

# Statistics - 4 - Representing data

## SOLUTIONS

### Section 1.

Q1

$$\bar{x} = \underline{33.7}$$

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

Q2,

$$\bar{x} = \underline{7.88}$$

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

Q3,

$$\bar{x}_{10} = 12 \quad s.d_{10} = 5.2 \quad n = 10$$

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

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

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

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

$$s.d = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2} \quad \textcircled{10} \quad 5.2 = \sqrt{\frac{\sum x^2}{10} - \left(\frac{120}{10}\right)^2}$$

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

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

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

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

$$s.d_T = \sqrt{\frac{3863 \cdot 4}{30} - \left(\frac{306}{30}\right)^2} = 4.9739 = \underline{4.97} \text{ (to 2dp)} \\ \text{or } \underline{5.01} \text{ (to 1dp)}$$

Q3. a,  $\bar{x} = 13$  (By symmetry!) or Calculator!  
 $s.d = 1.4142 = \underline{1.41}$  (to 2dp)

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

c,  $\bar{x} = 130$   $s.d = 1.41 \times 10$  (x 10)  
 $= \underline{14.1}$

d,  $\bar{x} = 29$   $s.d = 1.41 \times 2$  (x 2 + 3)  
 $= \underline{2.83}$   
 (careful of rounding)

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

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

c, mean = 130  $\sigma = 14.14$  (x 10)

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

Code!

## Section 2

$$Q_1, a, Q_2 = \frac{11}{2} = 5.5 = 6^{\text{th}} \text{ value} = 53 \quad \checkmark$$

$$Q_1 = \frac{11}{4} = 2.75 = 3^{\text{rd}} \text{ value} = 35 \quad \checkmark$$

$$Q_3 = 3 \times \frac{11}{4} = 8.25 = 9^{\text{th}} \text{ value} = 60 \quad \checkmark$$

(3)

$$b, \text{ i.i.} \Rightarrow Q_1 - 1.5 \times IQR \quad IQR = 60 - 35 = 25$$

$$\Rightarrow 35 - 1.5 \times 25 = -2.5$$

$47 - 2.5 \therefore$  no outlier at lower end.

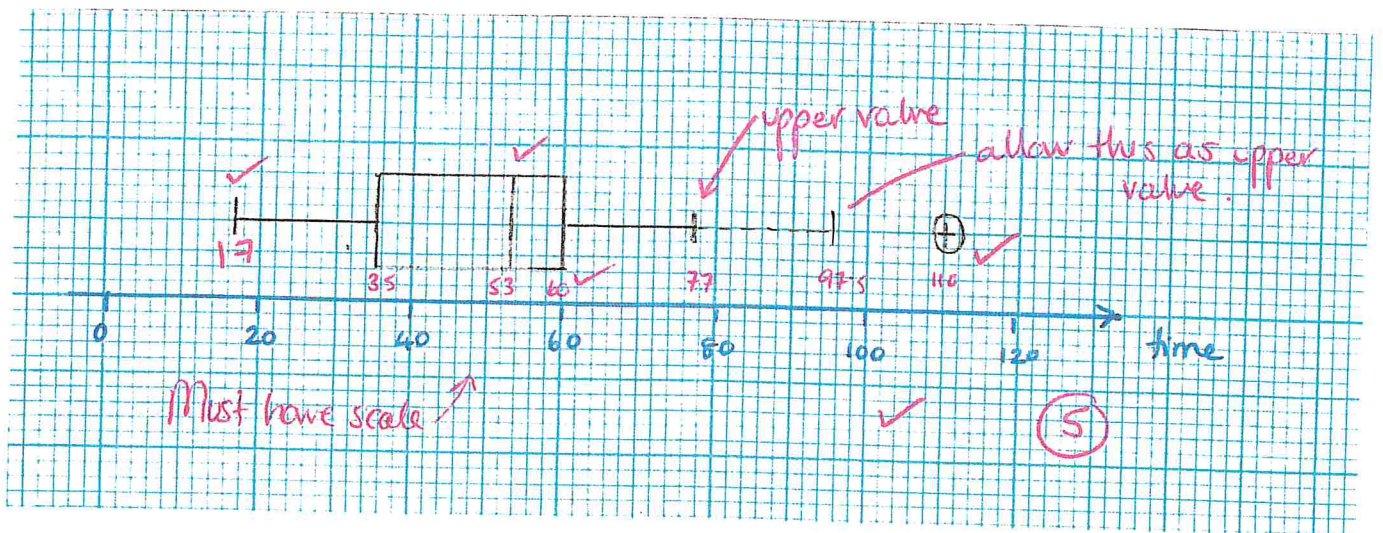
$$\text{ii.} \Rightarrow Q_3 + 1.5 \times IQR$$

$$60 + 1.5 \times 25 = 97.5$$

$110 > 97.5 \therefore$  is an outlier

(3)

Q1c,



Q2, a, 54 mins. (1)

b, 52 mins (1)

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

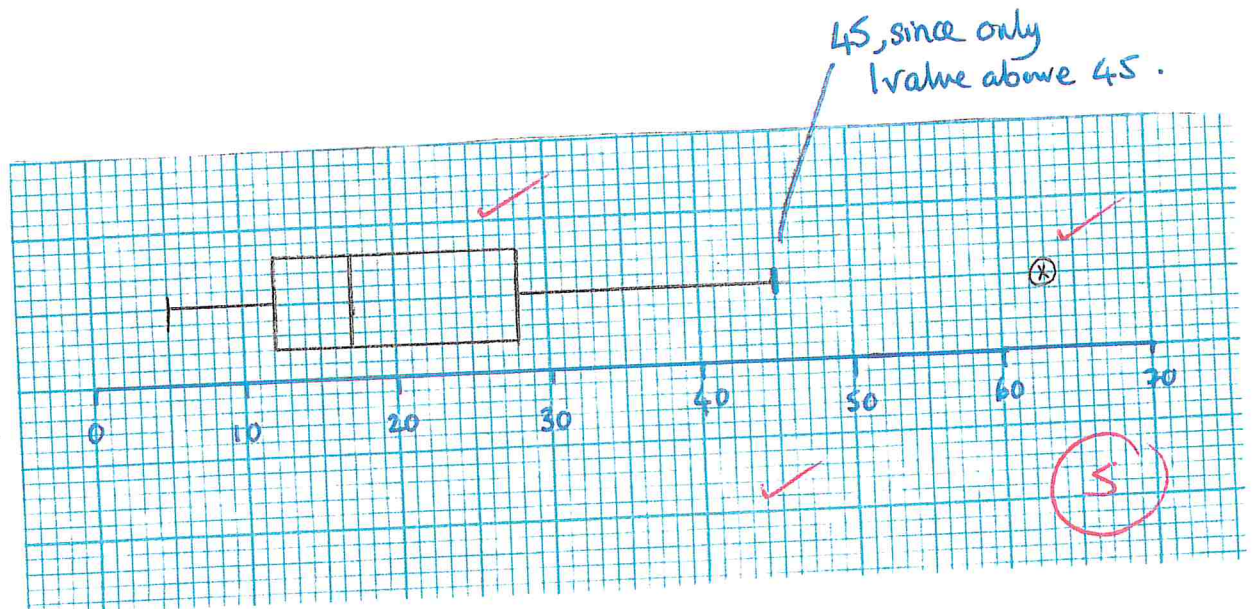
Q3,  $Q_1 = 12$   $Q_2 = 17$ ,  $Q_3 = 28$   $LV = 5$   $MV = 63$

a, Outlier check

$$Q_1 - 1.5(Q_3 - Q_1) = 12 - 1.5 \times 16 = -12 \quad \checkmark$$

$$Q_3 + 1.5(Q_3 - Q_1) = 28 + 1.5 \times 16 = 52 \quad \therefore$$

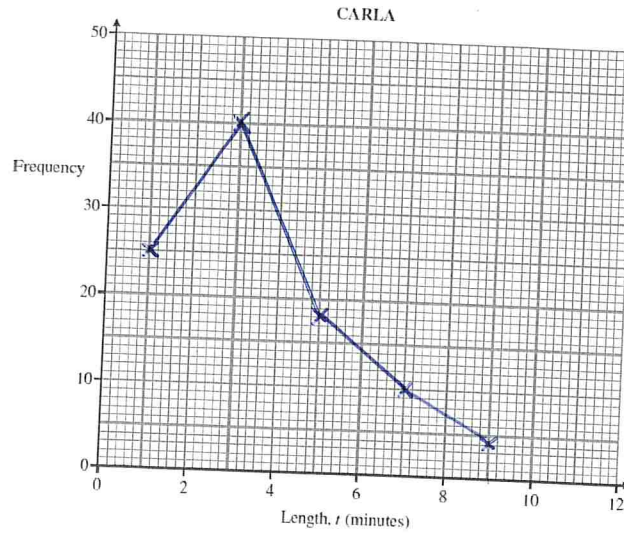
$\checkmark$  63 is an outlier



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.
- (Do not connect to x-axis)

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