

Name: _____

GCSE Statistics

Histograms

Total marks available: 63

Total marks achieved: _____

Instructions

- Use black ink or ball-point pen.
- Fill in the boxes at the top of this page with your name, Centre number and candidate number.
- Answer all questions.
- Answer the questions in the spaces provided
 - There may be more space than you need.
- Scientific calculators may be used.
- You must show all your working out with your answer clearly identified At the end of your solution.

Information

- The marks for each question are shown in brackets
 - use this as a guide as to how much time to spend on each question.

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.

Q1.

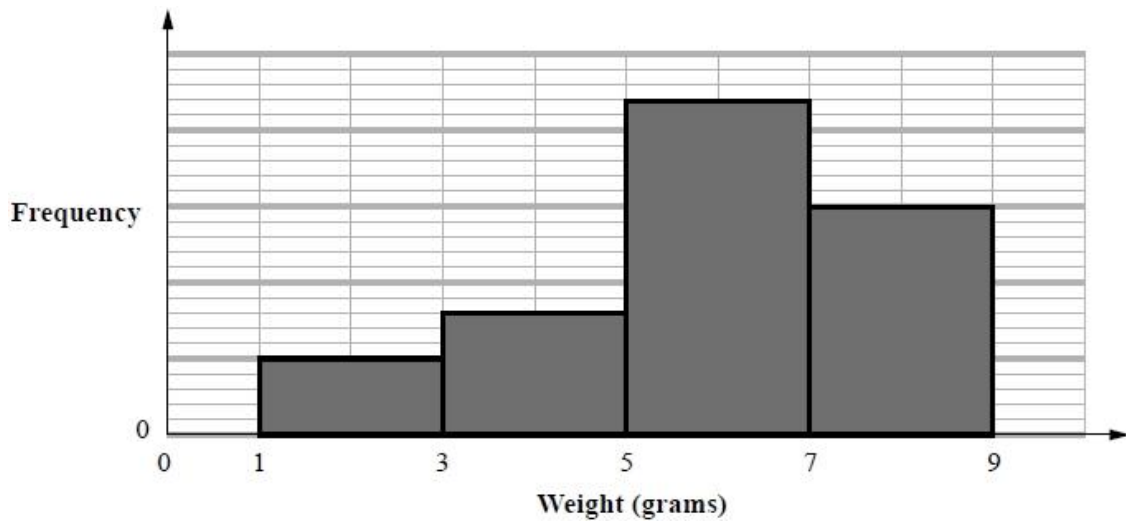
Tomoyo found the weight, in grams, of each of 100 cherries.

(a) Circle the **two** words from the list that best describe the data Tomoyo found.

- quantitative qualitative discrete continuous bivariate ordinal categorical

(2)

Tomoyo grouped the weights and she then drew this diagram for her results.



The incomplete frequency table shows some information about her results.

Weight (w grams)	Frequency
$1 \leq w < 3$	10
$3 \leq w < 5$	
$5 \leq w < 7$	
$7 \leq w < 9$	

(b) (i) complete the frequency column in the table.

(2)

(ii) Calculate an estimate of the mean weight of the 100 cherries.

..... 9

(3)

(Total for question = 7 marks)

Q2.

The table gives information about the birth weights (x kg) of 145 elephants.

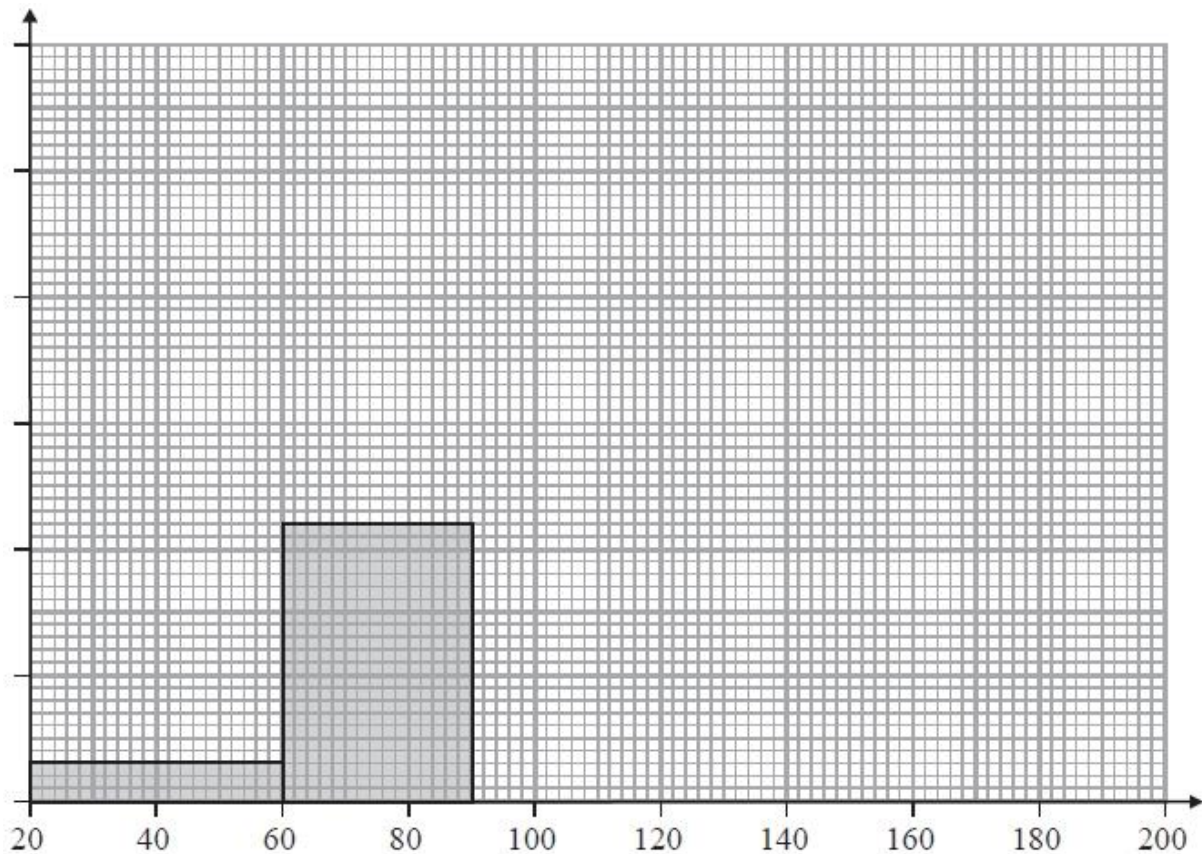
Birth weight (x kg)	Frequency (f)		
$20 < x \leq 60$	6		
$60 < x \leq 90$	33		
$90 < x \leq 110$	40		
$110 < x \leq 130$	37		
$130 < x \leq 160$	24		
$160 < x \leq 200$	5		

(Data source: www.elephant.se)

Complete the histogram for this information and label the axes.

You may use the extra columns in the table.

(4)



(Total for Question = 4 marks)

Q3.

The two tables show information about the race times of the 57 wheelchair competitors in the 2017 London Marathon.

The tables show information about the same data but the tables use different classes.

Table 1

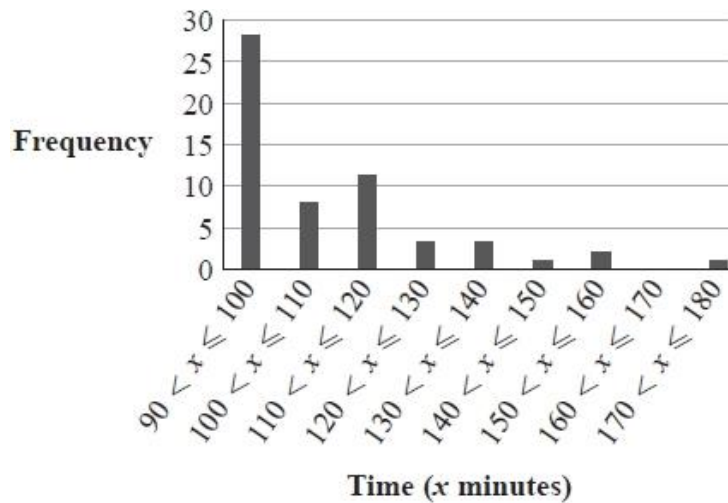
Time (x minutes)	Frequency
$90 < x \leq 100$	28
$100 < x \leq 110$	8
$110 < x \leq 120$	11
$120 < x \leq 130$	3
$130 < x \leq 140$	3
$140 < x \leq 150$	1
$150 < x \leq 160$	2
$160 < x \leq 170$	0
$170 < x \leq 180$	1

Table 2

Time (x minutes)	Frequency
$90 < x \leq 95$	22
$95 < x \leq 100$	6
$100 < x \leq 110$	8
$110 < x \leq 120$	11
$120 < x \leq 140$	6
$140 < x \leq 180$	4

(Source: www.virginmoneylondonmarathon.com)

Susan uses statistical software to produce this bar chart to represent the data in **Table 1**



Susan decides to draw a histogram for the data in **Table 2**

(a) Explain whether the bar chart or a histogram is the most appropriate way to represent the data.

Give reasons for your answer.

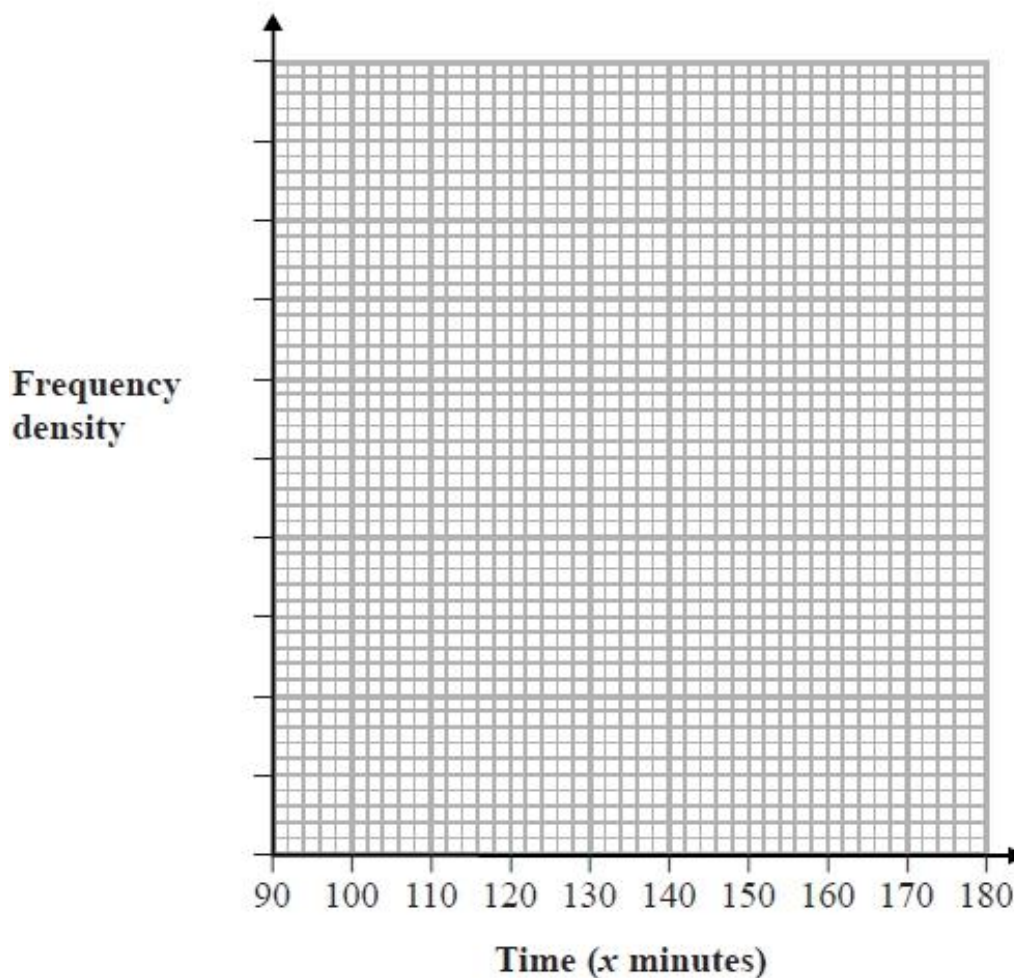
.....

.....

.....

(b) On the grid below, draw a histogram for the data in **Table 2**

Time (x minutes)	Frequency
$90 < x \leq 95$	22
$95 < x \leq 100$	6
$100 < x \leq 110$	8
$110 < x \leq 120$	11
$120 < x \leq 140$	6
$140 < x \leq 180$	4



(4)

Susan says that an estimate for the number of wheelchair competitors with a race time of between 105 minutes and 150 minutes is 22

(c) Show that Susan is correct.

(2)

(Total for question = 8 marks)

Q4.

The table gives information about the number of patients admitted to hospital in England in 2014 for different age groups.

Age (x years)	Number of patients admitted (10 000s)
$0 \leq x < 5$	190
$5 \leq x < 15$	200
$15 \leq x < 25$	280
$25 \leq x < 45$	500
$45 \leq x < 65$	360
$65 \leq x < 100$	385

Source: Adapted from Health & Social Care Information Centre

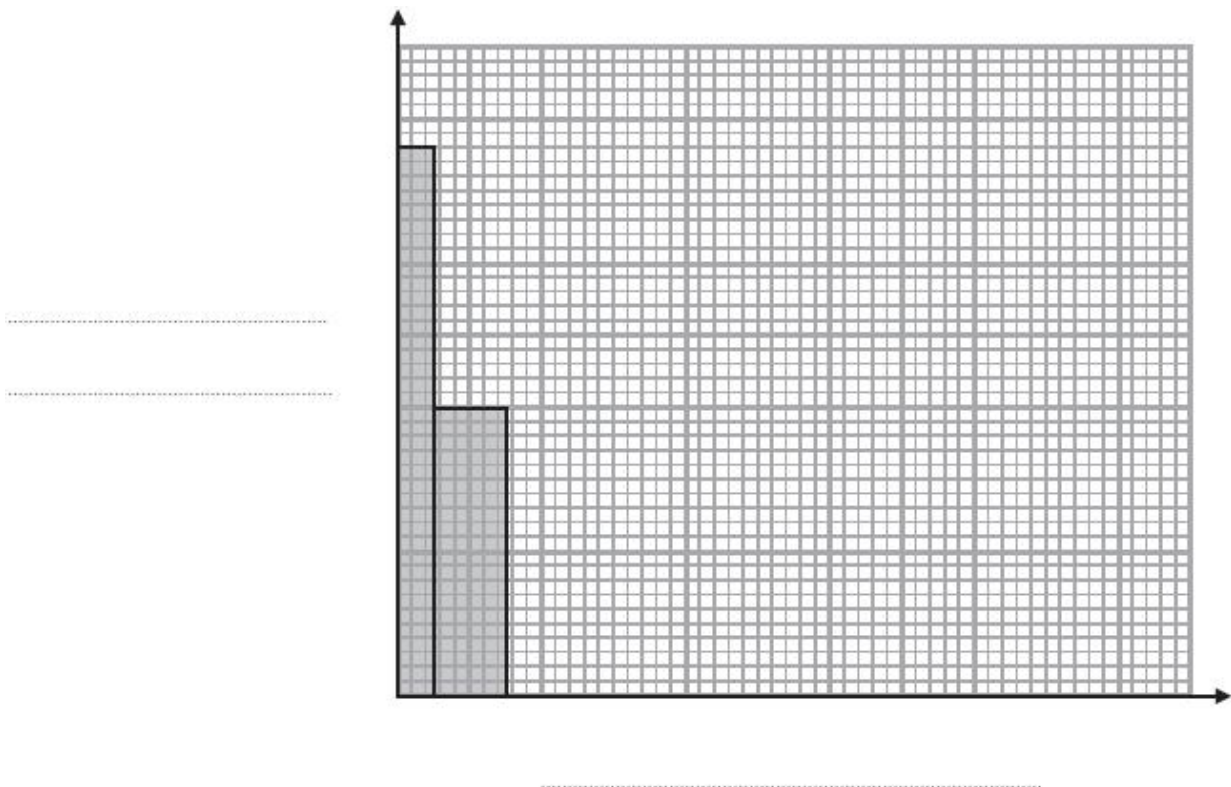
(a) Show that the class interval $0 \leq x < 5$ has a frequency density of 38

(1)

(b) On the grid below, complete the histogram for the information in the table.

The bars for the first two class intervals have already been drawn.

Label the axes.



(4)

(c) Find an estimate for the number of patients admitted to hospital who were aged 60 or over.

.....
(2)

(Total for question = 7 marks)

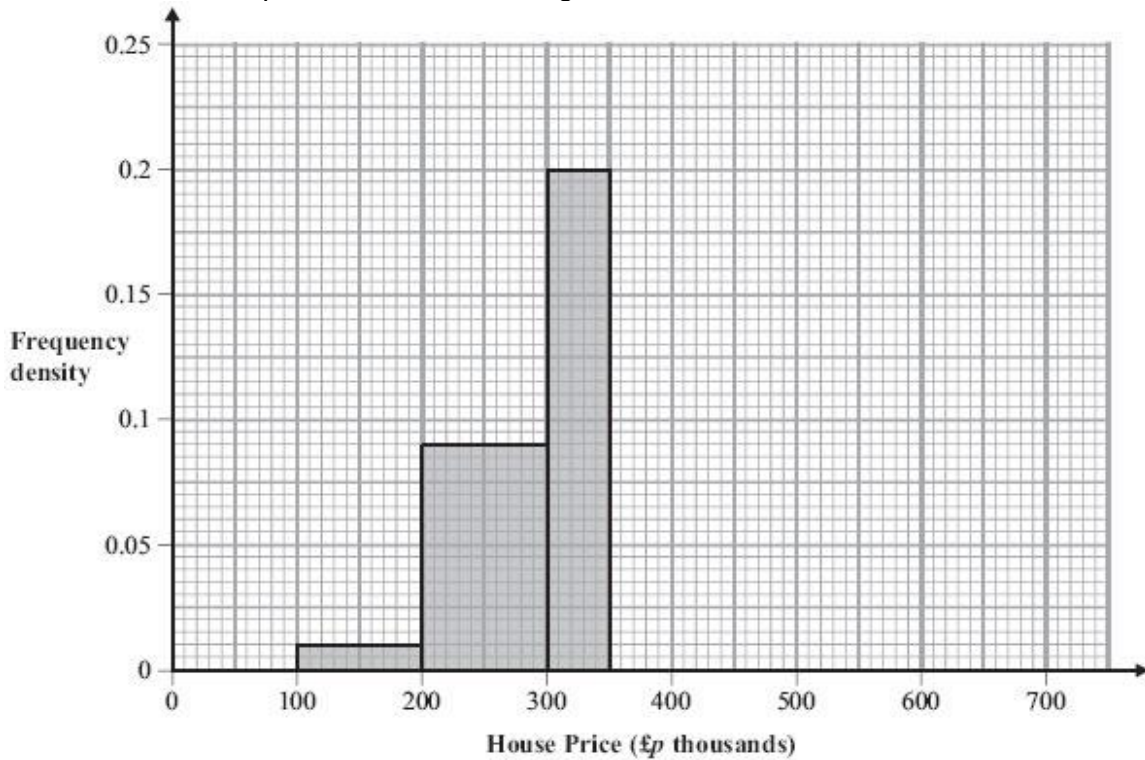
Q5.

The table gives some information about house prices in a small town.

House Price (£ p thousands)	Number of houses
$100 \leq p < 200$	1
$200 \leq p < 300$	9
$300 \leq p < 350$	10
$350 \leq p < 400$	8
$400 \leq p < 500$	2
$500 \leq p < 700$	3

(Data source: *rightmove.co.uk*)

Some of these data are represented on this histogram.



(a) Use the table to complete the histogram.

(3)

Jason thinks the **median** of these prices is £350000

Jason is wrong.

(b) Explain why.

.....

.....

.....

.....

.....

.....

(2)

The mean price of these houses is £340000

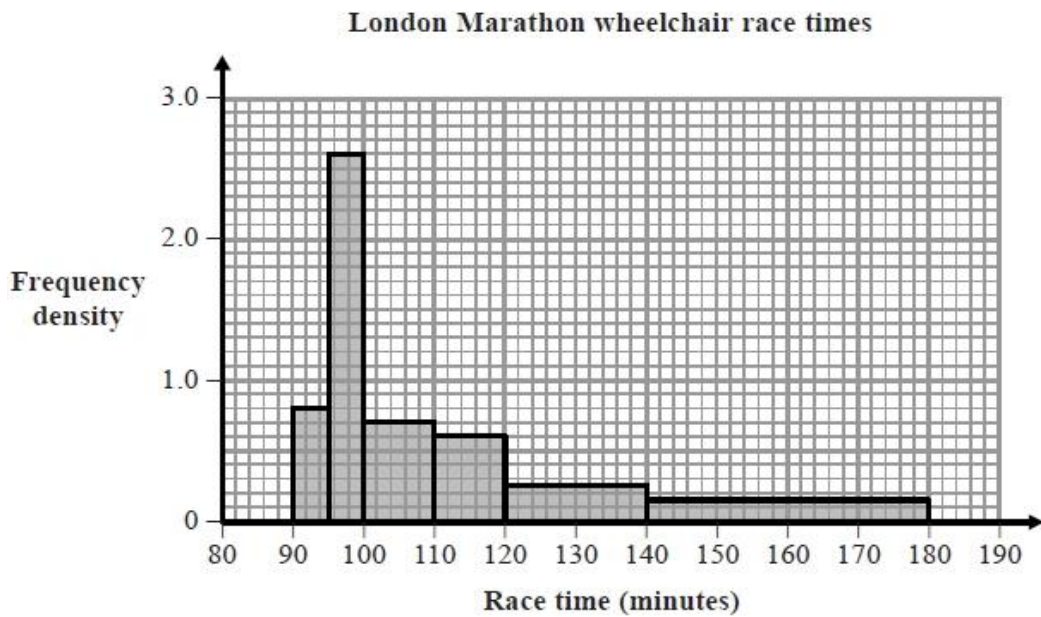
(c) Find an estimate for the number of these houses with a price less than the mean price.

.....
(2)

(Total for Question is 7 marks)

Q6.

The histogram shows information about the race times for the 41 wheelchair competitors in the 2014 London Marathon.



(Data source: virginmoneylondonmarathon.com)

(a) Describe the shape of the distribution.

.....
(1)

4 of the competitors had a race time in the class interval 90 minutes to 95 minutes.

(b) Show why the frequency density for this class interval is 0.8

(1)

(c) Calculate the number of competitors with a race time between 95 minutes and 110 minutes.

.....
(2)

(d) Use the histogram to work out an estimate for how long it took the first 21 competitors to finish the race.

..... minutes
(2)

David wants to predict the proportion of wheelchair competitors who will finish the New York Marathon in less than two hours.

(e) Explain whether or not it is sensible to use the London Marathon race times for the prediction.

.....
.....
.....
.....
.....

(2)

(Total for question = 8 marks)

Q7.

Susie is investigating the amount of time students at her school spend watching television. She selects a random sample of 50 students from her school and asks them the following question.

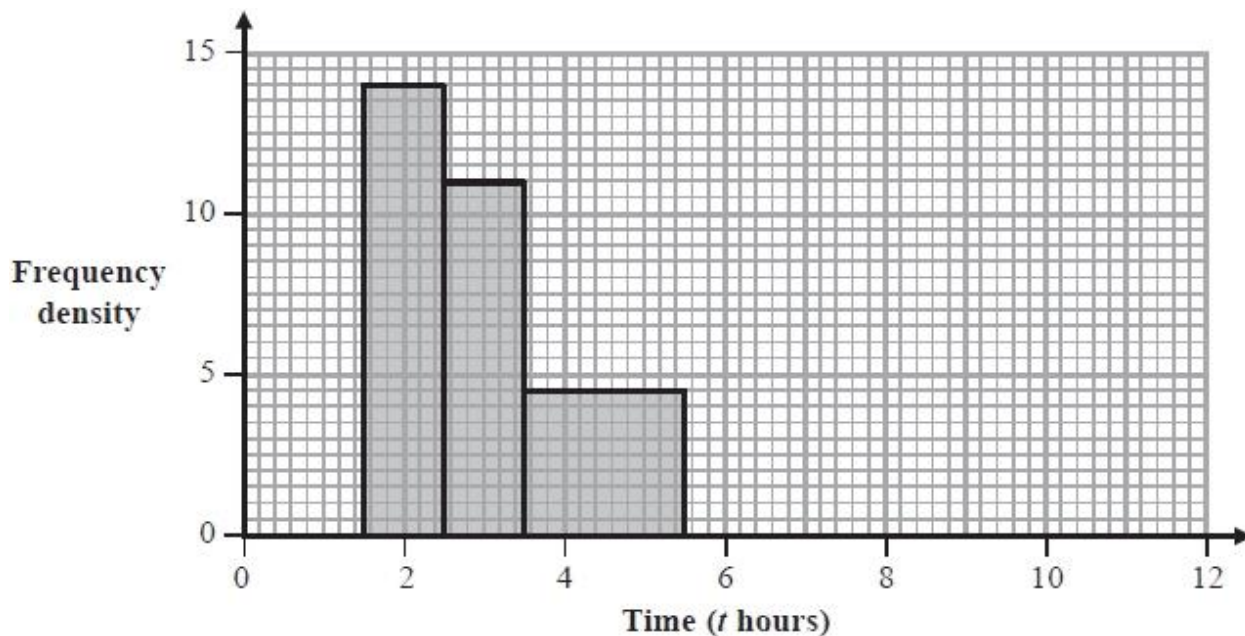
How much time, in hours to the nearest hour, did you spend watching television last Saturday?				
2	3	4 to 5	6 to 7	8 to 11
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Susie received answers from only 43 of the students.

The table shows a summary of her results.

Time (t hours)	Frequency
2	14
3	11
4 to 5	9
6 to 7	5
8 to 11	4

The incomplete histogram shows information about some of Susie's results.



(a) Give a reason to support using a histogram to represent the amounts of time students at her school spend watching television.

.....

.....

(b) Complete the histogram.

(3)

A student is selected at random from those students in the sample who watched less than 5.5 hours of television last Saturday.

(c) Estimate the probability that this student watched less than 2 hours of television last Saturday.

.....

(2)

(Total for question = 6 marks)

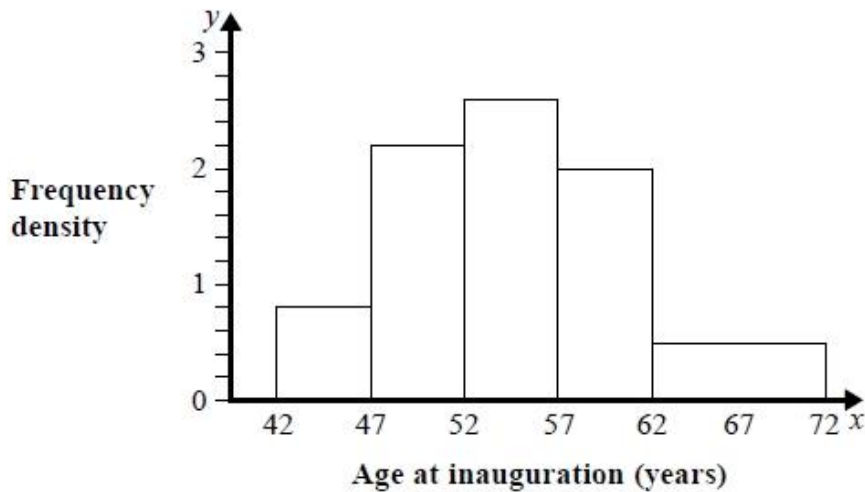
Q8. Peter thinks that the ages at inauguration of the presidents of the USA are normally distributed. He collects information about the ages at inauguration, in years, of 43 presidents of the USA from the internet.

The grouped frequency table gives information about his results.

Age at inauguration (x years)	Frequency (f)
$42 \leq x < 47$	4
$47 \leq x < 52$	11
$52 \leq x < 57$	13
$57 \leq x < 62$	10
$62 \leq x < 67$	5

(Source: *robinsonlibrary.com*)

Peter draws this histogram to show the information in the grouped frequency table.



(a) Calculate an estimate for the proportion of these 43 presidents whose age at inauguration was within 1 standard deviation of the mean.

Give your answer correct to 2 significant figures.

.....
(5)

Peter states that the age at inauguration of these presidents is normally distributed.

(b) Do you agree?

Use your answer to part (c) to justify your answer.

.....
.....

(1)
(Total for question = 6 marks)

Q9.

Owen collected data about the number of people per km² living in each of the 56 counties of England and Wales in 2015

Some information about Owen's data is shown in this incomplete table.

Number of people per km ² (x)	Number of counties
$0 \leq x < 200$	15
$200 \leq x < 400$	14
$400 \leq x < 600$	14
$600 \leq x < 1000$	r
$1000 \leq x < 6000$	s

(Source: *Office for National Statistics*)

(a) Work out the number of counties that have at least 600 people per km².

.....
(1)

(b) Use linear interpolation and the information in the table to find an estimate of the median number of people per km².

..... people per km²
(3)

(c) (i) Using the information in the table, explain how the mean of Owen's data would compare with the median of Owen's data.

.....
.....

(ii) Explain whether the mean or the median is the most appropriate measure of central tendency to use for Owen's data.

.....
.....
.....

(3)

Owen uses statistical software to draw a histogram to show the data he collected.

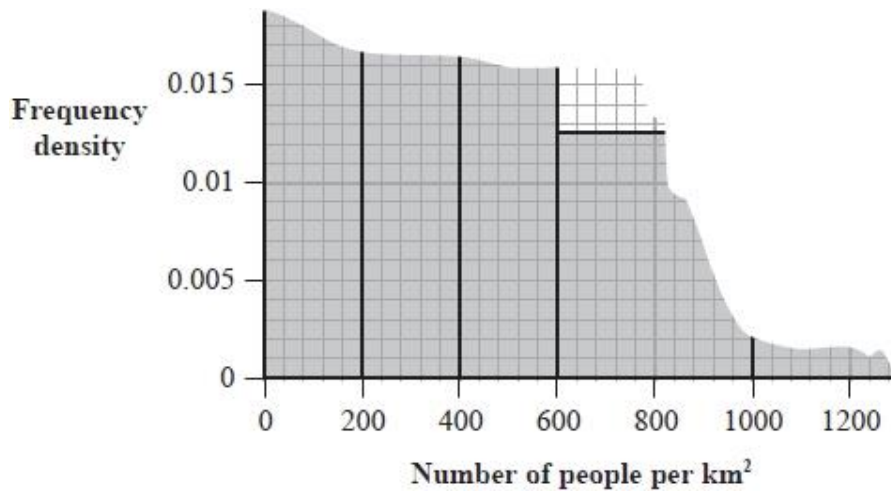
(d) Explain whether or not a histogram is appropriate.

.....
.....
.....

(1)

Owen prints his histogram but the page gets torn.

Here is what remains.



(e) Work out the value of r and the value of s in the table.

$r = \dots\dots\dots$ $s = \dots\dots\dots$
(2)

(Total for question = 10 marks)