

Grade 6 • Unit 8 • Assessment Guide

Data Sets and Distributions

For the Parent or Teacher — What to look for

Unit 8 is about understanding DISTRIBUTIONS — center, spread, and shape. Watch for whether students can choose the RIGHT measure (mean vs median) for the situation and articulate why.

Problem 1 — Statistical or not

- (a) NOT statistical. Single fixed answer (~330 m).
- (b) STATISTICAL. Different students → different heights.
- (c) NOT statistical. One pet, one name.
- (d) STATISTICAL. Different sixth-graders have different bedtimes.
- Strong: explains that statistical questions anticipate VARIABILITY.
- Weak: thinks any question with a number is statistical.

Problem 2 — Mean, median, range, MAD

- Mean: $(8+7+9+6+10+8)/6 = 48/6 = 8$.
- Median: (7,8,8,9 → middle two are 8,8) = 8.
- Range: $10 - 6 = 4$.
- MAD: deviations from 8 are 0, 1, 1, 2, 2, 0. Mean of deviations = $6/6 = 1$. So MAD = 1.
- Strong: clearly shows the deviation table; computes mean of deviations.
- Weak: takes deviations as signed (negative deviations included), gives MAD = 0; or computes range or variance by mistake.

Problem 3 — Compare two classes by MAD

- Class A is more CONSISTENT (lower MAD = scores cluster near the mean).
- Class B is more SPREAD OUT.
- Strong: explains MAD as the typical distance from the mean.
- Weak: thinks high MAD = high scores.

Problem 4 — Median, Q1, Q3, IQR

- Data (sorted): 4, 7, 8, 8, 10, 12, 13, 15, 18.
- $n=9$, median = 5th value = 10.
- Lower half (4,7,8,8): median = 7.5 (Q1).
- Upper half (12,13,15,18): median = 14 (Q3).
- $IQR = 14 - 7.5 = 6.5$.
- Watch for: students who include the median in BOTH halves (gives $Q1=8$, $Q3=13$).

Problem 5 — Same range, different IQR

- A small IQR with a large range means most data are clustered with some outliers in the tails. A large IQR with the same range means the middle 50% are also spread out.
- Example: {1, 5, 5, 5, 5, 5, 9} (range 8, IQR small) vs {1, 2, 4, 5, 6, 8, 9} (range 8, IQR larger).
- Strong: clear, concrete example.
- Weak: restates the question; no concrete example.

Problem 6 — Cat weights, box plot

- Min=7, Q1=9, Median=10, Q3=12, Max=15.
- Box plot: whiskers from 7 to 9 and 12 to 15; box from 9 to 12; line at 10.
- Strong: correctly identifies five-number summary AND draws box plot proportional to scale.
- Weak: confuses Q1 and Q3; draws box uneven; uses dot plot instead of box plot.

Problem 7 — Histogram vs box plot

- Histogram: useful for showing the SHAPE of a distribution (skew, modes) with continuous or grouped data.
- Box plot: useful for comparing distributions or summarizing spread (especially with outliers).
- Strong: gives concrete example for each ('Histogram for ages in years; box plot to compare boys' and girls' heights').
- Weak: just states 'when you want to see data'.

Problem 8 — Han's mean vs median

- Sorted: 4, 7, 7, 12, 50. Median = 7.
- Mean = $80/5 = 16$.
- The MEDIAN (7) is the better measure of center because the 50 is an OUTLIER pulling the mean up.
- Strong: identifies 50 as an outlier and articulates that the median is robust to outliers.
- Weak: defends the mean because 'it uses all the data'.
- ★ Outlier-driven choice between mean and median is the key Unit 8 judgment.

Problem 9 — Adding constant to all data

- Diego is WRONG. Mean of {2,4,6} = 4. After adding 5: {7,9,11}, mean = 9.
- The mean INCREASES by 5 — exactly the constant added.
- Strong: tests with data and articulates that the mean shifts by the constant.
- Weak: agrees without testing; or computes wrong mean.

Problem 10 — Original investigation

- Strong: a genuinely statistical question (with variability anticipated); reasonable plan for data collection; thoughtful choice of measure ('I'd use median for screen time because some kids might be outliers').
- Weak: question is not statistical; vague data plan; no measure justification.
- ★ Reflective synthesis question — the deepest test of statistical reasoning.

— end —