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.

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.

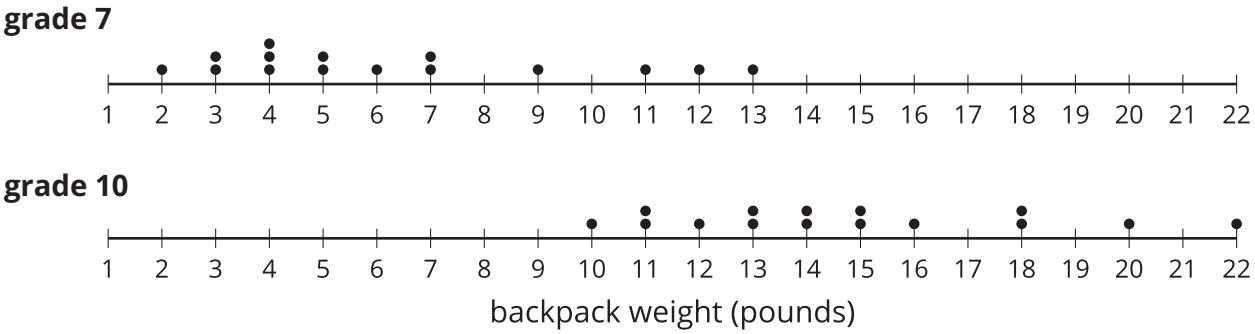
Set P	47	53	58	62
Set Q	37	43	68	72

CC BY NC Illustrative Mathematics, based on IM 6–8 Math, CC BY Open Up Resources.

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

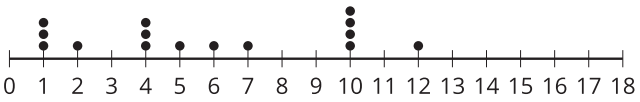
2. The mean weight of this sample of seventh-grade backpacks is 6.3 pounds. Do you think the mean weight of backpacks for all seventh-grade students is exactly 6.3 pounds?

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 random samples from the same population of seventh-grade students' backpack weights.

sample 1, mean: 5.8 pounds



sample 2, mean: 9.2 pounds



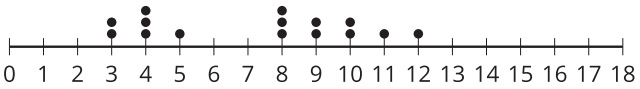
sample 3, mean: 5.5 pounds



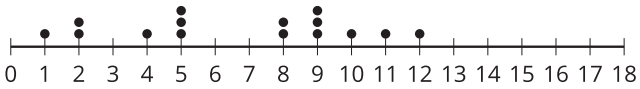
sample 4, mean: 7.3 pounds



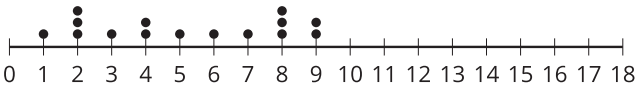
sample 5, mean: 7.2 pounds



sample 6, mean: 6.6 pounds



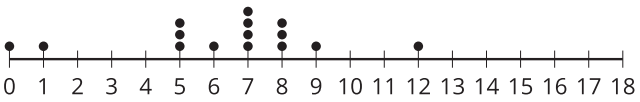
sample 7, mean: 5.2 pounds



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?

   b. Which sample has the least 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 the greatest and least sample means as a multiple of the MAD.

CC BY NC Illustrative Mathematics, based on IM 6–8 Math, CC BY Open Up Resources.

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

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?

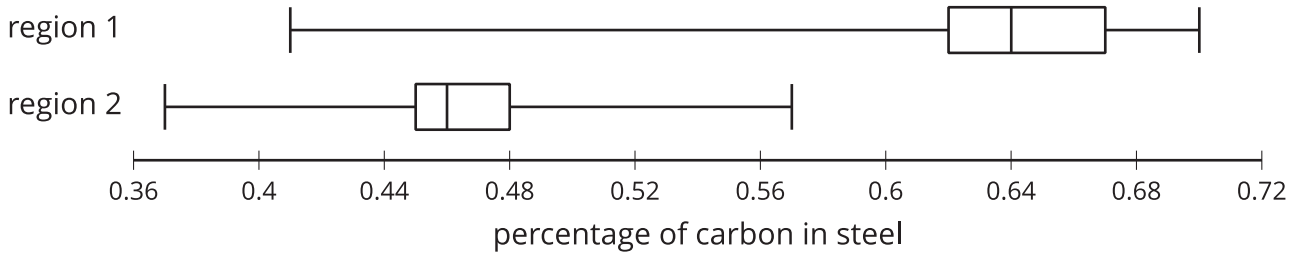
   b. Express the difference between these two sample means as a multiple of the larger of the MADs.

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.

CC BY NC Illustrative Mathematics, based on IM 6–8 Math, CC BY Open Up Resources.

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

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

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?

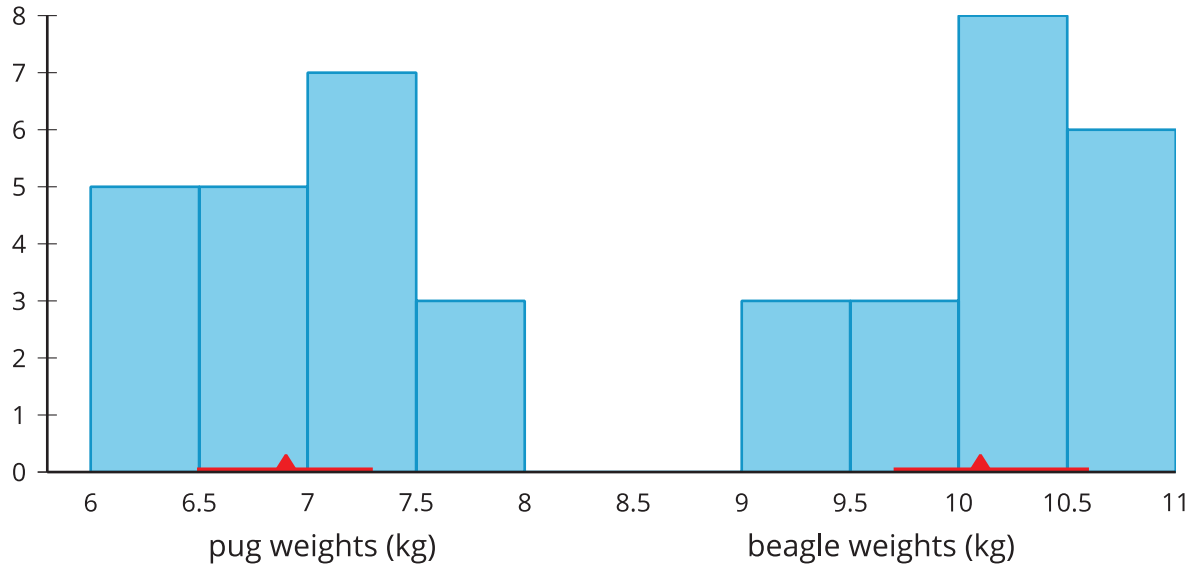
5. Express the difference between these two sample medians as a multiple of the larger interquartile range.

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.

CC BY NC Illustrative Mathematics, based on IM 6–8 Math, CC BY Open Up Resources.

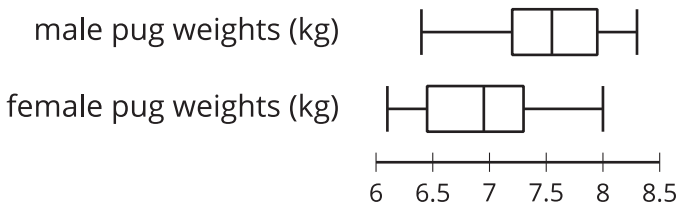
Lesson 18 Summary

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 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.

CC BY NC Illustrative Mathematics, based on IM 6–8 Math, CC BY Open Up Resources.