

## Solutions

---

Accept equivalent correct methods and supported interpretations. The marks check reasoning; they are not an official grade.

**2.1 D** – Distances from the shifted mean stay unchanged. A adds a shift to spread; B replaces spread by the shift; C treats addition as scaling. All confuse location with dispersion.

Correct response linked to the given case.

**2.2 A** – One standard deviation on both sides gives 90–110. B is the approximate 95% interval; C is not centred on the mean; D is only the upper half of the central one-SD interval.

Correct response linked to the given case.

**2.3 3** –  $18-15=3$ .

Correct response linked to the given case.

**2.4 False** – A third variable or selection can explain association.

Correct response linked to the given case.

**3.1** Under the stated normal model, about 68% lie in 74–86 and 95% in 68–92. Transformed mean is  $2(80) + 5 = 165$  and standard deviation  $2(6) = 12$ . These are model proportions, not guarantees for a finite sample.

Sound conclusion or proposed method. Relevant supporting steps or textual evidence. Accurate conditions, checks or qualification.

**3.2** Prediction is  $4(5) + 3 = 23$ . Residual  $e = y_{obs} - y_{pred} = 20 - 23 = -3$ . A positive slope gives association under the fit; confounding or selection can explain it without a direct cause.

Sound conclusion or proposed method. Relevant supporting steps or textual evidence. Accurate conditions, checks or qualification.

**3.3** Median  $Q_2 = (6+9)/2 = 7.5$ . Lower half 2,4,4,6 has  $Q_1 = 4$ ; upper half 9,11,12,16 has  $Q_3 = 11.5$ . Five-number summary is (2, 4, 7.5, 11.5, 16) and  $IQR = Q_3 - Q_1 = 7.5$ . A box plot displays these values. A category bar chart compares categorical counts; a scatterplot needs paired quantitative measurements. Neither directly displays this univariate five-number summary. The stated quartile convention prevents disagreement between other valid conventions.

Sound conclusion or proposed method. Relevant supporting steps or textual evidence. Accurate conditions, checks or qualification.

**3.4** Below 20 there are  $2+6=8$  of 10 observations, so 80%. The exact observations within bins are unknown, so no exact mean follows; midpoint estimates must be labelled estimates. A scatterplot represents pairs, and a fitted relation should describe its main shape with the unusual point considered. Joining every observation invents a path between separate cases and ignores the curved trend. Inspect actual axis scales and the outlier before fitting or extrapolating.

Sound conclusion or proposed method. Relevant supporting steps or textual evidence. Accurate conditions, checks or qualification.