Statistics - 5 - Representing data

SOLUTIONS

Section 1.

Q1 $\bar{x} = 33.7$

$S.d = 20.5526 = \underline{20.6}$ (to 1dp)

Q2, $\bar{x} = 7.88$

Variance $= 1.5829^2 = 2.50557 = \underline{2.51}$ (to 2dp)

Q3, $\bar{x}_{10} = 12 \quad S.d_{10} = 5.2 \quad n = 10$

$\bar{x}_{20} = 9.3 \quad S.d_{20} = 4.6 \quad n = 20$

$\Sigma x_{10} = \bar{x}_{10} \times n = 12 \times 10 = 120$

$\Sigma x_{20} = \bar{x}_{20} \times n = 9.3 \times 20 = 186$

$\therefore \quad \Sigma x_T = 306 \quad \therefore \bar{x}_T = \frac{306}{30} = \underline{10.2}$

$S.d = \sqrt{\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2}$ (10) $5.2 = \sqrt{\frac{\Sigma x^2}{10} - \left(\frac{120}{10}\right)^2}$

$\Sigma x_{10}^2 = 1710.4$

(20) $4.6 = \sqrt{\frac{\Sigma x^2}{20} - \left(\frac{186}{20}\right)^2}$

$\Sigma x_{20}^2 = 2153$

$\therefore \Sigma x_T^2 = 1710.4 + 2153 = 3863.4$

$$S.d_T = \sqrt{\frac{3863.4}{30} - \left(\frac{306}{30}\right)^2} = 4.9739 = \underline{4.97} \text{ (to 2dp)}$$

or $\underline{5.01}$ (to 1dp)

Q3. a, $\underline{\bar{x} = 13}$ (By symmetry!) or Calculator!

$s.d = 1.4142 = \underline{1.41}$ (to 2dp)

b, $\underline{\bar{x} = 113}$ $\underline{s.d = 1.41}$ (Shift of 100)

c, $\underline{\bar{x} = 130}$ $s.d = 1.41 \times 10$ ($\times 10$)

$= \underline{14.1}$

d, $\underline{\bar{x} = 29}$ $s.d = 1.41 \times 2$ ($\times 2 + 3$)

$= \underline{2.83}$

(careful of rounding)

Q4, a, $\underline{\bar{x} = 13}$, $\underline{\sigma = 1.41}$

Code!

b, mean = 113 $\sigma = 1.41$ ($+100$)

c, mean = 130 $\sigma = 141.4$ ($\times 10$)

d, mean = 29 ($13 \times 2 + 3$) $\sigma = 2.82$ ($\times 2$) ($\times 2 + 3$)

Section 2

Q1, a, $Q_2 = \frac{11}{2} = 5.5 = 6^{th}$ value $= 53$ ✓

$Q_1 = \frac{11}{4} = 2.75 = 3^{rd}$ value $= 35$ ✓

$Q_3 = 3 \times \frac{11}{4} = 8.25 = 9^{th}$ value $= 60$ ✓ (3)

$IQR = 60 - 35 = 25$

b, lb $\Rightarrow Q_1 - 1.5 \times IQR$

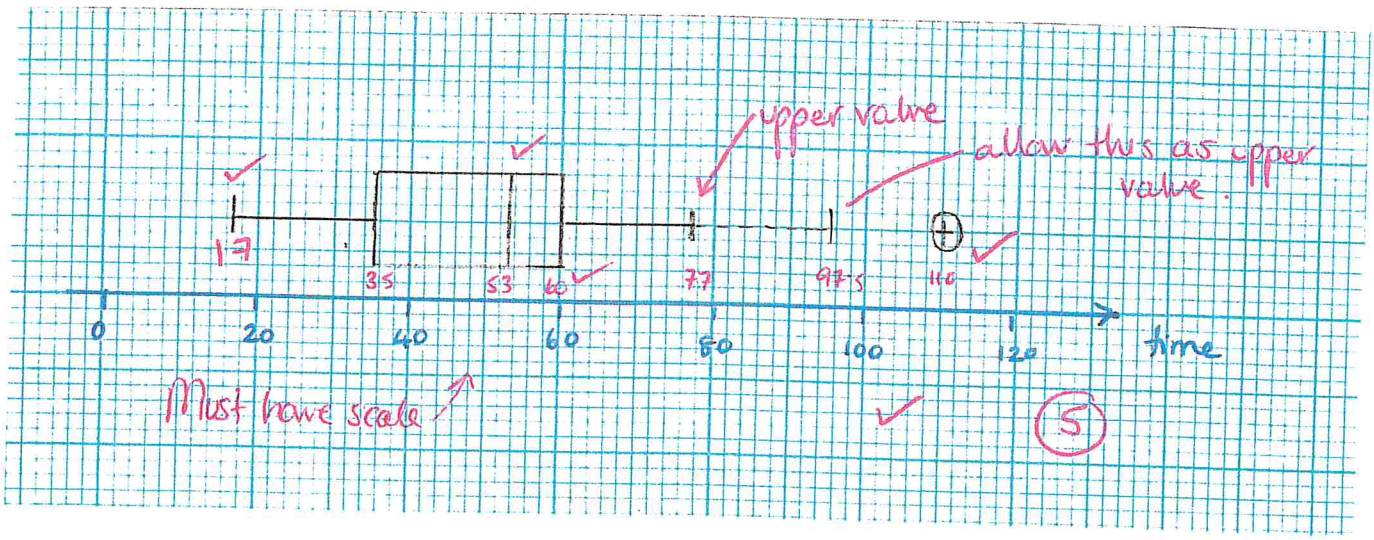
$\Rightarrow 35 - 1.5 \times 25 = -2.5$ ✓ $47 > -2.5$ ∴ no outlier at lower end.

ub. $\Rightarrow Q_3 + 1.5 \times IQR$

$60 + 1.5 \times 25 = 97.5$ ✓ $110 > 97.5$ ∴ is an outlier ✓

(3)

Q1c,



(5)

Q2, a, 54 mins. (1)

b, 52 mins (1)

c, The median for Gryffindor is lower than for Hufflefluff, so on average students in Gryffindor ran faster. (1)

The range for Gryffindor is larger than Hufflefluff, suggesting that the times is less consistent for Gryffindor students (1)

Q2d, Advantage: It is easy to see and compare the main areas e.g. Median, spread. (1)

Disadvantage: You cannot see the raw data. (1)

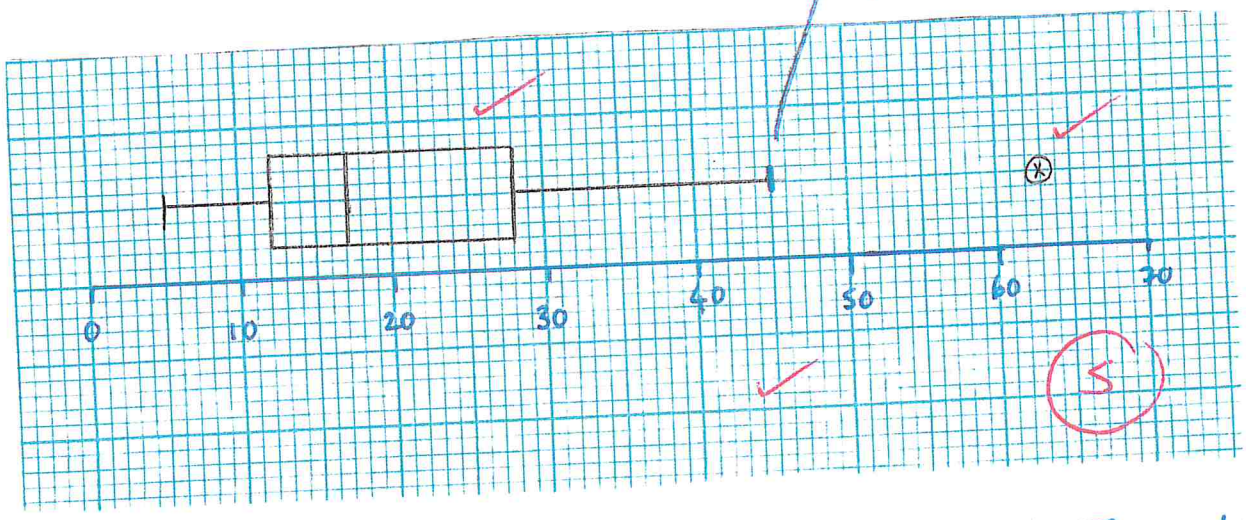
Q3, $Q_1 = 12$ $Q_2 = 17$, $Q_3 = 28$ LV = 5 MV = 63

a, Outlier check

$Q_1 - 1.5(28-12) = 12 - 1.5 \times 16 = -12$ ✓

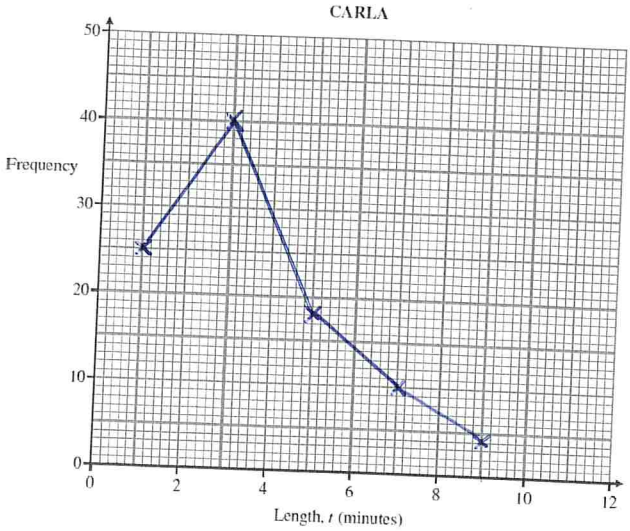
$Q_3 + 1.5(28-12) = 28 + 1.5 \times 16 = 52$ ∴ ✓ 63 is an outlier

45, since only 1 value above 45.



b, Generally flights are delayed by around 17 minutes, with a few being delayed by up to 52 mins ✓ Delays of over 52 mins are very unlikely. (2)

Q4, a,



① Points marked correctly
① Joined with straight line.
(① Do not connect to x-axis)

b, Modal class $2 < t \leq 4$