Comparing Populations Using Samples

Goals

- Calculate the difference between the mean or median of two samples from different populations, and express it as a multiple of the MAD or IQR.
- Interpret a pair of box plots, including the amount of visual overlap between the two distributions.
- Justify (orally and in writing) whether there is likely to be a meaningful difference between two populations based on a sample from each population.

Learning Targets

- I can calculate the difference between two medians as a multiple of the interquartile range.
- I can determine whether there is a meaningful difference between two populations based on a sample from each population.

Lesson Narrative

In this lesson, students compare two groups based on samples from each group. This builds on their work in earlier activities in which they compared populations based on their measures of center and variability.

Students see that if samples of two different populations have only a small difference between their measures of center relative to their variability, then we cannot say that there is a meaningful difference between the measures of center of the populations. Due to sampling variability, it is possible that the two populations may not be very different. However, if samples from two different populations have a large difference between their measures of center relative to their variability, then we can say that there is likely to be a meaningful difference between the measures of center of the two populations.

Student Learning Goal

Let's compare different populations using samples.

Lesson Timeline

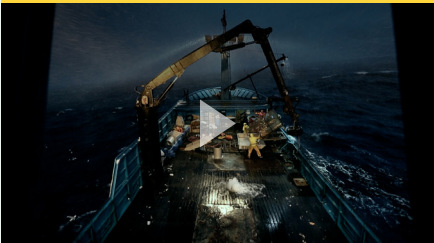| 5 min	10 min	15 min	15 min	10 min	5 min
Warm-up	Activity 1	Activity 2	Activity 3	Lesson Synthesis	Cool-down

Assessment

Inspire Math

Crabbing video

Go Online

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

ilclass.com/l/614247

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

Building on Student Thinking

For students who have a difficult time starting without calculating, help them to compare the values in the ones place for the first and third pairs of data.

Student Workbook

LESSON 18

Comparing Populations Using Samples
Let's compare different populations using samples.

Warm-up   Same Mean? Same MAD?

Without calculating, tell whether each pair of data sets have the same mean and whether they have the same mean absolute deviation.

1. Set A   1   3   3   5   6   8   10   14
   Set B   21   23   23   25   26   28   30   34

2. Set X   1   2   3   4   5
   Set Y   1   2   3   4   5   6

3. Set P   47   53   58   62
   Set Q   37   43   68   72

Warm-up

Same Mean? Same MAD?

5 min

Activity Narrative

This *Warm-up* reminds students of the meanings of mean and MAD by comparing two sets of data with similar values and asking whether they will have the same means or MADs or both.

Launch

Explain to students that the pairs of data sets are the following: A and B, X and Y, and P and Q.

Student Task Statement

Without calculating, tell whether each pair of data sets have the same mean and whether they have the same mean absolute deviation.

1. | Set A | 1 | 3 | 3 | 5 | 6 | 8 | 10 | 14 |
   |-------|---|---|---|---|---|---|----|----|
   | Set B | 21 | 23 | 23 | 25 | 26 | 28 | 30 | 34 |

   Sets A and B have different means but the same MADs.

2. | Set X | 1 | 2 | 3 | 4 | 5 | |
   |-------|---|---|---|---|---|---|
   | Set Y | 1 | 2 | 3 | 4 | 5 | 6 |

   Sets X and Y have different means and different MADs.

3. | Set P | 47 | 53 | 58 | 62 |
   |-------|----|----|----|----|
   | Set Q | 37 | 43 | 68 | 72 |

   Sets P and Q have the same means but different MADs.

Activity Synthesis

The purpose of the discussion is to bring out methods students used to notice whether the pairs of data sets have the same mean or the same MAD or both.

Poll the class for each pair of data sets as to whether they had the same mean, the same MAD, both, or neither.

After students have had a chance to register their vote, ask some students to explain their reasoning for their answer.

Activity 1
With a Heavy Load

10 min

Activity Narrative

Earlier, students compared two groups when data from both populations are known. In this activity, students compare two groups using only samples from each group. Students construct informal arguments to explain why the different samples come from populations that are meaningfully different or not.

Launch 👥

Arrange students in groups of 2.

Allow students 3–5 minutes quiet work time. Follow with partner and whole-class discussions.

Student Task Statement

Consider the question: Do tenth-grade students' backpacks generally weigh more than seventh-grade students' backpacks?

Here are dot plots showing the weights of backpacks for a random sample of students from these two grades:

grade 7

grade 10

backpack weight (pounds)
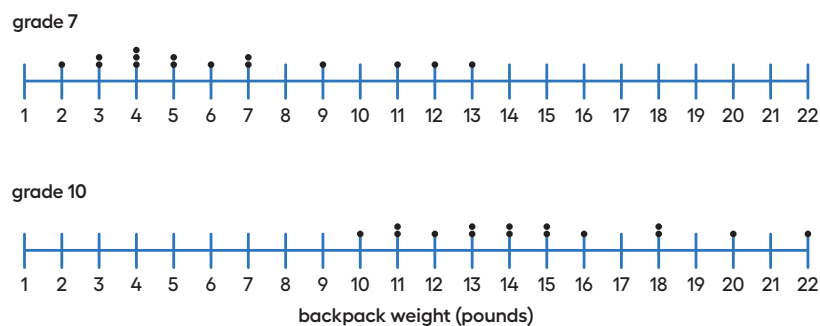
1. Do any seventh-grade backpacks in this sample weigh more than a tenth-grade backpack?

   Yes, three seventh-grade backpacks weigh more than the lightest tenth-grade backpack.

2. The mean weight of this sample of seventh-grade backpacks is 6.3 pounds. Do you think the mean weight of backpacks for all seventh-grade students is exactly 6.3 pounds?

   No. 6.3 pounds is the mean of a sample, and the population mean will probably be at least a little different.

3. The mean weight of this sample of tenth-grade backpacks is 14.8 pounds. Do you think there is a meaningful difference between the weight of all seventh-grade and tenth-grade students' backpacks? Explain or show your reasoning.

   Sample response: There is still probably a meaningful difference in the mean weights since there is very little overlap.

Student Workbook

Activity Synthesis

The purpose of the discussion is for students to think about how comparing groups by using data from samples differs from comparing groups when the population is known.

Ask partners to share their decision about whether the groups had a meaningful difference with the class.

Consider asking these questions for discussion:

💬 *"Compare what you did in this activity to how you compared two populations earlier in the unit."*

Earlier, we had data from the entire population, and here it is only a sample. That means there may be more hidden overlap between the groups when you only know some of the data, and you should be more careful about the comparison.

💬 *"Is the overlap of the data more important when you only have a sample or when you have data from the population? Explain your reasoning."*

It is more important when you only have a sample. If there is overlap with only some of the data, it's possible there is more overlap when we include more data from the population.

💬 *"Is it possible that the data in the two samples were drawn from population data that is identical?"*

It is unlikely but possible.

Activity 2

Do They Carry More?

15 min

Activity Narrative

In this activity, students look at different samples from the same population to see that their means are relatively close based on the MADs of the samples to motivate a general rule for determining whether two groups are meaningfully different. This concept can be reversed to say that if two samples have means that are not very close, then the samples likely came from populations that are quite different. A general rule is given to determine whether two populations are meaningfully different based on the mean and MAD from a sample of each.

Launch 👥

Keep students in groups of 2.

Allow students 5 minutes of partner work time, and then pause the class to assign samples and explain the general rule.

Ask students to pause after the third question in order to explain the general rule and assign a sample to each group. After all students have paused, assign each group 1 of the 10 samples to work with for the last two questions. Further, explain to students:

- As a general rule, if two populations have the same mean (or median) and similar variability, the sample means (or medians) should be within 2 MADs (or IQRs) of one another.

- If the sample means (or medians) are more than 2 MADs (or IQRs) apart, it is very likely that the population means (or medians) are different. We will say that there is a *meaningful difference* between the two population means (or medians).

- If the sample means (or medians) are less than or equal to 2 MADs (or IQRs) apart, it is more difficult to say that the two population means (or medians) are very different. In this case we will say that the samples do not provide evidence that the population means (or medians) differ.

Give students 5 more minutes of partner work time, and follow with a whole-class discussion.

Student Task Statement

Here are 10 random samples from the same population of seventh-grade students' backpack weights.
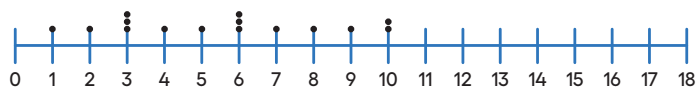
sample 1, mean: 5.8 pounds
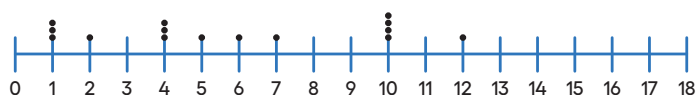
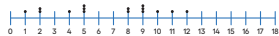**sample 2, mean: 9.2 pounds**
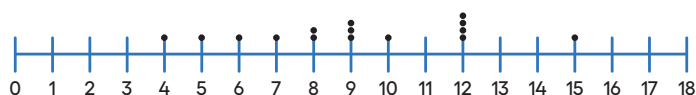
sample 3, mean: 5.5 pounds

Student Workbook

sample 4, mean: 7.3 pounds

sample 5, mean: 7.2 pounds

sample 6, mean: 6.6 pounds

sample 7, mean: 5.2 pounds

sample 8, mean: 5.3 pounds
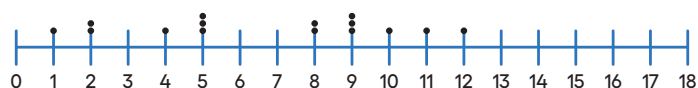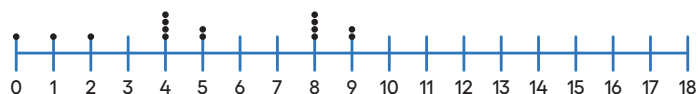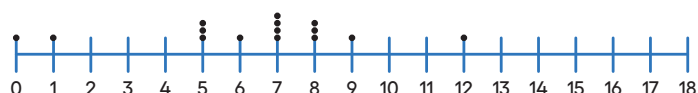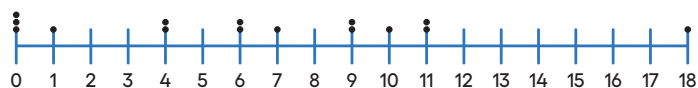

sample 9, mean: 6.3 pounds


sample 10, mean: 6.4 pounds

1. a. Which sample has the greatest mean weight?

Sample 2

b. Which sample has the least mean weight?

Sample 7

c. What is the difference between these two sample means?

4 pounds

Sample resasoning: 9.2 – 5.2 = 4

2. All of the samples have a mean absolute deviation of about 2.8 pounds. Express the difference between the greatest and least sample means as a multiple of the MAD.

$4 \approx 2.8 \cdot 1.43$ since $4 \div 2.8 \approx 1.43$

3. Are these samples very different? Explain or show your reasoning.

No, these samples are not very different. Sample reasoning: All of the samples came from the same population and are all within 2 MADs of one another.

4. A sample of tenth-grade students' backpacks has a mean weight of 14.8 pounds. The MAD for this sample is 2.7 pounds. Your teacher will assign you one of the samples of seventh-grade students' backpacks to use.

a. What is the difference between the sample means for the tenth-grade students' backpacks and the seventh-grade students' backpacks?

b. Express the difference between these two sample means as a multiple of the larger of the MADs.

Sample responses:

sample	1	2	3	4	5	6	7	8	9	10
difference in means	9	5.6	9.3	7.5	7.6	8.2	9.6	9.5	8.5	8.4
multiplier for the MAD	3.2	2	3.3	2.7	2.7	2.9	3.4	3.4	3.0	3.0

5. Do you think there is a meaningful difference between the weights of all seventh-grade and tenth-grade students' backpacks? Explain or show your reasoning.

Yes, there is a meaningful difference

Sample reasoning: The difference in the means is more than 2 MADs, so the weights of the two populations have a meaningful difference.

Student Workbook
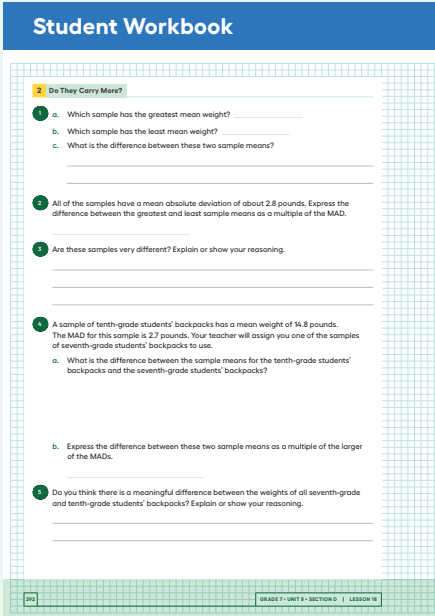
Activity Synthesis

The purpose of the discussion is for students to understand the general rule for determining if two samples suggest a meaningful difference between their populations.

Select at least one group assigned to each of the samples to share their responses to the last two questions and record for all to see. Note that all 10 samples from the seventh-grade students have means that are within 2 MADs of one another, but the mean from the tenth-grade-student sample is at least 2 MADs away from the mean of each of the seventh-grade-student samples.

Note that the general rule only has two possible outcomes: "There is a meaningful difference." or "There is not enough information to say there is a meaningful difference." If the means are less than 2 MADs apart, the general rule cannot say whether two samples were drawn from populations that contain identical data.

Ask students,

💬 *"Based only on the dot plots for the 10 samples, would you have guessed that they all might have come from the same population? Explain your reasoning."*

Maybe. There is a lot of overlap among all of the samples.

Activity 3

Steel from Different Regions

15 min

Activity Narrative

In this activity, students practice using the general rule developed in an earlier activity by estimating the measure of center for a population and comparing populations based on those estimates as well as the associated measure of variability. Students must construct viable arguments for their conclusions about whether there is a meaningful difference between the groups.

Launch 👥

Keep students in groups of 2.

Explain to students that different regions had different raw materials and techniques for constructing metal. One way of testing ancient metal is by looking at the carbon content in the steel. In some cases, this content could determine the region where the metal was made.

Ask students how the general rule from the previous activity might be adapted to use median and interquartile range (IQR) rather than mean and MAD.

Allow students 10 minutes of partner work time, and follow with a whole-class discussion.

Student Task Statement

When anthropologists find steel artifacts, they can test the amount of carbon in the steel to learn about the people that made the artifacts. Here are the box plots showing the percentage of carbon in samples of steel that were found in two different regions:
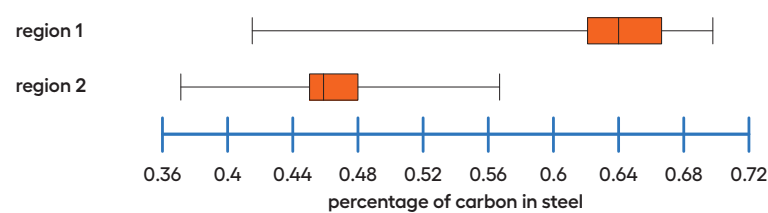
1. Is there any steel found in region 1 that has:

   a. more carbon than some of the steel found in region 2?

      Yes. Most of the steel from region 1 had more carbon in it than steel from region 2.

   b. less carbon than some of the steel found in region 2?

      Yes. Since the left end of the region 1 box plot overlaps with the box plot for region 2, there was at least 1 piece of steel that had less carbon in it than some of the steel from region 2.

2. Based only on the box plots, do you think there is a meaningful difference between all the steel artifacts found in regions 1 and 2?

   Sample response: Based on the box plots, there is some overlap, but the boxes look so far apart that I think there will be a meaningful difference.

3. Which sample has a distribution that is not approximately symmetric?

   The distribution for region 1 is not symmetric since it has the very long segment on the left and the short segment on the right.

4. A sample of artifacts known to come from region 1 has a median of 0.64% carbon in the steel and an interquartile range of 0.05%.

   A sample from region 2 has a median of 0.47% carbon in the steel and an IQR of 0.03%.

   What is the difference between the sample medians for these two regions?

   0.17%, since 0.64 – 0.47 = 0.17

5. Express the difference between these two sample medians as a multiple of the larger interquartile range.

   0.17 ≈ 0.05 · 3.4

6. The anthropologists who conducted the study conclude that there is a meaningful difference between the steel from these regions. Do you agree? Explain or show your reasoning.

   I agree with the anthropologists.

   Sample reasoning: There is evidence of a meaningful difference because the difference in sample medians is greater than 2 IQRs.

Access for Students with Diverse Abilities (Activity 3, Student Task)

Action and Expression: Internalize Executive Functions.
To support development of organizational skills in problem solving, chunk this task into more manageable parts. For example, check in with students after the first three questions analyzing the box plots.

Supports accessibility for: Organization, Attention

Student Workbook

Student Workbook

Activity Synthesis

The purpose of the discussion is for students to understand how to adapt the general rule for determining a meaningful difference between populations to median and IQR.

Consider asking these questions for discussion:

💬 *"Why did this problem use median and IQR instead of mean and MAD?"*

Since the distribution for region I is not symmetric, it makes more sense to use the median. Also the box plots will show the median and IQR, but there is not a good way to know the mean and MAD.

💬 *"Is there any overlap in the data from the two regions?"*

Yes. The smallest percentage of carbon from the region I is well below the median from region 2, while the typical percentage of carbon from region I is much greater than from region 2.

💬 *"On the box plot in the activity, draw a dot two IQRs above the median for region 2. Then draw a star two IQRs below the median for region 1. How do these help you see that there is a meaningful difference in the medians?"*

The dot is at 0.53 and the star is at 0.54. Since the median for region I is not below the dot nor is the median for region I above the star, there must be a meaningful difference.

💬 *"A piece of steel is found in a place between the two regions sampled. Would testing the percentage of carbon from this metal be useful in determining the region from which it came?"*

Yes. Since there is a meaningful difference in the percentage of carbon in the steel from the two regions, it should give a good indication which region created the metal.

Lesson Synthesis

Consider asking these discussion questions to emphasize the main ideas from this lesson:

💬 *"When is it useful to use a median rather than a mean?"*

It is useful when the distribution is not approximately symmetric.

💬 *"What values do you need to calculate from a sample to use the general rule for determining if the measures of center of two populations are meaningfully different?"*

The measure of center and measure of variation for each sample should be calculated to compare the groups.

💬 *"What is the general rule used to determine if the means of two populations are meaningfully different?"*

If the difference between the means for the two samples is greater than twice the greater of the MADs, then the means are meaningfully different.
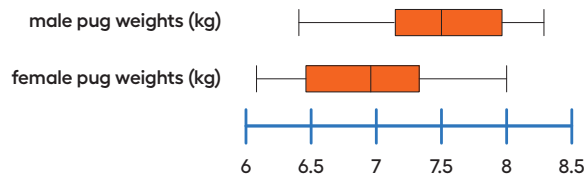
Lesson Summary

Sometimes we want to compare two different populations. For example, is there a meaningful difference between the weights of pugs and beagles? Here are histograms showing the weights for a sample of dogs from each of these breeds:
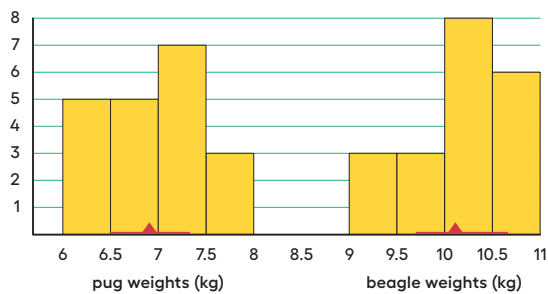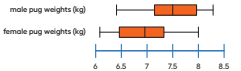

The red triangles show the mean weight of each sample, 6.9 kg for the pugs and 10.1 kg for the beagles. The red lines show the weights that are within 1 MAD of the mean. We can think of these as typical weights for the breed. These typical weights do not overlap. In fact, the distance between the means is 10.1 – 6.9, or 3.2 kg, over 6 times the larger MAD! So we can say there is a meaningful difference between the weights of pugs and beagles.

Is there a meaningful difference between the weights of male pugs and female pugs? Here are box plots showing the weights for a sample of male pugs and a sample of female pugs:

We can see that the medians are different, but the weights between the first and third quartiles overlap. Based on these samples, we would say there is not a meaningful difference between the weights of male pugs and female pugs.

In general, if the measures of center for two samples are at least two measures of variability apart, we say the difference in the measures of center is meaningful. Visually, this means the ranges of typical values do not overlap. If they are closer, then we don't consider the difference to be meaningful.

<div style="border:1px solid #ccc;">

Responding To Student Thinking

Points to Emphasize

If students struggle with using the mean and MAD to compare groups, revisit strategies for how to do this as opportunities arise over the next several lessons. For example, the practice problem referred to here can be used:

Unit 8, Lesson 19, Practice Problem 1

</div>

Cool-down   **5 min**

Teachers Watching Movies

Student Task Statement

Noah is interested in comparing the number of movies watched by students and teachers over the winter break. He takes a random sample of 10 students and 10 teachers and makes a dot plot of their responses.

students

movies watched over break

teachers

movies watched over break
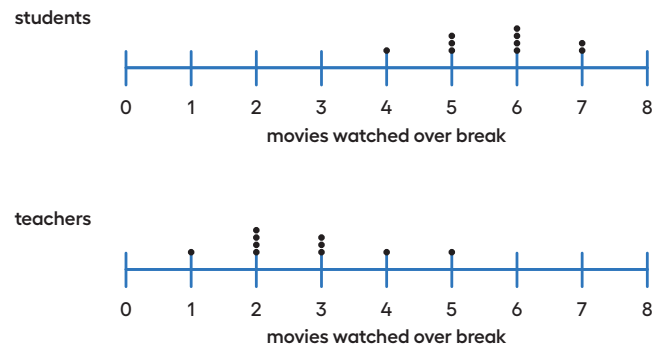
Noah then computes the measures of center and variability for each group:

- Students: mean: 5.7 movies, MAD: 0.76 movies

- Teachers: mean: 2.7 movies, MAD: 0.9 movies

Should Noah conclude that there is a meaningful difference in the mean number of movies watched over winter break between the two groups? Explain your reasoning.

Yes

Sample reasoning: Because the difference in the means is greater than 2 MADs, there is a meaningful difference in the mean number of movies watched. $(5.7 - 2.7) \div 0.9 \approx 3.33$

Practice Problems
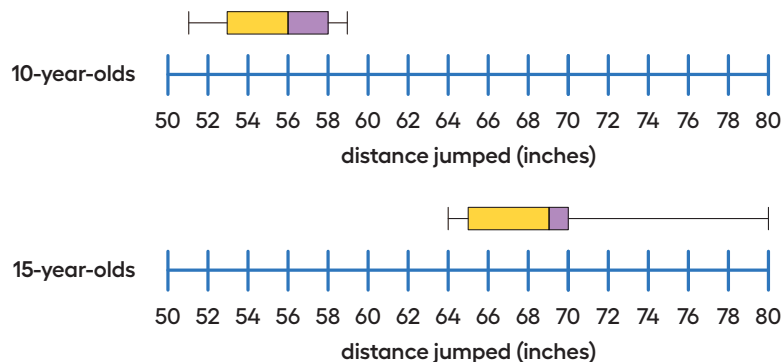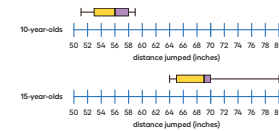
4 Problems

Problem 1

Lin wants to know if students in elementary school generally spend more time playing outdoors than students in middle school. She selects a random sample of size 20 from each population of students and asks them how many hours they played outdoors last week. Suppose that the MAD for each of her samples is about 3 hours.

Select **all** pairs of sample means for which Lin could conclude there is a meaningful difference between the two populations.

A. elementary school: 12 hours, middle school: 10 hours

B. elementary school: 14 hours, middle school: 9 hours

C. elementary school: 13 hours, middle school: 6 hours

D. elementary school: 13 hours, middle school: 10 hours

E. elementary school: 7 hours, middle school: 15 hours

Problem 2

These two box plots show the distances of a standing jump, in inches, for a random sample of 10-year-olds and a random sample of 15-year-olds.

Is there a meaningful difference in median distance for the two populations? Explain how you know.

Yes, the difference in medians is 13 inches. This difference is more than 2 IQRs (the IQR is 5, and 13 ÷ 5 = 2.6), so there is a meaningful difference in the median distances for 10-year-olds and 15-year-olds.

Problem 3

The median income for a sample of people from Chicago is about $60,000 and the median income for a sample of people from Kansas City is about $46,000, but researchers have determined there is not a meaningful difference in the medians. Explain why the researchers might be correct.

The medians differ by $14,000, but if the IQR is larger than about $7,000, there will not be a meaningful difference between the median salaries in the two cities.

Problem 4   from Unit 8, Lesson 17

A farmer grows 5,000 pumpkins each year. The pumpkins are priced according to their weight, so the farmer would like to estimate the mean weight of the pumpkins he grew this year. He randomly selects 8 pumpkins and weighs them. Here are the weights (in pounds) of these pumpkins:

| 2.9 | 6.8 | 7.3 | 7.7 | 8.9 | 10.6 | 12.3 | 15.3 |

a. Calculate the mean weight of the pumpkins the farmer grew.

8.975 pounds

This dot plot shows the mean weight of 100 samples of 8 pumpkins, similar to the sample above.
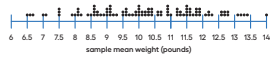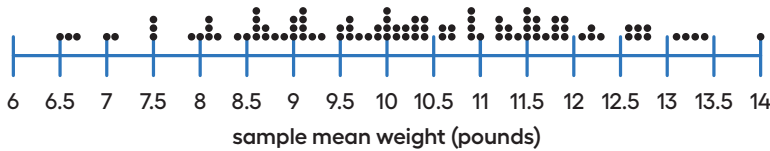
b. What appears to be the mean weight of the 5,000 pumpkins?

about 10 pounds

c. What does the dot plot of the sample means suggest about how accurate an estimate based on a single sample of 8 pumpkins might be?

The sample means range from about 6.5 to 14 pounds. If the actual population mean is about 10 pounds, this shows that a sample mean based on a sample of size 8 might not be very close to the actual population value.

d. What do you think the farmer might do to get a more accurate estimate of the population mean?

He could use a larger sample size.