

Centre Number						Candidate Number				
Surname										
Other Names										
Candidate Signature										

For Examiner's Use	
Examiner's Initials	
Question	Mark
1	
2	
3	
4	
5	
6	
TOTAL	



General Certificate of Education
Advanced Subsidiary Examination
June 2012

Statistics

SS02

Unit Statistics 2

Thursday 24 May 2012 9.00 am to 10.30 am

For this paper you must have:

- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed

- 1 hour 30 minutes

Instructions

- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- Write the question part reference (eg (a), (b)(i) etc) in the left-hand margin.
- You must answer each question in the space provided for that question. If you require extra space, use an AQA supplementary answer book; do **not** use the space provided for a different question.
- Do not write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information

- The marks for questions are shown in brackets.
- The maximum mark for this paper is 75.

Advice

- Unless stated otherwise, you may quote formulae, without proof, from the booklet.
- You do not necessarily need to use all the space provided.



J U N 1 2 S S 0 2 0 1

Answer **all** questions.

Answer each question in the space provided for that question.

- 1** Past experience shows that, when they play home matches, Newcastle Athletic football team win half the matches and lose 20% of the matches. They draw the remaining home matches. The team is awarded three points for a win, one point for a draw and no points for a loss.
- (a)** The table below shows the probability distribution of X , the number of points awarded to Newcastle Athletic after a home match. Complete the table. *(2 marks)*
- (b)** Calculate the mean and the standard deviation of the number of points awarded to Newcastle Athletic after a home match. *(5 marks)*
- (c)** State the probability that the number of points awarded to Newcastle Athletic after a home match is more than one standard deviation above the mean. *(1 mark)*

QUESTION
PART
REFERENCE

Answer space for question 1

(a)

x	0	1	3
$P(X = x)$			0.5



QUESTION
PART
REFERENCE

Answer space for question 1

A large rectangular area containing horizontal dotted lines for writing an answer.

Turn over ►



2 The table shows the number, in millions, of overseas visitors to the United Kingdom. It also shows the values of an appropriate n -point moving average.

Year	2007				2008				2009				2010
Quarter	1	2	3	4	1	2	3	4	1	2	3	4	1
Million visitors	6.8	8.6	9.6	7.4	7.2	8.4	9.2	6.8	6.2	8.0	8.8	7.0	x
Moving average	8.10		8.20	y	8.05	7.90	7.65	7.55	7.45	7.50	7.40		

- (a) (i) State the value of n . (1 mark)
- (ii) Calculate the value of y , representing the third value of the moving average. (2 marks)
- (iii) Calculate the value of x , representing the number, in millions, of overseas visitors during Quarter 1 of 2010. (2 marks)
- (b) The data from the table have been plotted in **Figure 1** on page 7, together with a regression line.
 - (i) Complete **Figure 1** by plotting the points representing the values of x and y . (2 marks)
 - (ii) Describe the variation and trend shown by the moving average. (2 marks)
 - (iii) Using the given regression line, estimate the second quarter seasonal effect. (3 marks)
 - (iv) Forecast the number of overseas visitors to the United Kingdom during the second quarter of 2010. Show the method that you have used to make the forecast. (3 marks)
 - (v) Comment on the likely accuracy of a forecast, made using the same data and same method, for the number of overseas visitors to the United Kingdom during the second quarter of 2012. (1 mark)

QUESTION PART REFERENCE

Answer space for question 2

.....

.....

.....

.....

.....

.....

.....



QUESTION
PART
REFERENCE

Answer space for question 2

A large rectangular area with horizontal dotted lines for writing.



Turn over ►

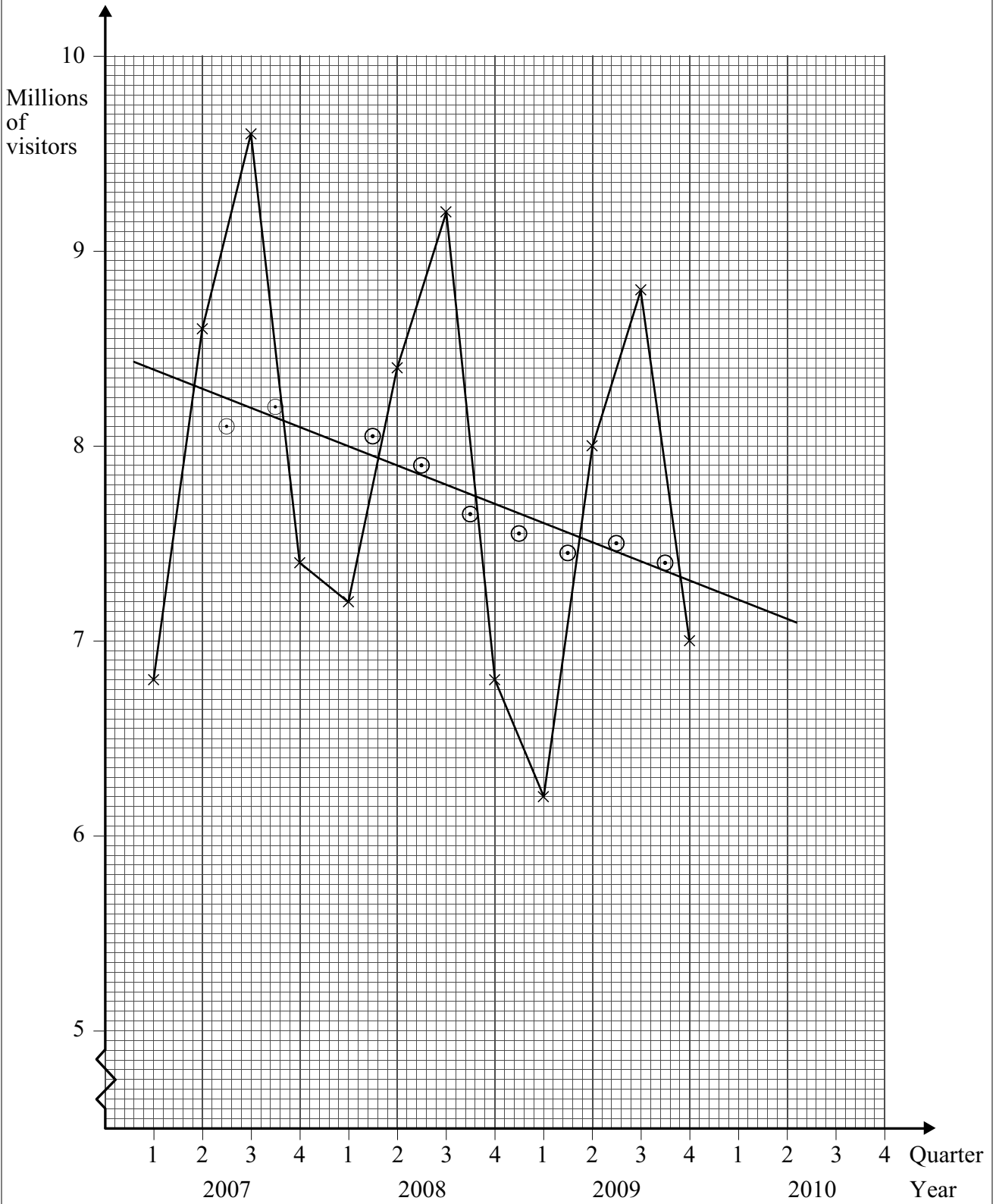
QUESTION
PART
REFERENCE

Answer space for question 2

A large rectangular area with horizontal dotted lines for writing an answer.



Figure 1



Turn over ►



- 3 The US Geological Survey detects and records earthquakes throughout the world. The table shows the data collected for the United States over the ten-year period 2000–2009.

**Number of earthquakes in the United States for 2000–2009
located by the US Geological Survey National Earthquake Information Center**

Year Magnitude	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
8.0 to 8.9	0	0	0	0	0	0	0	0	0	0
7.0 to 7.9	0	1	1	2	0	1	0	1	0	0
6.0 to 6.9	6	5	4	7	2	4	7	9	9	6
5.0 to 5.9	63	41	63	54	25	47	51	72	85	53
4.0 to 4.9	281	290	536	541	284	345	346	366	432	293
3.0 to 3.9	917	842	1535	1303	1362	1475	1213		1486	1491
2.0 to 2.9	660	646	1228	704	1336	1738	1145	1173	1573	2372
1.0 to 1.9	0	2	2	2	1	2	7	11	13	25
0.1 to 0.9	0	0	0	0	0	0	1	0	0	1
No magnitude	415	434	507	333	540	73	13	22	20	16
Total	2342	2261	3876	2946	3550	3685	2783	2791	3618	4257

Source: US Geological Survey, 2010

- (a) The number of earthquakes of magnitude 3.0 to 3.9 for the year 2007 is missing from the table. Evaluate this number. (2 marks)
- (b) Calculate the mean number of earthquakes per year classed as magnitude 6.0 or greater. (2 marks)
- (c) A geologist decides to use a Poisson distribution with mean 6.5 to model the number of earthquakes, X , classed as magnitude 6.0 or greater that occur in the United States during a year.
- Using this Poisson model, calculate the probability that, during a particular year in the United States, there will occur:
- (i) fewer than 3 earthquakes classed as magnitude 6.0 or greater; (1 mark)
- (ii) more than 10 earthquakes classed as magnitude 6.0 or greater; (2 marks)
- (iii) exactly 5 earthquakes classed as magnitude 6.0 or greater. (2 marks)



- (d) The geologist also decides to model the number of earthquakes, Y , classed as magnitude 6.0 or greater that occur in Canada during a year. This model will use a Poisson distribution with mean 3.1. Assume that X and Y are independent.
 - (i) State the mean for the Poisson distribution that could be used to model the total number of earthquakes classed as magnitude 6.0 or greater that occur in the United States and Canada during a year. *(1 mark)*
 - (ii) Hence, or otherwise, calculate the probability that **no** earthquakes classed as magnitude 6.0 or greater will occur in either the United States or Canada during a period of **three months**. *(2 marks)*
- (e) An earthquake classed as magnitude 7.0 or greater is often followed by several further earthquakes of slightly smaller magnitude. Using this information, comment on the appropriateness of using the Poisson distribution to model the number of earthquakes classed as magnitude 6.0 or greater. *(2 marks)*

QUESTION
PART
REFERENCE

Answer space for question 3

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Turn over ►



QUESTION
PART
REFERENCE

Answer space for question 3

A large rectangular area containing horizontal dotted lines for writing an answer.



QUESTION
PART
REFERENCE

Answer space for question 3

A large rectangular area with horizontal dotted lines for writing an answer.

Turn over ►



4 An exam board was developing a new examination for Year 11 pupils. It was intended that the mean mark obtained in the examination should be 50.

A specimen examination paper was produced. It was suggested that this examination paper should be taken by 80 pupils. Four Year 11 classes would be chosen, one class from each of four schools, and from each of these classes, 20 pupils would be chosen to take the examination paper.

(a) (i) Name this type of sampling. *(1 mark)*

(ii) Explain why data obtained by this sampling method would not give a reliable result for a test of the hypothesis that the mean mark of all Year 11 pupils on this examination was 50. *(1 mark)*

(b) The sampling method actually used by the exam board was to select Year 11 pupils at random from those registered with the exam board. The specimen examination paper was sent to each pupil's school, with a request that the pupil should take the examination and that the examination paper should then be returned to the exam board. The schools returned 81 papers.

The mean mark for the 81 papers returned was 52.1 and the standard deviation was 7.8. Carry out a hypothesis test, at the 1% level of significance, to investigate whether the mean mark for the specimen examination is 50. *(8 marks)*

(c) State, with a reason, whether either of the following facts affects the validity of your conclusion in part **(b)**.

(i) It is not known that the distribution of marks is normal. *(2 marks)*

(ii) The exam board actually sent examination papers to 100 pupils but only 81 were returned. *(2 marks)*

QUESTION
PART
REFERENCE

Answer space for question 4

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....



QUESTION
PART
REFERENCE

Answer space for question 4

A large rectangular area with horizontal dotted lines for writing an answer.

Turn over ►



- 5** A doctor in California routinely weighs patients who visit his surgery. As part of a health survey, he selects a random sample of 180 women aged 30–39 years whose weights have been recorded. **Figure 2**, printed on the opposite page, is a cumulative frequency diagram based on the data.
- (a) (i) State the number of women weighing less than 125 pounds. *(1 mark)*
- (ii) Calculate the percentage of women weighing more than 200 pounds. *(3 marks)*
- (b) The doctor selects a similar sample of men aged 30–39 years. **Figure 3**, also printed on the opposite page, is a box and whisker plot based on the data obtained.
- (i) Draw a box and whisker plot in the space provided on **Figure 3** to illustrate the data for the women as presented in **Figure 2**. You may assume that there are no outliers. *(5 marks)*
- (ii) Make **three** comments on the differences between the way in which the weights of the men are distributed compared with the way in which the weights of the women are distributed, as shown in the box and whisker plots. *(3 marks)*
- (c) Explain why the data obtained from the two samples may not reliably represent the population of Californian people aged 30–39 years. *(1 mark)*

QUESTION
PART
REFERENCE

Answer space for question 5

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....



Figure 2

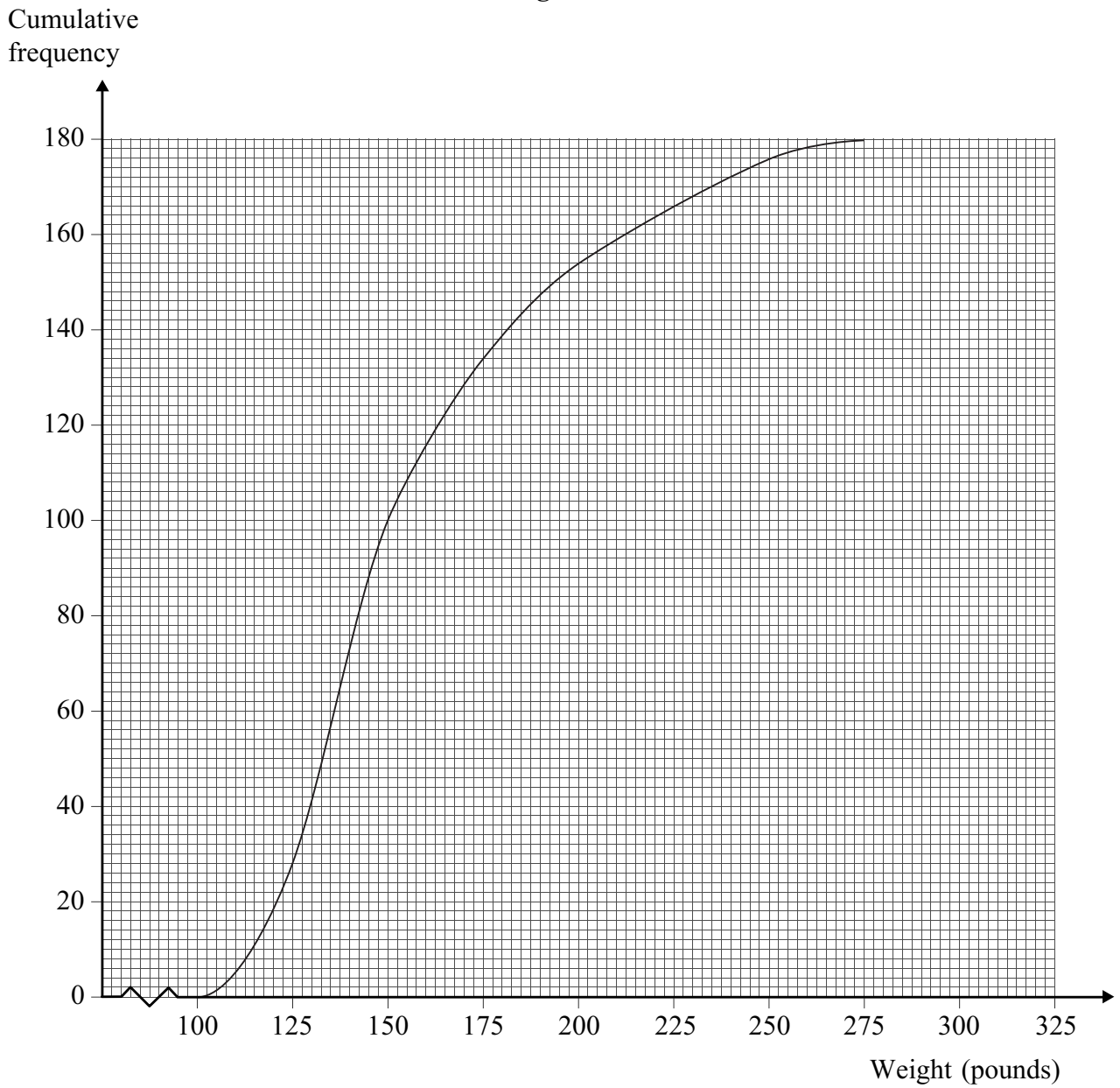
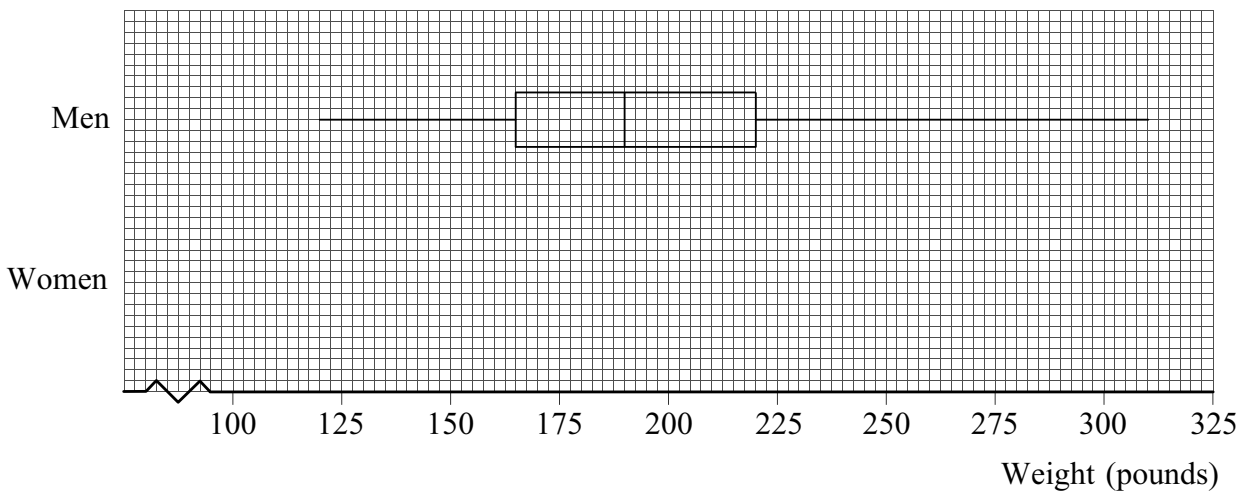


Figure 3



Turn over ►



QUESTION
PART
REFERENCE

Answer space for question 5

A large rectangular area with horizontal dotted lines for writing an answer.



QUESTION
PART
REFERENCE

Answer space for question 5

A large rectangular area with horizontal dotted lines for writing an answer.



6 A journalist wishes to examine the expenses claims of a sample of 10 local councillors. There are 80 councillors, of whom 48 belong to the Preservative Party and 32 belong to the Action Party. The journalist has a list of all the councillors' names in alphabetical order. She numbers the councillors from 00 to 79 and plans to select a simple random sample.

(a) Describe how she could use random number tables to obtain such a simple random sample of 10 councillors. *(4 marks)*

(b) The journalist in fact uses random numbers from **Table 1**.

Table 1

0 9 6 3 2	0 7 1 8 2	7 8 1 1 1	9 1 3 7 1
5 4 2 6 2	7 8 6 2 6	4 0 8 6 5	1 9 2 5 3

List the numbers obtained from **Table 1**, in the order that they are chosen, that will identify the councillors who will form the journalist's sample. Start with the figures **0 9** in bold and read along the row from left to right. *(2 marks)*

(c) When the journalist looks at the councillors in her random sample, she finds that there are nine from the Preservative Party and only one from the Action Party. The journalist does not want to be accused of bias in selecting her sample.

(i) She considers taking a systematic sample from the list of 80 councillors. Explain why a systematic sample from this list would not necessarily produce a sample of councillors which is more representative of the numbers of councillors in the two parties. *(2 marks)*

(ii) She decides instead to select a stratified sample. State how many councillors from each party should be contained in a stratified sample containing 10 councillors such that it is representative of the numbers of councillors in the two parties. *(2 marks)*

QUESTION
PART
REFERENCE

Answer space for question 6

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....



QUESTION
PART
REFERENCE

Answer space for question 6

A large rectangular area with horizontal dotted lines for writing an answer.

Turn over ►



QUESTION
PART
REFERENCE

Answer space for question 6

A large rectangular area containing horizontal dotted lines for writing an answer.

END OF QUESTIONS

Copyright © 2012 AQA and its licensors. All rights reserved.

