

## Section 2: Data presentation and interpretation

### Exercise level 3 solutions

1. (i) Winner's points: Range = 97, IQR = 27, SD = 20 (2 s.f.)  
 Loser's points: Range = 68, IQR = 34, SD = 19 (2 s.f.)  
 The winners have a larger range than the losers. This indicates that the spread of winning scores is larger than the spread of losing scores.  
 The losers have a larger IQR than the winners. This indicates that the spread of the middle 50% of the losing scores is larger than the spread of the middle 50% of the winning scores.  
 Winners have a larger SD than the losers, although these values only differ by around 1.
- (ii) The data shows that over 75% of winning scores are less than 100 points. Thus the midrange (which is 98.5 points) would seem to be unrepresentative. Similarly the mean will be affected by the larger values (by a lesser extent) and the mean of 84 points is quite close to the median of 78 points which is not affected by the higher scores. Since higher scores are a common feature of the game, it seems that the mean is the most representative.
- (iii) From the data, at least 25% of the losing scores must have been 0 points. If any other data values were repeated this often then it would not be possible to distinguish the median, upper quartile or maximum values. It is likely that the mode for the losers is zero, but it is not possible to say what the mode for the winner was.
- (iv) Winner's points:  $LQ - 1.5 \times IQR = 29.5$ ,  $UQ + 1.5 \times IQR = 137.5$ . There is at least one outlier as the largest recorded winning score was 147. There are no low outliers as the lowest score was 50.  
 Loser's points:  $LQ - 1.5 \times IQR = -51$ ,  $UQ + 1.5 \times IQR = 85$ . There are no outliers in the loser's scores data set.
- (v) The lower quartile will be the same as the median for the loser's points as there will be a quarter of the scores below this value, so  $LQ = 13.5$ .

The mean will be the average of the mean number of points, since the data sets have the same size, so mean = 51.6 (3 s.f.). More accurately, calculate the total number of the points of the winners and losers by multiplying their means by 618 to get the total number of winners' points to be 52042 and the total number of losers' points to be 11794, since the results must be integers. So the total of all the scores is 63834 and so the mean number of points per frame is 51.6 (3 s.f.) (the two results actually agree to 4 d.p.)

It is not possible to deduce the median exactly. However, the median must be between 50 and 68 as there is less than 50% of the data values below 50 and more

## Edexcel AS Maths Data 2 Exercise solutions

than 50% below 68.

- (vi) The total number of points in each frame would generally increase overall as every ball would be potted (and the chance of more points being scored from fouls). There would be fewer losing scores of 0 as the loser would have more opportunities to score some points so the spread of losing scores would also be likely to increase, but due to the fact there is a maximum number of points available the spread of the winning scores might decrease.

2. (i) Possible comments include

- more houses were sold in 2000 than in 2015
- the spread of house prices has increased significantly, looking at the range of prices, even excluding the one house in the highest price bracket
- there were no houses sold for less than £125,000 in 2015, and there were approximately 217 such houses sold in 2000
- the estimated mean house price, using mid-interval values, has increased from £153000 in 2000 to £468000 (3 s.f.) in 2015, so by just over 3 times.

- (ii) Assuming that the mid-interval value is representative of the houses in each group, the amount of stamp duty which could be expected in each group is as follows:

House prices in 2000, £x	Mid-interval value	Stamp duty	Frequency
$0 \leq x < 60000$	30000	0	12
$60000 \leq x < 100000$	80000	800	120
$100000 \leq x < 150000$	125000	1250	170
$150000 \leq x < 200000$	175000	1750	67
$200000 \leq x < 250000$	225000	2250	37
$250000 \leq x < 500000$	375000	11250	46
$500000 \leq x < 1000000$	750000	30000	1

Estimate of total stamp duty paid in 2000 = £1,056,500

House prices in 2015, £x	Mid-interval value	Stamp duty	Frequency
$125000 \leq x < 250000$	187500	1250	13
$250000 \leq x < 375000$	312500	5625	119
$375000 \leq x < 500000$	437500	11875	71
$500000 \leq x < 750000$	625000	21250	69
$750000 \leq x < 925000$	837500	31875	17
$925000 \leq x < 1500000$	1212500	65000	6
$1500000 \leq x < 5000000$	3250000	303750	1

(Note: No stamp duty is paid on the first £125,000, 2% on £125,000 to £250,000, 5% on £250,000 to £925,000, 10% on £925,000 to £1.5 million

and 12% above £1.5 million.)

Estimate of total stamp duty paid in 2015 = £4,230,625

## Edexcel AS Maths Data 2 Exercise solutions

So the estimated amount of stamp duty raised has increased by just over 4 times from 2000 to 2015. The estimated mean house price has increased 3 times. This indicates that the changes to stamp duty are keeping pace with the increase in house prices.

The assumption that the mid-interval value is representative of the prices in each group may not be very representative – under the system in 2000, in particular, house prices were often set just below the threshold for higher stamp duty charges.

3. (i) Looking at the data for January, when there were the most observations, the range is over 50g which is over 50% of the weight of the lightest birds. The evidence seems to support this claim. However we don't know how the weight of any individual bird varies.

Evidence inconclusive

- (ii) The blackbirds measured have the highest median and maximum weights in December and January. However there is no evidence that this allows them to survive winter.

Evidence inconclusive

- (iii) The lightest blackbirds are measured in March and April. This might mean that babies were being caught and measured, so these might have significantly lower weights. The lowest median weight observed is in August when the juveniles might be joining the wider population – however there is very little data behind the August result – only 6 blackbirds were ever measured in August.

No evidence

- (iv) In most months there is a slightly larger difference between the median and LQ than between the UQ and the median, e.g. Feb, Apr, May, June, Oct, Nov and Dec. So most months there actually is a slight negative skew.

No evidence

- (v) Based on this data if we captured a blackbird in this garden weighing 140 grams or more we could guess that it is Jan or Dec. Thus there are circumstances where weighing a bird captured in this garden could suggest the time of year.

Evidence supports

- (vi) Potential evidence for this could come from the number of captures in each month. The fewest captures are in Aug and Nov which are certainly not spring months. However we do not know if capturing effort is maintained across the year.

Evidence inconclusive

- (vii) Based on this evidence, blackbirds with the ability to fly into a particular garden in the East Midlands never weigh below 80 grams.

Evidence supports

## Edexcel AS Maths Data 2 Exercise solutions

(viii) The answer is similar to (vi), although captures are lower in these months we cannot deduce from the data why this is so.  
Evidence inconclusive

(ix) One reason why blackbirds could be spending time in the garden is to catch food. There are more captures in winter months which could indicate they are looking for food. The blackbirds caught in winter tend to weigh more than those at other times of year but this could be due to babies maturing over time. Indeed the higher weights could indicate an abundance of suitable food.  
Evidence inconclusive