



Summarizing data in context

Putting the whole domain together

- Today: describe a real data set completely, and justify your choices.

Minutes read last night, 12 students

- 10, 15, 15, 20, 20, 20, 25, 25, 30, 30, 35, 95

Turn & talk: what is a typical amount of reading here?

A summary says **FIVE** things

- How many · what was measured and its units · center · variability · shape and anything striking.

Mean $28\frac{1}{3}$, median $22\frac{1}{2}$

- The mean is higher than 9 of the 12 actual values.
- One student read 95 minutes — 60 more than anyone else.

The mean is correct. It is also misleading.

Let the SHAPE choose the measure

- Outlier or skew → median and IQR.
- Roughly symmetric → the mean is fine.

Range 85 vs IQR $12\frac{1}{2}$ — same problem, same fix.

Your turn

- Quiz scores: 8, 9, 9, 10, 10, 10, 11, 11, 12, 12.
- Mean or median? Thumbs up when you decide.

Say it as a sentence

- “Twelve students read last night. A typical student read about $22\frac{1}{2}$ minutes, with the middle half between $17\frac{1}{2}$ and 30. One student read 95 minutes, far more than anyone else.”

Name the outlier. Hiding it misrepresents the data.

Now you

- → In-class practice sheet