UNIVERSITY OF CAMBRIDGE INTERNATIONAL EXAMINATIONS
General Certificate of Education Ordinary Level

CANDIDATE NAME

CENTRE NUMBER

CANDIDATE NUMBER

0192736882

STATISTICS **4040/12**

Paper 1 **October/November 2013**

2 hours 15 minutes

Candidates answer on the question paper.

Additional Materials: Pair of compasses
Protractor

READ THESE INSTRUCTIONS FIRST

Write your Centre number, candidate number and name on all the work you hand in.
Write in dark blue or black pen.
You may use a soft pencil for any diagrams or graphs.
Do not use staples, paper clips, highlighters, glue or correction fluid.
DO **NOT** WRITE IN ANY BARCODES.

Answer **all** questions in Section A and not more than **four** questions from Section B.
If working is needed for any question it must be shown below that question.
The use of an electronic calculator is expected in this paper.

At the end of the examination, fasten all your work securely together.
The number of marks is given in brackets [ ] at the end of each question or part question.

This document consists of **19** printed pages and **1** blank page.

Section A [36 marks]

Answer **all** of the questions 1 to 6.

For Examiner's Use

1 Seven statistical measures are

mean,
median,
mode,
range,
interquartile range,
variance
and standard deviation.

In each of the following situations, one of these measures is to be found by the person described. State the appropriate measure in each case.

(i) A doctor finds the most common age of her patients.

.....[1]

(ii) An athlete who competes in the 100 metres sprint finds the difference between his slowest and quickest practice times.

.....[1]

(iii) A graduate who seeks employment with a company finds a measure of central tendency for the salaries of the company's employees. The company has twenty employees, of whom three are managers earning salaries very much higher than the other employees.

.....[1]

(iv) A teacher finds a measure of dispersion for the scores of her pupils in a test, in which no pupil scored an exceptionally high mark, and no pupil scored an exceptionally low mark.

.....[1]

(v) A biologist finds a measure of dispersion for the growth of twelve plants over a period of three months. Two plants have been attacked by insects and have grown very much less than the others.

.....[1]

(vi) A sociologist finds a measure of central tendency for the first names given to the male babies born in a hospital over a period of six months.

.....[1]

2 A large keep fit class for women is held at a sports club once every week. The manager of the club asks the class instructor to select a sample of size 10 from the class.

(i) State the method of sampling used if the instructor decides to select

(a) the first 10 women to arrive at the class,

.....[1]

(b) women at regular intervals from the class register.

.....[1]

The sample is required to obtain responses to a proposal to change the time of the class from Monday evening to Monday afternoon. For class members the only items of data presently available to the instructor are name and age.

(ii) State, and justify, two other items of data relating to class members which the instructor needs to know when selecting the sample in order to avoid bias in responses. You are **not** required to describe how the sample is selected.

.....

.....

.....

.....

.....

.....

.....

.....[4]

3 In a photographic equipment store a record was kept of the number of cameras sold each day. The values, for eleven consecutive days, were as follows.

For Examiner's Use

6 0 8 2 1 6 0 9 6 4 1

For these values find

(i) the mode,

.....[1]

(ii) the mean, correct to one decimal place,

.....[2]

(iii) the median.

.....[2]

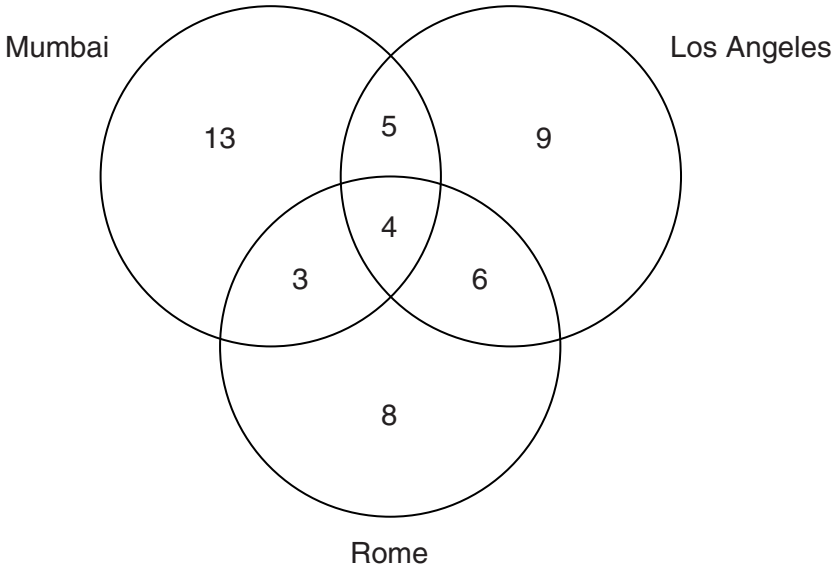
The values recorded for the next three days were x, $x + 1$ and $x + 2$.

(iv) If the median for the entire fourteen-day period was the same as the median for the first eleven days, find x.

$x =$ .....[1]

4 The diagram below shows the number of actors at a film festival who have worked in one or more of the cities Mumbai, Los Angeles and Rome.

For Examiner's Use



(i) Find the number of actors who have worked in Mumbai.

.....[1]

(ii) Interpret the value 6 in the diagram.

.....

.....[1]

A journalist selects one of these actors at random for interview.

Find the probability of selecting an actor who has worked in

(iii) Mumbai or Los Angeles or both,

.....[2]

(iv) Los Angeles and Rome,

.....[1]

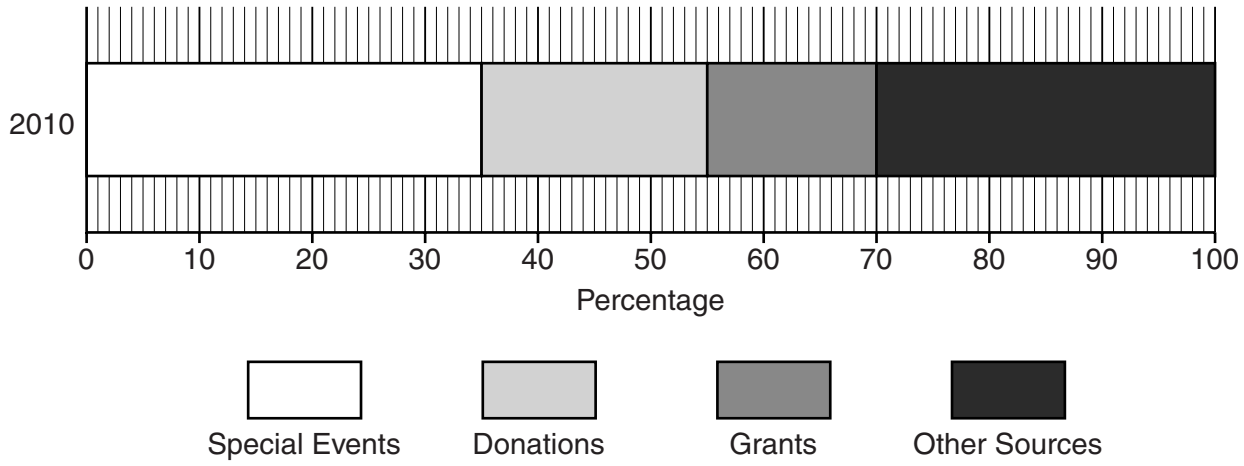
(v) Rome, given that the actor has worked in Mumbai and Los Angeles.

.....[1]

5 In this question you are not required to draw any charts.

For Examiner's Use

A charity, Camfam, classifies the income it receives under the headings Special Events, Donations, Grants, and Other Sources. In Camfam's report for 2010, the following percentage bar chart was given, which represents a total income of $80 million.



(i) Find the income which Camfam received in 2010 from Grants.

$ .....[2]

(ii) If a pie chart were to be drawn to represent this information, find the angle which would represent the sector for Special Events.

.....[2]

Camfam's total income in 2011 was $60 million.
Two pie charts, one for 2010 and one for 2011, are to be presented together in a new report.

(iii) Find, in its simplest terms, the ratio of the area of the chart representing 2010 to the area of the chart representing 2011.

.....[2]

6 The following table is to show the distance, in kilometres, between any two of the five towns *A*, *B*, *C*, *D* and *E*.

For Examiner's Use

	A	*B*	*C*	*D*	*E*
A		42			
B					
C					
D	20	39			
E	36	18		25	

For example, the distance between *B* and *D* is 39 km.

(i) **Complete the table** using the following information.

(a) The distance between *B* and *C* is 10 km more than the distance between *D* and *E*.

[1]

(b) The distance between *A* and *C* is two thirds of the distance between *A* and *E*.

[1]

(c) The distance between *A* and *B* is twice the distance between *C* and *E*.

[1]

(d) *C* is 19 km further from *D* than *B* is from *E*.

[1]

Dimitri lives in town *A*, but has one friend in each of the towns *D* and *E*. He makes a journey in which he leaves his home, visits each of these friends once, and then returns home.

(ii) Find the distance which Dimitri travels to complete the journey.

..... km [2]

Section B [64 marks]

For Examiner's Use

Answer not more than **four** of the questions 7 to 11.

Each question in this section carries 16 marks.

7 In this question the fertility rate of a population is defined as the number of births per 1000 females.

The table below gives information about the female population and age group fertility rates in a particular city for the year 2012, together with the standard population of the area in which the city is situated.

Age group of females	*Births*	*Population of females in age group*	*Age group fertility rate*	*Standard population of females* (%)
Under 20		2900	50	18
20 – 29		4500	184	22
30 – 39		5250	136	25
Over 39		5800	15	35

(i) Calculate, to 1 decimal place, the standardised fertility rate for the city.

.....[4]

(ii) Calculate the number of births for each age group and insert the values in the table above.

[2]

(iii) Calculate, to 1 decimal place, the crude fertility rate for the city.

For Examiner's Use

.....[4]

There are equal numbers of males and females in the city and in the standard population. The standardised and crude death rates for the city in 2012 were 8.5 and 7.8 per thousand of the population respectively.

(iv) Using **one** of these values, and any other appropriate values from parts **(i)**, **(ii)** and **(iii)**, find the increase in the population of the city in 2012 due to births and deaths.

.....[5]

It is not possible to obtain an accurate measure of population increase or decrease in a city from information on births and deaths alone.

(v) State what additional information is required.

.....

.....[1]

8 In a large residential building there are 120 apartments, of which 50 are private apartments (owned by the residents) and 70 are company apartments (owned by the company which constructed the building).

For Examiner's Use

If two apartments are chosen at random, find the probability of choosing

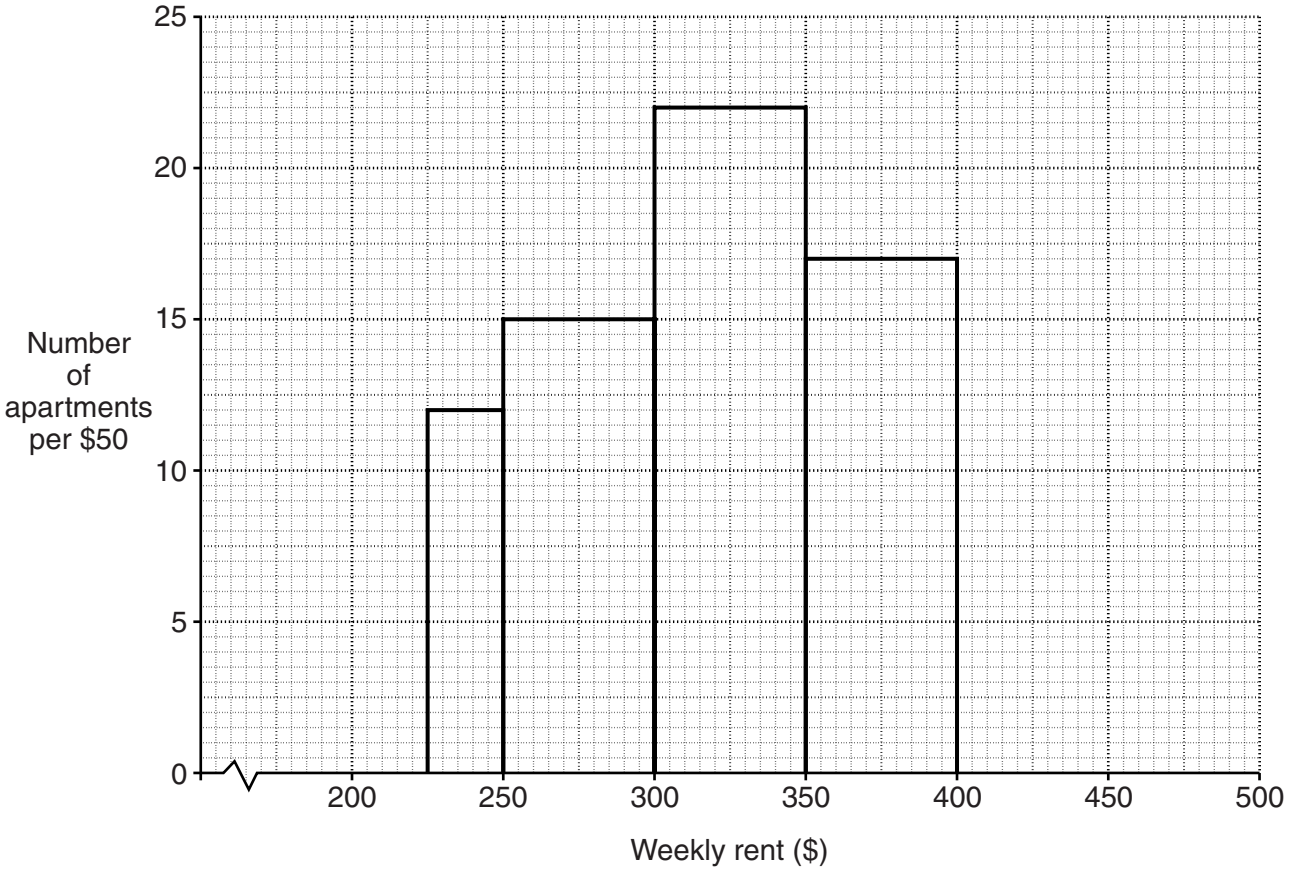
(i) two private apartments,

.....[2]

(ii) at least one company apartment.

.....[2]

The weekly rents, in dollars, charged on the company apartments are represented in the histogram below, from which one rectangle, representing the \$400 to under \$500 class, has been omitted.



Use the histogram to find the number of company apartments for which the weekly rent was

For Examiner's Use

(iii) from \$250 to under \$400,

.....[2]

(iv) from \$225 to under \$250.

.....[2]

There were 10 company apartments for which the weekly rent was from \$400 to under \$500.

(v) Complete the histogram by drawing on the grid the rectangle representing the \$400 to under \$500 class.

[1]

(vi) Write down the term used to describe the \$300 to under \$350 class.

.....[1]

The private apartments are of three different sizes. There are 24 apartments with three rooms, 14 with four rooms, and 12 with five rooms.
A safety expert, conducting a survey on the use of smoke detectors, chooses three private apartments at random.

(vii) If the apartments chosen have 12 rooms in total, find the probability that the apartments are all of the same size.

.....[6]

9 The mid-day temperature at a particular location in a city was measured every day throughout the year 2010. The following table summarises the results obtained.

Temperature (°C)	*Number of days*	*Cumulative frequency*
0 – under 5	8	
5 – under 10	25	
10 – under 15	52	
15 – under 20	81	
20 – under 25	79	
25 – under 30	68	
30 – under 35	37	
35 – under 40	15	

(i) Complete the cumulative frequency column in the above table. [2]

(ii) Plot the cumulative frequencies on the grid opposite, joining the points by a smooth curve. [3]

(iii) Use your graph to estimate

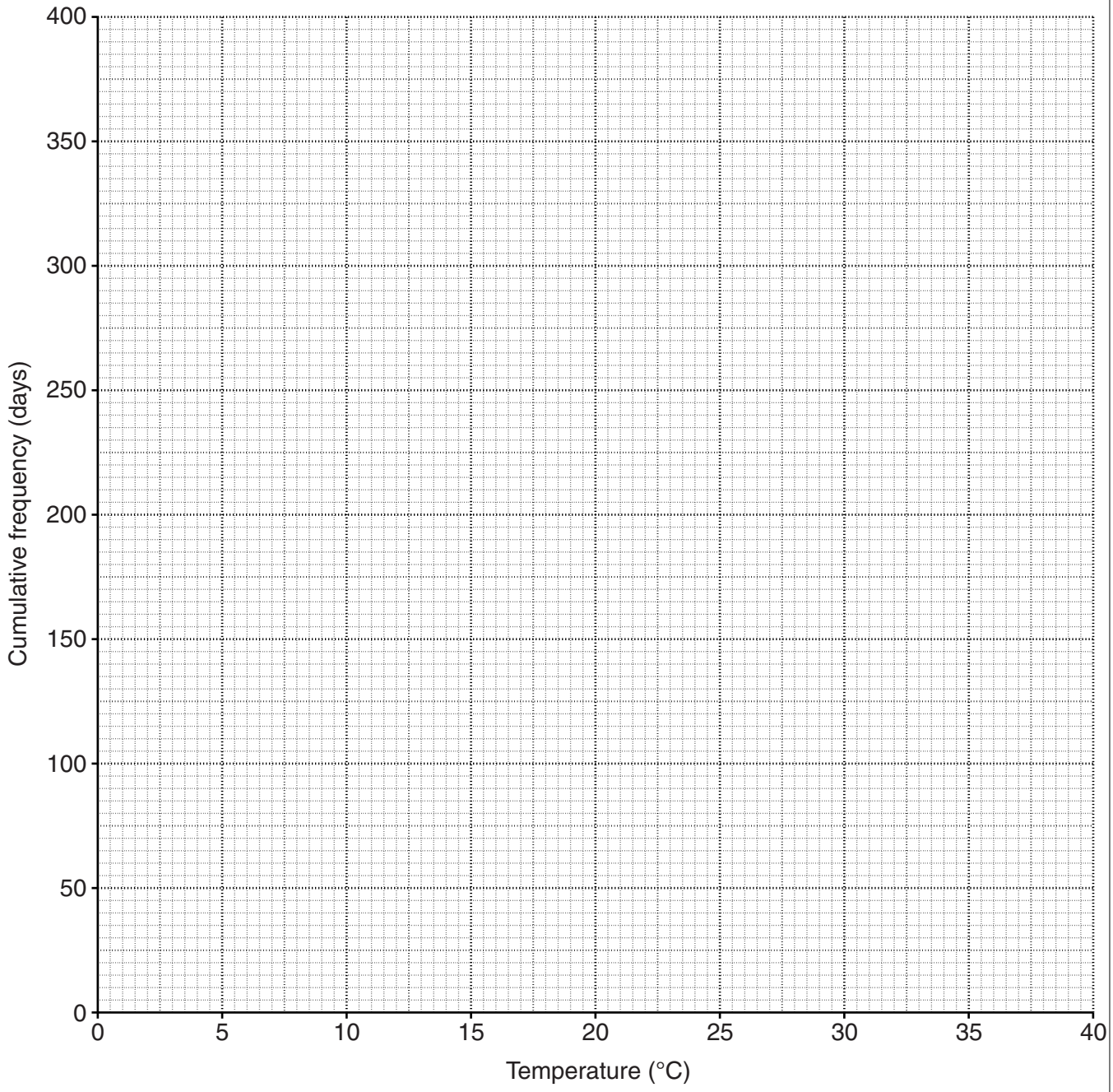
(a) the median of these temperatures,

..... °C [1]

(b) the interquartile range of these temperatures.

..... °C [4]

For Examiner's Use



When the results were obtained, a scientist predicted that, because of climate change, temperatures in the city would increase at the rate of 0.5 °C every ten years.
Assume that this prediction is accurate.

For Examiner's Use

For this particular location,

(iv) use your answers to part **(iii)** to estimate, for the year 2050,

(a) the median of the mid-day temperatures,

.....°C [2]

(b) the interquartile range of the mid-day temperatures,

.....°C [1]

(v) use your graph to estimate, for the period 2010 to 2050, the increase in the number of days with a mid-day temperature of more than 36 °C.

.....[3]

BLANK PAGE

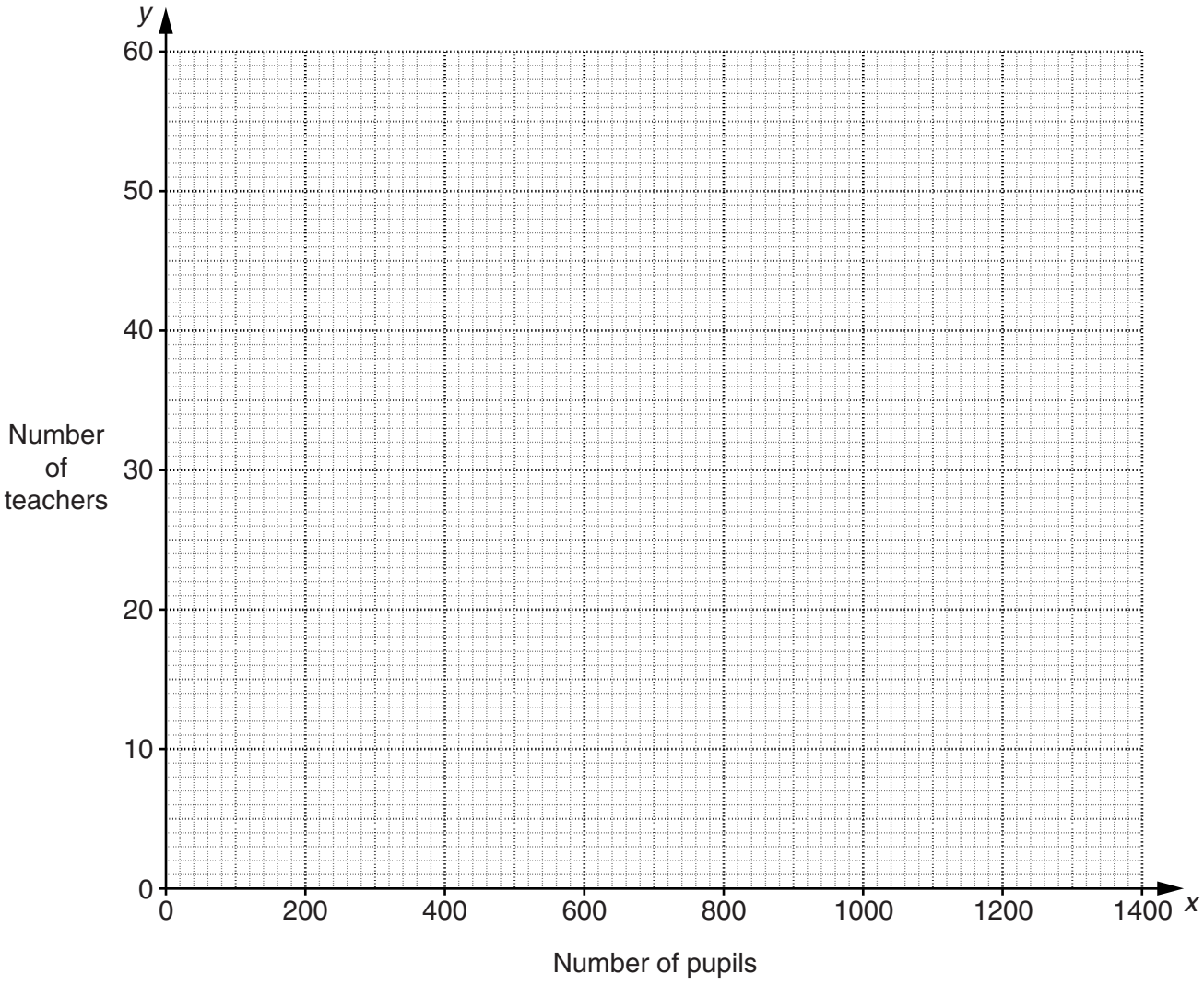
[Turn over for Question 10]

10 Emilie, a student teacher, conducted research on the number of pupils and the number of teachers in the schools in the town of Astra, where she lives. The schools supplied the following data.

For Examiner's Use

School	A	B	C	D	E	F	G	H
Number of pupils, x	760	1219	927	470	1361	628	381	1085
Number of teachers, y	29	44	33	34	52	24	16	40

(i) Plot these data on the grid below.



[2]

The data have an overall mean of (853.875, 34) and an upper semi-average of (1148, 42.25).

For Examiner's Use

(ii) Show how the value 1148 is calculated.

[2]

(iii) Find the lower semi-average.

.....[2]

(iv) **Without plotting the averages, and without drawing the line,** find the equation of the line of best fit in the form $y = mx + c$.

.....[3]

(v) Explain briefly why the value of c which you have found in part **(iv)** might give you cause for concern.

.....

.....[1]

Emilie discovered later that the data supplied by one of the schools gave, incorrectly, the total number of people employed by the school, and not the number of teachers.

(vi) Ignoring the point representing the school which supplied incorrect data, draw, by eye, on the grid in part **(i)**, a line of best fit through the remaining seven points. [1]

(vii) **Use the line you have drawn in part (vi)** to find its equation in the form $y = mx + c$.

.....[3]

Emilie repeated the research for schools in the nearby town of Belport, for which she found the equation of the line of best fit to be $y = 0.0431x + 1.72$.

For Examiner's Use

(viii) Using this equation, and your answer to part **(vii)**, state in which of the two towns a pupil might choose to be educated, if free to choose. Explain your answer briefly.

.....

.....

.....[2]

11 **(i)** Give one advantage and one disadvantage of forming a large set of data into a grouped frequency distribution.

For Examiner's Use

Advantage.....

.....

Disadvantage.....

.....[2]

The presenter of a radio programme, in which recordings of popular songs are played, plans his programme. For each song chosen he writes down the song length, in terms of time, in minutes, taken to play the song. The following table summarises the song lengths.

Song length (minutes)	*Number of songs*	
2.8 – under 3.2	3	
3.2 – under 3.4	5	
3.4 – under 3.6	9	
3.6 – under 3.8	8	
3.8 – under 4.0	7	
4.0 – under 4.2	4	
4.2 – under 4.6	2	

(ii) Estimate, in minutes, the mean and standard deviation of these song lengths. Give your answers to 3 significant figures.

Mean = ...

Standard deviation = ...[8]

Information about five of the presenter's earlier programmes is shown below.

For Examiner's Use

Programme	*Number of songs played*	*Mean of song lengths* (minutes)	*Standard deviation of song lengths* (minutes)
P	38	3.70	0.339
Q	39	3.52	0.328
R	42	3.69	0.294
S	37	3.83	0.305
T	38	3.74	0.291

(iii) State in which of the programmes P, Q, R, S or T, songs were generally

(a) shortest in length,

.....[1]

(b) most similar in length.

.....[1]

All the presenter's programmes are three hours in duration. Songs are not played continuously throughout each programme; for some of the time the presenter talks about the songs and the singers.

A listener switched on programme P at a random time during its transmission.

(iv) Find the probability that a song was not being played at that moment.

.....[4]