

Grade 7 - Comparing Two Populations - Lesson

Common Core 7.SP.B.3, 7.SP.B.4 · maths-pdf.fr/en/

To compare two data sets, compare their centers with the mean or median and their spreads with the mean absolute deviation (MAD) or the interquartile range (IQR). Dot plots and box plots show how much the distributions overlap. A difference between centers matters more when it is large compared with the variability of the data.

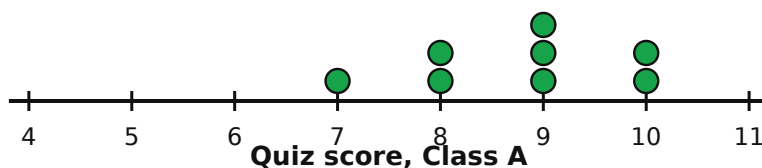
Grade 7 · Common Core State Standards: **7.SP.B.3, 7.SP.B.4**

Key facts

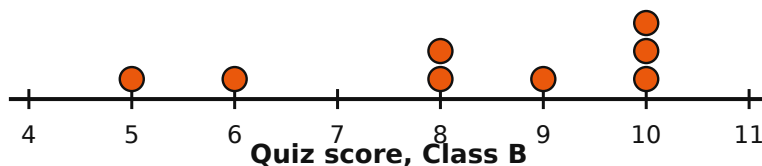
- Center: mean or median. Spread: MAD, range, or IQR.
- MAD is the mean of the distances from the mean.
- A box plot shows the minimum, Q1, median, Q3, and maximum; $IQR = Q3 - Q1$.
- Express the difference in centers as a multiple of the variability.
- Random samples support informal inferences about populations.

How do you compare the centers of two data sets?

Use the **mean** or the **median** to describe the center. Two quiz score sets can have different centers: Class A has a mean of 8.75 and Class B has a mean of 8.25.



Class A scores (mean 8.75): the dots are bunched close to the center.



Class B scores (mean 8.25): the dots are more spread out.

How do you measure how spread out the data are?

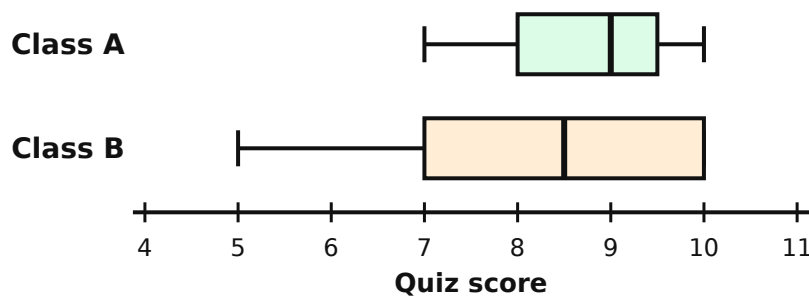
Mean absolute deviation (MAD)

The **MAD** is the mean distance of the data values from the mean. Find the mean, find each distance from it, and average those distances. A small MAD means the data are close together.

For Class A, the distances from 8.75 average 0.81. For Class B, the MAD is 1.5, so Class B is more spread out.

How do box plots help?

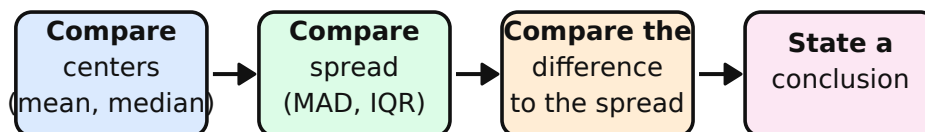
A box plot shows the **five-number summary**: minimum, first quartile Q1, median, third quartile Q3, and maximum. The **IQR** is $Q3 - Q1$, the width of the box.



On box plots, compare the medians, the lengths of the boxes (IQR), and how much the boxes overlap.

How do you decide if a difference is meaningful?

Compare the difference between the centers with the spread. If the difference is much larger than the typical spread, the groups really are different. If the distributions overlap a lot, the difference may be small. For example, a gap of 8 points between means with a MAD of 4 is a difference of $8 \div 4 = 2$ MADs.



Always connect the numbers to the story.

These ideas also let you make an informal inference: if random samples from two populations show a clear gap compared to their spread, you can say the populations probably differ, but

you should check that the samples are random and large enough.

Questions about comparing two populations

What is the mean absolute deviation?

The MAD is the average distance of the data values from the mean. For 4, 6, 8, 10, 12 the mean is 8 and the MAD is 2.4.

How do you find the interquartile range?

Order the data, find Q1 (median of the lower half) and Q3 (median of the upper half), and subtract: $IQR = Q3 - Q1$.

When should I use the median instead of the mean?

Use the median when the data have extreme values or are skewed, because the median is not pulled by outliers.

How can I tell if two data sets overlap?

Compare their ranges and boxes on the plots. If much of one box lies inside the range of the other, the distributions overlap a lot.

What does it mean to express a difference as a multiple of the MAD?

Divide the difference between the means by the MAD. A difference of 8 with MAD 4 is 2 times the typical spread.