Surname

Other Names

Candidate Signature

AQA

General Certificate of Secondary Education
Higher Tier
June 2011

Statistics
Written Paper

43101H

H

Monday 27 June 2011   9.00 am to 11.00 am

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

Time allowed
- 2 hours

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.
- Do all rough work in this book. Cross through any work you do not want to be marked.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 100.
- You may ask for more answer paper, graph paper and tracing paper. These must be tagged securely to this answer booklet.
- You are expected to use a calculator where appropriate.

Advice
- In all calculations, show clearly how you work out your answer.

You may need to use the following formulae:

Mean of a frequency distribution $= \dfrac{\sum fx}{\sum f}$

Mean of a grouped frequency distribution $= \dfrac{\sum fx}{\sum f}$, where x is the mid-interval value.

Standard deviation for a set of numbers $x_1, x_2, \ldots, x_n$ having a mean value of $\bar{x}$ is given by

$$\sqrt{\dfrac{\sum (x - \bar{x})^2}{n}} \quad \text{or} \quad \sqrt{\dfrac{\sum x^2}{n} - \bar{x}^2}$$

Standard deviation for a frequency distribution

$$\sqrt{\dfrac{\sum f(x - \bar{x})^2}{\sum f}} \quad \text{or} \quad \sqrt{\dfrac{\sum fx^2}{\sum f} - \bar{x}^2}$$

The same formula applies to the standard deviation of a grouped frequency distribution where x is the mid-interval value.

Spearman's rank correlation coefficient $= 1 - \dfrac{6\sum d^2}{n(n^2 - 1)}$

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

1   Jenny is doing a survey on people (tenants) who rent flats.
She uses two rental companies 'Letsmove' and 'Supaflat'.
Her hypothesis is

    "Tenants who rent with 'Letsmove' make fewer complaints than those who use 'Supaflat'."

1 (a)   Give **two** different reasons why Jenny should take a sample of tenants and not take a census.

Reason 1 ....

....

Reason 2 ....

....

(2 marks)

1 (b)   Jenny decides to take a sample using stratification.

1 (b) (i)   Name one category which she could use to stratify her sample.

Answer ....   *(1 mark)*

1 (b) (ii)   Give a reason why she might choose this category.

....

....

(1 mark)

1 (c)   Write a question which Jenny could use to find out the number of complaints a tenant had made.

....

....

....

(3 marks)

7

Turn over ▶

1 (d) Jenny is considering which data collection method to use from

telephone   door-to-door interviews   Internet survey.

Which method would you choose from this list?

Data collection method ....

1 (d) (i) Give **one** advantage of your method over the other two methods.

....

....
(1 mark)

1 (d) (ii) Describe a problem Jenny may have using your data collection method.

....

....
(1 mark)

2 Eva works in Spain for a travel company.
Her job is to help tourists on holiday.
The Venn diagram shows the **percentages** of tourists she has helped over the years
who have lost passports (P), tickets (T), money (M) or a combination of these.
She also helped people for other reasons.
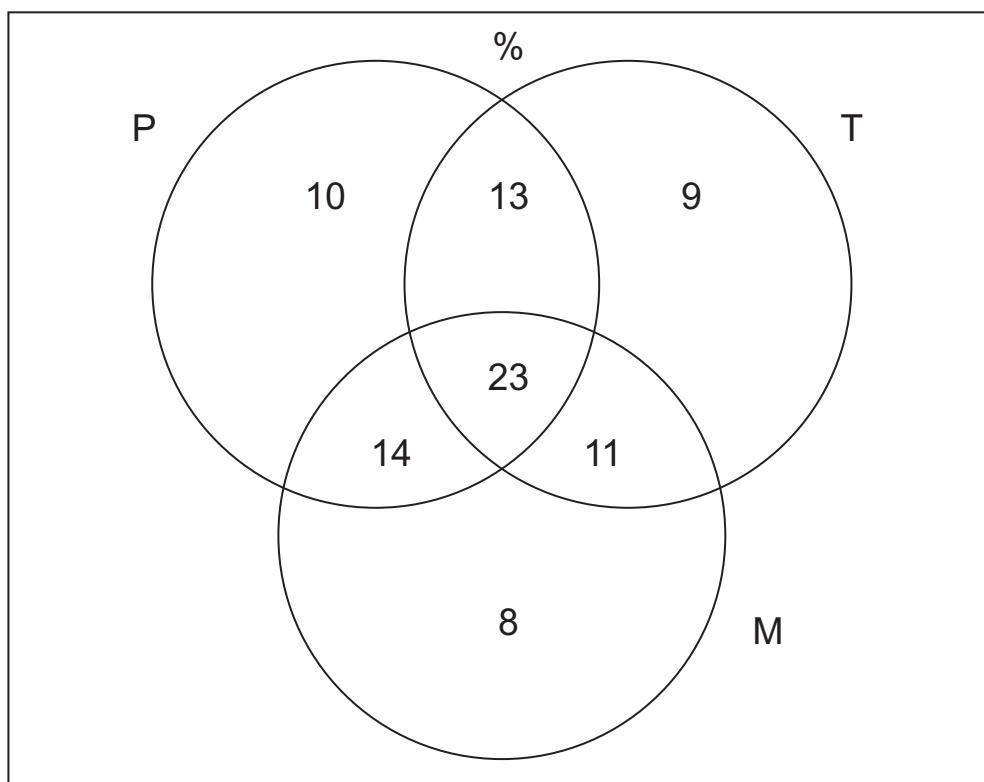
2 (a)   Explain what the 23 in the diagram represents.

....

....

(1 mark)

2 (b)   Work out the percentage of people Eva helped who had **not** lost their passports, tickets or money.

....

....

Answer .... %   *(2 marks)*

2 (c)   One person helped by Eva is chosen at random.

2 (c) (i)   What is the probability that this person lost their passport but **not** their tickets or money?

Answer ....   *(1 mark)*

2 (c) (ii)   What is the probability that this person lost their tickets and money but had **not** lost their passport?

....

Answer ....   *(1 mark)*

2 (d)   In 2010 Eva helped 4100 people.

Estimate the number who had lost their passport and money but had **not** lost their tickets.

....

....

Answer ....   *(2 marks)*

Turn over for the next question

Turn over ▶

3 A hotel is interviewing for a new receptionist.
Two of the interview panel observe and record whether candidates make good eye-contact when speaking.

3 (a) The data collected showed that differences occurred due to inter-observer bias.

Explain what this means in this context.

....

....

(1 mark)

3 (b) Give **one** way of reducing or avoiding inter-observer bias in this situation.

....

....

(1 mark)

3 (c) Name a possible extraneous variable which may have affected these results.

Answer ....   *(1 mark)*

4 The tables show information about the populations of two areas.

Eastbourne

Date	Number age 85+	Number all ages
June-01	4000	89 800
June-02	3900	91 000
June-03	3700	92 000
June-04	3700	93 100
June-05	3900	94 200
June-06	4000	94 900
June-07	4100	95 600
June-08	4200	96 100

North Lincolnshire

Date	Number age 85+	Number all ages
June-01	2800	153 000
June-02	2800	154 100
June-03	2800	155 900
June-04	2800	157 400
June-05	3000	158 200
June-06	3300	159 000
June-07	3400	159 400
June-08	3500	160 300

http://www.neighbourhood.statistics.gov.uk

4 (a) Write down **one** similarity between the data for the two areas.

....

....

(1 mark)

4 (b) Write down **one** difference between the data for the two areas.

....

....

(1 mark)

4 (c) Harry is investigating the hypothesis

'The population of Eastbourne is older on average than the population of North Lincolnshire'.

4 (c) (i) Give a reason why the data could support Harry's hypothesis.

....

....

(1 mark)

4 (c) (ii) Give a reason why the data does **not** prove Harry's hypothesis.

....

....

(1 mark)

7

Turn over ▶

5 The following is a summary of birth statistics in England and Wales for 1998-2008.

Year	Number of births	Sex ratio (male births per 1000 female births)	Mean age of mother at childbirth (years)	Percentage of births outside marriage	Percentage of births to non-UK born mothers
1998	635 901	1051	28.3	37.8	13.6
1999	621 872	1055	28.4	38.9	14.3
2000	604 441	1050	28.5	39.5	15.5
2001	594 634	1050	28.6	40.0	16.5
2002	596 122	1055	28.7	40.6	17.7
2003	621 469	1051	28.8	41.4	18.6
2004	639 721	1054	28.9	42.2	19.5
2005	645 835	1049	29.1	42.8	20.8
2006	669 601	1047	29.1	43.5	21.9
2007	690 013	1057	29.3	44.3	23.2
2008	708 711	1050	29.3	45.3	24.1

Source: *ONS 2009*

5 (a) Which year showed the largest increase in the number of births in England and Wales for 1999 - 2008?

....

....

....

Answer .... *(2 marks)*

5 (b) Make **two** distinct comments based on the figures in the column headed 'sex ratio (male births per 1000 female births).'

Comment 1 ....

....

....

Comment 2 ....

....

....

(2 marks)

5 (c) (i)   Suggest a possible reason for the increase in the 'mean age of mother at childbirth' over the period.

....

....

(1 mark)

5 (c) (ii)   Give a reason why the mean may **not** be the best measure of average to use in this case.

....

....

....

(1 mark)

5 (d)   Use the table to identify a factor which could explain the increase in births over the period.

....

....

....

(1 mark)

5 (e)   The table represents secondary data.

Give **two** advantages of using this type of data.

Advantage 1   ....

....

....

Advantage 2   ....

....

....

(2 marks)

6 The comparative pie charts below show details of UK waste disposal in 2000 and 2008.

2000

Recycling   Incineration

36°

270°

Landfill

2008

Incineration

42°

Recycling

192°

Landfill
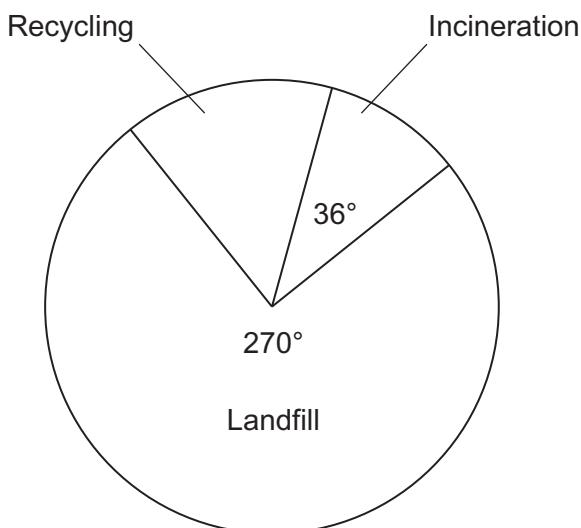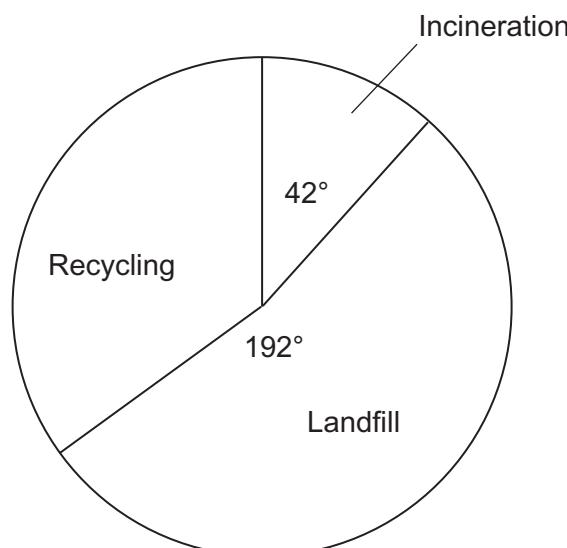
(Source: DEFRA Statistical Summary 2009)

In 2000, the **total** amount of UK waste disposal was **24.6 million tonnes**.

6 (a) Calculate the total amount of waste recycled in 2000.

....

....

....

....

....

Answer .... million tonnes   *(3 marks)*

6 (b) The radii of the two pie charts shown are 3 cm and 3.3 cm respectively.

Calculate the total amount of waste recycled in 2008.

....

....

....

....

....

....

Answer .... million tonnes   *(5 marks)*

6 (c) State **two** changes in the pattern of waste disposal from 2000 to 2008.

Change 1 ....

....

....

Change 2 ....

....

....

(2 marks)

Turn over for the next question

___
10

Turn over ▶

7 The table shows the number of times a sample of Year 11 students checked their e-mails last Tuesday.

Number of times checked e-mails	Number of students	
0	20	
1	23	
2	16	
3	11	
4	9	
5	0	
6	5	
7	4	
8	2	
Total	**90**	

7 (a) On the grid below, draw a cumulative frequency step polygon for the data.

Cumulative frequency

Number of times checked e-mails
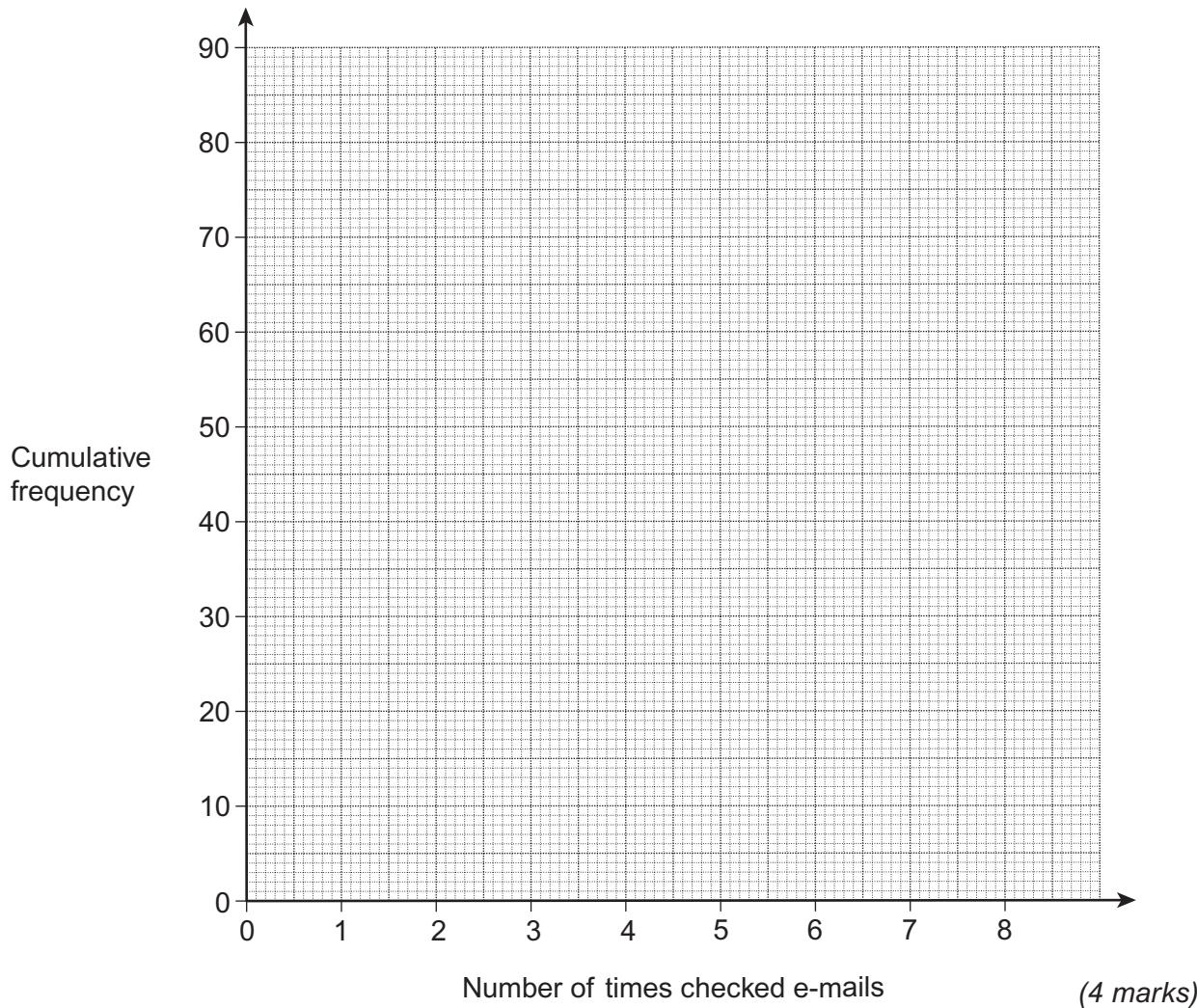
(4 marks)

7 (b) (i) Work out the median

Answer .... *(1 mark)*

7 (b) (ii) Work out the range between the 3rd and 8th decile.

....

....

Answer .... *(3 marks)*

7 (c) Three of the Year 11 students were selected at random from the group.
Find the probability that last Tuesday

7 (c) (i) the first student chosen checked their e-mails six times

....

Answer .... *(1 mark)*

7 (c) (ii) both the first two students chosen checked their e-mails fewer than five times each

....

....

....

Answer .... *(3 marks)*

7 (c) (iii) between them, the three students chosen checked their e-mails only once in total.

....

....

....

....

....

....

Answer .... *(4 marks)*

16

Turn over ▶

There are no questions printed on this page

**DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED**

8 The table shows the annual cost of Sunita's home contents insurance.
It also shows some of the chain base index numbers for the annual cost.

Year	2003	2004	2005	2006	2007	2008
Annual cost (£)	150	162	168	172	176	204
Chain base index number		108.0	103.7			

8 (a) Calculate the chain base numbers for the years 2006, 2007 and 2008 and enter them in the table.

Give your answers to 1 decimal place.

....

....

....

....

....

....
(4 marks)

8 (b) Which year had the smallest % increase in insurance cost?

....

Answer .... *(1 mark)*

8 (c) The geometric mean for the chain base index numbers is 106.3 to 1 decimal place.

Show clearly how this figure has been calculated.

....

....

....
*(2 marks)*

Turn over ▶

9   Andy works as delivery driver for a parcel distribution company.
Last Friday he recorded the weights of each of 100 parcels he delivered that day.
The results were summarised in the table.

Weight of parcel, x (kg)	Number of parcels	
$0 \leqslant x < 1$	12	
$1 \leqslant x < 3$	33	
$3 \leqslant x < 4$	18	
$4 \leqslant x < 7$	15	
$7 \leqslant x < 10$	12	
$10 \leqslant x < 14$	10	

Andy drew an **incorrect** histogram, shown below, to illustrate the data.

Histogram to show weights of parcels
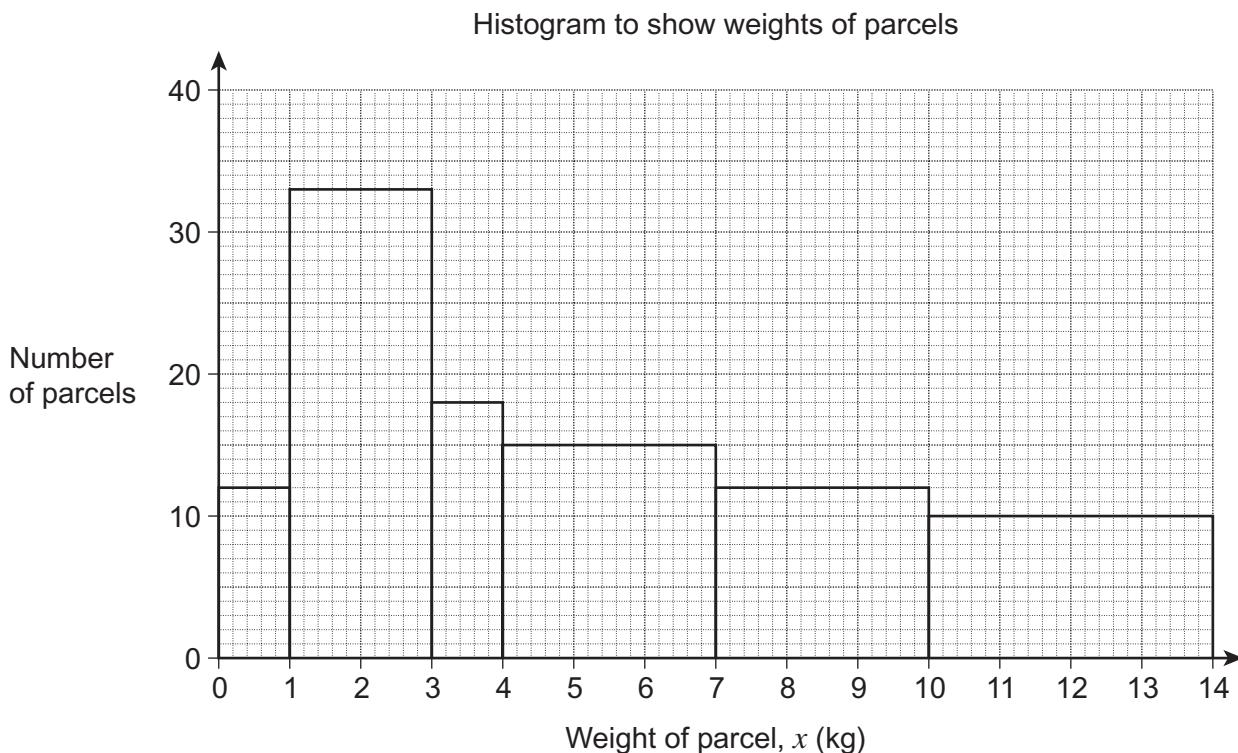
9 (a)   What has Andy done wrong?

...

....

(1 mark)

9 (b)   On the grid below draw a correct histogram to represent the data.

....

....

....

....

....

(6 marks)

Weight of parcel, x (kg)
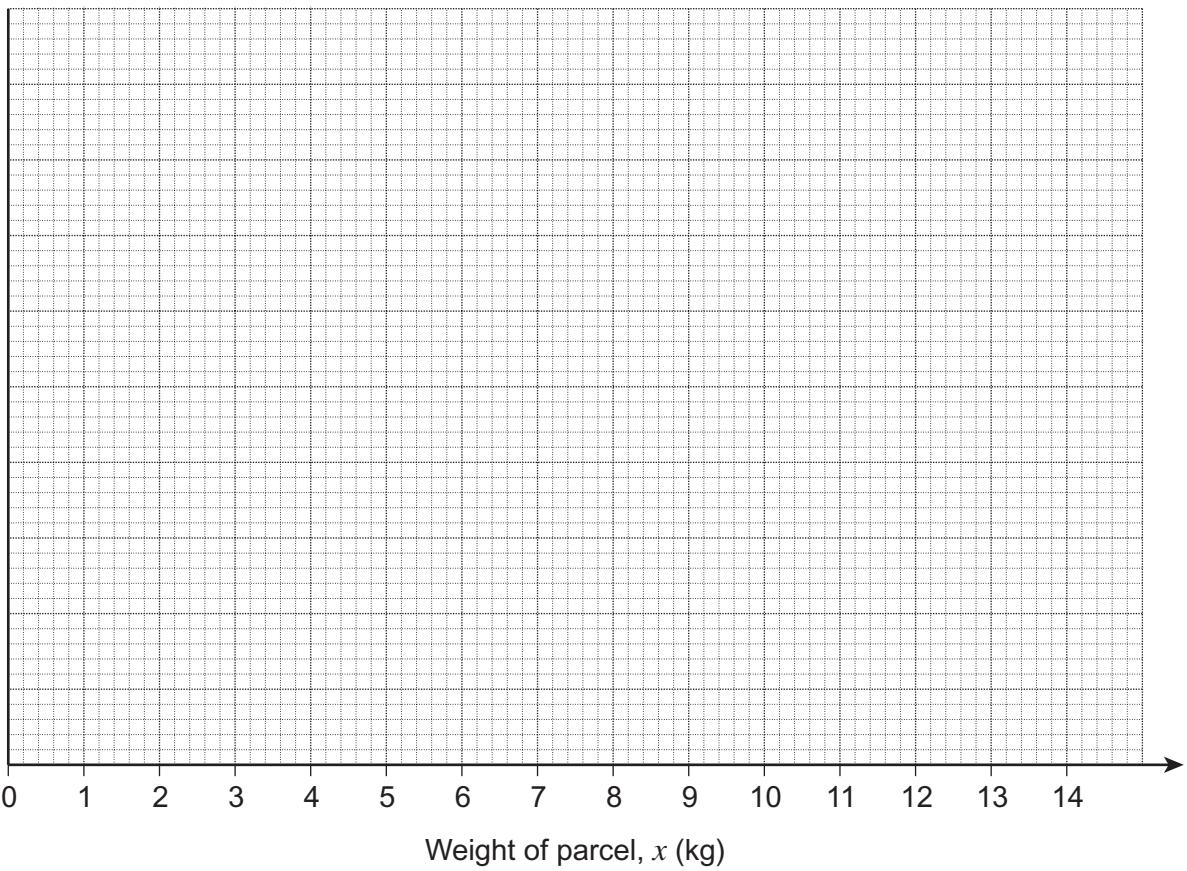
9 (c)   Describe the skewness of this distribution.

....

(1 mark)

9 (d)   Estimate the number of parcels Andy delivered last Friday weighing less than 5 kg.

....

....

....

Answer ....   *(2 marks)*

10

Turn over ▶

9 (e) Last Friday, equal numbers of parcels delivered by two other drivers, Bill and Tracy, were also weighed.
The weights of the parcels delivered by each driver were **Normally** distributed with mean and standard deviation as shown.

	Mean (kg)	**Standard deviation (kg)**
Bill	13	1.5
Tracy	12	3.5

9 (e) (i) A parcel, from one of the deliveries made by either Bill or Tracy, weighed 14 kg.

Use standardised scores to decide if this parcel is more likely to have been delivered by Bill or Tracy.

....

....

....

....

....

....

....

Answer .... *(4 marks)*

9 (e) (ii) Is it likely that a parcel weighing 8.3 kg could have been delivered by Bill?
You **must** show working to support your answer.

....

....

....

....

....

(3 marks)

There are no questions printed on this page

**DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED**

10 The following table shows the weekly household expenditure per person for three main ethnic groups (White, Asian and Black) in the UK for 2005/2006.

Weekly household expenditure per person by ethnic group in the UK 2005/2006

	£ per week		
	White	**Asian**	**Black**
Food and non-alcoholic drinks	19.30	12.70	13.50
Alcoholic drinks, tobacco and narcotics	13.50	1.10	1.70
Clothing and footwear	9.90	8.10	8.40
Housing (net), fuel and power	17.40	15.60	21.40
Household goods and services	5.10	8.70	8.70
Health	2.40	1.10	1.30
Transport	26.30	17.40	18.80
Communication	4.90	4.10	3.20
Recreation and culture	25.50	10.90	13.90
Education	2.30	4.50	6.10

Source Adapted From: *Expenditure and food survey, Office for National Statistics 2007*

As part of her Statistics assignment Sally has to measure how closely the expenditure patterns of White and Asian persons agree.

10 (a) She decides to use Spearman's rank correlation coefficient for this purpose.

You are given $\sum d^2 = 41.5$

Show clearly that the rank correlation coefficient between White and Asian pairs is + 0.748

....

....

....

....

....

....

(3 marks)

10 (b) The equivalent value for Spearman's rank correlation coefficient between Black and Asian pairs is + 0.973

Use the information in the table to explain why there is a stronger association between Black and Asian expenditure patterns than between White and Asian expenditure patterns.

Reason 1 ....

....

....

Reason 2 ....

....

....

(2 marks)

5

Turn over for the next question

Turn over ▶

11 A large department store employs additional part-time sales staff on Saturdays. The manager wants to know whether the number of part-time staff employed increases daily takings.

The scatter diagram below shows the number of part-time staff employed and takings over 10 successive Saturdays.

Part-time staff employed, x
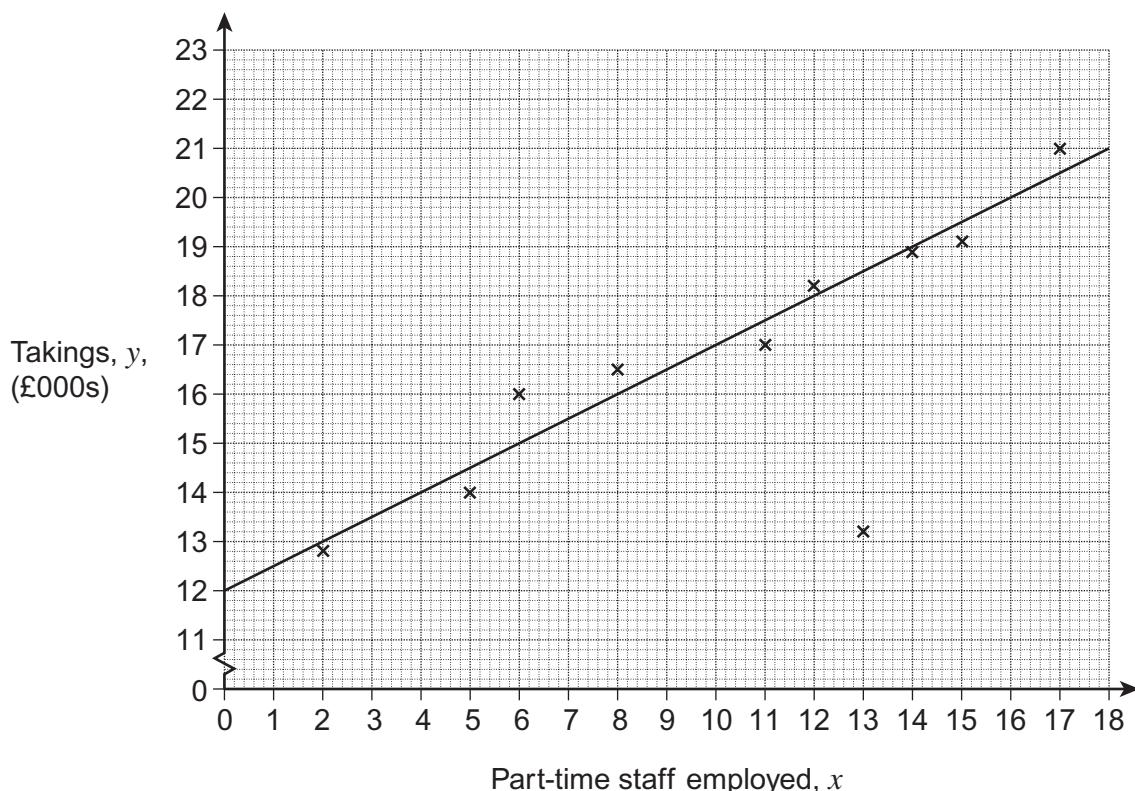
11 (a) On one of the Saturdays the store's car park was closed to customers.

11 (a) (i) How many part-time staff do you think were employed on that Saturday?

Answer .... *(1 mark)*

11 (a) (ii) Justify your choice.

....

....

(1 mark)

11 (b) A line of best fit has been drawn on the scatter diagram.

Work out the equation of this line in the form $y = mx + c$

....

....

....

....

....

....

Answer .... *(4 marks)*

11 (c) The manager comments that

'increasing the number of part-time staff will result in an increase in Saturday takings.'

State **two** reasons why this may **not** be a valid conclusion to draw based on the data provided.

Reason 1 ....

....

Reason 2 ....

....

(2 marks)

Turn over for the next question

8

Turn over ▶

12 The following table summarises the results of a survey into attitudes to road transport and the environment in 2008.

Attitudes to road transport and the environment

| | **Percentages** | | | | |
	X	**Agree**	**Y**	**Disagree**	**Disagree Strongly**
For the sake of the environment, car users should pay higher taxes	3	12	15	47	23
The Government should build more motorways to reduce traffic congestion	8	43	22	22	5
Driving one's car is too convenient to give up for the sake of the environment	8	37	24	21	10
Building more roads just encourages more traffic	13	38	18	25	6

Source: British Social Attitudes Survey, National Centre for Social Research 2009

12 (a) Complete the two missing headings in the table.

Answer **X** ....

Y ....

(2 marks)

12 (b) Explain why two conclusions that may be drawn from this survey appear to be contradictory.

....

....

....

....

(1 mark)

Do not write outside the box

12 (c)   The survey was undertaken using multistage sampling.

Give **one** advantage and **one** disadvantage of using this sampling method.

Advantage ....

....

Disadvantage ....

....

(2 marks)

END OF QUESTIONS

2 5

5

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED

There are no questions printed on this page

**DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED**