NAME DATE PERIOD

Unit 8, Lesson 18: Comparing Populations Using Samples

Let's compare different populations using samples.

18.1: 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.

set A	1	3	3	5	6	8	10	14

set B	21	23	23	25	26	28	30	34

set X	1	2	3	4	5

set Y	1	2	3	4	5	6

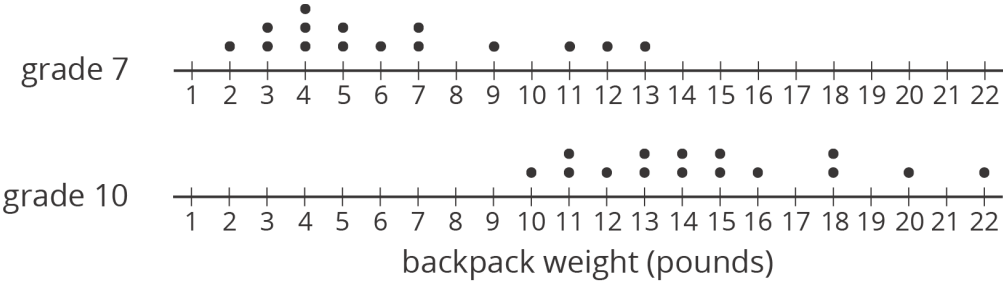
set P	47	53	58	62

set Q	37	43	68	72

18.2: With a Heavy Load

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:



1. Did any seventh-grade backpacks in this sample weigh more than a tenth-grade backpack?

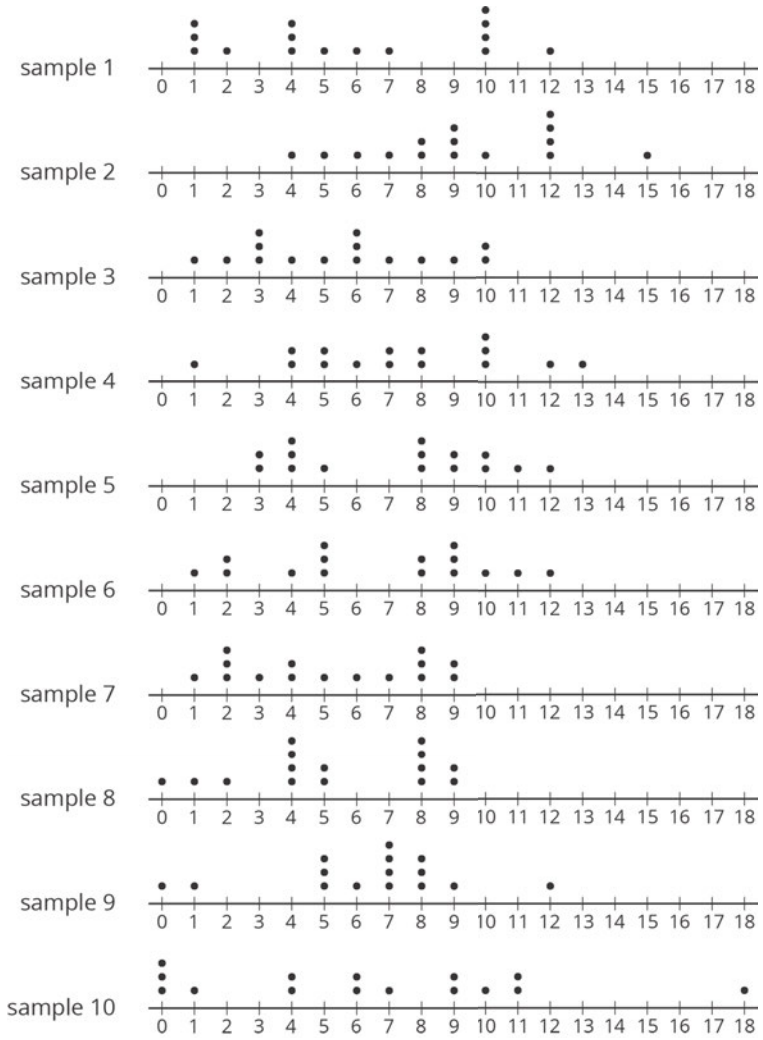
NAME DATE PERIOD

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?

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.

18.3: Do They Carry More?

Here are 10 more random samples of seventh-grade students' backpack weights.



sample number	mean weight (pounds)
1	5.8
2	9.2
3	5.5
4	7.3
5	7.2
6	6.6
7	5.2
8	5.2
9	6.3
10	6.4

NAME DATE PERIOD

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

   b. Which sample has the lowest mean weight?

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

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

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

Remember our sample of tenth-grade students' backpacks had 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.

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

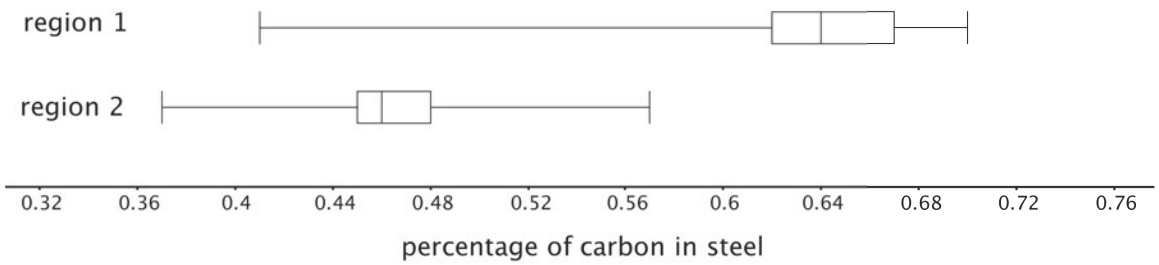
   b. Express the difference between these two sample means as a multiple of the larger of the MADs.

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.

NAME DATE PERIOD

18.4: Steel from Different Regions

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 some box plots showing the percentage of carbon in samples of steel that were found in two different regions:



1. Was there any steel found in region 1 that had:

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

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

2. Do you think there is a meaningful difference between all the steel artifacts found in regions 1 and 2?

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

	sample median (%)	IQR (%)
region 1	0.64	0.05
region 2	0.47	0.03

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

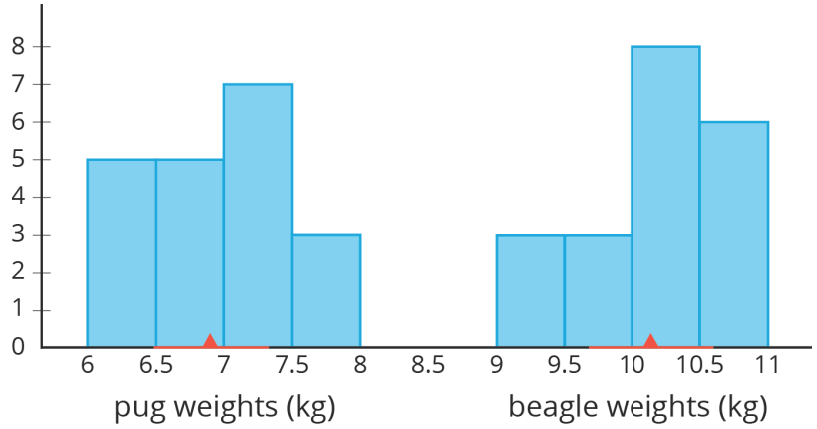
5. Express the difference between these two sample medians as a multiple of the larger interquartile range.

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

NAME DATE PERIOD

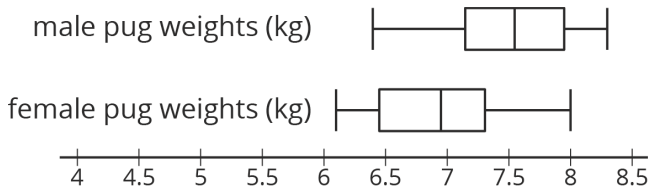
Lesson 18 Summary

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 and 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 range of typical values does not overlap. If they are closer, then we don't consider the difference to be meaningful.

NAME DATE PERIOD

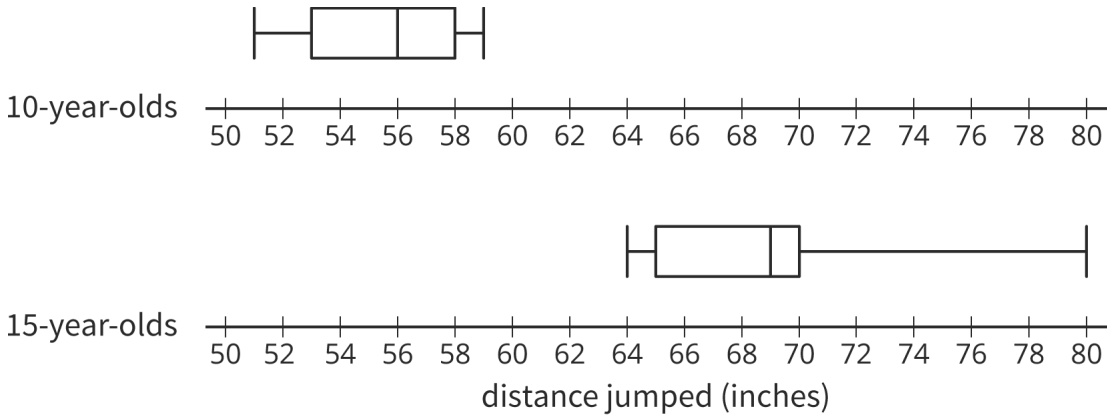
Unit 8, Lesson 18: Comparing Populations Using Samples

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

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

NAME DATE PERIOD

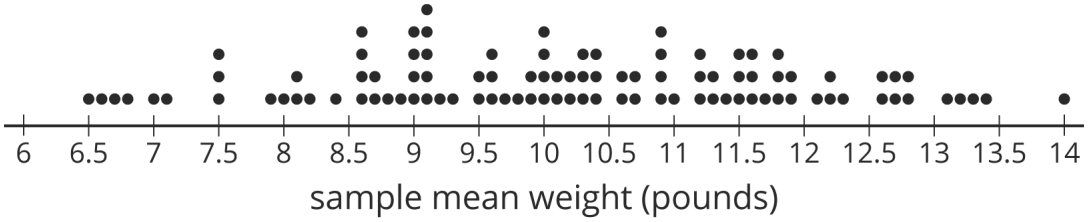
3. The median income for a sample of people from Chicago is about $60,000 and the median income for a sample of people from Kansas City is about $46,000, but researchers have determined there is not a meaningful difference in the medians. Explain why the researchers might be correct.

4. 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:

weight (pounds)	2.9	6.8	7.3	7.7	8.9	10.6	12.3	15.3

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

This dot plot shows the mean weight of 100 samples of eight pumpkins, similar to the one above.



b. What appears to be the mean weight of the 5,000 pumpkins?

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?

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

(from Unit 8, Lesson 17)