Lesson 10: Finding and interpreting the mean as the balance point

Goals

- Calculate and interpret (orally and in writing) distances between data points and the mean of the data set.
- Interpret diagrams that represent the mean as a "balance point" for both symmetrical and non-symmetrical distributions.
- Represent the mean of a data set on a dot plot and interpret it in the context of the situation.

Learning Targets

- I can describe what the mean tells us in the context of the data.
- I can explain how the mean represents a balance point for the data on a dot plot.

Lesson Narrative

In the previous lesson, students interpreted the mean as a fair-share value—i.e., what each group member would have if all the values are distributed such that all members have the same amount. In this lesson, students use the structure of the data to interpret the mean as the balance point of a numerical distribution. They calculate how far away each data point is from the mean and study how the distances on either side of the mean compare.

Students connect this interpretation to why we call the mean a **measure of the centre** of a distribution and, through this interpretation, begin to see how the mean is useful in characterising a "typical" value for the group. Students continue to practise calculating the mean of a data set and interpreting it in context.

Addressing

- Recognise that a measure of centre for a numerical data set summarises all of its values with a single number, while a measure of variation describes how its values vary with a single number.
- Giving quantitative measures of centre (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered.

Instructional Routines

- Stronger and Clearer Each Time
- Co-Craft Questions
- Which One Doesn't Belong?

MEI Mathematics Education Innovation

Student Learning Goals

Let's look at another way to understand the mean of a data set.

10.1 Which One Doesn't Belong: Division

Warm Up: 5 minutes

This warm-up encourages students to analyse the structure and value of expressions, and to connect them to the process of calculating a mean. Each expression has one obvious reason it does not belong, however, there is not one single correct answer.

As students discuss in small groups, listen for ideas related to finding the mean of a data set. Highlight these ideas during whole-class discussion.

Instructional Routines

- Which One Doesn't Belong?

Launch

Arrange students in groups of 2–4. Display the expressions for all to see. Give students 1 minute of quiet think time, and ask them to indicate when they have noticed one expression that does not belong and can explain why. When the minute is up, ask them to share their thinking with their small group, and then, together, find at least one reason each expression doesn't belong.

Student Task Statement

Which expression does not belong? Be prepared to explain your reasoning.

$$\frac{8+8+4+4}{4}$$

$$\frac{10+10+4}{4}$$

$$\frac{9+9+5+5}{4}$$

$$\frac{6+6+6+6+6}{5}$$

Student Response

Answers vary. Sample responses:

$\frac{8+8+4+4}{4}$ doesn't belong because it is the only expression where each number in the numerator is a multiple of the denominator.

$\frac{9+9+5+5}{4}$ doesn't belong because it is the only expression with a value of 7.

$\frac{10+10+4}{4}$ doesn't belong because it is the only expression where the number of terms in the numerator is not the same as the value of the denominator.

$\frac{6+6+6+6+6}{5}$ doesn't belong because it is the only expression with 5 as the denominator.

Activity Synthesis

Ask each group to share one reason why a particular expression does not belong. Record and display the responses for all to see. After each response, ask the class if they agree or disagree. Since there is no single correct answer to the question, attend to students' explanations and ensure the reasons given are reasonable. If students give unsubstantiated claims, ask them to substantiate them.

At the end of the discussion, ask students which expression or expressions most represent how they would find the mean of a data set. There is a reason each expression (other than C) could represent how they would find the mean of a data set, however, highlight reasoning about the number of terms in the numerator being the same as the value of the denominator (e.g., there are 5 terms in the numerator and the denominator is 5).

10.2 Travel Times (Part 1)

15 minutes

In this activity, students explore the idea of the mean as a **measure of centre** of all the values in the data, using a dot plot to help them visualise this idea. Students determine the distance between each data point and the mean, and notice that the sum of distances to the left is equal to the sum of distances to the right. In this sense, the mean can been seen as "balancing" the sets of points with smaller values than it and those with larger values. They make use of the structure to calculate the distance between each data point and another point that is *not* the mean to see that the sums on the two sides are not equal. The idea of the mean as a *measure of centre* of a distribution is introduced in this context.

As students work and discuss, identify those who could articulate why the mean can be considered a balancing point of a data set.

Instructional Routines

- Stronger and Clearer Each Time

Launch

Arrange students in groups of 2. Give students 5 minutes to complete the first two questions with a partner, and then 5 minutes of quiet work time to complete the last two questions. Follow with a whole-class discussion.

Action and Expression: Internalise Executive Functions. Chunk this task into more manageable parts to support students who benefit from support with organisational skills. Check in with students within the first 3-5 minutes of work time to see how they calculate

the distance between each point and 11. If necessary, remind students that distances are positive.
Supports accessibility for: Organisation; Attention

Anticipated Misconceptions

Some students might write negative values for distances between the mean and points to the left of the mean. They might recall looking at distances between 0 and numbers to the left of it in a previous unit and mistakenly think that numbers to the left of the mean would have a negative distance from the mean. Remind students that distances are always positive; the answer to "How far away?" or "How many units away?" cannot be a negative number.

Student Task Statement

Here is the data set from an earlier lesson showing how long it takes for Diego to walk to school, in minutes, over 5 days. The mean number of minutes is 11.

12 7 13 9 14

1. Represent Diego's data on a dot plot. Mark the location of the mean with a triangle.
2. The mean can also be seen as a **measure of centre** that balances the points in a data set. If we find the distance between every point and the mean, add the distances on each side of the mean, and compare the two sums, we can see this balancing.
    a. Record the distance between each point and 11 and its location relative to 11.

time in minutes	distance from 11	left of 11 or right of 11?
12		
7		
13		
9		
14		

    b. Sum of distances left of 11:__________ Sum of distances right of 11:__________

    c. What do you notice about the two sums?

3. Can another point that is *not* the mean produce similar sums of distances?
4. Let's investigate whether 10 can produce similar sums as those of 11.
    a. Complete the table with the distance of each data point from 10.

time in minutes	distance from 10	left of 10 or right of 10?
12		
7		
13		
9		
14		

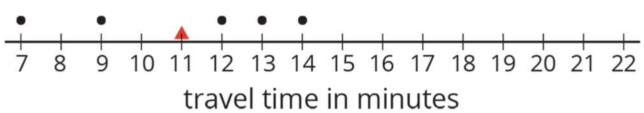
b. Sum of distances left of 10:__________ Sum of distances right of 10:__________

c. What do you notice about the two sums?

5. Based on your work so far, explain why the mean can be considered a balance point for the data set.

Student Response

1.



2.

time in minutes	distance from 11	left of 11 or right of 11?
12	1	right
7	4	left
13	2	right
9	2	left
14	3	right

a. The sum of the distances to the left of 11 is $4 + 2 = 6$.

b. The sum of the distances to the right of 11 is $1 + 2 + 3 = 6$.

c. The two sums are equal.

4.

time in minutes	distance from 10	left of 11 or right of 10?
12	2	right
7	3	left
13	3	right
9	1	left
14	4	right

a. The sum of the distances to the left of 10 is $3 + 1 = 4$.

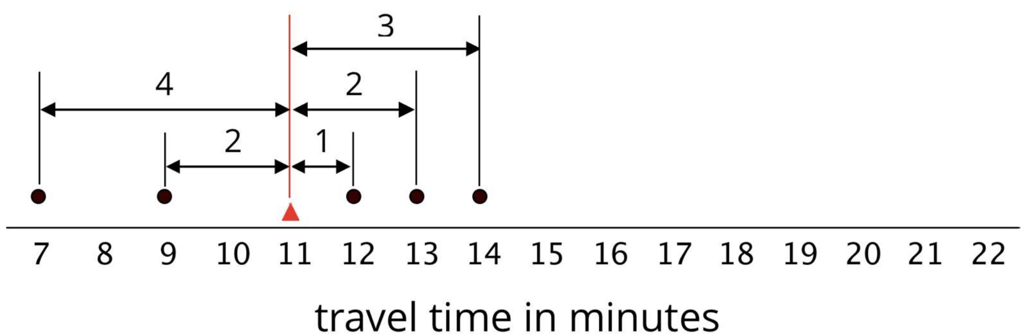
b. The sum of the distances to the right of 10 is $2 + 3 + 4 = 9$.

c. The two sums are not equal.

5. The sum of distances to the left of the mean is equal to the sum of distances to the right of the mean, so the mean balances the data values that are larger and those that are smaller. If a number is not the mean of the data set, then the sum of distances to the left and the sum of the distances to the right of it are not equal.

Activity Synthesis

Select a couple of students to share their observations on the distances between Diego's mean travel time and other points. To facilitate discussion, display this dot plot (with the distances labelled) for all to see. Discuss how the sums of distances change when different points are chosen as a reference from which deviations are measured.



Ask:

- "How are the sums of distances to mean (11 minutes) and the sums of distances to another point other than the mean (e.g., 10 minutes) different?"
- "If you choose another point or location on the number line, would it produce equal sums of distances to the left and to the right?"

Highlight the idea that only the mean could produce an equal sum of distances. Remind students that they have previously described centres of data sets. Explain that the mean is used as a *measure of centre* of a distribution because it balances the values in a data set. Because data points that are greater than the mean balance with those that are less than the mean, the mean is used to describe what is typical for a data set.

Writing, Speaking, Listening: Stronger and Clearer Each Time. Use this routine to give students a structured opportunity to revise and refine their initial response to "Based on your work so far, explain why the mean can be considered a balance point for the data set." Display prompts for feedback that will help students strengthen their ideas and clarify their language. For example, "Can you include details about the sum of the distances from the

balancing point that are important?" and "What did you mean when you said . . .?" Invite students to build on to their initial response by incorporating some of the feedback. This will help students refine their interpretation of the mean as a balancing point of a data set before the whole-group discussion.
Design Principle(s): Support sense-making

10.3 Travel Times (Part 2)

15 minutes

This activity serves two key purposes: to reinforce the idea of the mean as a balance point and a measure of centre of a distribution, and to introduce the idea that distances of data points from the mean can help us describe variability in data, which prepares students to think about mean absolute deviation in the next lesson. Students also practise calculating mean of a distribution and interpreting it in context.

Unlike in previous activities, students are given less scaffolding for finding both the mean and the sums of distances from the mean. As students work, notice those who may need additional prompts to perform these tasks. Also listen for students' explanations on what a larger mean tells us in this context. Identify those who can clearly distinguish how the mean differs from deviations from the mean.

Instructional Routines

- Co-Craft Questions

Launch

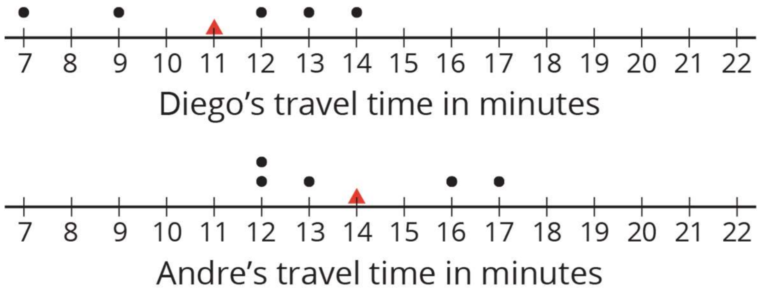
Keep students in groups of 2. Give students 5 minutes of quiet work time to complete the first set of questions and then 2–3 minutes to discuss their responses with their partner before working on the second set of questions together.

The term "variation" is used in student text for the first time. If needed, explain that it has a similar meaning as "variability" and refers to how different or alike the data values are.

Representing, Conversing: Co-craft Questions. Before revealing the task, display the image of Diego's and Andres's dot plots, and only the first sentence of the problem statement. Ask students to write a list of mathematical questions that could be asked about what they see. Invite students to share their questions with a partner before selecting 2–3 to share with the class. Listen for questions that use the terms 'mean', 'spread', or 'centre' and highlight where these are represented in the dot point. This helps students use mathematical language related to representing distributions of data sets and to understand the context of this problem prior to be asked to reason about the different quantities in the situation.
Design Principle(s): Cultivate conversation; Maximise meta-awareness

Student Task Statement

1. Here are dot plots showing how long Diego's trips to school took in minutes—which you studied earlier—and how long Andre's trips to school took in minutes. The dot plots include the means for each data set, marked by triangles.

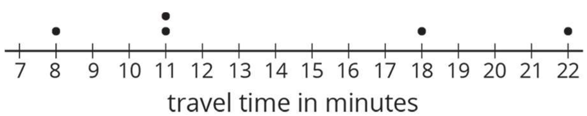


2.

a. Which of the two data sets has a larger mean? In this context, what does a larger mean tell us?

b. Which of the two data sets has larger sums of distances to the left and right of the mean? What do these sums tell us about the variation in Diego’s and Andre’s travel times?

3. Here is a dot plot showing lengths of Lin’s trips to school.



4.

a. Calculate the mean of Lin’s travel times.

b. Complete the table with the distance between each point and the mean as well whether the point is to the left or right of the mean.

time in minutes	distance from the mean	left or right of the mean?
22		
18		
11		
8		
11		

c. Find the sum of distances to the left of the mean and the sum of distances to the right of the mean.

d. Use your work to compare Lin’s travel times to Andre’s. What can you say about their average travel times? What about the variability in their travel times?

Student Response

2.

a. Andre’s data set has a larger mean, since Andre's mean is 14 minutes, and Diego’s mean is 11 minutes. In the context of this problem, this implies Andre's average travel time is longer than Diego’s average.

b. The sum of distances to the left and right of the mean for Diego’s data set is $(11-7)+(11-9)+(12-11)+(13-11)+(14-11)=4+2+1+2+3=12$. The sum of distances for Andre’s data set is $2(14-12)+(14-13)+(16-14)+(17-14)=2\times 2+1+2+3=10$. Diego’s sum is greater than Andre’s implying that the variability in of Diego’s travel times is greater.

4.

a. The mean is $\frac{8+11+11+18+22}{5}=\frac{70}{5}=14$

time in minutes	distance from the mean	left or right of the mean?
22	8	right
18	4	right
11	3	left
8	6	left
11	3	left

b. Sum of distances to the left: $3+3+6=12$.

c. Sum of distances to the right: $8+4=12$.

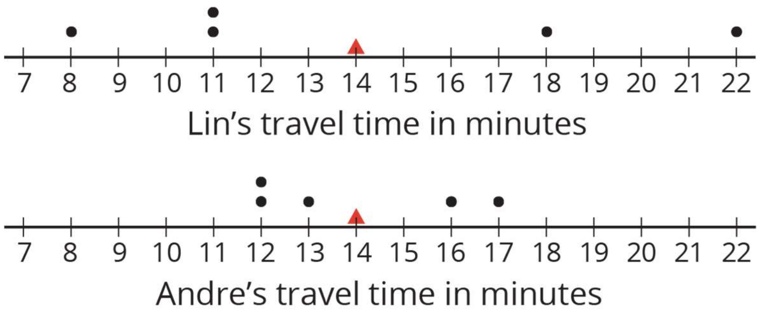
d. Answers vary. Sample response: Lin’s average travel time is the same as Andre’s (both have a mean of 14 minutes), and these are both greater than Diego’s. The sum of the distances of data points from Lin's mean is much larger than Andre’s and Diego’s sums. I think this implies that Lin’s travel times are much more varied than Andre’s and Diego’s travel times.

Activity Synthesis

The two big ideas to emphasise during discussion are: what the means tell us in this context, and what the sums of distances to either side of each mean tell us about the travel times.

Select a couple of students to share their analyses of Diego and Andre’s travel times. After each student explains, briefly poll the class for agreement or disagreement. If one or more students disagree with an analysis, ask for their reasoning and alternative explanations.

Then, focus the conversation how Lin and Andre’s travel times compare. Display the dot plots of their travel times for all to see.



Discuss:

- "How do the data points in Lin's dot plot compare to those in Andre's?"
- "How do their means compare? How do their sums of distances from the mean compare?"
- "What do the sums of distances tell us about the travel times?"
- "If more than half of Lin's data points are far from the mean of 14 minutes, is the mean still a good description of her typical travel time? Why or why not?"

Students should see that larger distances from the mean suggest greater variability in the travel times. Even though both students have the same average travel time (both 14 minutes), Lin's travel times are much more varied than Andre's. A couple of Lin's travel times are a lot longer or shorter than 14 minutes. Overall, her data points are within 6–8 minutes of the mean. For Andre, all of his data points are within 3 minutes of the mean.

The last discussion question prepares students to think about a different way to measure the centre of a distribution in upcoming lessons.

Lesson Synthesis

In this lesson, we learn that the mean can be interpreted as the balance point of a distribution.

- "How does the mean balance the distribution of a data set?"
- "How can a dot plot help us make sense of this interpretation?"
- "Could another value—besides the mean—balance a data distribution? How can we tell?"

We also learn that the mean is used as a **measure of centre** of a distribution, or a number that summarises the centre of a distribution.

- "Why might it make sense for the mean to be a number that describes the centre of a distribution?"

- "In earlier lessons, we had used an estimate of the centre of a distribution to describe what is typical or characteristic of a group. Why might it make sense to use the mean to describe a typical feature of a group?"

10.4 Text Messages

Cool Down: 5 minutes

Student Task Statement

The three data sets show the number of text messages sent by Jada, Diego, and Lin over 6 days. One of the data sets has a mean of 4, one has a mean of 5, and one has a mean of 6.

Jada

4 4 4 6 6 6

Diego

4 5 5 6 8 8

Lin

1 1 2 2 9 9

1. Which data set has which mean? What does this tell you about the text messages sent by the three students?
2. Which data set has the greatest variability? Explain your reasoning.

Student Response

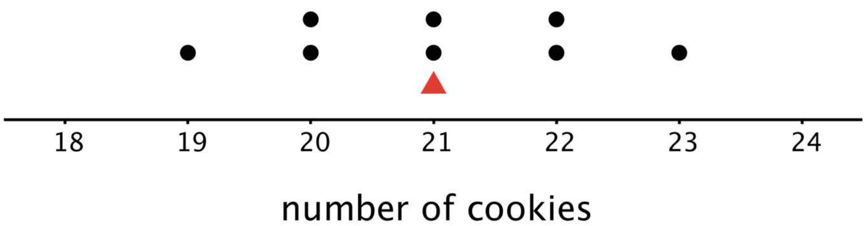
1. Jada's mean is 5, since $\frac{4+4+4+6+6+6}{6} = \frac{30}{6} = 5$. Diego's mean is 6, since $\frac{4+5+5+6+8+8}{6} = \frac{36}{6} = 6$. Lin's mean is 4, since $\frac{1+1+2+2+9+9}{6} = \frac{24}{6} = 4$. On average, Diego sent the most text messages per day, and Lin sent the fewest text messages per day.
2. Answers vary. Sample response: Lin's data had the highest variability. All data points lie far away from the mean; the sum of the differences is the largest.

Student Lesson Summary

The mean is often used as a **measure of centre** of a distribution. This is because the mean of a distribution can be seen as the "balance point" for the distribution. Why is this a good way to think about the mean? Let's look at a very simple set of data on the number of cookies that each of eight friends baked:

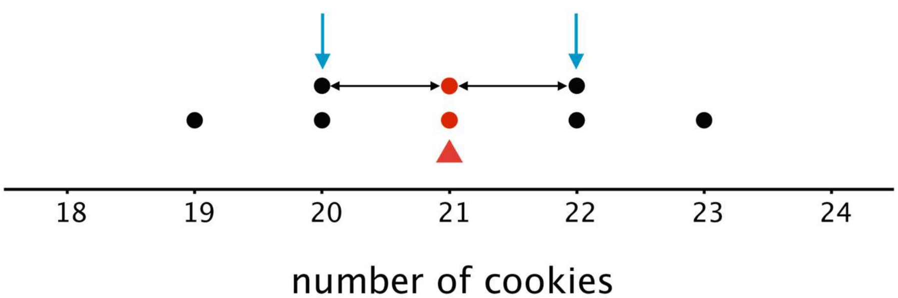
19 20 20 21 21 22 22 23

Here is a dot plot showing the data set.

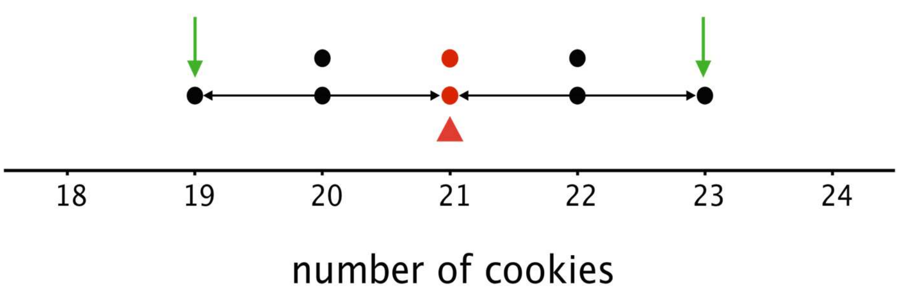


The distribution shown is completely symmetrical. The mean number of cookies is 21, because $(19 + 20 + 20 + 21 + 21 + 22 + 22 + 23) \div 8 = 21$. If we mark the location of the mean on the dot plot, we can see that the data points could balance at 21.

In this plot, each point on either side of the mean has a mirror image. For example, the two points at 20 and the two at 22 are the same distance from 21, but each pair is located on either side of 21. We can think of them as balancing each other around 21.



Similarly, the points at 19 and 23 are the same distance from 21 but are on either side of it. They, too, can be seen as balancing each other around 21.



We can say that the distribution of the cookies has a centre at 21 because that is its balance point, and that the eight friends, on average, baked 21 cookies.

Even when a distribution is not completely symmetrical, the distances of values below the mean, on the whole, balance the distances of values above the mean.

Glossary

- measure of centre

Lesson 10 Practice Problems

Problem 1 Statement

On school days, Kiran walks to school. Here are the lengths of time, in minutes, for Kiran's walks on 5 school days:

16 11 18 12 13

a. Create a dot plot for Kiran's data.

b. Without calculating, decide if 15 minutes would be a good estimate of the mean. If you think it is a good estimate, explain your reasoning. If not, give a better estimate and explain your reasoning.

c. Calculate the mean for Kiran's data.

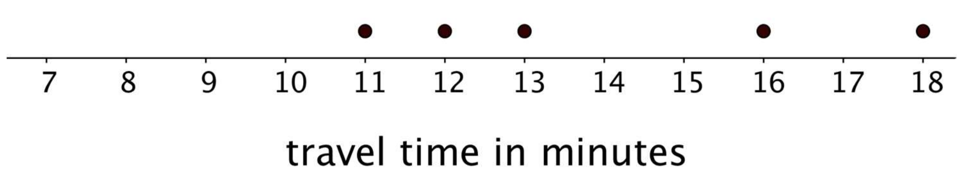
d. In the table, record the distance of each data point from the mean and its location relative to the mean.

time in minutes	distance from the mean	left or right of the mean?
16		
11		
18		
12		
13		

e. Calculate the sum of all distances to the left of the mean, then calculate the sum of distances to the right of the mean. Explain how these sums show that the mean is a balance point for the values in the data set.

Solution

a.



b. Answers vary. Sample response: 15 minutes seems to be a little too high an estimate for the mean, because (looking at the dot plot) the sum of the distances between the mean and the points to its left seems to be greater than the sum of the distances between it and the points to its right. A better estimate would be 14 minutes, because the sums of distances to its left and the right would be more balanced.

c. $\frac{11+12+13+16+1}{5} = 14.$

d.

time in minutes	distance from mean	direction: left or right of mean?
16	2	right
11	3	left
18	4	right
12	2	left
13	1	left

e. The sum of the distances to the left: $3 + 2 + 1 = 6$. The sum of the distances to the right: $4 + 2 = 6$. The sum of distances on the left of the mean is equal to the sum of distances to the right of the mean, which tells us that the data values are balanced on the mean.

Problem 2 Statement

Noah scored 20 points in a game. Mai's score was 30 points. The mean score for Noah, Mai, and Clare was 40 points. What was Clare's score? Explain or show your reasoning.

Solution

70 points. Reasoning varies. Sample reasoning:

- Clare would need to have a score that would be 30 points to the right of the mean score of 40. This score would balance the 30 points at the left of the mean.
- If the mean score was 40 points, Noah’s score was 20 points short and Mai’s was 10 points short. Clare’s score must be 30 points above the mean so that when the points are distributed each person’s share is 40 points.

Problem 3 Statement

Compare the numbers using >, <, or =.

a. -2 _____ 3

b. $-|12|$ _____ $|15|$

a. 3 ____ -4

b. $|15|$ ____ $|-12|$

a. -7 ____ -11

b. -4 ____ $|5|$

Solution

a. $<$

b. $<$

c. $>$

d. $>$

e. $>$

f. $<$

Problem 4 Statement

a. Plot $\frac{2}{3}$ and $\frac{3}{4}$ on a number line.

b. Is $\frac{2}{3} < \frac{3}{4}$, or is $\frac{3}{4} < \frac{2}{3}$? Explain how you know.

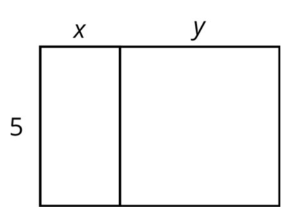
Solution

a. The number line should show $\frac{2}{3}$ to the left of $\frac{3}{4}$, and both closer to 1 than to 0.

b. $\frac{2}{3} < \frac{3}{4}$ because $\frac{2}{3}$ is to the left of $\frac{3}{4}$ on the number line.

Problem 5 Statement

Select **all** the expressions that represent the total area of the large rectangle.



a. $5(x + y)$

b. $5 + xy$

c. $5x + 5y$

d. $2(5+x+y)$

e. $5xy$

Solution ["A", "C"]