Lesson 11: Comparing Groups

Goals

- Calculate the mean and mean absolute deviation for a data set, and interpret (orally) these measures.
- Compare and contrast (orally and in writing) populations represented on dot plots in terms of their shape, center, spread, and visual overlap.
- Justify (in writing) whether two populations are “very different” based on the difference in their means expressed as a multiple of the mean absolute deviation.

Learning Targets

- I can calculate the difference between two means as a multiple of the mean absolute deviation.
- When looking at a pair of dot plots, I can determine whether the distributions are very different or have a lot of overlap.

Lesson Narrative

In this lesson, students review measures of center and variability from grade 6. They also work at deciding whether or not two distributions are very different from each other (MP3). This lesson introduces the idea of expressing the difference between the centers of two distributions as a multiple of a measure of variability as a way to help students make this determination (MP2). For the problems in this lesson, the populations under study are small and the data for the entire populations are known. In future lessons, students will revisit this calculation as a way to decide whether there is a meaningful difference between two populations given data from only a sample of each population.

Alignments

Addressing

- 7.SP.B: Draw informal comparative inferences about two populations.
- 7.SP.B.3: Informally assess the degree of visual overlap of two numerical data distributions with similar variabilities, measuring the difference between the centers by expressing it as a multiple of a measure of variability. For example, the mean height of players on the basketball team is 10 cm greater than the mean height of players on the soccer team, about twice the variability (mean absolute deviation) on either team; on a dot plot, the separation between the two distributions of heights is noticeable.

Instructional Routines

- Anticipate, Monitor, Select, Sequence, Connect
- MLR2: Collect and Display

- MLR8: Discussion Supports
- Notice and Wonder

Student Learning Goals

Let's compare two groups.

11.1 Notice and Wonder: Comparing Heights

Warm Up: 5 minutes

The purpose of this warm-up is to collect ideas that will be useful in the discussions in this lesson. While students may notice and wonder many things about these images, the methods of showing that the volleyball team is much taller than the gymnastic team are the important discussion points.

Addressing

- 7.SP.B

Instructional Routines

- Notice and Wonder

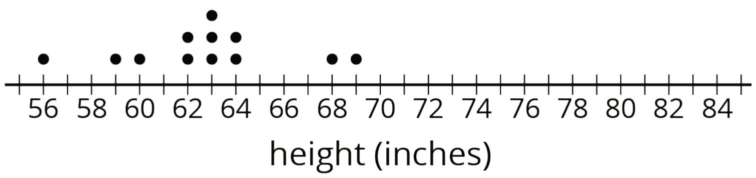
Launch

Arrange students in groups of 2. Tell students that they will look at the dot plots, and their job is to think of at least one thing they notice and at least one thing they wonder. Display the dot plots for all to see. Ask students to give a signal when they have noticed or wondered about something. Give students 1 minute of quiet think time, and then 1 minute to discuss the things they notice with their partner, followed by a whole-class discussion.

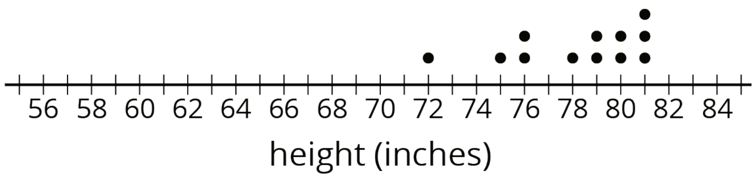
Student Task Statement

What do you notice? What do you wonder?

women's gymnastics team



men's volleyball team



Student Response

Things students may notice:

- The volleyball team is much taller than the gymnastics team.
- The gymnastics team's heights are all under 70 inches tall.
- The volleyball team's heights are all over 70 inches tall.

Things students may wonder:

- How much taller are the volleyball players?
- Are players on volleyball teams usually that much taller than gymnasts or is this just for these two teams?
- What are the heights of the female volleyball players?

Activity Synthesis

Ask students to share the things they noticed and wondered. Record and display their responses for all to see. If possible, record the relevant reasoning on or near the image. After each response, ask the class if they agree or disagree and to explain alternative ways of thinking, referring back to the images each time. If the definitive difference in height does not come up during the conversation, ask students to discuss this idea.

The next activity looks more closely at comparing these data sets. It is not necessary to have students calculate anything (mean, median, MAD, IQR) yet.

11.2 More Team Heights

15 minutes
In this activity, students are asked to compare the heights of two groups of people. The wording of the questions allows for multiple interpretations and any reasonable answer should be accepted (MP1). This activity also provides an opportunity to remind students of how to analyze dot plots as well as how to calculate the measures of center and variability of the data.

Addressing

- 7.SP.B.3

Instructional Routines

- MLR8: Discussion Supports

Launch

Keep students in groups of 2.

Display the dot plots from the warm-up activity and help students see that the data sets given in their books or devices match the numbers shown in the dot plots.

For the problem addressing the tennis and badminton teams, you may suggest each student create a dot plot of one of the groups and then compare with their partner.

Allow students 10 minutes of partner work time followed by a whole-class discussion.

Access for Students with Disabilities

Action and Expression: Internalize Executive Functions. Chunk this task into more manageable parts. After students have solved the first problem, check-in with either select groups of students or the whole class. Invite students to share the strategies they have used so far as well as any questions they have before continuing.
Supports accessibility for: Organization; Attention

Student Task Statement

1. How much taller is the volleyball team than the gymnastics team?
    - Gymnastics team's heights (in inches) : 56, 59, 60, 62, 62, 63, 63, 63, 64, 64, 68, 69
    - Volleyball team's heights (in inches): 72, 75, 76, 76, 78, 79, 79, 80, 80, 81, 81, 81
2. Make dot plots to compare the heights of the tennis and badminton teams.
    - Tennis team's heights (in inches): 66, 67, 69, 70, 71, 73, 73, 74, 75, 75, 76
    - Badminton team's heights (in inches): 62, 62, 65, 66, 68, 71, 73

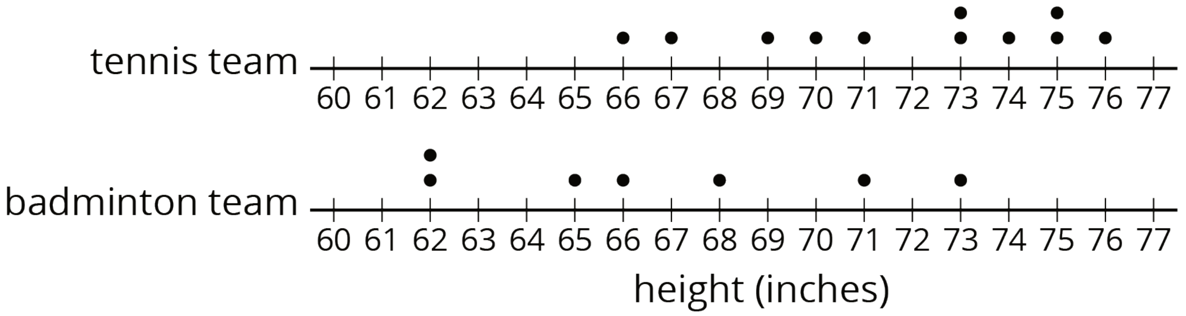
    What do you notice about your dot plots?
3. Elena says the members of the tennis team were taller than the badminton team. Lin disagrees. Do you agree with either of them? Explain or show your reasoning.

Student Response

1. Answers vary. Sample responses:
    - The total height of the volleyball team is 185 inches, or about 15 feet 5 inches, greater than the total height of the gymnastics team, because I added up everyone on each team and $938 - 753 = 185$.
    - The mean height of the volleyball team is about 15.4 inches greater than the mean height of the gymnastics team, because $753 \div 12 = 62\frac{3}{4}$, $938 \div 12 = 78\frac{1}{6}$, and $78\frac{1}{6} - 62\frac{3}{4} = 15\frac{5}{12}$, or about 15.4 inches.
    - The median height of the volleyball team is 16 inches greater than the median height of the gymnastics team, and this may be more important than the difference in the means because the shape of the distribution for the volleyball team is not symmetric.
    - The tallest person on the volleyball team is 12 inches taller than the tallest person on the gymnastics team.

- The shortest person on the volleyball team is 16 inches taller than the shortest person on the gymnastics team.
- The tallest person on the volleyball team is 25 inches taller than the shortest person on the gymnastics team.
- The shortest person on the volleyball team is 3 inches taller than the tallest person on the gymnastics team.

2. The center of the distribution as well as the minimum and maximum heights for the tennis team are all greater than the center, minimum, and maximum for the badminton team; however, there is a lot of overlap between the two distributions. Unlike with the gymnastics and volleyball teams, the tallest person on the badminton team is taller than the shortest person on the tennis team.



   - The total height of the tennis team is 322 inches, or 26 feet 10 inches, greater than the total height of the badminton team; however, it does not make sense to compare the totals, because the tennis team had more people.
   - The mean height of the tennis team is about 5 inches greater than the mean height of the badminton team.
   - The median height of the tennis team is 7 inches greater than the median height of the badminton team.
   - The tallest person on the tennis team is only 3 inches taller than the tallest person on the badminton team.
   - The shortest person on the tennis team is 4 inches taller than the shortest person on the badminton team.
   - The tallest person on the tennis team is 14 inches taller than the shortest person on the badminton team.
   - The shortest person on the tennis team is actually 7 inches shorter than the tallest person on the badminton team, because there is so much overlap between the two distributions.

3. Answers vary. Sample responses:
   - I agree with Elena because the center of the distribution for the tennis team is greater than the center for the badminton team.

- I agree with Lin because the two distributions overlap so much. Some of the people on the badminton team were taller than some of the people on the tennis team. The difference in centers is not big enough to matter. If you mixed the people from both teams together into one group, you would not be able to determine who was a member of each team.

Activity Synthesis

The purpose of this discussion is for students to think about ways we could approach comparing two groups as well as have an opportunity to review dot plots, measures of center, and measures of variation from prior grades.

Ask, "What are some ways we can compare groups of things?" At this stage, students are only expected to informally compare the groups. Although a consistent "general rule" for comparing groups will be introduced in later lessons, this activity is about getting a general idea that some groups (like the gymnastics and volleyball teams) have a rather clear difference while others (like the tennis and badminton teams) may be more alike.

Ask students about the *distribution* of the data shown in the dot plots. Make sure to highlight the shape, center, and spread. Review how to find the **mean** as a measure of the center of a data set. Review how to calculate the **mean absolute deviation (MAD)** as a measure of the variability of a data set. Students may mention **median** and **interquartile range (IQR)** as other ways to measure center and variability. Although median and IQR are not needed in this activity, it may be useful to review how to calculate those values as well. Both measures of center and variability will be used later in the unit.

Introduce the idea of judging how much two data sets overlap.

Access for English Language Learners

Speaking: MLR8 Discussion Supports. Use this routine to support whole-class discussion. For each observation that is shared, ask students to restate what they heard using precise mathematical language. Consider providing students time to restate what they hear to a partner, before selecting one or two students to share with the class. Ask the original speaker if their peer was accurately able to restate their thinking. Call students' attention to any words or phrases that helped clarify the original statement. This will provide more students with an opportunity to produce language as they interpret the reasoning of others.
Design Principle(s): Support sense-making

11.3 Family Heights

15 minutes
Following the review from the previous activity, students are asked to use these calculations to compare two groups more formally. Students are shown one quantifiable method of determining

whether the two groups are relatively close or relatively very different in the discussion following the activity involving describing the difference of the measures of center as a multiple of the variability (MP2). The important idea for students to grasp from this activity is that the measures of center and measures of variability of the groups work together to give an idea of how similar or different the groups are.

As students work to compare heights of the two families, monitor for students who:

1. Create dot plots of the data and look at the overlap visually.
2. Compute measures of center to compare the groups numerically.
3. Also compute measures of variability to compare the measures of center for each group accounting for the spread of values.

Addressing

- 7.SP.B.3

Instructional Routines

- Anticipate, Monitor, Select, Sequence, Connect
- MLR2: Collect and Display

Launch

Keep students in groups of 2.

Introduce students to the idea that both the measures of center as well as the measures of variability are important when comparing data sets by asking students about these contexts.

- "Two groups of adults have mean weights that are different by 10 pounds. Are the two groups very different in weight?"
- "Two groups of 8 year olds have weights that are different by 10 pounds. Are the two groups very different in weight?"
- "Two groups of birds have weights that are different by 3.2 ounces (0.2 pounds). Are the two groups very different in weight?"

Without knowing more about the exact numbers, it seems that the adults might not be that different. A typical adult might weigh 180 pounds and, depending on diet and activity, even one person's weight might go up or down by 10 pounds in a few weeks. It might be hard to tell these two groups apart even with the 10 pound difference in means.

On the other hand, 8 year olds typically weigh about 50 pounds, so a 10 pound difference can mean a lot more. A group with an average weight of 50 pounds is 25% heavier than a group with an average weight of 40 pounds.

For the birds, it might be hard to say how different the groups are unless we have more information. If the two groups contain the same type of bird (e.g., parrots), we might expect the weights to be fairly consistent, so a 3.2 ounce difference could be a lot. If the two groups each contain a variety of types (from hummingbirds up to emus), the 3.2 ounce difference might not be very much.

Based on these examples, it helps to know not only the means (or medians), but also the variability of the groups. We saw this in the previous activity with the overlap in the dot plots.

For the work in this activity, tell students to try to give some values to help back up their comparison of the two groups.

Allow students 1 minute of quiet think time to examine ways of approaching the problem, followed by partner work time and a whole-class discussion.

Access for Students with Disabilities

Representation: Internalize Comprehension. Activate or supply background knowledge of measures of centers. Allow students to use calculators to ensure inclusive participation in the activity.
Supports accessibility for: Memory; Conceptual processing

Access for English Language Learners

Conversing: MLR2 Collect and Display. Use this routine to capture the language students use as they compare the heights of these two families. Circulate and listen to student talk during small-group and whole-class discussion. Record the words and phrases students use on a display for all to see. Invite students to borrow from—or add more language to—the display throughout the remainder of the lesson. This will help students read and use mathematical language during their paired and whole-group discussions.
Design Principle(s): Optimize output (for comparison); Maximize meta-awareness

Student Task Statement

Compare the heights of these two families. Explain or show your reasoning.

- The heights (in inches) of Noah’s family members: 28, 39, 41, 52, 63, 66, 71
- The heights (in inches) of Jada’s family members: 49, 60, 68, 70, 71, 73, 77

Student Response

The mean height of Noah's family members is about 51.43 inches. The mean height of Jada's family members is about 66.86 inches, which is greater than for Noah's family. The difference between these means is $66.86 - 51.43$, or 15.4 inches.

There is also some overlap in the heights of the two families. The tallest person in Noah's family is 22 inches taller than the shortest person in Jada's family.

> ### Are You Ready for More?
>
> If Jada's family adopts newborn twins who are each 18 inches tall, does this change your thinking? Explain your reasoning.

Student Response

Answers vary. Sample responses:

- Yes. After the addition of the new babies, the mean for Jada's family drops to 56 inches, which is much closer to the mean height of Noah's family (about 51 inches).
- No. With the addition of the new babies, the mean does not make sense to use as a measure of center for Jada's family anymore; we should use the median. The median of Noah's family is 52 inches, and the new median for Jada's family is 68 inches, and there is still a large difference between the two groups.

Activity Synthesis

The purpose of this discussion is for students to begin to be more formal in their comparison of data from different groups. In particular, a general rule is established that will be used in this unit.

Select students to share their approaches in the sequence outlined in the Activity Narrative.

The heights of Noah's and Jada's families overlap more than the heights of the gymnastics and volleyball teams. The difference in means for the two families is about the same as between the two teams.

- The difference between the mean heights of the volleyball and gymnastics teams is $78.17 - 62.75$, or 15.42 inches.
- The difference between the mean heights of Jada's and Noah's families is $66.86 - 51.43$, or 15.43 inches. The variability in heights for the families is greater than the variability in heights for the teams.

Explain: The difference between the means is not enough information to know whether or not the data sets are very different. One way to express the amount of overlap is to divide the difference in means by the (larger) mean absolute deviation.

Demonstrate for students how to do this calculation for the volleyball and gymnastics teams:

- The difference in means is more than 6 times the measure of variability, because $15.42 \div 2.46 \approx 6.3$.

Leave the calculation for the two teams displayed. Ask students to do the calculation for Jada's and Noah's families.

- The difference in means is a little more than 1 time the measure of variability, because $15.43 \div 13.22 \approx 1.2$.

As a general rule, we will consider it a large difference between the data sets if the difference in means is *more than twice* the mean absolute deviation. If the mean absolute deviation is different for each group, use the larger one for this calculation.

For students who ask why twice the MAD is used rather than some other value, defer the question for later in the unit. A later lesson titled Do They Carry More? will address this question in more detail.

11.4 Track Length

Optional: 10 minutes
In this activity, students continue to review the meaning of mean and mean absolute deviation as well as practice using the method they were shown in the previous activity to compare multiple groups based on their means and measure of variability. Students begin by matching information about a set of data to its dot plot and calculating the mean and mean absolute deviation for the remaining dot plots, then they compare the data sets pairwise using the mean and MAD values (MP3). In the discussion, students are introduced to a way to add extra information to dot plots to visualize the general rule given in the previous activity.

Addressing

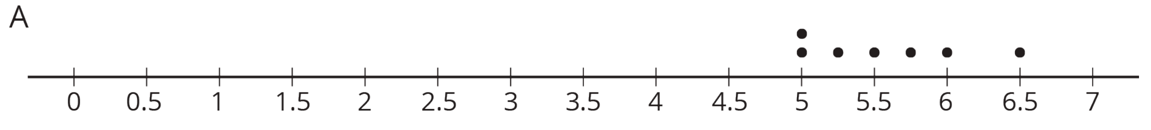
- 7.SP.B.3

Launch

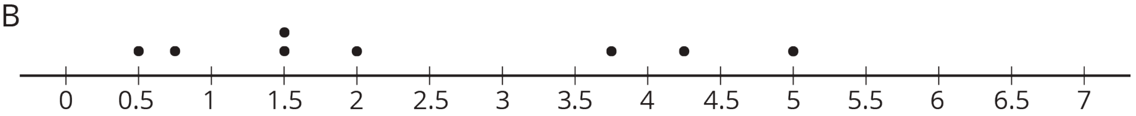
Keep students in groups of 2. Give students 5 minutes of partner work time followed by a whole-class discussion.

Student Task Statement

Here are three dot plots that represent the lengths, in minutes, of songs on different albums.



B



C

1. One of these data sets has a mean of 5.57 minutes and another has a mean of 3.91 minutes.

    a. Which dot plot shows each of these data sets?

    b. Calculate the mean for the data set on the other dot plot.

2. One of these data sets has a mean absolute deviation of 0.30 and another has a MAD of 0.44.

    a. Which dot plot shows each of these data sets?

    b. Calculate the MAD for the other data set.

3. Do you think the three groups are very different or not? Be prepared to explain your reasoning.

4. A fourth album has a mean length of 8 minutes with a mean absolute deviation of 1.2. Is this data set very different from each of the others?

Student Response

1. a. Dot plot A has a mean of 5.57 minutes, and dot plot C has a mean of 3.91 minutes.

    b. Dot plot B has a mean of 2.41 minutes.

2. a. Dot plot C has a MAD of 0.30 minutes, and dot plot A has a MAD of 0.44 minutes.

    b. Dot plot B has a MAD of 1.44 minutes.

3. Since dot plot B has such a large mean absolute deviation, it is hard to say that it is very different from dot plot C, but it is very different from dot plot A. Since the MADs are much smaller for the other two albums, it is easier to say that the length of the tracks from dot plot C and dot plot A are very different.

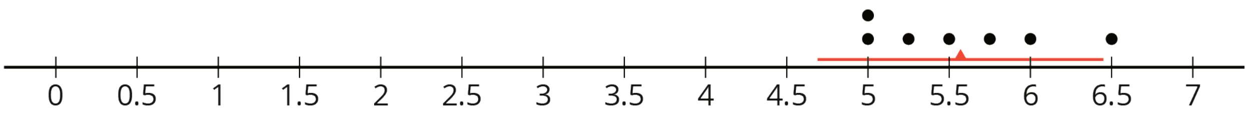
4. The large mean for the fourth album makes it very different from the others even with a large MAD.

Activity Synthesis

The purpose of the discussion is to help students visualize the calculations they performed.

Ask, “Before calculating any of the values, would you have guessed that dot plots B and C would not be very different, but A would be very different from the other two? Explain your reasoning.” (Yes, since there is some overlap between the data of dot plots B and C, but the data in A does not overlap very much with B and not at all with C.)

Demonstrate a technique for students to see the the general rule on the dot plots using dot plot A. Mark the mean of the data set on the dot plot with a triangle, then draw a line segment from the triangle so that its length is 2 MADs in each direction. On dot plot A, draw a triangle at 5.57 and a line segment from 4.69 ($5.57 - 2 \cdot 0.44$) to 6.45 ($5.57 + 2 \cdot 0.44$) as in the image here.



Tell students to draw the triangle for the mean and segment representing 2 MADs in each direction on all three dot plots. Ask students how we might see the general rule from these pictures. (If any of the means (triangles) are within 2 MADs (the line segment) of the mean of another group, then there is not a large difference between the two groups.) Ask them to compare their visual to their answers from the activity.

Lesson Synthesis

Consider asking these questions for discussion:

- “What does a dot plot tell you?”
- “What are some measures of center, and how are they calculated?”
- “Why is a measure of center useful for comparing two groups?”
- “Why is a measure of variability also needed when comparing two groups?”
- “What is the general rule we will use to determine whether two groups have a large difference or not?”

11.5 Prices of Homes

Cool Down: 5 minutes

Addressing

- 7.SP.B.3

Student Task Statement

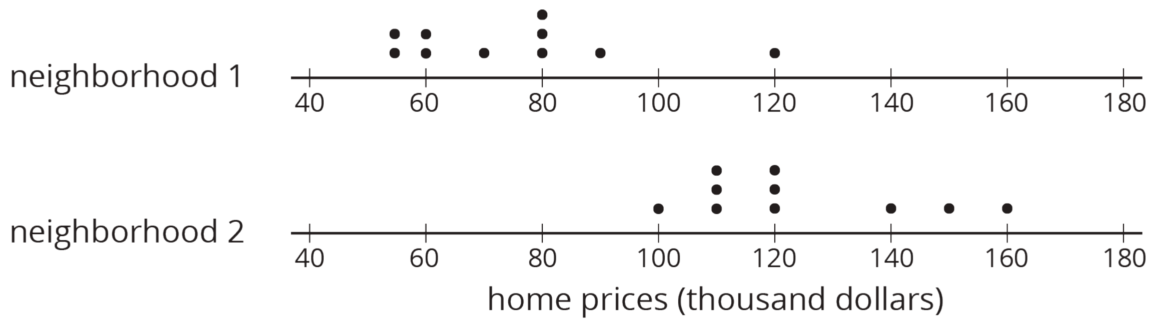
Noah's parents are interested in moving to another part of town. They look up all the prices of the homes for sale and record them in thousands of dollars.

neighborhood 1

80	55	80	120	60	90	60	80	55	70

neighborhood 2

110	120	160	110	100	110	140	150	120	120



Find the mean and MAD for each of the neighborhoods. Then decide whether the two groups are very different or not.

Student Response

Neighborhood 1: Mean = 75. MAD = 15.

Neighborhood 2: Mean = 124. MAD = 15.6.

They are very different since the difference in means is 49, which is more than 3 times the larger MAD.

Student Lesson Summary

Comparing two individuals is fairly straightforward. The question "Which dog is taller?" can be answered by measuring the heights of two dogs and comparing them directly. Comparing two groups can be more challenging. What does it mean for the basketball team to generally be taller than the soccer team?

To compare two groups, we use the distribution of values for the two groups. Most importantly, a measure of center (usually **mean** or **median**) and its associated measure of variability (usually **mean absolute deviation** or *interquartile range)* can help determine the differences between groups.

For example, if the average height of pugs in a dog show is 11 inches, and the average height of the beagles in the dog show is 15 inches, it seems that the beagles are generally taller. On the other hand, if the MAD is 3 inches, it would not be unreasonable to find a beagle that is 11 inches tall or a pug that is 14 inches tall. Therefore the heights of the two dog breeds may not be very different from one another.

Glossary

- mean
- mean absolute deviation (MAD)

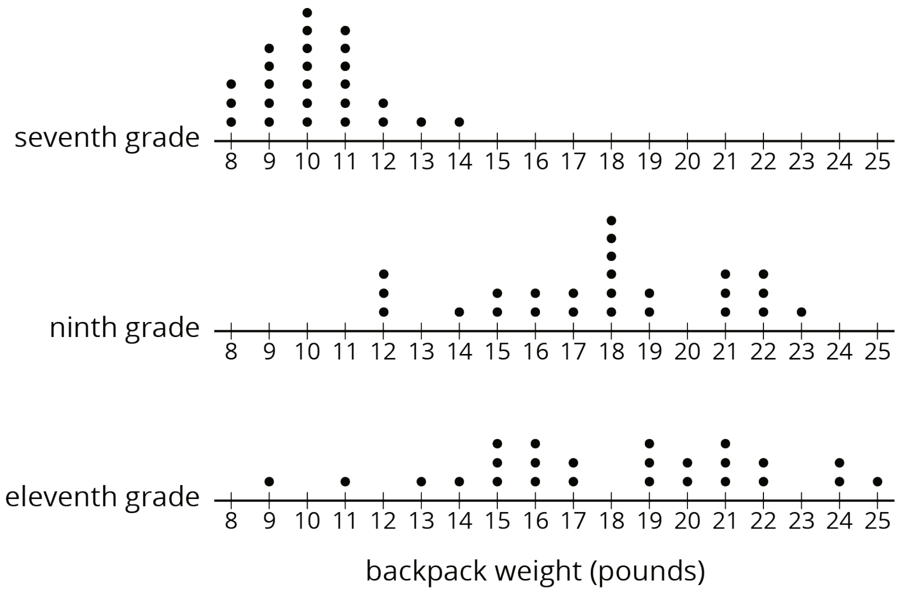
- median

Lesson 11 Practice Problems

Problem 1

Statement

Compare the weights of the backpacks for the students in these three classes.



Solution

Answers vary. Sample response:

The backpacks of the seventh graders tend to weigh less than the backpacks of the ninth graders. A typical weight for the seventh graders' packs is about 10 pounds, compared to a typical weight of about 18 pounds for the ninth graders' packs. The weights were also less variable for the seventh graders than the ninth graders. Similar things can be said when comparing seventh and eleventh graders' backpacks.

The distribution of weights for the ninth graders and the eleventh graders are similar, but the eleventh graders had a slightly larger spread. Both distributions are centered at around 18 pounds, and backpack weights varied quite a bit from student to student in both of these grades.

Problem 2

Statement

A bookstore has marked down the price for all the books in a certain series by 15%.

a. How much is the discount on a book that normally costs $18.00?

b. After the discount, how much would the book cost?

Solution

a. \$2.70, because $18 \cdot 0.15 = 2.7$.

b. \$15.30, because $18 - 2.7 = 15.3$.

(From Unit 4, Lesson 11.)

Problem 3

Statement

Match each expression in the first list with an equivalent expression from the second list.

A. $6(x + 2y) - 2(y - 2x)$

B. $2.5(2x + 4y) - 5(4y - x)$

C. $4(5x - 3y) - 10x + 6y$

D. $5.5(x + y) - 2(x + y) + 6.5(x + y)$

E. $7.9(5x + 3y) - 4.2(5x + 3y) - 1.7(5x + 3y)$

1. $10(x - y)$
2. $10(x + y)$
3. $10x + 6y$
4. $10x - 6y$

Solution

- A: 2
- B: 1
- C: 4
- D: 2
- E: 3

(From Unit 6, Lesson 22.)

Problem 4

Statement

Angles C and D are complementary. The ratio of the measure of Angle C to the measure of Angle D is $2 : 3$. Find the measure of each angle. Explain or show your reasoning.

Solution

Angle C: 36°, Angle D: 54°. The two angle measures must add to 90° and since $\frac{2}{5}$ of 90° comes from angle C, it must have measure 36°. Then Angle D comes from $90 - 36 = 54$.

(From Unit 7, Lesson 2.)