

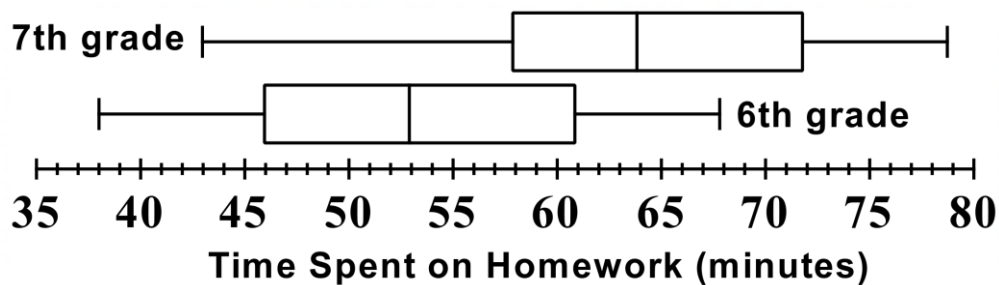
COMPARING SAMPLES

N-GEN MATH[®] 7



Statistics is often used to compare two **populations** by comparing two **samples** taken from those populations. Sophisticated techniques involving **probability** are used to justify whether two samples come from different populations. You will explore those in future courses. In this lesson we will do so more informally.

Exercise #1: A middle school took a **sample** of 25 6th grade students and 25 7th grade students to see if 7th grade students spend more time on homework than 6th grade students. The results are shown in the double box plot below.



- (a) How much greater is the median amount of time spent by 7th graders compared to 6th graders?
- (b) State the IQR of time that 6th graders and 7th graders spend on homework.
- (c) What do your answers to (b) tell you about the variation of the middle half of both data sets?
- (d) By how much do the samples **overlap**? A small amount or a large amount?
- (e) How does the difference between the two medians compare to their interquartile ranges? What does this tell you in terms of making an **inference** about the **populations**?



Unfortunately, the tools we have now in statistics do not allow us to **definitively** say whether two **populations** differ enough to be detected by samples. But we can do it informally.

Exercise #2: A study was done to answer the question: “Are cars more fuel efficient than trucks?” A sample of 20 cars and a sample of 20 trucks were taken and the gas mileage of each vehicle (how many miles can be driven per gallon of fuel) was recorded. The results are shown below.

Trucks:

Mean = 17 miles per gallon

MAD = 4 miles per gallon

Cars:

Mean = 29 miles per gallon

MAD = 5 miles per gallon

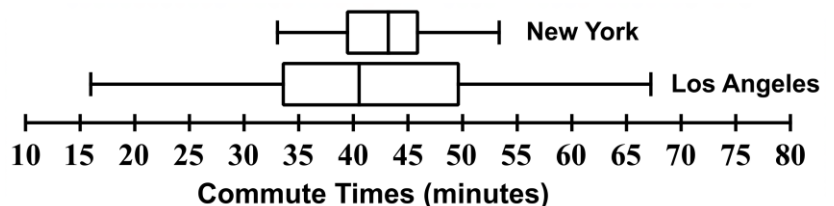
- (a) How much higher was the mean fuel efficiency for cars than for trucks? (b) What can you say about the variation in the two data sets?
- (c) Based on your answer to (a) and the mean absolute deviations, how confident would you be in concluding that the average fuel efficiency for all cars is greater than the fuel efficiency for all trucks (i.e., the populations)?

INFORMAL RULE FOR INFERRING FROM SAMPLES TO POPULATIONS

If the **difference in the medians** of the samples is **greater than their interquartile range** or the **difference in the means** of the samples is **greater than twice their mean absolute deviation**, you can say that one **population** mean (or median) is likely greater than the other.

Exercise #3: The amount of time it takes to get to work is called the commute time. Samples of 50 commute times for Los Angeles and New York are shown below.

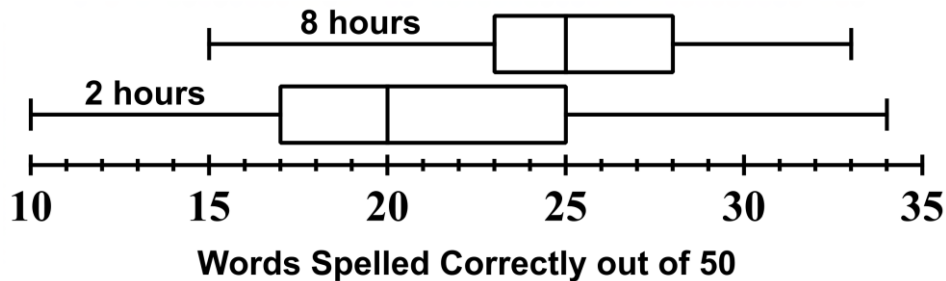
Based on the box plot, can you conclude that the median commute time for all New York City residents is greater than that for residents of Los Angeles?



COMPARING SAMPLES
N-GEN MATH[®] 7 HOMEWORK

USING YOUR MATH

1. A study was done to see whether sleep made a difference in a person's ability to spell words. Scientists had one group of 30 people sleep only 2 hours in a night while another group of 30 people slept 8 hours. Each group was then given a spelling test with 50 words on it. The results are shown in the box plot below.



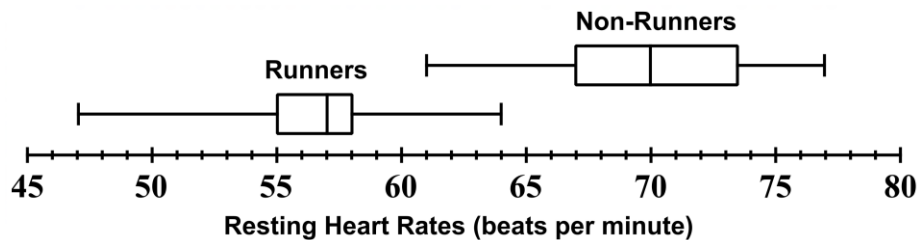
- (a) How much greater was the median for the 8-hour group than the 2-hour group?
- (b) Which of the two samples had a greater variation in the middle half of the data? Justify.
- (c) How does the difference in the sample medians compare to the interquartile ranges of the two samples?
- (d) Based on your answer to (c), can you infer from these samples that the population of all people who only get 2 hours of sleep would have a lower median than all people who get 8 hours of sleep?



2. Your **resting heart rate** is how many beats per minute your heart makes when you are not exercising. For her science fair project, Camila asked the following question:

“Is the resting heart rate of people who run regularly lower than that of those who do not?”

She collected resting heart rates from 40 people who run regularly and from 40 people who do not run regularly. The results are shown below.



- (a) Is the amount of overlap between the two samples a lot, a little, or none at all?

- (b) What is the difference in the median resting heart rates of the two groups?

- (c) How confident would you be in stating as the answer to the question:

“The typical resting heart rate for those who run is lower than for those who don’t run.”

Explain.

3. Eli was trying to answer the question: “Are 25-year-old men on average taller than boys who are 17 years old?” He measures the heights of 30 boys 17 years old and 30 men who are 25 years old. The mean height for those who were 17 was 68.2 inches and the mean height for those who were 25 was 69.7 inches. Both samples had a mean absolute deviation of around 5 inches.

From this information, can Eli conclude that on average men who are 25 years old are taller than boys who are 17 years old? Explain.

