the data set.

"What does a larger IQR tell us?"

A wider spread around the center of data, more variability in the middle half of the data set where typical values tend to be.

"Can a data set have a large range and a small IQR?"

Yes, Data set B could be described this way.

If not mentioned by students, explain that the IQR plays a role similar to that of the mean absolute deviation (MAD). It tells us how different and spread out the data values are, but instead of measuring the average distance of data values from the mean, it measures the span of the middle half of the data.

Create a display of important terms and vocabulary for students to reference throughout the unit. Include the terms "quartile," "range," and "interquartile range (IQR)."

Lesson Synthesis

"What are the quartiles for a numerical data set?"
Numbers that show where we split the data up so it is in quarters.

"What is the relationship between the quartiles and the median?"
The second quartile is also the median.

"What is the Interquartile range (IQR)? What does it mean?"

The IQR is the difference between the third and first quartile. It is a measure of the variability or spread of the data. It tells us how much "space" the middle half of the data occupies.

"Compare MAD and IQR. How are they alike? How are they different?"

They both provide information on the distribution of a set of data. MAD works with the mean while IQR works with the median. MAD considers all the data values and tells us the average distance between each data value and the mean, and IQR focuses on the middle half of the data and tells us how widely distributed it is.

Lesson Summary

Earlier we learned that the mean is a measure of the center of a distribution and the MAD is a measure of the variability (or spread) that goes with the mean. There is also a measure of spread that goes with the median. It is called the interquartile range (IQR).

Finding the IQR involves splitting a data set into fourths. Each of the three values that splits the data into fourths is called a **quartile**.

- The median, or second quartile (Q2), splits the data into two halves.
- The first quartile (Q1) is the middle value of the lower half of the data.
- The third quartile (Q3) is the middle value of the upper half of the data.

For example, here is a data set with 11 values.

12	19	20	21	22	33	34	35	40	40	49
		Q1			Q2			Q3		

- The median, or second quartile (Q2), splits the data into two equal halves. For this data set, the median is 33.
- The first quartile (Q1) is the middle value of the lower half of the data. For this data set, the first quartile is 20. It is the median of the numbers that are less than 33.
- The third quartile (Q3) is the middle value of the upper half of the data. For this data set, the third quartile 40. It is the median of the numbers that are greater than 33.

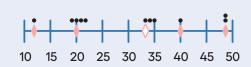
The difference between the maximum and minimum values of a data set is the **range**. For this data set, the range is 37 because 49 - 12 = 37.

The difference between Q3 and Q1 is the **interquartile range (IQR).** For this data set, the IQR is 20 because Q3 – Q1 = 40 – 20. Because the distance between Q1 and Q3 includes the middle two-fourths of the distribution, the values between those two quartiles are sometimes called the *middle half of the data*.

Responding To Student Thinking

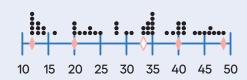
Points to Emphasize

If students struggle with finding or interpreting the IQR or median, revisit both ideas in this lesson: Unit 8, Lesson 16 Box Plots The bigger the IQR, the more spread out the middle half of the data values are. The smaller the IQR, the closer together the middle half of the data values are. This is why we can use the IQR as a measure of spread.



A *five-number summary* can be used to summarize a distribution. It includes the minimum, first quartile, median, third quartile, and maximum of the data set. For the previous example, the five-number summary is 12, 20, 33, 40, and 49. These numbers are marked with diamonds on the dot plot.

Different data sets can have the same five-number summary. For instance, here is another data set with the same minimum, maximum, and quartiles as the previous example.



57

63

76

Cool-down

How Far Can You Throw?

5 mir

Student Task Statement

Diego wondered how far sixth-grade students could throw a heavy ball. He decided to collect data to find out. He asked 10 friends to throw the ball as far as they could and measured the distance from the starting line to where the ball landed. The data shows the distances he recorded in feet.

40 40 47 49 50 53 55

1. Find the median and IQR of the data set.

The median is 51.5 feet. $(50 + 53) \div 2 = 51.5$.

The IQR is IO, because QI is 47, Q3 is 57, and 57 - 47 = 10.

- **2.** On a later day, he asked the same group of 10 friends to throw a ball again and collected another set of data. The median of the second data set is 49 feet, and the IQR is 6 feet.
 - **a.** Did the 10 friends, as a group, perform better (throw farther) or worse in the second round compared to the first round? Explain how you know.

Worse

Sample reasoning: The median of the second data set is 49 feet, which is 2.5 feet lower than in the first round.

b. Were the distances in the second data set more variable or less variable compared to those in the first round? Explain how you know.

Less variable

Sample reasoning: The IQR of the second data set is smaller, so the values are less spread out.

Practice Problems

6 Problems

Problem 1

Suppose that there are 20 numbers in a data set and that they are all different.

a. How many of the values in this data set are between the first quartile and the third quartile?

10

There are 5 numbers in each quartile, and there are two quartiles in between the first and third quartiles.

b. How many of the values in this data set are between the first quartile and the median?

5

The median is the second quartile. The first quartile to the second comprises one-fourth of the data in the data set.

Problem 2

In a word game, 1 letter is worth 1 point. This dot plot shows the scores for 20 common words.



a. What is the median score?

6.5 points

b. What is the first quartile (Q1)?

5

c. What is the third quartile (Q3)?

8

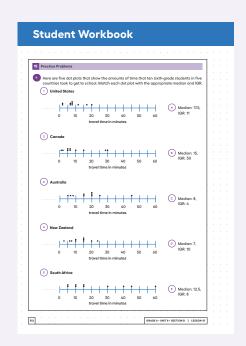
d. What is the interquartile range (IQR)?

3

Problem 3

Mai and Priya each bowl 10 games and record the scores. Mai's median score is 120, and her IQR is 5. Priya's median score is 118, and her IQR is 15. Whose scores probably have less variability? Explain how you know.

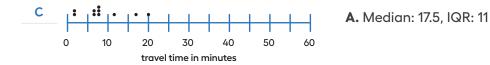
Sample explanation: Mai's IQR is smaller, so her scores probably vary less than Priya's scores.



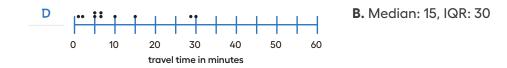
Problem 4

Here are five dot plots that show the amounts of time that ten sixth-grade students in five countries took to get to school. Match each dot plot with the appropriate median and IQR.

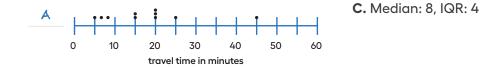
1. United States



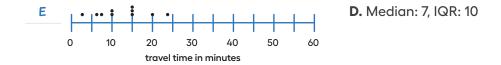
2. Canada



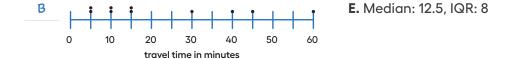
3. Australia



4. New Zealand



5. South Africa



Problem 5

from Unit 7, Lesson 12

Draw and label an appropriate pair of axes and plot the points. A = (10, 50), B = (30, 25), C = (0, 30), D = (20, 35)

Check student work to ensure that they made reasonable choices about axes and scale that allowed them to clearly plot all the points.

Problem 6

from Unit 6, Lesson 7

There are 20 pennies in a jar. If 16% of the coins in the jar are pennies, how many coins are there in the jar? 125

Sample response: $20 \div 0.16 = 125$

