

Lesson 15.7: Comparing Populations

Essential Question

How can you compare data sets that represent two populations?

Find the median, IQR, Mean, and MAD for the data sets.

1. 5, 7, 8, 8, 12, 12, 16, 19, 21

Median = 12 IQR = $17.5 - 7.5 = 10$

Mean = 12 MAD = $4.\bar{4}$

$\frac{5+7+8+8+12+12+16+19+21}{9}$ $\frac{7+5+4+4+0+0+4+7+9}{9}$

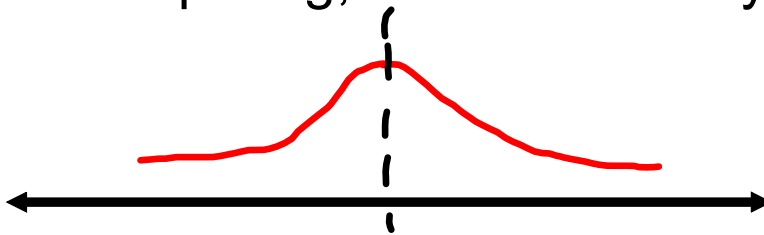
2. 56, 57, 57, 62, 65, 65, 65

Median = 62 IQR = $65 - 57 = 8$

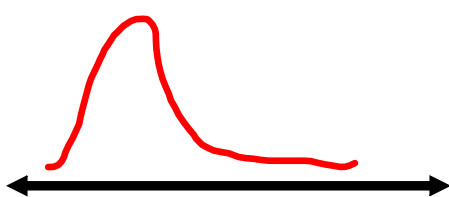
Mean = 61 MAD ≈ 3.7

$\frac{56+57+57+62+65+65+65}{7}$ $\frac{5+4+4+1+4+4+4}{7}$

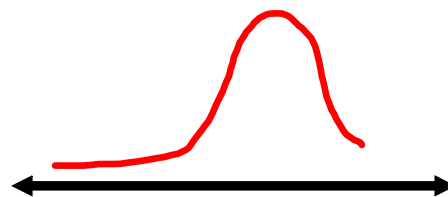
Symmetric distribution: Mean and MAD
(when comparing, both must be symmetric)



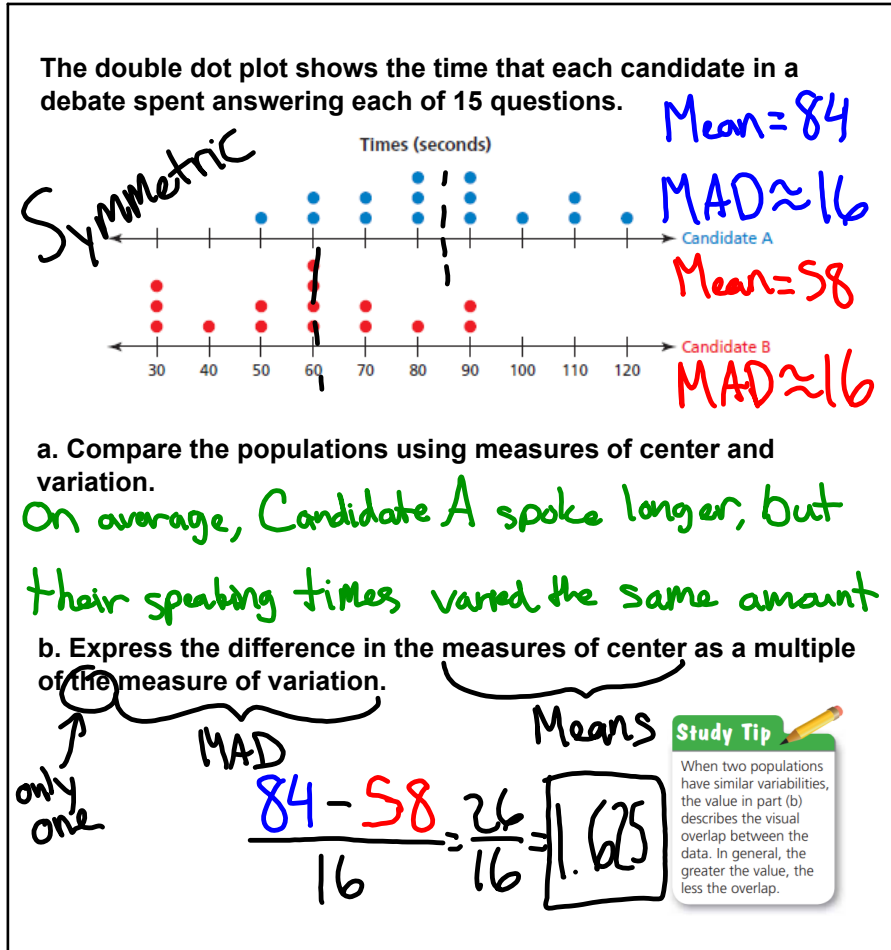
Skewed distribution: Median and IQR



Skewed Right



Skewed Left

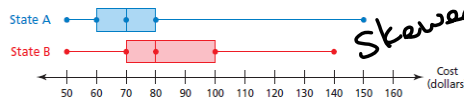


1. **WHAT IF?** Each value in the dot plot for Candidate A increases by 30 seconds. How does this affect the answers in Example 1? Explain.

Candidate A's mean would be 114 (84+30), but the MAD would stay the same. The difference between the means would be larger, meaning less overlap. $\left(\frac{114-58}{16} = \frac{56}{16} = 3.5\right)$

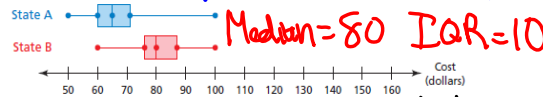
You want to compare the costs of speeding tickets in two states.

a. The double box-and-whisker plot shows a random sample of 10 speeding tickets issued in two states. Compare the samples using measures of center and variation. Can you use this to make a valid comparison about speeding tickets in the two states? Explain.



Median = 70 On average, ticket cost more
 IQR = 20 and vary more in state B.
 Median = 80 However, not enough data +
 IQR = 30 large variation → not reliable

b. The double box-and-whisker plot shows the medians of 100 random samples of 10 speeding tickets for each state. Compare the variability of the sample medians to the variability of the sample costs in part (a).



Less variability → more reliable
 c. Make a conclusion about the costs of speeding tickets in the two states.

Average ticket cost is higher in State B.