



Interpreting box and whisker plots (CSSI)

Key focus

This objective builds on the CSSI framework (Centre, Spread, Shape, Interesting features) by teaching the 75%/50% rule for informal inference: if one group's median sits outside the interquartile range (the middle 50%) of the other group, that is evidence of a real difference between the groups; if it sits inside, the apparent difference could just be normal sampling variation. Every task in this LO applies this rule directly, alongside reading the five-number summary, comparing centre and spread, and describing shape. The beta textbook reference is Ex 30.02, pg 570.

Suggested teaching sequence

1. Recap reading the five-number summary (min, LQ, median, UQ, max) from a box plot, and calculating the IQR as $UQ - LQ$.
2. Show a two-group comparison and model comparing centre (which median is higher) and spread (which IQR is smaller/more consistent).
3. Introduce the 75%/50% rule directly: shade or highlight one group's IQR, and check whether the other group's median falls inside or outside it.
4. Contrast a clear case (one median well outside the other's IQR – real difference) against an overlapping case (one median inside the other's IQR – no strong evidence of a difference).
5. Discuss a borderline case where a median sits just past the edge of the other's IQR, and establish that the rule is a simple threshold test, not a precise statistical measure.
6. Ask why the rule uses the IQR rather than the full range: the range is easily stretched by a single extreme value, while the IQR reflects where the typical middle 50% of data actually sits.
7. Extend to three or more groups: have students apply the rule pairwise to work out which pairs show real evidence of a difference and which don't.
8. Critique a claim that overstates a small, rule-inside difference ("it's definitely better") as well as a claim that understates a clear, rule-outside difference, reinforcing that the rule's verdict should match the language used.

Core examples to model

- Clear difference: Class B's median (45) sits outside Class A's IQR (20 to 40) – evidence of a real difference.
- No strong evidence: Team Y's median (32) sits inside Team X's IQR (25 to 35) – consistent with normal variation.
- Borderline: a median sitting just 2 units past the other group's upper quartile still counts as "outside" under the rule's simple threshold.
- Outlier resistance: a single outlier at 90 does not change a box plot's median or IQR, because both depend on the position of the main body of data, not on extreme values.

Watch for

- Assuming any higher median means a “real” or “definite” difference – always check the rule (is the higher median inside or outside the other group’s IQR) before making that claim.
- Confusing spread (IQR) with the full range – the range can be dominated by a single extreme value, while the IQR reflects the typical middle 50%.
- Treating the rule as a precise statistical test – it’s a simple informal threshold, useful for a first judgement, not a substitute for a formal hypothesis test.
- Forgetting that shape (symmetrical vs. skewed) is read from the median’s position within the box and the relative whisker lengths, not from the IQR value alone.
- Assuming a box plot’s median and IQR will shift when an outlier is added or removed – they are resistant to a single extreme value, unlike the mean.

Checks for understanding

- Can students read the five-number summary and calculate the IQR from a box plot?
- Can they compare two groups’ centre and spread, and describe shape from the median’s position and whisker lengths?
- Can they correctly apply the 75%/50% rule – checking whether one group’s median is inside or outside the other’s IQR?
- Can they explain why the rule uses the IQR rather than the full range?
- Can they critique a conclusion that overstates a rule-inside (not clearly different) result, or understates a rule-outside (clearly different) one?
- Can they explain why a box plot’s five-number summary is resistant to a single outlier?
- Can they apply the rule pairwise across three or more groups to identify which pairs show real evidence of a difference?