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

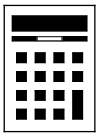
H

Higher tier Paper 1

Time allowed: 1 hour 45 minutes

Materials

For this paper you must have:
- a calculator
- mathematical instruments.



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	
TOTAL	

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

Do not write outside the box

1 Two fair spinners, each numbered 1 to 8, are spun.
The numbers they land on are added up.

What is the probability the total is 16?

[1 mark]

$\frac{1}{4}$ $\frac{1}{16}$ $\frac{1}{32}$ $\frac{1}{64}$

1

2 Here is the definition of a term used in sampling.

'Those who are actually available to be part of a survey or investigation.'

Circle the term being defined.

[1 mark]

sample frame census group population

1

3 Which of these data lists

is bi-modal
and
has the mean double the median?

Circle the letter of your answer.

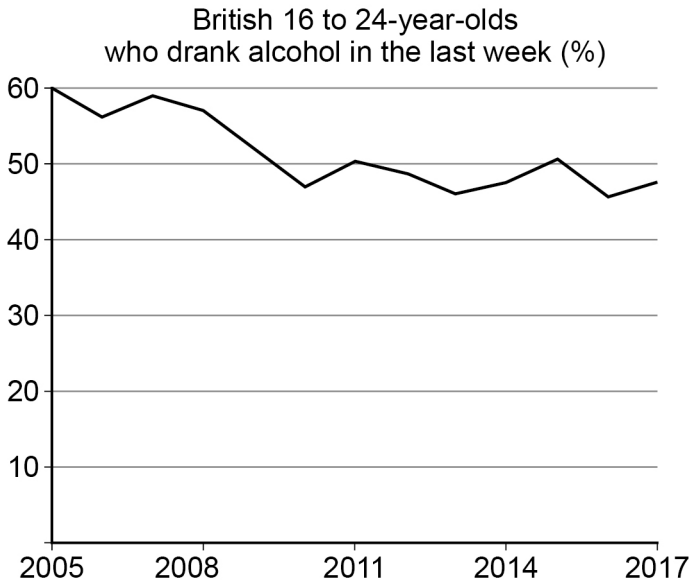
[1 mark]

A	4	4	6	6	30
B	2	4	6	6	42
C	0	6	6	24	24
D	0	0	6	6	38

1

Do not write outside the box

4 This graph was seen on the BBC News App.



Source: ONS

Circle the letter of the statement for the graph which is **definitely** true.

[1 mark]

A 2005 had the highest % of 16 to 24-year-olds drinking since 2000.

B The % of 16 to 24-year-olds drinking has fallen every year from 2005.

C Most 16 to 24-year-olds in 2017 did not drink alcohol in the last week.

D Most 16 to 24-year-olds in 2017 do not drink alcohol at all.

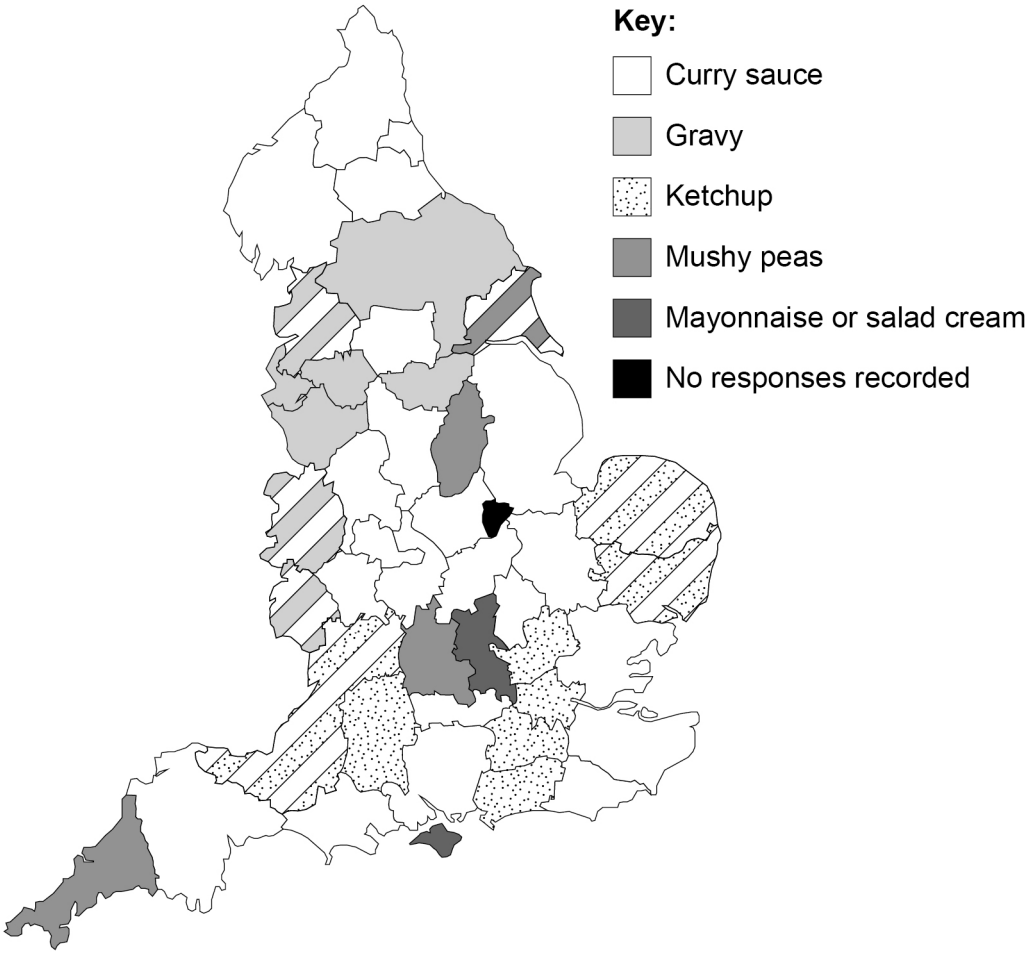
1

Turn over for the next question

Do not write outside the box

5 A sample of 670 adults in England were asked what side order they preferred at fish and chip shops.

A striped area indicates two equally popular side orders.



5 (a) Based on area of land, which is the most popular side order?

[1 mark]

Answer _____

5 (b) Give **two** reasons why your answer to **(a)** might not be the side order that **most** people eating fish and chips in England prefer.

[2 marks]

1 _____

_____

2 _____

_____

3

Do not write outside the box

6 Tom is doing a statistical study into the amount of homework received by Year 7 and Year 11 students in his school.

6 (a) Write down a hypothesis Tom could use.

[1 mark]

6 (b) State the population of his study.

[1 mark]

6 (c) Tom wants a sample of Year 7 students and a sample of Year 11 students to complete a questionnaire for him.

He considers these three sampling methods for Year 7 students.

Method A

Number all the students in Year 7.
Obtain 30 random numbers.
Ask the students whose random numbers come up to complete the questionnaire.

Method B

Wait outside the dinner hall.
Ask the first 30 Year 7 students he sees to complete the questionnaire.

Method C

Choose three Year 7 students from each of the 10 maths sets.
Ask these students to complete his questionnaire.

Name and compare the merits of each sampling method.

Make a reasoned choice of which method Tom should use.

[7 marks]

Do not write outside the box

Do not write outside the box

6 (d) One of Tom's questions is,

'How much homework do you receive?'

Write down **two** problems with this question.

[2 marks]

Problem 1 _____

_____

Problem 2 _____

_____

6 (e) Tom improves his questionnaire and collects his data.

He finds that,

- on average Year 7 have five hours of homework per week
- on average Year 11 have eight hours of homework per week.

Write a possible conclusion for Tom.

[1 mark]

_____

_____

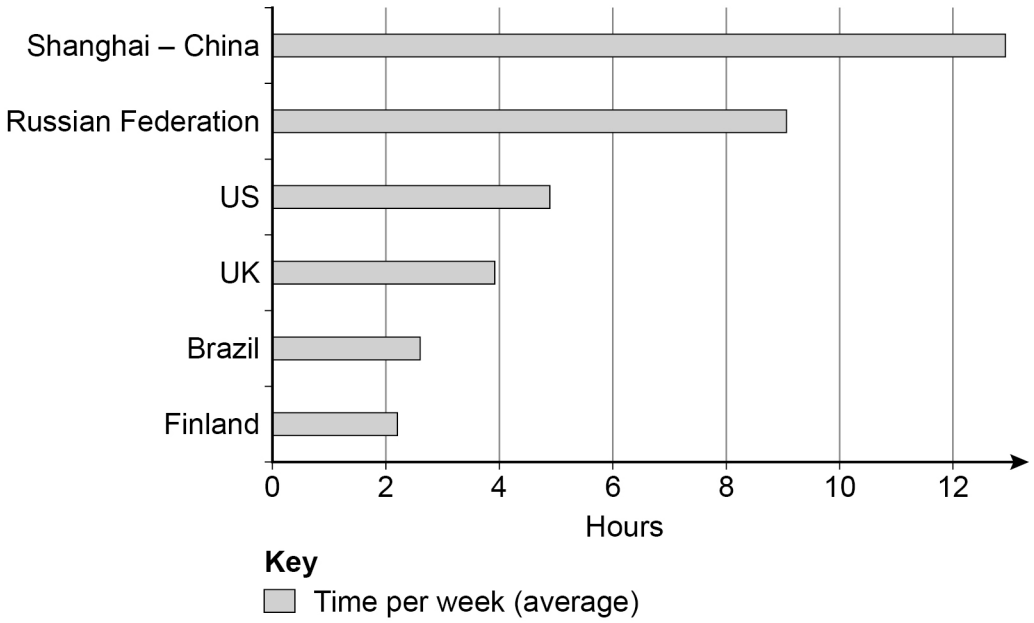
Question 6 continues on the next page

Do not write outside the box

6 (f) Tom wonders how this compares with other schools in the UK and schools in other countries.

He finds this chart on the internet but it has no source.

How much time do 15-year-olds spend on homework?



Use the chart to compare Tom's Year 11 results of an average of 8 hours homework per week with those for other schools in the UK and with other countries.

[2 marks]

Tom's school and other UK schools _____

_____

_____

Tom's school and schools in other countries _____

_____

_____

6 (g) Why are Tom's data and the internet data not completely comparable?

[1 mark]

_____

_____

Do not write outside the box

6 (h) Is Tom's data or the internet data more reliable?

Give a reason for your answer.

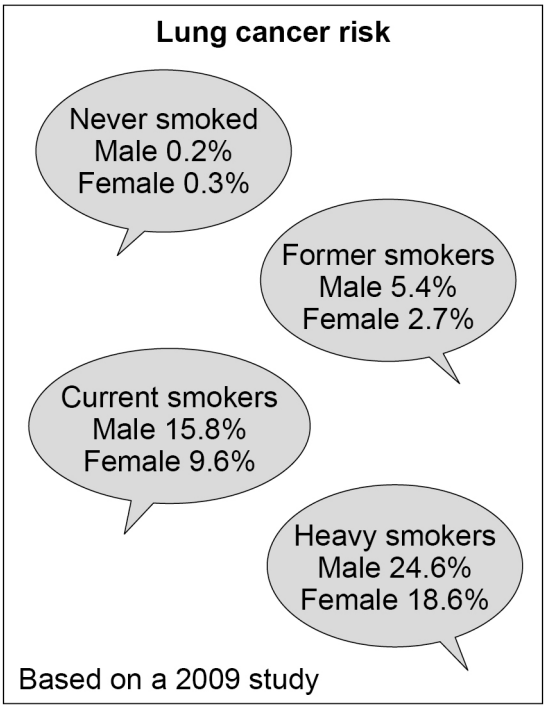
[1 mark]

16

Turn over for the next question

Do not write outside the box

7 This picture gives the risk of getting lung cancer and how smoking changes that risk.



7 (a) How many out of every 1000 women who have **never smoked** are expected to get lung cancer?

[2 marks]

Answer _____

7 (b) How many times greater is the risk of getting lung cancer for a man who is a **heavy smoker** compared with a man who has **never smoked**?

[2 marks]

Answer _____

Do not write outside the box

7 (c) Give **two** reasons why the picture might not be that helpful for someone to estimate their risk of lung cancer now.

[2 marks]

1 _____

_____

_____

2 _____

_____

_____

6

Turn over for the next question

Do not write outside the box

8 Rachel and Mitch own a café.

They want to decide whether more staff are needed at certain times.

On one day, Rachel records how long customers wait to be served at specific times **after 9am**.

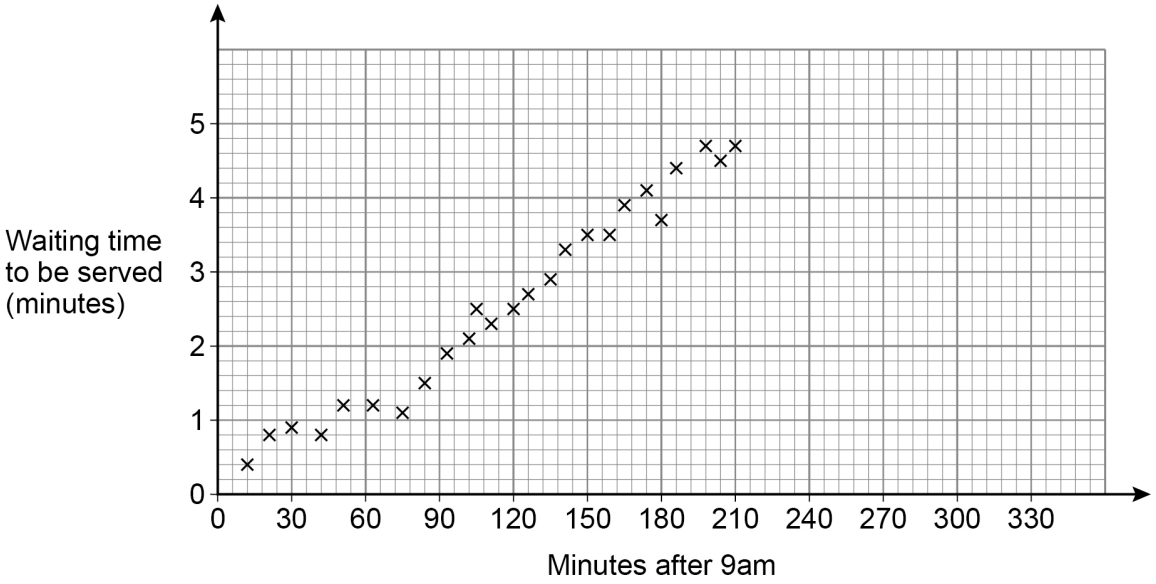
8 (a) Write down the independent and dependent variables.

[2 marks]

Independent variable _____

Dependent variable _____

8 (b) The scatter diagram shows the results for 25 customers Rachel sampled.



8 (b) (i) Which of these is a possible equation for a regression line for the data shown?

Circle your answer.

[1 mark]

$y = -0.06 + 0.689x$

$y = -0.06 + 0.023x$

$y = 4.8 - 0.689x$

$y = 4.8 - 0.023x$

Do not write outside the box

8 (b) (ii) Here is some information about the results sampled for later in the day.

- After 12.30pm but before 1.30pm all five customers sampled waited between 4 and 5 minutes.
- From 1.30pm, the six results showed a strong negative correlation.

On the scatter diagram, show possible results for these additional 11 sampled customers.

[2 marks]

8 (b) (iii) Mitch decides to employ an extra person for a two-hour time period.

Which time period would you suggest?

[1 mark]

Answer _____ to _____

8 (c) Lucy is a statistician who visits the café.

She identifies a problem with Rachel's data collection strategy and offers a solution.

Describe the problem and the solution Lucy may have suggested.

[2 marks]

Problem _____

_____

_____

Solution _____

_____

_____

8

Turn over for the next question

Do not write outside the box

9 A tech company's website offers rewards according to how many days you visit it.

The table shows the number of rewards achieved by a sample of 500 customers.

Number of rewards	% of sample
1	52.8
2	27.2
3	10.4
4	7.6
5	2

9 (a) Show, with working that, for this sample,

median = interquartile range

[3 marks]

Do not write outside the box

9 (b) The tech company wants to change the reward system so that

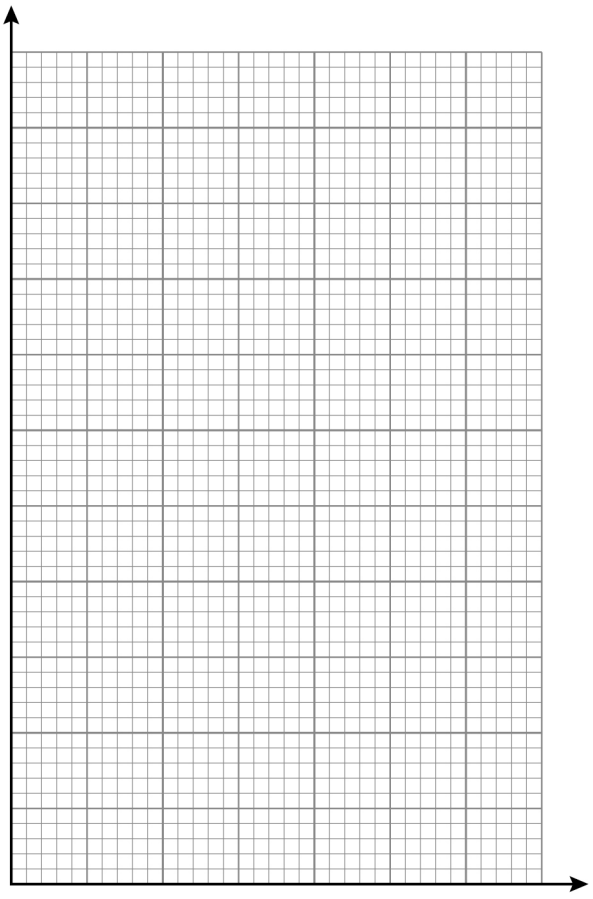
- the median number of rewards is increased
- the interquartile range is increased.

They make some changes and take a new sample of 500 customers.

Number of rewards	Number in sample
1	190
2	145
3	90
4	60
5	15

Use the grid to draw an **appropriate** cumulative frequency graph for the data.

[4 marks]



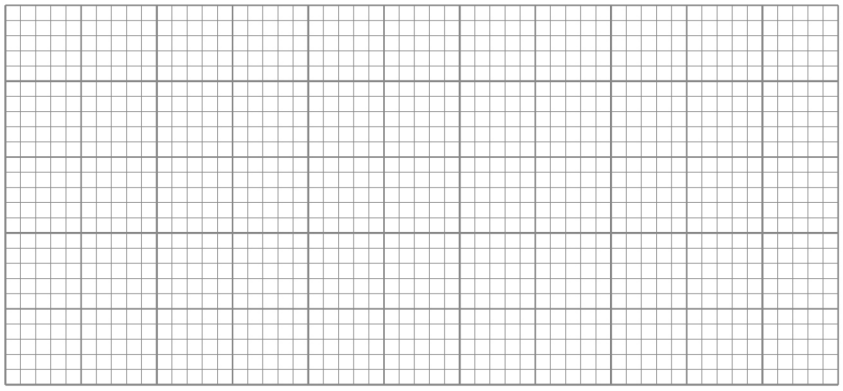
Question 9 continues on the next page

Do not write outside the box

9 (c) Use **parts (a)** and **(b)** to draw, on the same grid, box plots for the two samples.

When drawing the box plot do **not** check for outliers.

[4 marks]



9 (d) Use calculations to confirm that there are no outliers in the **new** sample of 500 customers.

You **must** show your working.

[3 marks]

Do not write outside the box

9 (e) Has the company achieved each of its targets stated in **part (b)**?

You **must** give a reason for each target.

[2 marks]

Target 1 – median _____

Target 2 – interquartile range _____

16

Turn over for the next question

Do not write outside the box

10 Fifteen people apply to take part in an 'IronMan' triathlon.

Before they can compete they must prove they are fit enough to attempt the course.

Angelina collects data on each competitor's,

- age
- blood pressure
- resting heart rate.

10 (a) Angelina thinks that finishing position in the triathlon (1st, 2nd, 3rd, etc) and resting heart rate will show positive correlation.

Write a hypothesis for her to investigate.

[1 mark]

10 (b) Once the race is run, Angelina notes the order in which the competitors finished, from 1st to 15th position.

She also ranks the resting heart rate data from lowest to highest.

She calculates that the value of $\sum d^2$ = 60, where d is the difference in the ranks of the finishing position and the resting heart rates.

Show that the value of Spearman's Rank Correlation Coefficient (SRCC) is 0.89 (to 2 decimal places).

Use SRCC = $1-\frac{6\sum d^2}{n(n^2-1)}$

[3 marks]

Do not write outside the box

10 (c) Write a possible conclusion to your hypothesis in **part (a)**.

[1 mark]

5

Turn over for the next question

Do not write outside the box

11 A large forest contains an unknown number of squirrels.

Fynn is asked to estimate the number of squirrels in the forest.

He catches 50 and tags them before releasing them back into the forest.

Two weeks later, he catches 40 more squirrels and finds that 11 have a tag.

11 (a) (i) Give **one** reason why Fynn waits two weeks before catching the 40 squirrels.

[1 mark]

11 (a) (ii) Give **one** reason why Fynn doesn't wait a lot longer than two weeks.

[1 mark]

11 (b) Calculate an estimate of the number of squirrels in the forest.

[3 marks]

Answer _____

5

Do not write outside the box

Turn over for the next question

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED

Do not write outside the box

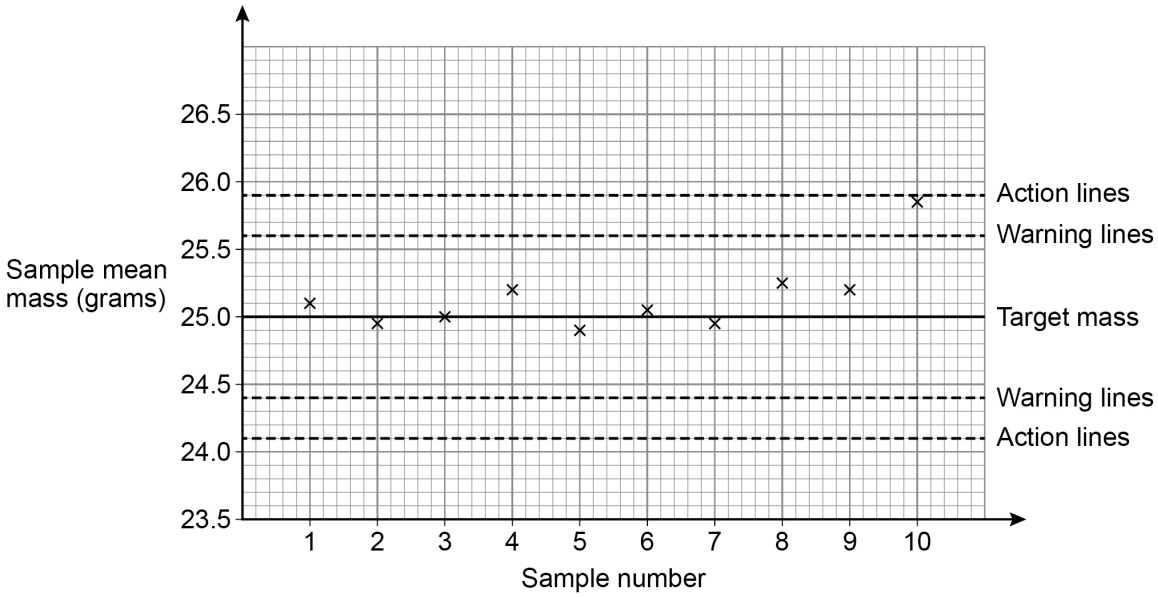
12 In a factory, three different machines, **A**, **B** and **C**, put crisps into bags.

Filled bags are labelled as having a mass of 25g.

Every two days, Luka, the quality control manager, obtains a sample of filled bags from each machine and calculates the sample mean.

12 (a) The control chart for 10 consecutive large samples for **Machine A** is shown below, together with warning and action limits.

Machine A



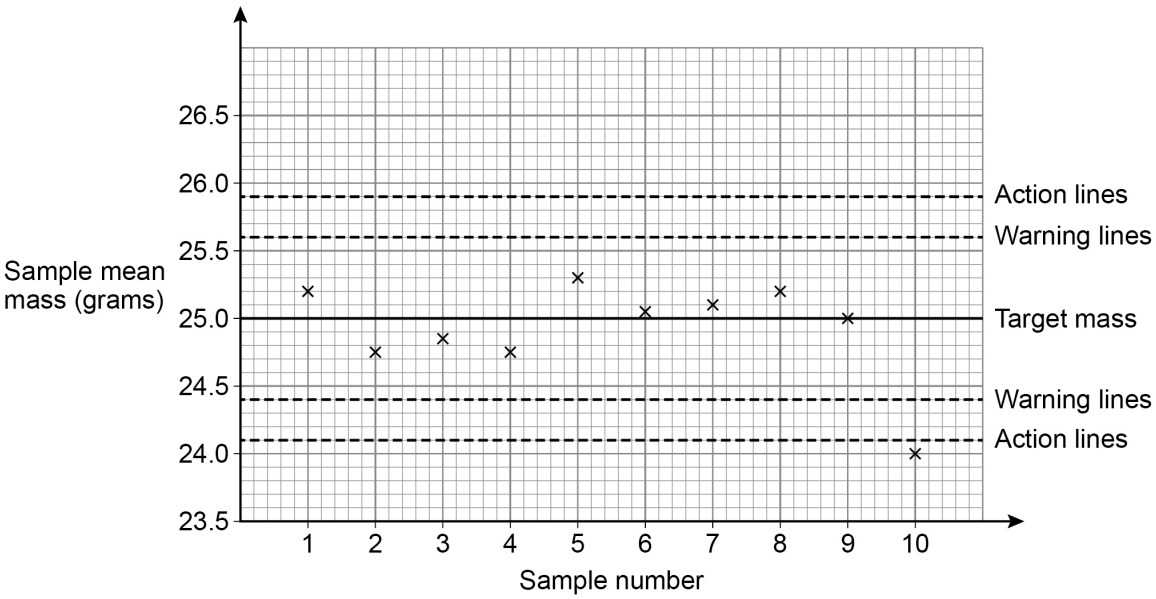
What should Luka do, if anything, with **Machine A**, based on this chart?

[1 mark]

Do not write outside the box

12 (b) The control chart for 10 consecutive large samples for **Machine B** is shown below, together with warning and action limits.

Machine B



What should Luka do, if anything, with **Machine B**, based on this chart?

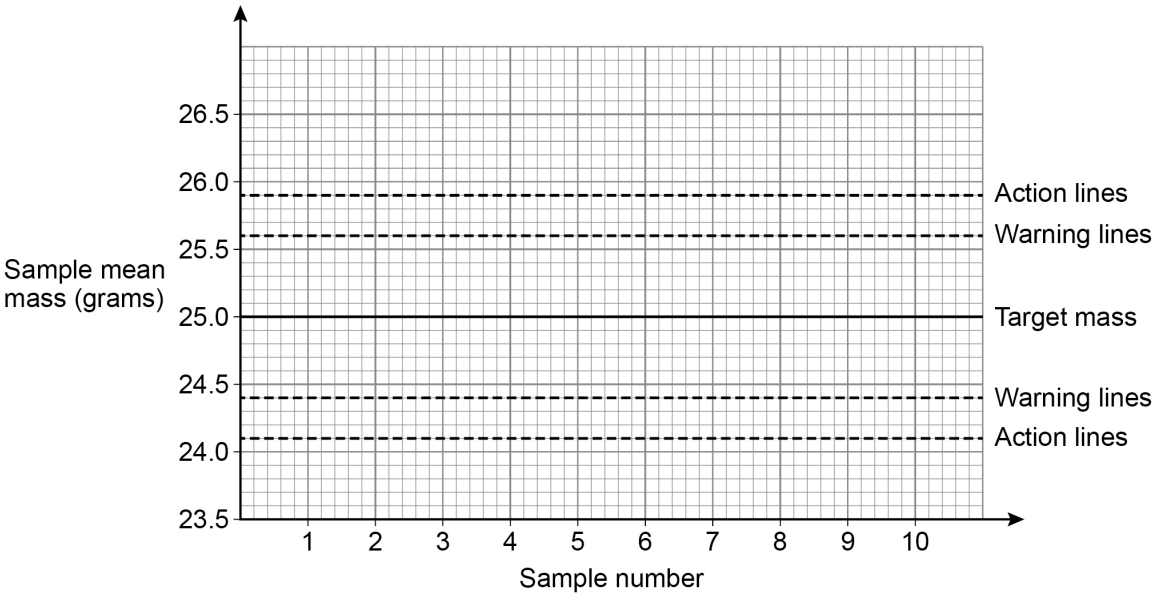
[1 mark]

12 (c) Ten large samples for **Machine C** show it to be working normally.

Complete a possible control chart for **Machine C** on the graph below.

[1 mark]

Machine C



3

Do not write outside the box

13 Maisy has her house valued each year.

The table shows the percentage increase in its value, written as a multiplier, over each of the last 5 years.

Year	2016	2017	2018	2019	2020
Multiplier	1.067	1.013	1.008	1.005	1.126

13 (a) Maisy wants to know the average percentage increase over the 5 years.

She works out the arithmetic mean.

$$\frac{6.7 + 1.3 + 0.8 + 0.5 + 12.6}{5} = 4.4\%$$

Average percentage increase = 4.4% to one decimal place.

13 (a) (i) Maisy has used the wrong method.

Name the measure she should have calculated.

[1 mark]

13 (a) (ii) Show that the correct average percentage increase is 4.3% to one decimal place.

[2 marks]

Do not write outside the box

13 (a) (iii) In 2020 Maisy's house was worth £200 000.

If the house continues to grow in value at the same average rate, how much will it be worth in 2024?

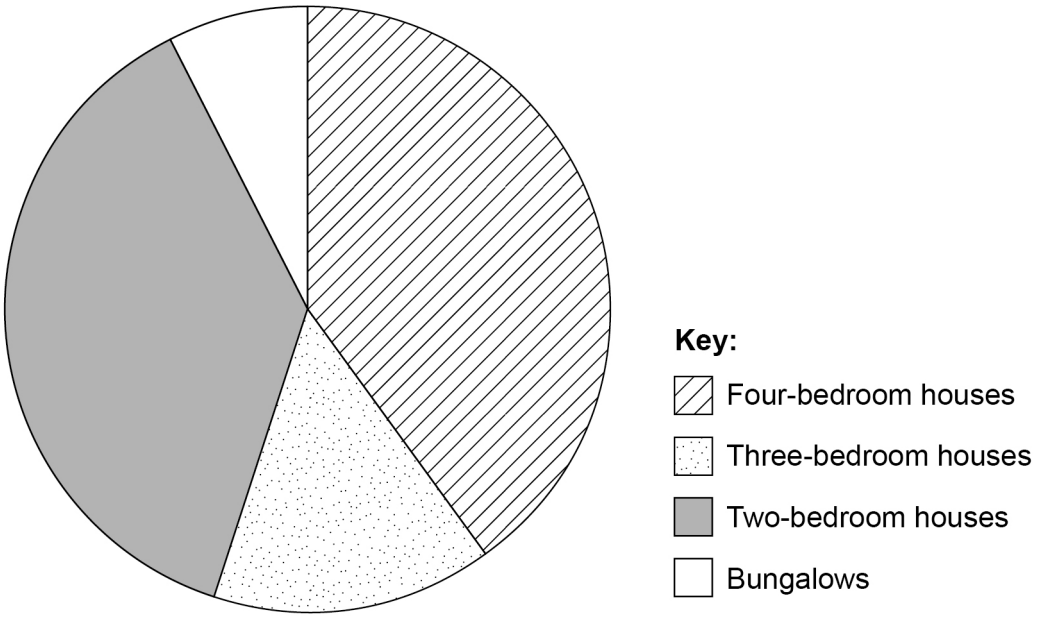
[2 marks]

Answer **£** _____

Question 13 continues on the next page

Do not write outside the box

13 (b) A new housing estate with 80 houses is built near Maisy's house.
She sees this pie chart in her local paper showing the types of houses in the estate.



Vicky buys a bungalow on this new estate.

What proportion of the houses on this new estate are bungalows?

[2 marks]

Answer _____

Do not write outside the box

13 (c) Maisy lives on an older housing estate.

On **Maisy's estate** there are,

- 60 four-bedroom houses
- 38 three-bedroom houses
- 22 two-bedroom houses
- 40 bungalows.

In the space below, draw a fully labelled comparative (proportional) pie chart to represent the types of houses on Maisy's estate.

[6 marks]

Maisy's estate

Question 13 continues on the next page

Do not write outside the box

13 (d) Make **one** comparison between the proportions of one of the types of houses on Maisy's estate with those on the new estate.

[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.**