Please check the examination details below before entering your candidate information

Candidate surname	Other names

Centre Number	Candidate Number

Pearson Edexcel Level 1/Level 2 GCSE (9–1)

Monday 12 June 2023

Afternoon (Time: 1 hour 30 minutes)

Paper reference **1ST0/1F**

Statistics
PAPER 1
Foundation Tier

You must have:
Ruler graduated in centimetres and millimetres, protractor, pair of compasses, pen, HB pencil, eraser, scientific calculator.

Total Marks

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 total mark for this paper is 80.
- 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.

Turn over ▶

R72893A

N:1/1/1/

Pearson

Answer ALL questions.

Write your answers in the spaces provided.

You must write down all the stages in your working.

1 The table shows the average heights, to the nearest cm, of Adult Males and Adult Females who were born in the year 1996 in some countries of the world.

Country	Average height (cm)	
	Adult Males	Adult Females
Philippines	163	150
Latvia	181	170
Italy	178	165
Zimbabwe	169	158
Australia	180	166

(Source: *Ourworldindata.org*)

(a) Write down the average height of Adult Males in Italy.

..... cm

(1)

(b) Write down the country in the table in which the Adult Females have the greatest average height.

.....

(1)

Afzal thinks that the country in the table with the greatest difference in average heights between Adult Males and Adult Females is Italy.

(c) Is Afzal correct?
Give a reason for your answer.

(2)

(d) Using the table, compare the average height of Adult Males in Australia, with the average height of Adult Males in Zimbabwe.

(1)

Afzal suggests drawing a time series graph to represent the data in the table.

(e) Explain whether or not this is an appropriate graph to use.

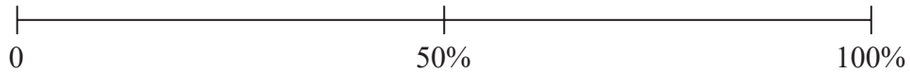
(1)

(Total for Question 1 is 6 marks)

2 Jonathan has a fair 6-faced dice which has the numbers 1, 2, 3, 4, 5 and 6 on its faces.

Jonathan rolls his dice once.

(a) On the probability scale below, mark with a cross (×) the probability that the dice will land on an odd number.



(1)

Kasia has a fair 8-faced dice which has the numbers 1, 2, 3, 4, 5, 6, 7 and 8 on its faces.

(b) Circle the word from the list below that best describes the likelihood that the dice lands on a 9

impossible unlikely evens likely certain

(1)

Kasia rolls her dice 80 times.

(c) Work out the number of times you would expect her dice to land on a 5

.....

(2)

Jonathan is going to roll his dice once.
Kasia is going to roll her dice once.

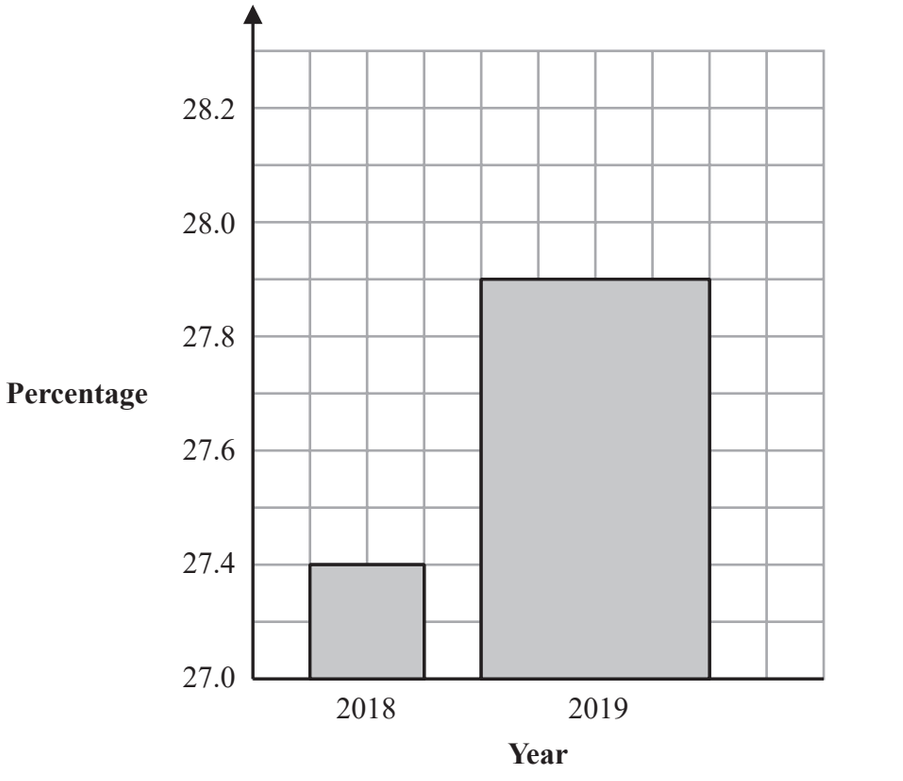
(d) Is Jonathan more likely to roll a 6 than Kasia?
You should justify your answer.

(2)

(Total for Question 2 is 6 marks)

3 A newspaper reporter completed some research into the percentage of household waste that was recycled in the London Borough of Hackney in 2018 and 2019

The newspaper reporter drew the following bar chart to show this information.



(Source: *consultation.hackney.gov.uk*)

Give **two** reasons why the bar chart could be misleading or wrong.

(Total for Question 3 is 2 marks)

BLANK PAGE

Turn over for next question

4 A town council is proposing to build a new leisure centre.
Michelle is going to carry out a survey to find out what all the people in the town think of the proposal.

Michelle thinks that she should take a sample rather than a census.

(a) Give **two** reasons why Michelle might think this.

(2)

Michelle plans to use the electoral register as the sampling frame.

(b) (i) Explain what you understand by the term sampling frame.

(1)

(ii) Give one problem Michelle may have using the electoral register as the sampling frame.

(1)

Michelle intends to conduct a pilot study.

(c) Give **two** reasons why it is a good idea to conduct a pilot study.

(2)

Michelle is writing a plan for her investigation into people's views on the leisure centre proposal.

(d) Write down what Michelle should include in her plan.
You should include each of the following

- a sampling method
- a question she could ask in her questionnaire
- a statistical diagram she could use to show the results of the survey.

Explain why each of the things you have written down is appropriate.

(6)

(Total for Question 4 is 12 marks)

5 Jim is investigating the relationship between air temperature and altitude.

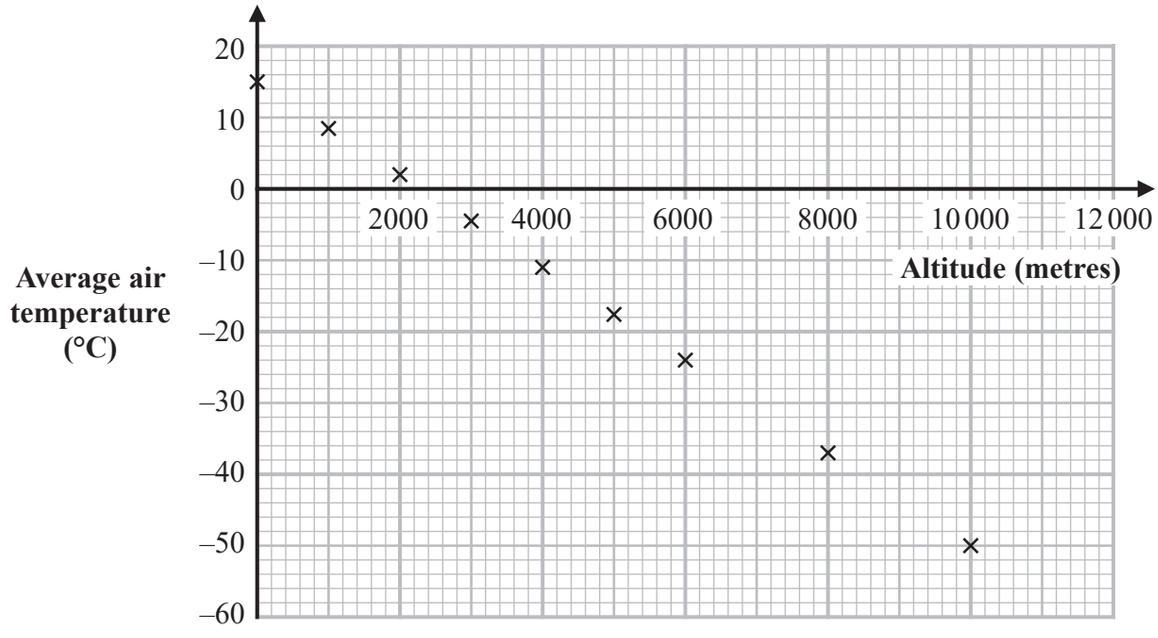
He has found data on the internet showing the average air temperature at different altitudes on one particular day.

The table shows the results he collected.

Altitude (metres)	0	1000	2000	3000	4000	5000	6000	8000	10 000
Average air temperature (°C)	15.0	8.5	2.0	–4.5	–11.0	–17.5	–24.0	–36.9	–49.9

(Source: adapted from *www.engineeringtoolbox.com*)

He uses this information to draw the scatter diagram.



(a) Describe and interpret the correlation shown in the scatter diagram.

(2)

(b) Draw a line of best fit on the scatter diagram.

(1)

(c) Use the scatter diagram and your line of best fit to predict the average air temperature at an altitude of 4400 metres.

.....°C

(1)

Jim wants to predict the average air temperature at an altitude of 11 000 metres.

(d) Discuss whether or not it is appropriate to use the line of best fit to make his prediction.

(2)

(Total for Question 5 is 6 marks)

6 David asked 15 of his friends about the number of pets they each have. Here is the data he collected.

0 0 0 0 0 1 1 2 2 2 4 4 4 4 8

(a) Circle the word in the list below that describes this type of data.

continuous qualitative discrete grouped

(1)

(b) Write down the modal number of pets.

.....

(1)

(c) Find the median number of pets.

.....

(1)

(d) State which average, the mode or the median, best represents these data. Give a reason for your answer.

(1)

(e) Find the interquartile range of the number of pets.

.....

(2)

Wanda asked some of her friends about the number of pets they each have.
The table below is a summary of the data she collected.

Lower quartile	Median	Upper quartile
1	3	6

(f) Compare the distribution of the numbers of pets for David with the distribution of the numbers of pets for Wanda.
Give **two** comparisons and interpret each of your comparisons.

(4)

Wanda recorded the highest number of pets as 15
She says that this must be an outlier and concludes that it should be removed from her data.

(g) (i) Give one reason why Wanda's conclusion may be appropriate.

(1)

(ii) Give one reason why Wanda's conclusion may **not** be appropriate.

(1)

(Total for Question 6 is 12 marks)

7 The table gives information about the ages of people on the electoral register in the West Midlands in December 2018

Age	17 years old	18 years old and older
Number of people	28 152	4 146 375

(Source: *Office for National Statistics*)

A researcher wanted to find out information about voting intentions in the West Midlands.

He sent a questionnaire to a sample of 10 000 people on the electoral register in the West Midlands stratified by age of voter.

Describe how the researcher would have carried out this stratified sampling.

You should show any calculations that you use.

Discuss the appropriateness of this stratified sample.

(Total for Question 7 is 5 marks)

BLANK PAGE

Turn over for next question

8 The following is an extract from part of a row of a random number list.

68236 35335 71329

(a) Use the random number list to complete the table for the first 5 random 2-digit numbers.

68	.....	.....	.....	.....

(2)

The most common blood type in the United Kingdom is O+

The percentage of people in the United Kingdom with O+ blood type is 38%

Asha uses a simulation method to estimate how many donors would be needed to find exactly 3 donors with O+ blood type.

Asha is going to use the following 2-digit numbers for her simulation.

Blood type	O+	Not O+
Random numbers	00 – 37	38 – 99

(b) Explain why this is an appropriate way to allocate the random numbers.

(1)

Asha runs trials using her simulation method.
The result of each trial is the number of random numbers used until Asha gets exactly 3 donors with O+ blood type.
The table below shows the results of her first 4 trials.

Trial	1	2	3	4
Result	7	5	8	4

The set of random numbers used by Asha to complete the fifth trial are shown below.

60 13 12 86 73 10 98 95 43 46

(c) Using this set of random numbers, find the result for the fifth trial.
You must make it clear how you obtain your answer.

(2)

Asha finds the mean of her 5 results and decides that the results of her simulation are sufficient to predict the number of donors needed to find at least 3 with O+ blood type in the next blood donation session.

(d) Explain whether the method that Asha uses to predict the number of donors required is appropriate.

(2)

(Total for Question 8 is 7 marks)

9 Kyle is investigating the heights and the weights of professional basketball players.

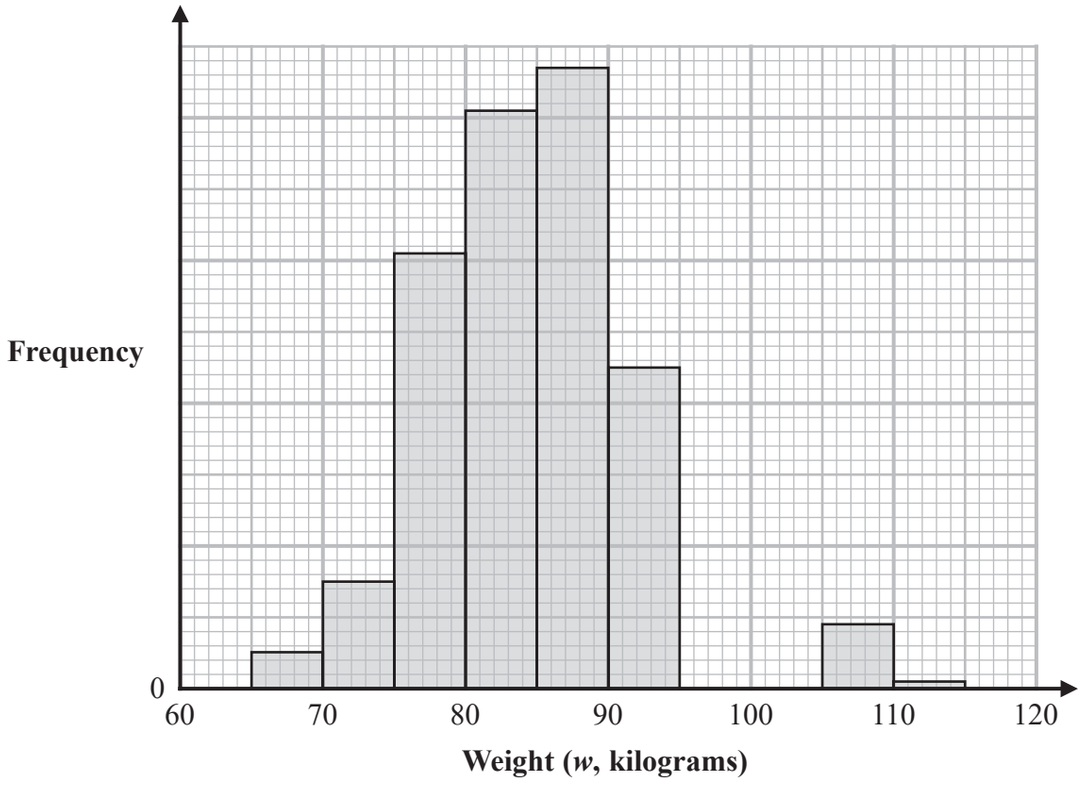
He found the weight, in kilograms, of some professional basketball players from 1950 to 1959

(a) Circle the word in the list below that describes weight, in kilograms, as a type of data.

discrete continuous ordinal categorical

(1)

The incomplete histogram and incomplete grouped frequency table give information about the weights, in kilograms, of the professional basketball players from 1950 to 1959



(Source: *www.kaggle.com*)

Weight (w kilograms)	Frequency
$65 < w \leqslant 70$	5
$70 < w \leqslant 75$	15
$75 < w \leqslant 80$	61
$80 < w \leqslant 85$	81
$85 < w \leqslant 90$	
$90 < w \leqslant 95$	
$95 < w \leqslant 100$	35
$100 < w \leqslant 105$	14
$105 < w \leqslant 110$	9
$110 < w \leqslant 115$	1

(b) Use the information in the histogram to complete the table.

(2)

(c) Use the information in the table to complete the histogram.

(2)

Kyle also drew a histogram for the weights of professional basketball players from 2000 to 2009
This histogram was negatively skewed.

(d) Interpret the negative skew of the weights of professional basketball players from 2000 to 2009

(1)

Kyle also collected data about the heights of professional basketball players from 1950 to 1959 and the heights of professional basketball players from 2000 to 2009

The grouped frequency table below gives information about the heights of professional basketball players from 2000 to 2009

Height (h centimetres)	**Frequency**		
$170 < h \leqslant 180$	12		
$180 < h \leqslant 190$	146		
$190 < h \leqslant 200$	175		
$200 < h \leqslant 210$	323		
$210 < h \leqslant 220$	146		
$220 < h \leqslant 230$	8		
Total	**810**		

(Source: *www.kaggle.com*)

The estimate of the mean height for professional basketball players from 1950 to 1959 is calculated to be 190.9 cm to one decimal place.

(e) (i) Calculate an estimate of the mean height of basketball players from 2000 to 2009

..... cm

(3)

(ii) Comment on how the mean height of professional basketball players has changed between the two sets of data.

(1)

(Total for Question 9 is 10 marks)

BLANK PAGE

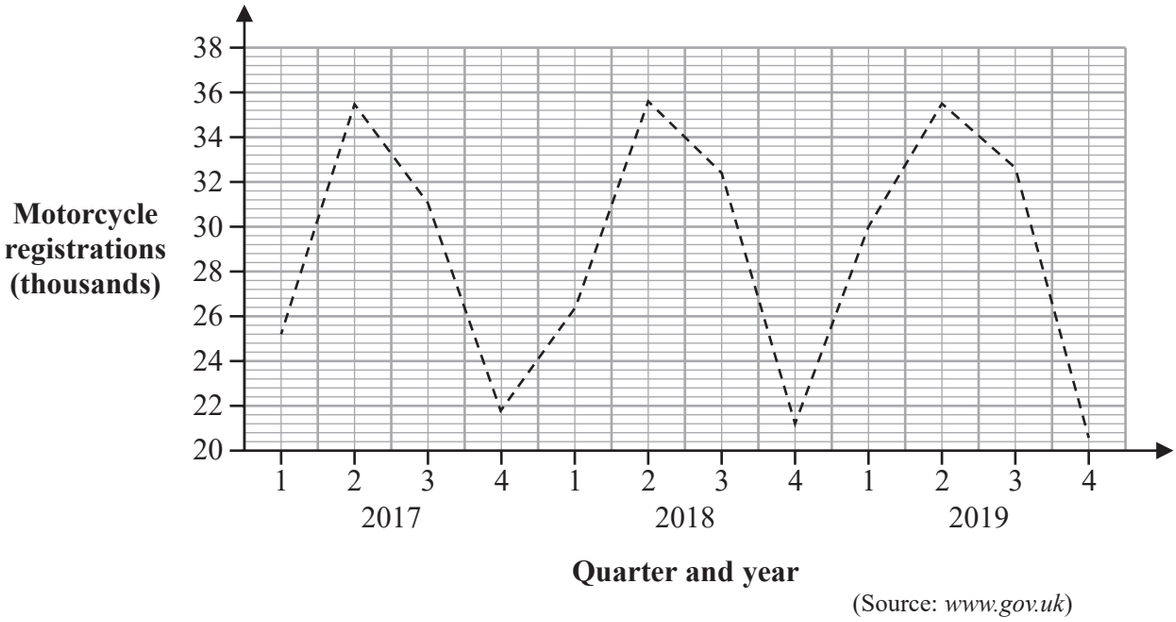
Turn over for next question

10 Claire is investigating sales of different types of vehicle over time.
She plans to collect data on the numbers of motorcycles first registered in the UK over time.

(a) Write down a suitable hypothesis for this investigation.

(1)

The time series graph shows some information about the numbers of motorcycles first registered in the UK from 2017 to 2019



(Source: *www.gov.uk*)

(b) Identify and interpret one example of seasonal trend shown by the time series graph.

(2)

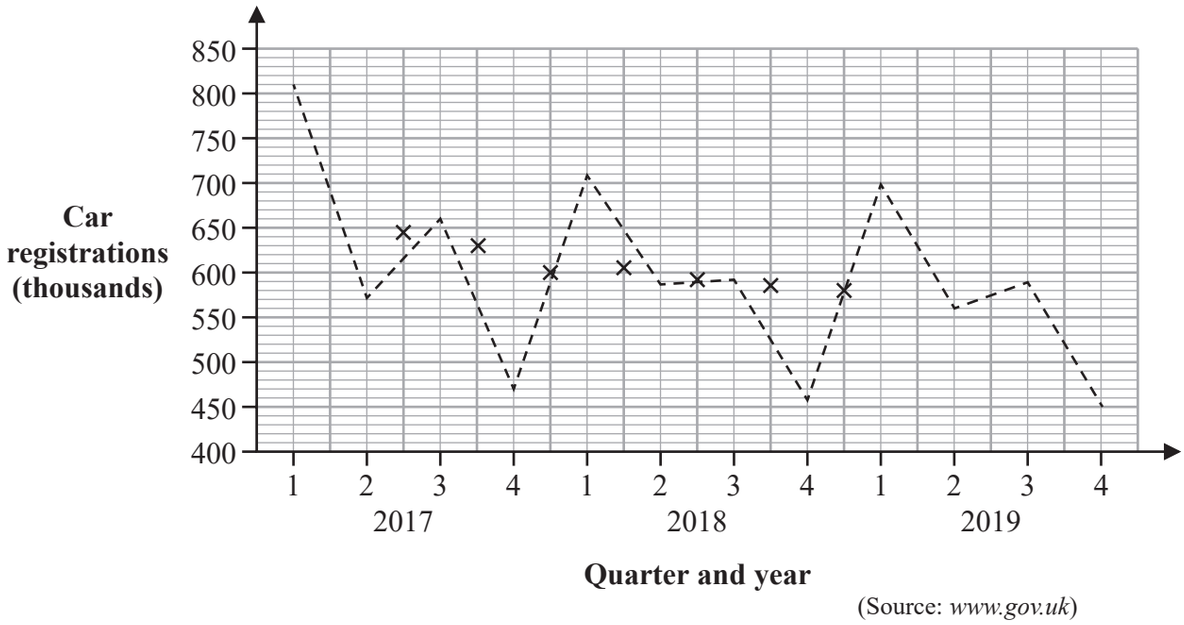
Claire calculated 4-point moving averages for the information shown in the time series graph.

(c) Explain why this is appropriate.

(1)

Claire also collected data on the numbers of cars first registered in the UK from 2017 to 2019

The time series graph shows some information about the numbers of cars first registered in the UK from 2017 to 2019 together with the first seven 4-point moving averages.



(Source: *www.gov.uk*)

(d) Compare the seasonal trend shown for the numbers of motorcycles first registered in the UK with the seasonal trend for the numbers of cars first registered in the UK.

(1)

The last three 4-point moving averages (thousands) for the number of cars registered in the UK from 2017 to 2019 are

576.0 575.3 573.9

(e) Plot these three moving averages on the time series graph and draw a trend line.

(3)

(f) Describe and interpret the trend in the numbers of cars first registered in the UK from 2017 to 2019

(2)

(Total for Question 10 is 10 marks)

11 A fair 3-sided spinner numbered 1, 2 and 3 and a fair 4-faced dice numbered 1, 2, 3 and 4 are used in a game.

To play the game, a player spins the spinner once and rolls the dice once. The total score is found by adding the number the spinner lands on and the number the dice lands on.

(a) Complete the sample space diagram to show all the possible total scores.

4-faced dice

		1	**2**	**3**	**4**
3-sided spinner	**1**	2	3	.....	.....
	2	3	.....	.....	.....
	3	.....	.....	.....	.....

(2)

To win the game a player needs to get a total score of at least 6
Chloe plays the game once.

(b) Find the probability that Chloe does **not** win the game.

.....

(2)

(Total for Question 11 is 4 marks)

TOTAL FOR PAPER IS 80 MARKS

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA