2021 Assessment resources
GCSE Statistics

SEC and Data – Higher

Answers and commentaries

The question numbers in this resource reflect the question numbers from the original papers and match the question numbers in the corresponding 2021 assessment materials.

Question 5(a)

5   Marcus is planning a Driver Safety course.

He wants to give the people attending the course a questionnaire to complete.

5 (a)   Marcus wants to know how far each person usually drives in a week.

Write a closed question that Marcus could ask to find out this information.
Include a response section.

[3 marks]

Student A

5 (a)   Marcus wants to know how far each person usually drives in a week.

Write a closed question that Marcus could ask to find out this information.
Include a response section.

[3 marks]

How far do you usually drive per week?

☐ 0 - 20 miles   ☐ 21 - 40 miles   ☐ 41 - 60 miles   ☐ 61 + miles

Commentary

A suitable question with a time frame. Options boxes are acceptable allowing rounding of decimal response to obtain exhaustiveness.

3 marks

Student B

5 (a) Marcus wants to know how far each person usually drives in a week.

Write a closed question that Marcus could ask to find out this information.

Include a response section.

[3 marks]

How far do you usually drive in a week?

☐ I don't drive ☐ Up to 20 miles ☐ Up to 50 miles

☐ Up to 100 miles ☐ 100 miles +

Commentary

A suitable question with a time frame. Option boxes are exhaustive but not exclusive with many overlaps.

2 marks

Student C

5 (a) Marcus wants to know how far each person usually drives in a week.

Write a closed question that Marcus could ask to find out this information.

Include a response section.

[3 marks]

How far do you normally drive a week in miles.

$0 \leqslant 10$, $10 \leqslant 20$, $20 \leqslant 30$, $30 \leqslant 40$, $40 \leqslant 50$,

$50 <$

Commentary

A suitable question with a time frame. The poorly expressed options do seem to show exhaustiveness but can only earn a maximum of 2 marks (see Additional Guidance in the mark scheme)

2 marks

Question 5(b)(i)

5 (b)   Marcus also wants to know whether people regularly drive faster than the speed limit.

He plans to collect the information using this method.

> He asks each person to secretly throw a dice.
>
> The person then answers as follows:
>
> - if the person gets an odd number, they answer 'Yes'
> - if the person gets an even number, they truthfully answer the question,
>   'Do you regularly drive faster than the speed limit?'

5 (b) (i)   Why does Marcus use this method?

[1 mark]

Student A

5 (b) (i)   Why does Marcus use this method?

[1 mar

so more people answer truthfully because it's a sensitive question.

Commentary

Appropriate reference to this question being sensitive.

1 mark

Student B

5 (b) (i)   Why does Marcus use this method?

[1 mar

keeps it anonymous

Commentary

Not considered an appropriate answer as the responses could be anonymous anyway.

0 marks

Student C

5 (b) (i) Why does Marcus use this method?

[1 m]

It makes results more reliable, because
if he asked without the dice, people would
lie.

Commentary
A good response references the effect of using the dice.
1 mark

Question 5(b)(ii)

5 (b) (ii) Marcus collects data from 100 people using this method.

60 people give the answer 'Yes'.

Marcus says,

"60% of these people regularly drive faster than the speed limit."

Explain why Marcus is wrong.

[1 mark]

Student A

5 (b) (ii) Marcus collects data from 100 people using this method.

60 people give the answer 'Yes'.

Marcus says,

"60% of these people regularly drive faster than the speed limit."

Explain why Marcus is wrong.

[1 mark]

Some people may have been forced to say yes, due
to throwing an odd number.

Commentary
'Forced to say yes' would have been enough to imply that the dice roll had caused the 'yes'
1 mark

Student B

5 (b) (ii) Marcus collects data from 100 people using this method.

60 people give the answer 'Yes'.

Marcus says,

"60% of these people regularly drive faster than the speed limit."

Explain why Marcus is wrong.

[1 mark]

The people that rolled even number could also drive over the speed limit.

Commentary
Student B does not address the 'false' yes responses coming from the odd number rolls.
0 marks

Question 1
Please see the mark scheme

Question 2
Please see the mark scheme

Question 4
Please see the mark scheme

Question 10(a) and 10(b)
Please see the mark scheme

Question 10(c)(i)
No examples available

Commentary
The expected response is to reference the control group given a comparison to what happens for plants who are not given the plant food.

Question 10(c)(ii)
No examples available

Commentary
There were two possible answers here. Students could reference A and comment that plants should be watered the same in both groups or students could reference C and comment that all plants should receive the same amount of sunlight.

Question 16(a)

16 In a golf tournament, players take part in several rounds of golf.

Players try to complete the course taking as few golf strokes as possible.

Justin wants to compare the number of strokes taken by the players in the first two rounds of a tournament.

He collects data for the top 50 players.

Justin's hypothesis is,

In which of the first two rounds will players take the fewer strokes on average?

16 (a) What mistake has Justin made when writing his hypothesis?

[1 mark]

Student A

16 (a) What mistake has Justin made when writing his hypothesis?

[1

He His hypothesis is a question, rather than a statement.

Commentary
Correct reference to the fact that a hypothesis should not be in the form of a question.
1 mark

Student B

16 (a)   What mistake has Justin made when writing his hypothesis?

[1

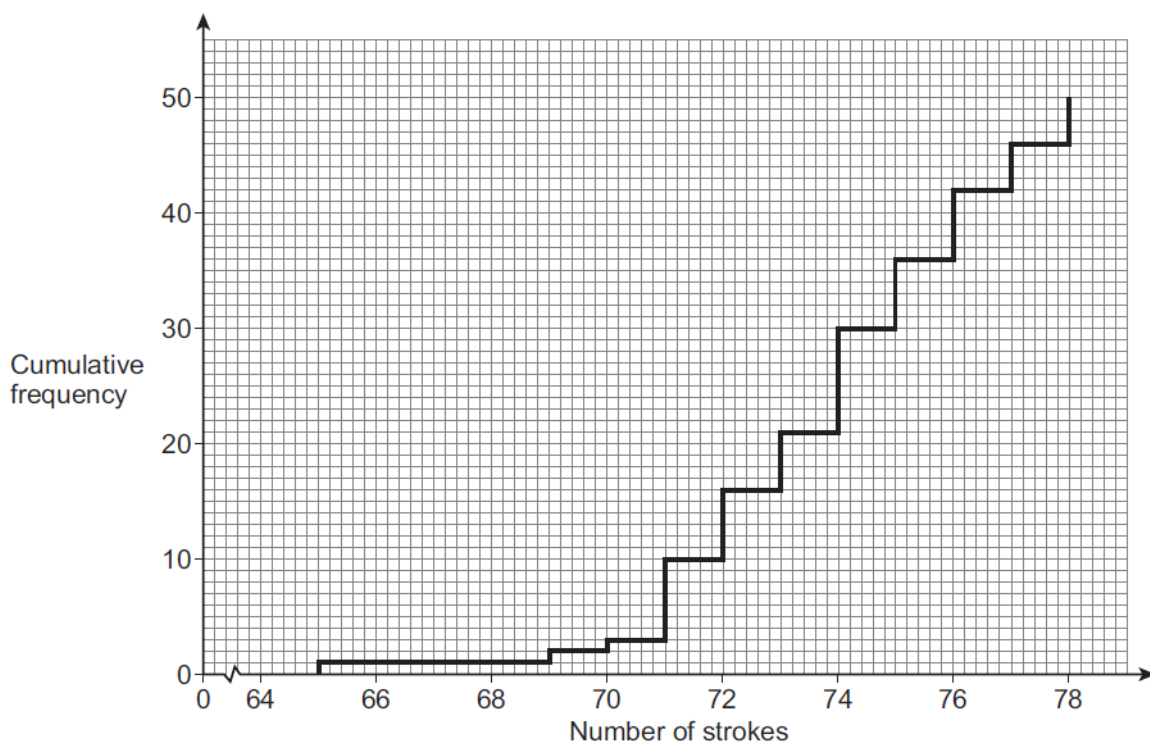He hasn't stated his sample frame in the hypothesis.

Commentary

Not a correct response.
0 marks

Question 16(b)

Justin draws a cumulative frequency step polygon to show the results for players in **Round 1**

16 (b)   Explain why a cumulative frequency step polygon is an appropriate graph for the data.

[1 mark]

Student A

16 (b)   Explain why a cumulative frequency step polygon is an appropriate graph for the **S7**
[1

You can only have a whole number of people and a whole number
of strokes. A line graph would suggest you could partial amount of
both of these.

Commentary
The first sentence is sufficient, with the second sentence deemed irrelevant.
1 mark

Student B

16 (b)   Explain why a cumulative frequency step polygon is an appropriate graph for the **S5**
[1

The data is discrete

Commentary
Correct answer.
1 mark

Question 16(c)

16 (c)   Work out the percentage of players who took 72 strokes or fewer for **Round 1**

[2 marks]

_____

_____ %

Student A

16 (c)   Work out the percentage of players who took 72 strokes or fewer for **Round 1**

[2

$^{16}/_{50} = 0.32$

_____32_____ %

Commentary
Fully correct answer.
2 marks

Student B

16 (c)   Work out the percentage of players who took 72 strokes or fewer for **Round 1**

$16 \div 72 = ^2/_9 \times 100 = 22.2\%$

_____22.2_____ %

Commentary
16 is correct for one mark with the subsequent work to change it into a percentage being incorrect.
1 mark

Question 16(d)
Please see the mark scheme

Question 16(e)

16 (e) The lowest number of strokes taken in **Round 1** is 65

Show by calculation that this value is an outlier.

[3 marks]

Student A

16 (e) The lowest number of strokes taken in **Round 1** is 65

Show by calculation that this value is an outlier.

[3 marks]

$$76-72=4$$
$$72-(4\times1.5)=66$$
$$65<66$$

Commentary
A fully correct response, where it is shown that 65 is below the point where outliers are considered to begin.
3 marks

Student B

16 (e) The lowest number of strokes taken in **Round 1** is 65

Show by calculation that this value is an outlier.

[3 marks]

$$76-72=4$$
$$4\times1.5=6$$
$$72-74-6=68 \quad 72-6=66$$

Commentary
This is an incomplete response – the student has not shown that 65 is below the boundary value.
2 marks

Student C

16 (d) Complete this table summarising the number of strokes taken by players in **Round 1**

[1 mark]

Median	Lower quartile	Upper quartile
74	72	75

UQ = 75
LQ = 72

$13/4 = 3.25 \rightarrow 3$

$\underset{\text{IQR}}{\downarrow}$

Inter quartile range = 3

Lowest value = 65   Highest value = 78

range = 13

$72 + 3 = 75$

16 (e) The lowest number of strokes taken in **Round 1** is 65

Show by calculation that this value is an outlier.

[3 marks]

Lowest outlier < LQ − 1·5 × 3

$72 - (1.5 \times 3) = 67.5$

65 is below 67.5 so is an outlier value.

Commentary

Student C has incorrect obtained an Upper Quartile value of 75 in part (d).

They have then used that value completely correctly throughout part (e), fulfilling all the requirements of the question with their value. We follow through their previous value and award full marks.
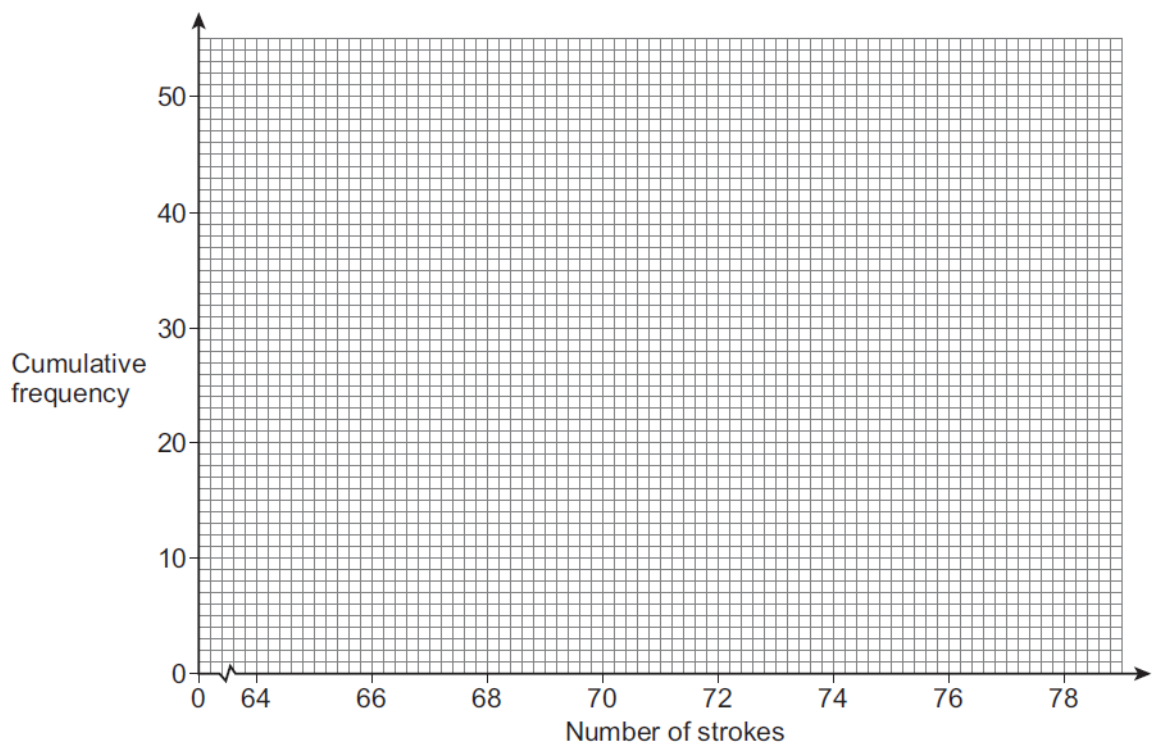
3 marks

Question 16(f)

The table shows a summary of the number of strokes taken by the same players in **Round 2**

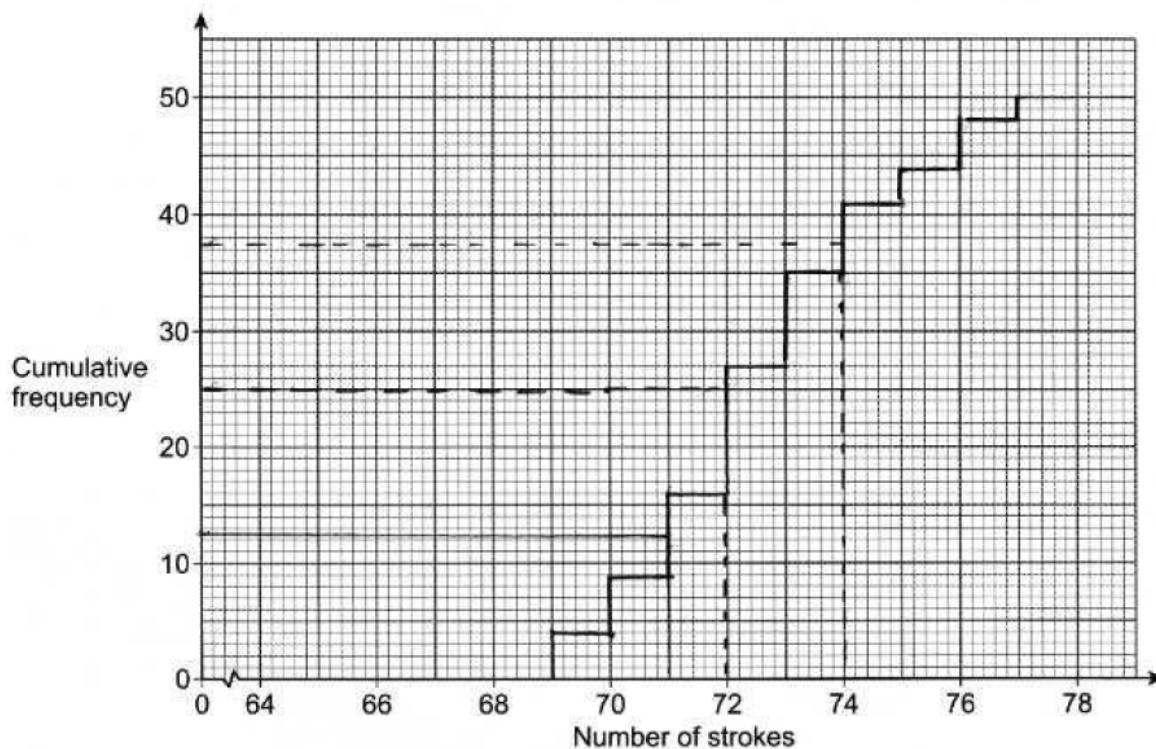| Number of strokes	Frequency	Cumulative frequency
69	4	
70	5	
71	7	
72	11	
73	8	
74	6	
75	3	
76	4	
77	2	

16 (f)   Draw a cumulative frequency step polygon to show the results for **Round 2**

[3 marks]

Student A

Number of strokes	Frequency	Cumulative frequency
69	4	4
70	5	9
71	7	16
72	11	27
73	8	35
74	6	41
75	3	44
76	4	48
77	2	50

Draw a cumulative frequency step polygon to show the results for **Round 2**
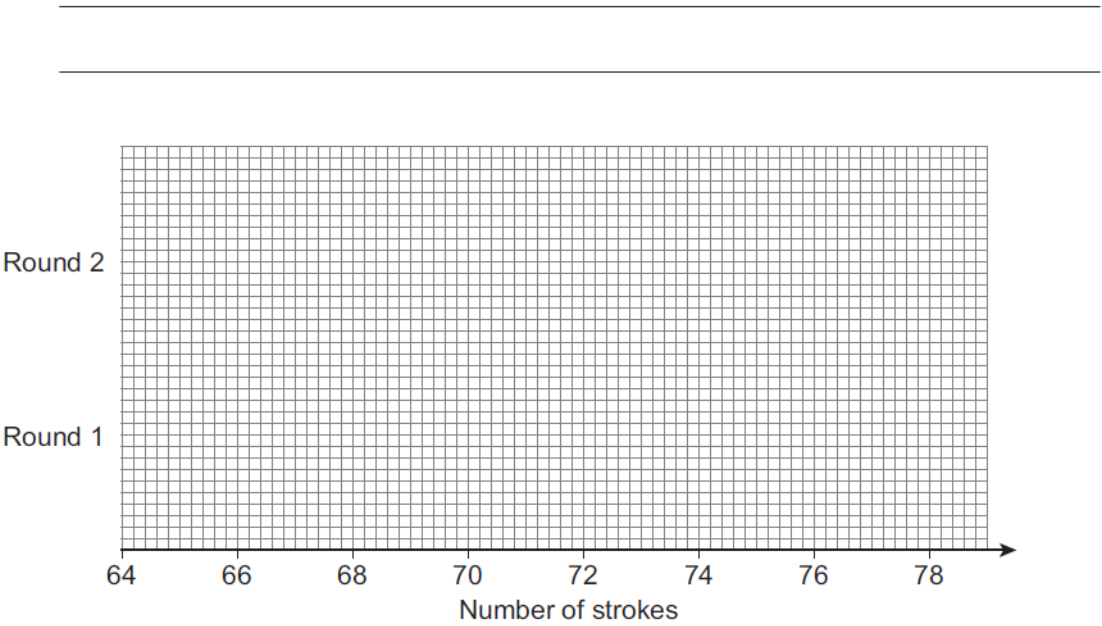

Commentary
A fully correct response.
3 marks

Question 16(g)

16 (g)   Draw separate box plots, on the grid below, for the number of strokes in **Round 1** and
Round 2

Mark clearly the outlier for **Round 1**

[4 marks]

_____

_____

Student A

16 (g) Draw separate box plots, on the grid below, for the number of strokes in **Round 1** and **Round 2**

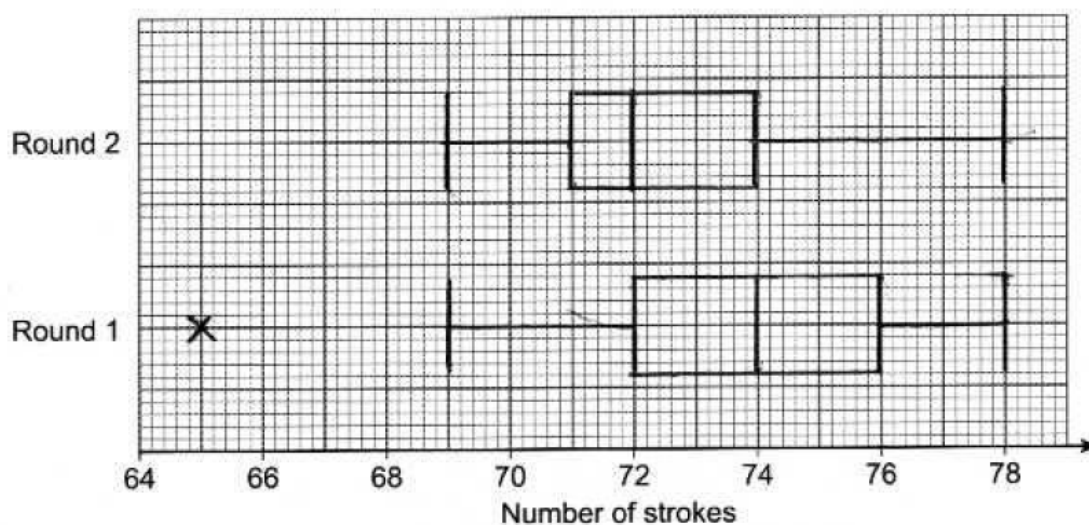Mark clearly the outlier for **Round 1**

[4 marks]

$$LQ = 71 \qquad Median = 72 \qquad UQ = 74$$

Number of strokes

Commentary
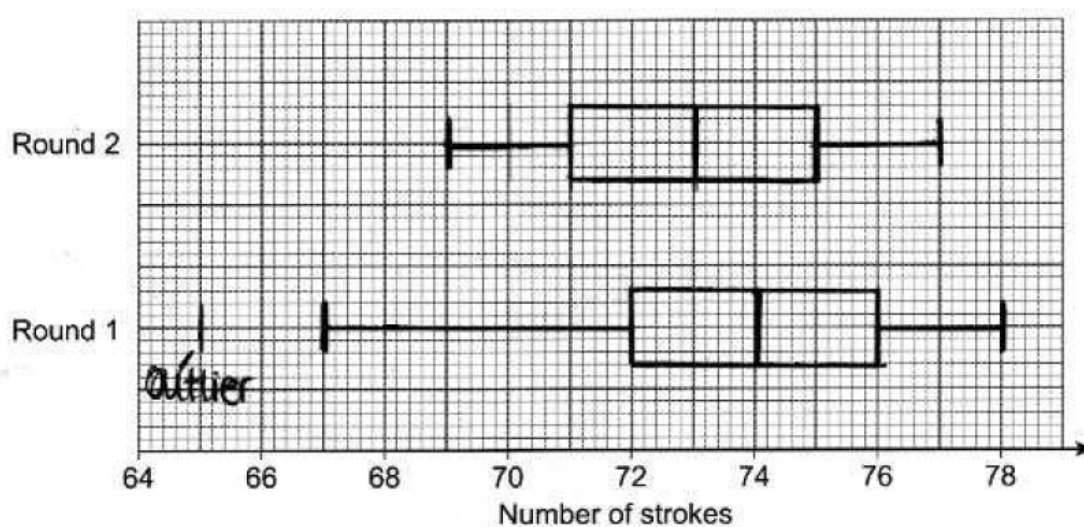Just one error for the Round 2 maximum value.
3 marks

Student B

16 (g) Draw separate box plots, on the grid below, for the number of strokes in **Round 1** and **Round 2**

Mark clearly the outlier for **Round 1**

[4 marks]

LQ : 71   UQ : 72   Median : 73

Commentary

Errors on both box plots but can score 1 mark for getting the box correct for Round 1.
1 mark

Question 16(h)

16 (h) Compare statistically the number of strokes taken for **Round 1** and **Round 2**

[2 marks]

_____

_____

_____

_____

_____

_____

Student A

16 (h) Compare statistically the number of strokes taken for **Round 1** and **Round 2**

[2 marks]

Overal Round 1 had more strokes than Round 2 as
the median is 74 for round one and 72 for
round 2. Round 1 all also had the higher
number for most strokes but lowest for
the least strokes. all together the box is placed
over by 1 and has a bigger IQR by 1.

Commentary

At this level, the 'comparison' of the medians is insufficient to score – needs to be more of an interpretation. The second comment about the maximum is also not a statistical comparison of worth at this level.

0 marks

Student B

16 (h) Compare statistically the number of strokes taken for **Round 1** and **Round 2**

[2 marks]

In round ~~two~~ one there was a higher range for the number of strokes taken than in round 2. The number of strokes taken in round two was more consistent than in round one. Round ~~2~~ 1 the no. of strokes is more positively ~~swett~~ skewed ~~and in round two it is more negatively skewed.~~

Commentary

The interpretation of the range is appropriate for this level, though work on the IQR would probably be preferable. There are no marks for comparison of skew.

1 mark

Student C

16 (h) Compare statistically the number of strokes taken for **Round 1** and **Round 2**

[2 marks]

On average less strokes were taken in round 2 as the median and quartiles are lower. The number of strokes for round 2 is positively skewed whereas there is no skew for round 1. The range and IQR is lower for round 2 then round 1 so the values are less widely spread.

Commentary

The comments on skew can be ignored, the other comments on median, range and IQR are interpreted well enough to be awarded the marks.

2 marks

Question 16(i)

16 (i) Write down a factor that could explain the difference between the number of strokes in the two rounds.

[1 mark]

_____

Student A

16 (i) Write down a factor that could explain the difference between the number of strokes in the two rounds.

[1 mark]

people had already played so had now gotten there practice therefore lowering strokes

Commentary
One of many appropriate responses.
1 mark

Student B

16 (i) Write down a factor that could explain the difference between the number of strokes in the two rounds.

[1 mark]

Weather conditions, e.g. wind.

Commentary
One of many appropriate responses.
1 mark