Describing a distribution in context

A dot plot does not answer a question by itself. How do you say in a few sentences what the data shows, so a reader who never saw the plot can picture it and check it?

WORDS TO KNOW

Context: what was measured, on whom, how, in what unit, and how many values there are.

Shape: symmetric, skewed left or right, uniform, one peak or two.

Center and spread: a matching pair, such as the median with the IQR, or the mean with the MAD.

Fences: $Q_1 - 1.5 \times \text{IQR}$ and $Q_3 + 1.5 \times \text{IQR}$.

Outlier: a value beyond a fence. A value exactly on a fence is not an outlier.

Gap: a stretch of the scale with no values. **Clusters** are groups with gaps between them.

KEY IDEAS

1. A description has five pieces

Piece	What it says
Context	what was measured, on whom, how, the unit, and how many
Shape	read from the display
Center and spread	one matching pair, chosen for the shape
Unusual features	outliers tested with the fences, gaps, clusters
The answer	a last sentence back to the question

Every number carries its unit. "11" means nothing; "11 cm, snout to tail tip" does.

2. The shape picks the pair

Skewed data, or data with outliers: use the **median with the IQR**. They stay with the bulk of the values.

Roughly symmetric data with no outliers: use the **mean with the MAD**. They use every value.

Two clusters: one center can land in the gap, so describe each cluster.

The mean and median are a cross-check: when they are well apart, the data probably leans toward the mean.

3. Test before you say "outlier"

A value that looks far away is often inside the fences. Work out the fences first, then name each outlier with its value and unit. Give a reason for it only as a possibility.

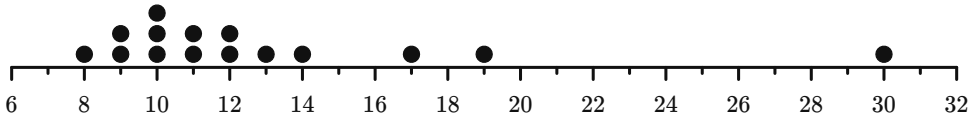
WATCH OUT

- **Keep the pair together.** The MAD is a distance from the mean, so it does not belong next to the median.
- **Numbers can all be right and still mislead.** A description with the wrong pair, or a missing outlier, misleads.
- **Say how many and how measured.** A shape from 6 values is far less sure than one from 60.
- **Name the skew by the tail.** Values piled up at the high end with a tail to the left are skewed left.

Free sample from math.wiki for classroom use. Do not post. math.wiki

WORKED EXAMPLES

Example 1. Nerea measured 15 lizards, snout to tail tip, to the nearest cm. Mean 13 cm, median 11 cm, $Q_1 = 10$ cm, $Q_3 = 14$ cm. Describe the lengths.



Length in centimeters, snout to tail tip

1. Shape: the dots crowd from 8 to 14, then thin out to 17, 19 and, far off, 30. Skewed right. The mean above the median agrees.
2. Pair: skewed, so the median and the IQR. The IQR is 14 − 10 = 4 cm.
3. Fences: 1.5 × 4 = 6, so 10 − 6 = 4 and 14 + 6 = 20 cm. The 30 is beyond 20, an outlier. The 19 is inside, so it is not. No lizard lies between 19 and 30 cm, a gap.

Answer: "Nerea measured 15 lizards from snout to tail tip, to the nearest centimeter. The lengths are skewed right. The median is 11 cm, and the middle half are 10 to 14 cm long, an IQR of 4 cm. One lizard, at 30 cm, is an outlier, with a gap from 19 to 30 cm. A typical lizard here is about 11 cm long."

Example 2. Miroslav counted the routes, out of 10, that each of 13 beginners finished. 10, 9, 10, 7, 10, 8, 2, 9, 6, 10, 8, 5, 10. Describe the counts.

1. Sort: 2, 5, 6, 7, 8, 8, 9, 9, 10, 10, 10, 10, 10. The median is the 7th value, 9 routes.
2. The six values on each side give $Q_1 = \frac{6+7}{2} = 6.5$ and $Q_3 = \frac{10+10}{2} = 10$. The IQR is 3.5 routes.
3. Shape: the counts pile up at 9 and 10, against the ceiling of 10, and trail left to 2. Skewed left. So use the median and the IQR.
4. Lower fence: 6.5 − 1.5 × 3.5 = 1.25. The 2 is inside it, so it is not an outlier.

Answer: "Miroslav counted the routes finished, out of 10, by 13 beginners. The counts are skewed left. The median is 9 routes, and the middle half finished 6.5 to 10 routes, an IQR of 3.5 routes. The lowest count, 2, is not an outlier. A typical beginner finished about 9 of the 10 routes."

Free sample from math.wiki for classroom use. Do not post. math.wiki