AQA

Please write clearly in block capitals.

Centre number | | | | | | Candidate number | | | | |

Surname _____

Forename(s) _____

Candidate signature _____

I declare this is my own work.

GCSE
STATISTICS

H

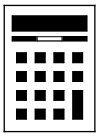
Higher tier Paper 2

Time allowed: 1 hour 45 minutes

Materials

For this paper you must have:

- a calculator
- mathematical instruments.
- a copy of the Data Sheet.



Instructions

- Use black ink or black ball-point pen. Draw diagrams in pencil.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- You must answer the questions in the spaces provided. Do not write outside the box around each page or on blank pages.
- If you need extra space for your answer(s), use the lined pages at the end of this book. Write the question number against your answer(s).
- Do all rough work in this book. Cross out any work you do not want to be marked.

Information

- The marks for the questions are shown in brackets.
- The maximum mark for this paper is 80.
- You may ask for more answer paper and graph paper. These must be tagged securely to this answer booklet.

For Examiner's Use	
Question	Mark
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
TOTAL	

Answer **all** questions in the spaces provided.

Do not write outside the box

1 Amy collected data about the number of pups born in guinea pig litters.

The data are given below.

2 4 5 8 6 6 4 3 4

Use the data to estimate the probability that her guinea pig 'Caramel' will have 6 pups in her litter.

Circle your answer.

[1 mark]

$\frac{1}{9}$ $\frac{2}{9}$ $\frac{2}{7}$ 2

1

2 Which of the following is a measure of the **change** in the cost of goods and services?

Circle your answer.

[1 mark]

Standardised score Gross domestic product

Average seasonal effect Consumer price index

1

3 Four values have a median of 10.

Three of the values are 6, 10 and 10.

Circle the value that the 4th number could **not** be.

[1 mark]

100 12 10 6

1

Do not write outside the box

4 Which of these is **not** a characteristic of the Normal distribution?
Circle the letter of your answer.

[1 mark]

A The distribution is symmetric

B The distribution is bell-shaped

C mean = median = mode

D mean = standard deviation

1

Turn over for the next question

Do not write outside the box

5 Pierre is collecting data about people and their visits to the cinema.

He hands out questionnaires for people to fill in after they have been to a cinema.

5 (a) One of the questions is about age.

Tick (✓) a box to indicate your age, a (years).

☐ $a < 18$

☐ $18 < a \leqslant 40$

☐ $40 < a \leqslant 60$

☐ $a > 60$

Give **two** criticisms of Pierre's question about age.

[2 marks]

Criticism 1 _____

_____

_____

Criticism 2 _____

_____

_____

Do not write outside the box

5 (b) Pierre also asks this question about regularity of cinema visits.

Tick (✓) how often you visit the cinema.

☐ never

☐ once

☐ twice or more

Give **two** criticisms of this question.

[2 marks]

Criticism 1 _____

_____

_____

Criticism 2 _____

_____

_____

4

Turn over for the next question

Do not write outside the box

6 (a) The total number of cars on the road in the UK in 2017 was 32 000 000.
The number of cars stolen in the UK in 2017 was 86 000.

Calculate the risk of a car being stolen in 2017.
Give your answer as a percentage.

[2 marks]

Answer _____ %

6 (b) The risk of a car being stolen in 2013 was 0.21%.

6 (b) (i) Compare the risk of a car being stolen in 2013 with the risk in 2017.

[1 mark]

6 (b) (ii) There were 30 900 000 cars in the UK in 2013.

Calculate an estimate of the number of cars that were stolen in 2013.

[2 marks]

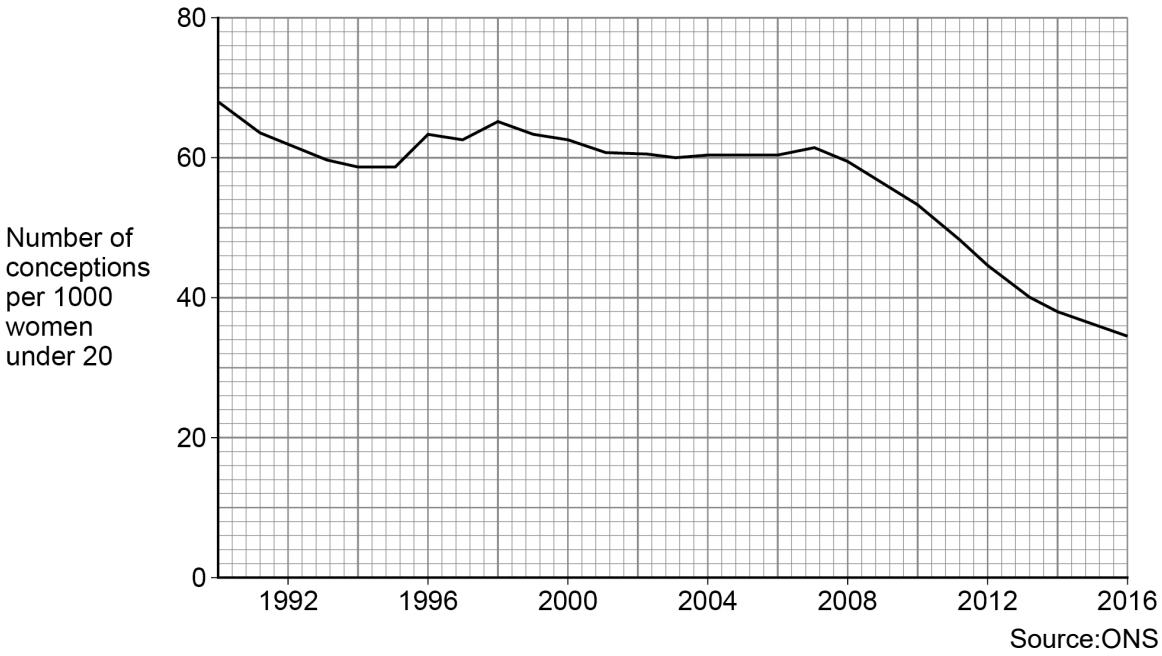
Answer _____

5

Do not write outside the box

7 The graph shows the number of conceptions per 1000 women **under 20** in England and Wales.

A conception is when a woman becomes pregnant.



Source:ONS

7 (a) (i) Make **two** comments about the patterns in the data.

[2 marks]

Comment 1 _____

_____

Comment 2 _____

_____

7 (a) (ii) Give **one** possible reason for the overall trend in the data.

[1 mark]

_____

Question 7 continues on the next page

Do not write outside the box

7 (b) Draw lines to connect the statements with whether they are likely to be correct or not.

[2 marks]

Less than 5% of women under 20 became pregnant in 2012	Definitely correct
There were fewer than 40 babies born to women under 20 in 2016	Probably correct
Fewer women **under 21** became pregnant in 2016 compared to 1992	Definitely incorrect

5

Do not write outside the box

8 Quin is organising a 'Dungeons and Dragons' gaming session for 36 players.
He needs to sort the players, at random, into 3 equal-sized groups.

Explain in detail how Quin could use dice to sort the players at random.

[3 marks]

3

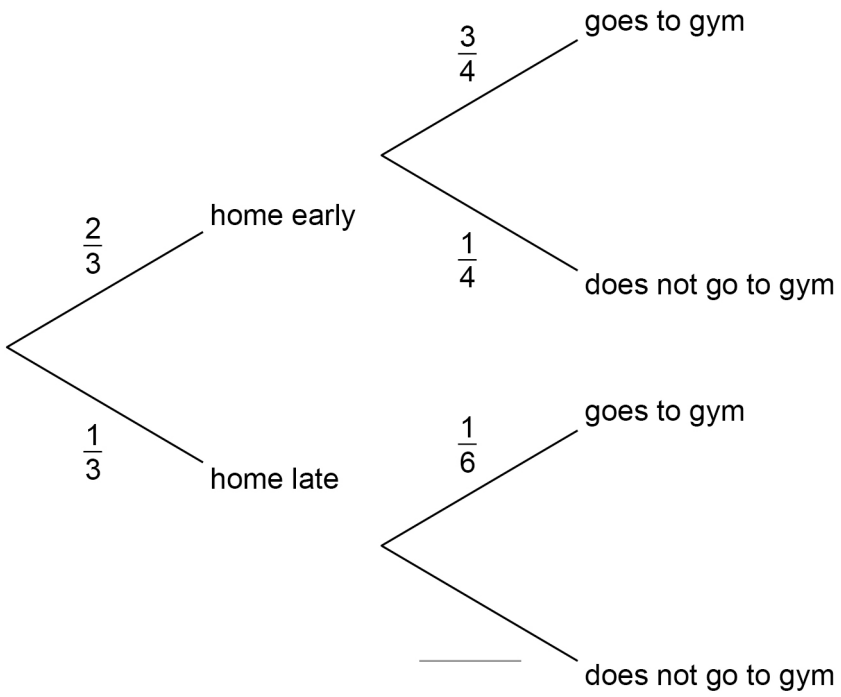
Turn over for the next question

Do not write outside the box

9 Darcey sometimes goes to the gym after work.

The probability of going to the gym is affected by whether she arrives home early or late.

The probabilities are shown in the tree diagram.



9 (a) Write the missing probability on the tree diagram.

[1 mark]

9 (b) What does the probability of $\frac{3}{4}$ represent in this context?

[1 mark]

Do not write outside the box

9 (c) Calculate the probability that, on a randomly chosen work day, Darcey will get home late and go to the gym.

[2 marks]

Answer _____

9 (d) Next year Darcey will work 225 days.

She only goes to the gym on a work day.

Estimate the number of times Darcey will go to the gym next year.

[4 marks]

Answer _____

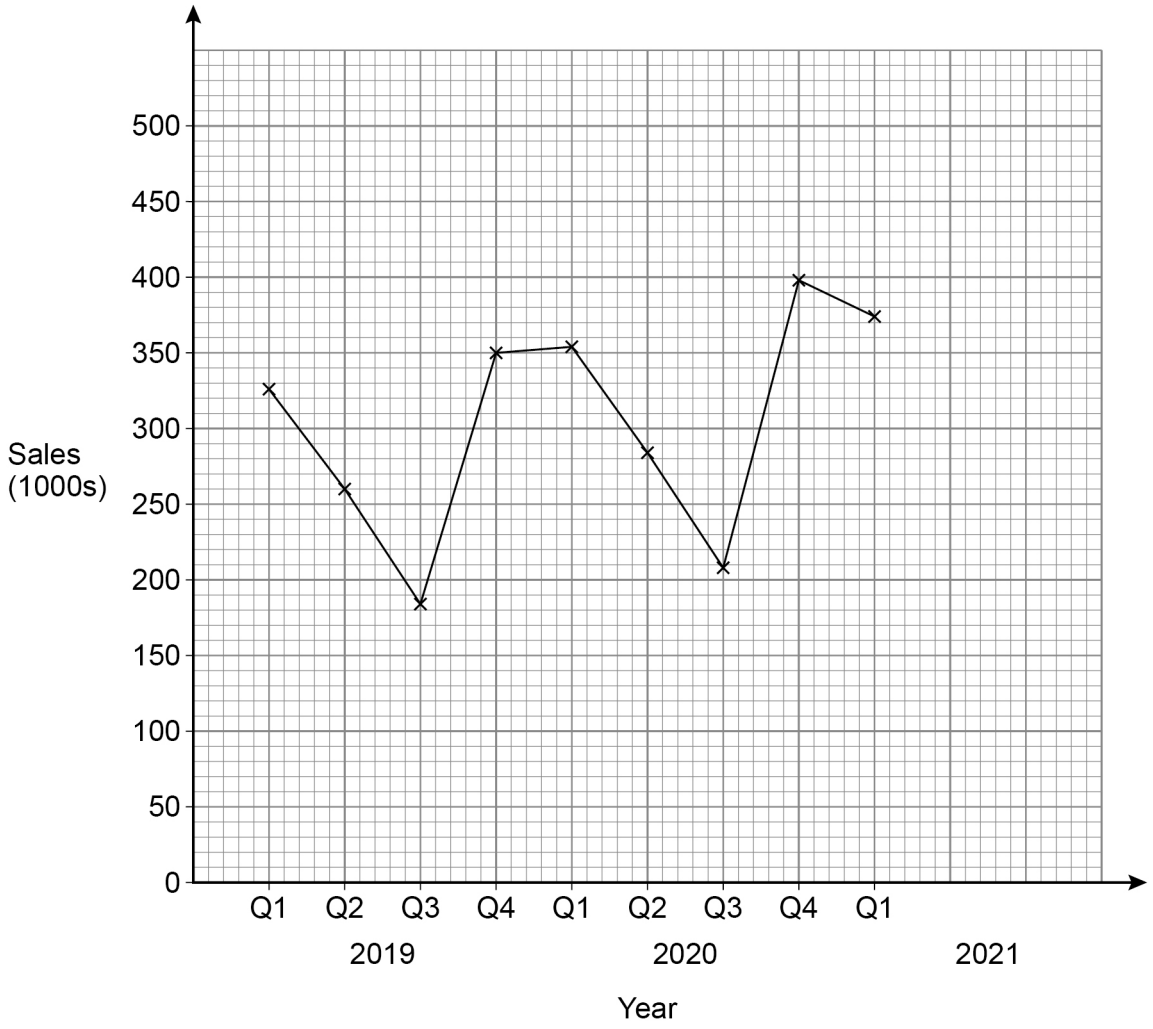
8

Turn over for the next question

Do not write outside the box

10 Sales of apple crumble made by the company Aunt Elsie for quarterly periods are shown in the table and on the time series graph.

Year	**2019**				**2020**				**2021**
Quarter	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1
Sales (1000s)	326	260	184	350	354	284	208	398	374



10 (a) Simeon wants to look at the trend in sales based on these data.

He concludes that 4-point moving averages would be appropriate in this situation.

Give a reason why Simeon is correct.

[1 mark]

Reason _____

_____

Do not write outside the box

10 (b) Some of the 4-point moving averages are in this table with the original data.

Year	2019				2020				2021
Quarter	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1
Sales (1000s)	326	260	184	350	354	284	208	398	374
4-point moving average (1000s)			280	287	293	299			

Calculate the two remaining 4-point moving averages.

Put your answers in the table.

[3 marks]

10 (c) Plot **all** the 4-point moving averages on the graph on the previous page.

[2 marks]

10 (d) Draw an appropriate trend line on the graph.

[1 mark]

10 (e) Describe the trend shown by the graph and your line in **part (d)**.

[1 mark]

Question 10 continues on the next page

Do not write outside the box

10 (f) (i) Calculate the mean seasonal variation (effect) for Quarter **2**.

You **must** complete the table to show your working.

[3 marks]

	Sales (1000s)	**Trend line value (1000s)**	**Seasonal variation (1000s)**
2019 Q2			
2020 Q2			
		Mean seasonal variation =	

Answer _____ thousand

10 (f) (ii) Estimate the number of apple crumbles sold by Aunt Elsie in Quarter 2 of 2021.

[2 marks]

Answer _____ thousand

13

Do not write outside the box

11 In this question,

- SRCC is Spearman's rank correlation coefficient.
- PMCC is product moment correlation coefficient.

Tick (✓) the appropriate box for each statement.

[3 marks]

	True	False
If SRCC has a value of +1, PMCC must also have a value of +1		
If PMCC has a value of +1, SRCC must also have a value of +1		
If PMCC has a value of – 1, the data must form a straight line on a scatter diagram		
If SRCC has a value of – 1, the data must form a straight line on a scatter diagram		

3

Turn over for the next question

Do not write outside the box

12 In 2018, Vales High School had a budget of £3.8 million.

The percentage of the budget spent on different areas is given in the table.

Item	Staffing	Energy/Other	Services/Goods
Percentage (3sf)	87.2%	8.30%	4.50%

12 (a) Calculate the amount of money spent on staffing in 2018.

Give your answer in thousands of pounds, to the nearest thousand.

[2 marks]

Answer £ _____ thousand

12 (b) Using 2018 as the base year, the index numbers for each area for 2020 are as follows.

Item	Staffing	Energy/Other	Services/Goods
Index (2018 as base)	102.3	109.2	99.0

12 (b) (i) Calculate the weighted index number for 2020 costs, using 2018 as base.

[3 marks]

Answer _____

Do not write outside the box

12 (b) (ii) Hence calculate the budget required to meet the increased costs for 2020.

Give your answer to three significant figures.

[2 marks]

Answer £ _____

7

Turn over for the next question

Do not write outside the box

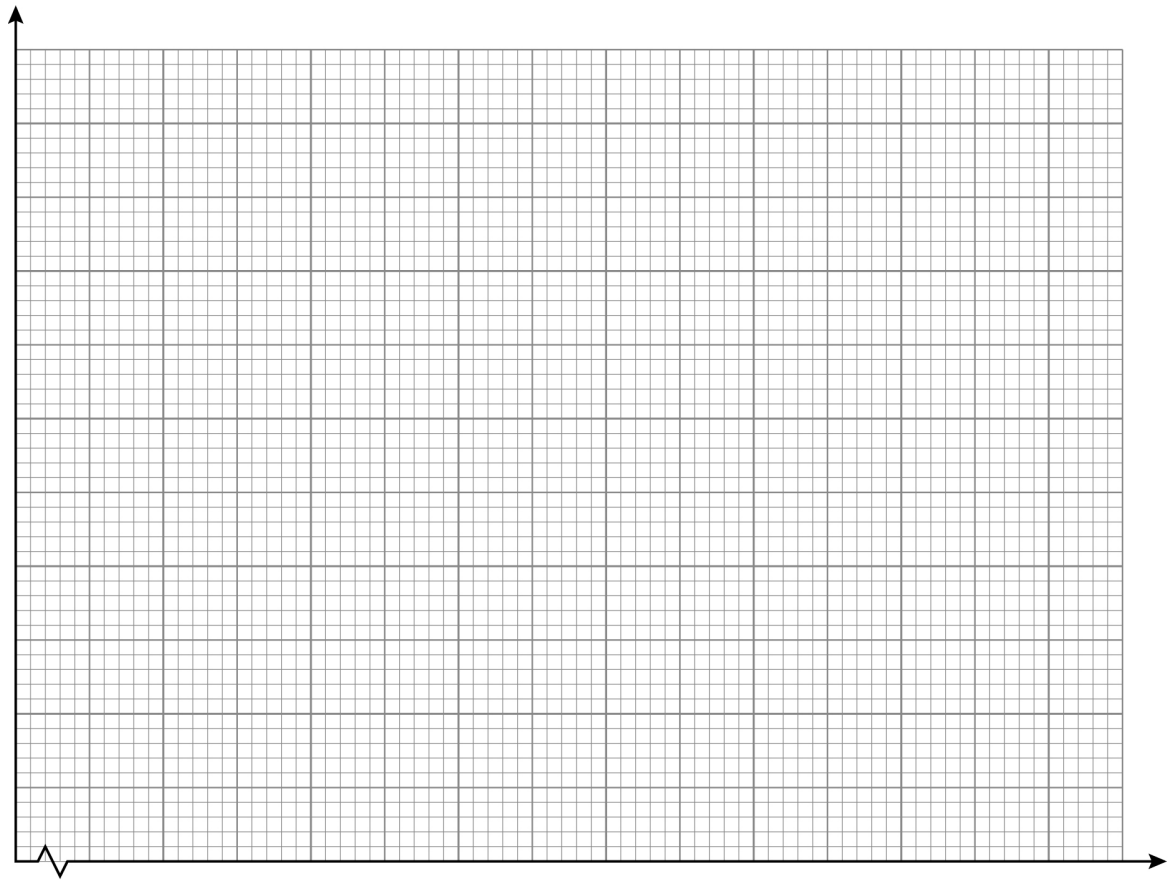
13 Bob records the maximum and minimum temperatures in his garden each day.

The table summarises the maximum temperatures for days in June over the last 5 years.

Temperature, t (°C)	**Frequency**		
$14 \leqslant t < 19$	45		
$19 \leqslant t < 20$	20		
$20 \leqslant t < 22$	29		
$22 \leqslant t < 25$	27		
$25 \leqslant t < 30$	20		
$30 \leqslant t < 40$	9		

13 (a) Draw a histogram using the grid below.

[4 marks]



Do not write outside the box

13 (b) From the frequency table, the estimated mean is 21.6 °C and the estimated standard deviation is 5.0 °C (both given to one decimal place).

13 (b) (i) The mean of the individual maximum temperatures for the 150 June days is 21.9 °C.

Tick (✓) the correct box.

[1 mark]

☐ On average, the actual temperatures are above the midpoint for the group they are in.

☐ On average, the actual temperatures are equal to the midpoint for the group they are in.

☐ On average, the actual temperatures are below the midpoint for the group they are in.

13 (b) (ii) The highest temperature recorded over the 150 June days was 37.9 °C.

Assuming the data is from a Normal distribution, use the estimated mean and estimated standard deviation to determine whether 37.9 °C is a statistical outlier.

You **must** show your working.

[2 marks]

Question 13 continues on the next page

Do not write outside the box

13 (c) Bob says that over half of the June days had a maximum temperature under 21 °C.

Evaluate Bob's statement.

[3 marks]

10

Do not write outside the box

14 England will play Australia in five cricket matches.

Before each match the England captain tosses a coin.

The Australian captain calls 'Heads' or 'Tails' and whoever wins the toss chooses to bat or bowl first.

14 (a) Work out the **exact** probability that the same captain wins all five tosses.

[2 marks]

Answer _____

14 (b) Give **two** reasons why the number of times out of five that the Australian captain wins the toss can be modelled by a Binomial distribution.

[2 marks]

1 _____

2 _____

4

Turn over for the next question

Do not write outside the box

15 You will need the **Data Sheet** to answer this question.

Mark and Fiona carry out roadside safety checks on trucks.

For each truck, they record,

- the number of wheels
- its length
- its mass.

15 (a) Circle the name given to the data that they collect.

[1 mark]

bivariate multivariate secondary qualitative

15 (b) Mark records data from trucks travelling on one part of the A1 road.

He records data from,

- the first 20 trucks with 6 or fewer wheels
- the first 20 trucks with more than 6 wheels.

His data are given on the **Data Sheet**.

15 (b) (i) State **one** problem with the data Mark has collected.

Suggest a solution to deal with this problem.

[2 marks]

Problem _____

_____

Solution _____

_____

15 (b) (ii) Mark wants to use his data to estimate the mean mass of trucks using this part of the A1 road.

Explain why the data Mark has collected is **not** likely to be suitable for this purpose.

[1 mark]

_____

_____

_____

Do not write outside the box

15 (c) Fiona carries out her checks on two roads, the A2 and the A229.

15 (c) (i) Some summary statistics for the lengths of trucks she checks on the **A2** are shown.

mean	10.20 metres
median	9.18 metres
standard deviation (s.d.)	2.90 metres

Calculate the skew of the data.

Use skew = $\frac{3(\text{mean} - \text{median})}{\text{s.d.}}$

[2 marks]

Answer _____

15 (c) (ii) Fiona says,

> "The data show positive skew, so the trucks below median length have more variable lengths than the trucks above median length."

Has Fiona interpreted the skew correctly?

Tick (✓) a box.

Yes ☐ No ☐

Explain your answer.

[1 mark]

Turn over ►

Do not write outside the box

15 (d)

$$\text{s.d.} = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

Fiona checks the lengths of **30** trucks on the **A229**.

The lengths, x (metres), of these 30 trucks are summarised by

$$\sum x = 267.12 \qquad \sum x^2 = 2538.52 \qquad \text{skew} = 0.43$$

Compare statistically the length of trucks checked on the A229 with those checked on the A2.

You **must** show your working.

[6 marks]

15 (e) Give a reason why the standard deviation may not be the best measure of spread to summarise Fiona's data.

[1 mark]

14

END OF QUESTIONS

Do not write outside the box

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED

Do not write outside the box

Question number	**Additional page, if required.** **Write the question numbers in the left-hand margin.**

Do not write outside the box

Question number	**Additional page, if required.** **Write the question numbers in the left-hand margin.**

Do not write outside the box

Question number	**Additional page, if required.** **Write the question numbers in the left-hand margin.**