edexcel
advancing learning, changing lives

Sample Assessment Materials

Edexcel GCSE in Statistics (2ST01)

Inside this Sample Assessment Materials pack you'll find:

- Accessible papers to help you and your students prepare for the assessment
- Clear and concise mark schemes to let you know what the examiners are looking for
- Supported controlled assessment information including sample tasks and assessment criteria to share with students.

Issue 3
February 2010

A PEARSON COMPANY

Welcome to the GCSE 2009 Statistics Sample Assessment Materials

Issue 2
As a result of feedback from centres we have made changes.
This version is Issue 2 and key changes are indicated by a sideline.

The sample assessment materials have been developed to give you and your students a flavour of the actual exam papers and mark schemes so they can experience what they will encounter in their live assessments. They feature:

- **Accessible papers** using a mixture of question styles. We've worked hard to ensure the papers are easy to follow with an encouraging tone so that the full range of students can show what they know.

- **Clear and concise mark schemes** for each paper outlining what examiners will be looking for in the assessments, so you can use the sample papers with students to help them prepare for the real thing.

- **Supported controlled assessment,** including sample controlled assessment materials to show you the sort of activity students will undertake. Used in conjunction with the guidance in the Teacher's Guide, these samples will help you manage the controlled assessment in your centre and help students achieve their best.

Our GCSE 2009 Statistics qualification will be supported better than ever before. Keep up to date with the latest news and services available by visiting our website: **www.edexcel.com/gcse2009**

Contents

General Marking Guidance **2**

Unit 1F: Foundation Tier

Sample Assessment Material 5

Sample Mark Scheme 29

Unit 1H: Higher Tier

Sample Assessment Material 41

Sample Mark Scheme 65

Unit 2: Sample Controlled Assessment Material

Theme: Estimation 81

Theme: Human Body 97

Assessment Criteria 113

 © Edexcel Limited 2008

General Marking Guidance

- All candidates must receive the same treatment. Examiners must mark the first candidate in exactly the same way as they mark the last.
- Mark schemes should be applied positively. Candidates must be rewarded for what they have shown they can do rather than penalised for omissions.
- Examiners should mark according to the mark scheme not according to their perception of where the grade boundaries may lie.
- There is no ceiling on achievement. All marks on the mark scheme should be used appropriately.
- All the marks on the mark scheme are designed to be awarded. Examiners should always award full marks if deserved, i.e. if the answer matches the mark scheme. Examiners should also be prepared to award zero marks if the candidate's response is not worthy of credit according to the mark scheme.
- Where some judgement is required, mark schemes will provide the principles by which marks will be awarded and exemplification may be limited.
- When examiners are in doubt regarding the application of the mark scheme to a candidate's response, the team leader must be consulted.
- Mark schemes will indicate within the table where, and which strands of QWC, are being assessed. The strands are as follows:

  i) ensure that text is legible and that spelling, punctuation and grammar are accurate so that meaning is clear

  ii) select and use a form and style of writing appropriate to purpose and to complex subject matter

  iii) organise information clearly and coherently, using specialist vocabulary when appropriate.
- Crossed out work should be marked UNLESS the candidate has replaced it with an alternative response.
- **Types of mark**

  - M marks: method marks
  - A marks: accuracy marks
  - B marks: unconditional accuracy marks (independent of M marks)

 © Edexcel Limited 2008

- **Abbreviations**

  cao – correct answer only
  ft – follow through
  isw – ignore subsequent working
  SC - special case
  oe – or equivalent (and appropriate)
  dep – dependent
  indep - independent

- **No working**
  If no working is shown then correct answers normally score full marks.
  If no working is shown then incorrect (even though nearly correct) answers score no marks.

- **With working**
  If there is a wrong answer indicated on the answer line always check the working in the body of the script (and on any diagrams), and award any marks appropriate from the mark scheme.
  If it is clear from the working that the "correct" answer has been obtained from incorrect working, award 0 marks.
  Any case of suspected misread loses A (and B) marks on that part, but can gain the M marks.
  If working is crossed out and still legible, then it should be given any appropriate marks, as long as it has not been replaced by alternative work.
  If there is a choice of methods shown, then no marks should be awarded, unless the answer on the answer line makes clear the method that has been used.
  If there is no answer on the answer line then check the working for an obvious answer.

- **Follow through marks**
  Follow through marks which involve a single stage calculation can be awarded without working since you can check the answer yourself, but if ambiguous do not award.
  Follow through marks which involve more than one stage of calculation can only be awarded on sight of the relevant working, even if it appears obvious that there is only one way you could get the answer given.

- **Ignoring subsequent work**
  It is appropriate to ignore subsequent work when the additional work does not change the answer in a way that is inappropriate for the question: eg. incorrect cancelling of a fraction that would otherwise be correct.
  It is not appropriate to ignore subsequent work when the additional work essentially makes the answer incorrect eg algebra.
  Transcription errors occur when candidates present a correct answer in working, and write it incorrectly on the answer line; mark the correct answer.

 © Edexcel Limited 2008

- **Probability**
  Probability answers must be given a fractions, percentages or decimals. If a candidate gives a decimal equivalent to a probability, this should be written to at least 2 decimal places (unless tenths).
  Incorrect notation should lose the accuracy marks, but be awarded any implied method marks.
  If a probability answer is given on the answer line using both incorrect and correct notation, award the marks.
  If a probability fraction is given then cancelled incorrectly, ignore the incorrectly cancelled answer.

- **Linear equations**
  Full marks can be gained if the solution alone is given on the answer line, or otherwise unambiguously indicated in working (without contradiction elsewhere). Where the correct solution only is shown substituted, but not identified as the solution, the accuracy mark is lost but any method marks can be awarded.

- **Parts of questions**

  Unless allowed by the mark scheme, the marks allocated to one part of the question CANNOT be awarded in another.

 © Edexcel Limited 2008

Write your name here

Surname	Other names

Centre Number | Candidate Number

Edexcel GCSE

Statistics

Unit 1 – Paper 1F

Foundation Tier

Sample Assessment Material

Time: 1 hour 30 minutes

Paper Reference

5ST1F/01

Ruler graduated in centimetres and millimetres, protractor, compasses, pen, HB pencil, eraser, electronic 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** the questions.
- Answer the questions in the spaces provided – *there may be more space than you need.*

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.*
- Questions labelled with an **asterisk** (*) are ones where the quality of your written communication will be assessed. – *you should take particular care with your spelling, punctuation and grammar, as well as the clarity of expression, on these questions.*

Advice

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

Turn over ▶

N35541A

©2008 Edexcel Limited.

3/1

N 3 5 5 4 1 A 0 1 1 6

edexcel advancing learning, changing lives

GCSE Statistics 1ST0

Foundation Tier Formulae

You must not write on this page.
Anything you write on this page will gain NO credit.

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.

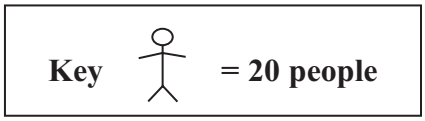
 © Edexcel Limited 2008

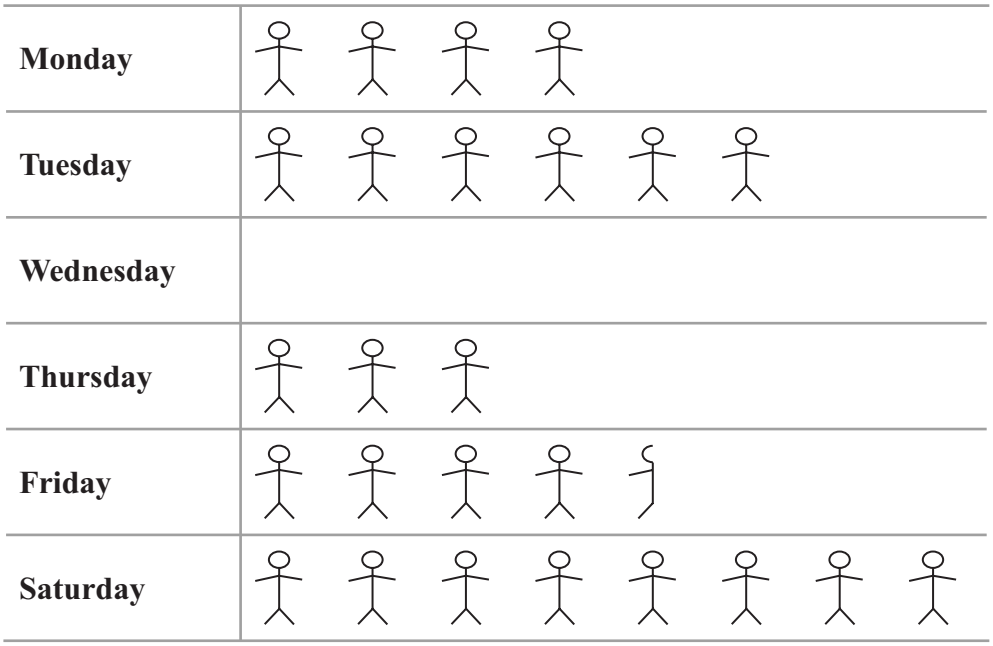
Answer ALL questions.

You must write down all stages of your working.

1 The pictogram shows the number of people who visited Woodley Post Office in one week.

The pictogram is not complete.





Forty people visited the post office on Wednesday.

(a) Complete the pictogram for Wednesday.

(1)

(b) Write down the day on which the most people visited the post office.

(1)

.....

(c) Write down the number of people who visited the post office on **Friday**.

(1)

.....

(Total for Question 1 = 3 marks)

2 Arla Foods employs 79 drivers.

The table gives information about these 79 drivers.
The table is not complete.

	Day shift	**Night shift**	**Total**
Tanker drivers	22	8	30
Van drivers	37	12	
Total			79

(Source: Arla Foods, Wythenshawe)

(a) Complete the table.

(2)

The Ministry of Transport checks the number of hours worked by one of these drivers picked at random.

(b) Write down the probability that this driver will be

(i) a tanker driver,

(1)

.....

(ii) a van driver who works on the night shift.

(1)

.....

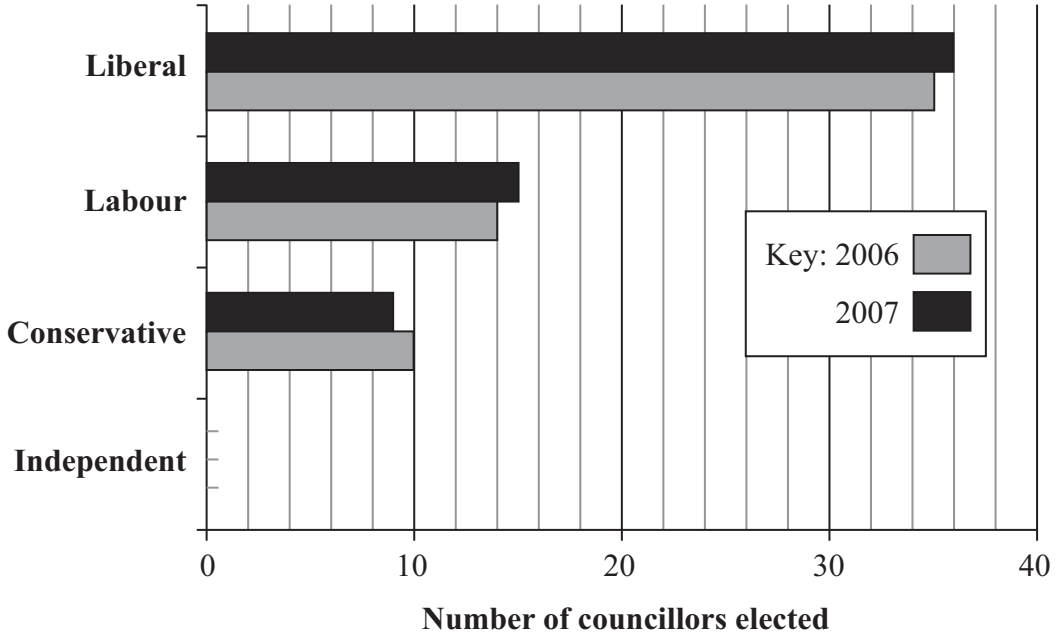
(Total for Question 2 = 4 marks)

 © Edexcel Limited 2008

BLANK PAGE

 © Edexcel Limited 2008

3 The chart gives information about councillors elected to Stockport Council in 2006 and in 2007.
It shows the number of councillors elected in each of the political groups.
The chart is incomplete.



(Source: Stockport Council)

There were 4 independent councillors in 2006.
There were 3 independent councillors in 2007.

(a) Use this information to complete the chart.
(2)

(b) Write down the name of the political group with the highest number of councillors elected in 2006 and 2007.
(1)

.....

(c) Write down the names of the **two** political groups with **fewer** councillors elected in 2007 than in 2006.
(1)

.....

.....

(d) Work out the total number of Conservative and Independent councillors elected in 2007.
(2)

.....

 © Edexcel Limited 2008

(e) Use the chart to comment on the changes to the political groups in Stockport Council over these two years.

(2)

.....

.....

.....

(Total for Question 3 = 8 marks)

 © Edexcel Limited 2008

4 The table shows the total rainfall for each quarter year from 2005 to 2007 at Eastbourne. The table is incomplete.

Year	Quarter	Rainfall (mm)	4-Point moving Average
2005	1	119	
	2	106	
			(119+106+153+234)/4 = 153
	3	153	
			(106+153+234+155)/4 = 162
	4	234	
			(153+234+155+110)/4 = 163
2006	1	155	
			(234+155+110+166)/4 = 166
	2	110	
			(155+110+166+271)/4 = 176
	3	166	
			(110+166+271+232)/4 = 195
	4	271	
			(166+271+232+150)/4 = 205
2007	1	232	
	2	150	
	3	151	
	4	187	

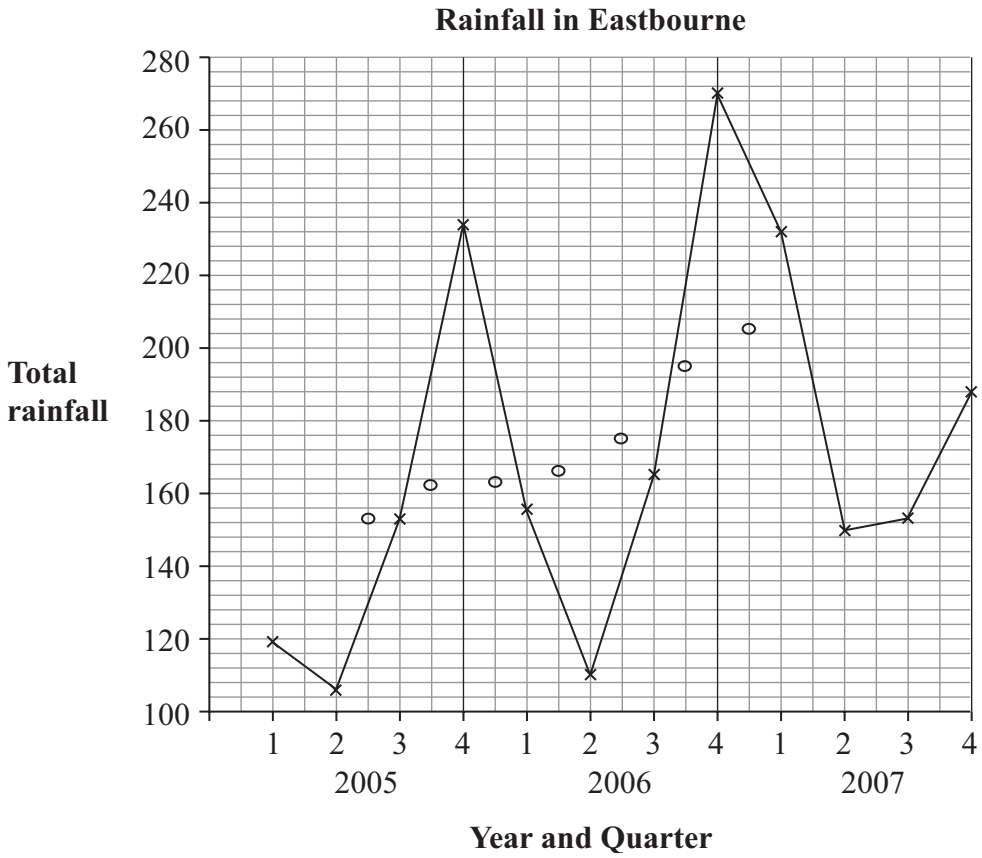
(Source: Eastbourne Borough Council)

 © Edexcel Limited 2008

(a) The first seven 4-point moving averages have been worked out to the nearest whole number.

(i) Complete the table to show the last **two** 4-point moving averages. **(2)**

(ii) Plot these **two** moving averages on the time series graph. **(2)**



Ali is planning to visit Eastbourne.
She wants to go when there is least rainfall.

(b) In which quarter does Eastbourne have the least rainfall? **(1)**

.....

(c) Draw a trend line on the graph. **(1)**

(d) Describe the trend in the rainfall from 2005 to 2007. **(1)**

.....

(Total for Question 4 = 7 marks)

5. Sarah is interested in how people travel to work.
She does a survey of all the people employed by her company.

She records the following data for each person

A The type of transport they used.
B The distance they travelled to work.
C The number of times they were late for work.

(a) Put a cross in **one** box in each table below to show which data can be described as

(i) continuous data

(1)

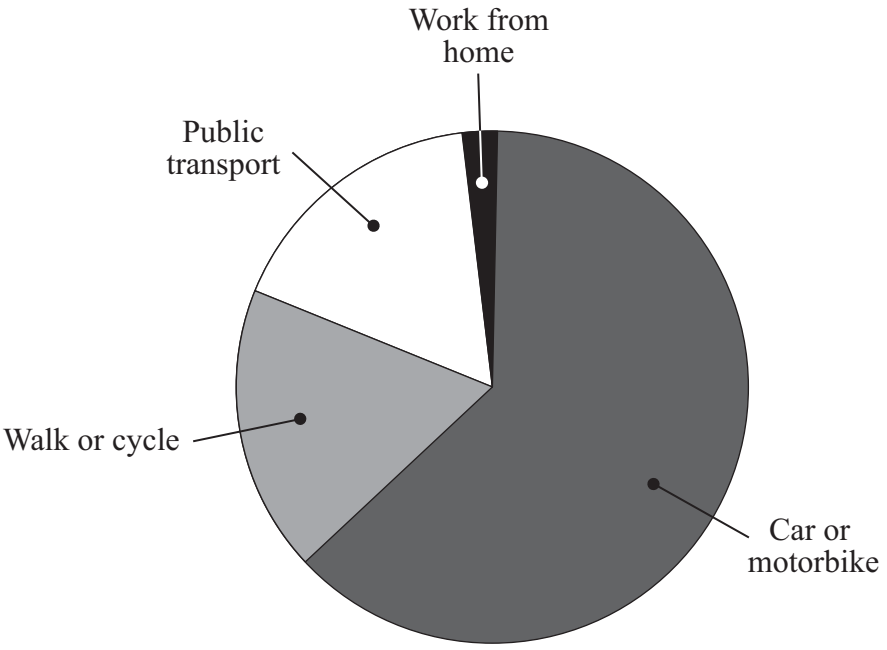
The type of transport they used.	☒
The distance they travelled to work.	☒
The number of times they were late for work.	☒

(ii) qualitative data.

(1)

The type of transport they used.	☒
The distance they travelled to work.	☒
The number of times they were late for work.	☒

The pie chart shows some information about how people in the UK travelled to work in



2007.

(Source: Department for Environment, Food and Rural Affairs)

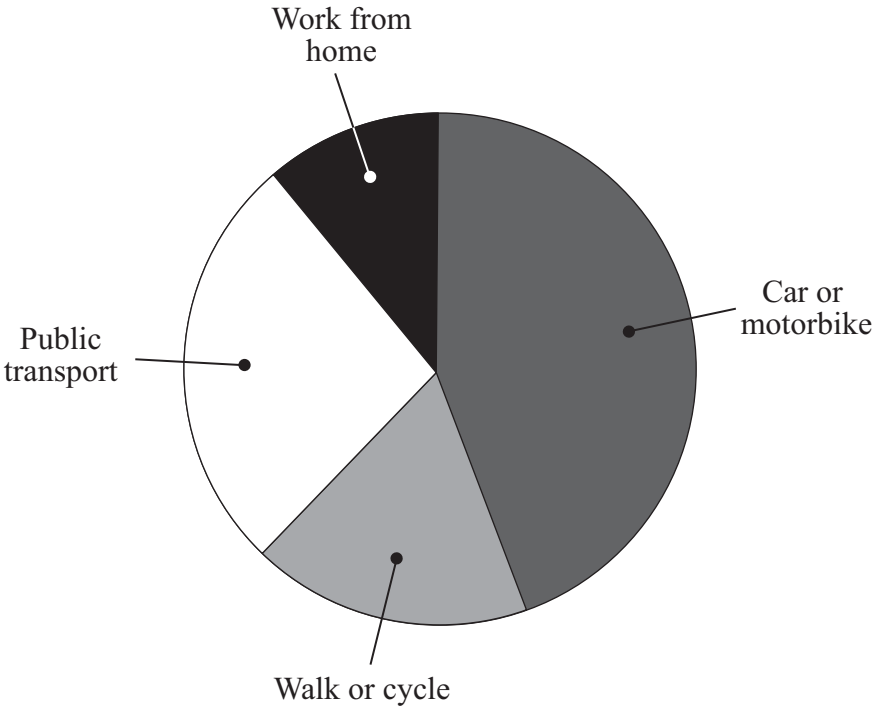
 © Edexcel Limited 2008

(b) Write down the most common method of travelling to work.

(1)

.....

The pie chart below shows information about how people from Sarah's company travel to work.



(c) Use the pie charts to compare how people travel to work at Sarah's company with how people travel to work in the whole of the UK.

(2)

.....

.....

.....

.....

.....

(Total for Question 5 = 5 marks)

6 The data below shows the average life expectancy, to the nearest year, of females in 11 countries in the Southern hemisphere.

31 39 56 61 62 71 71 71 72 45 69

(Source: World Health Organisation)

(a) Use these data to complete the stem and leaf diagram below.

(3)

3	1
4	
5	
6	
7	

(b) Work out the median of these data.

(1)

.....

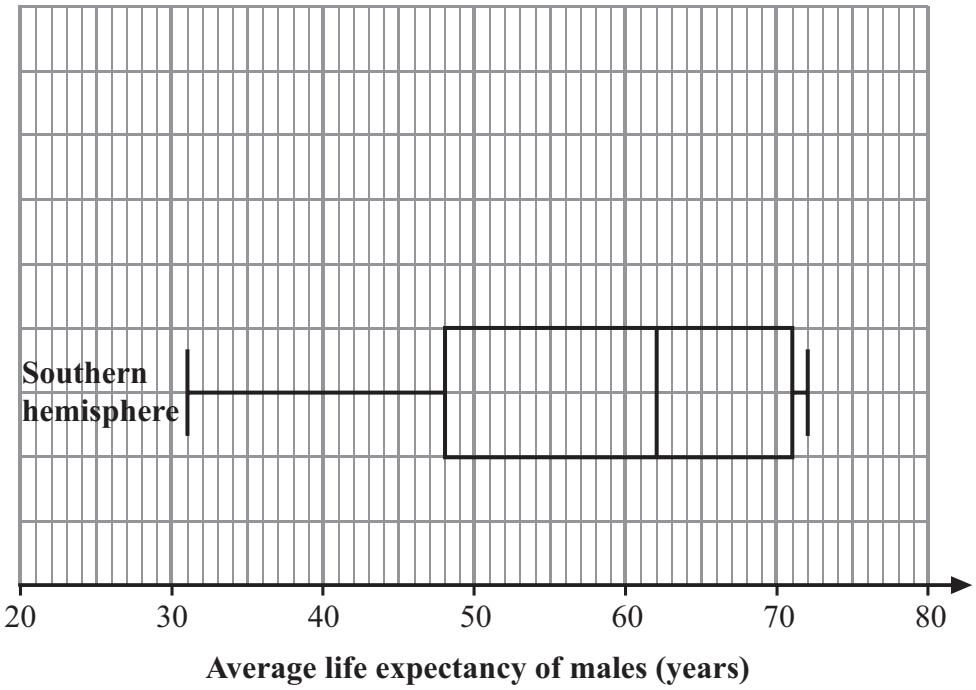
(c) Work out the range of these data.

(1)

.....

 © Edexcel Limited 2008

The box plot shows information about the average life expectancy of males in some countries in the **Southern hemisphere**.



(Source: World Health Organisation)

The table gives information about the average life expectancy, in years, of males in some countries in the **Northern hemisphere**.

Minimum	Lower quartile	Median	Upper quartile	Maximum
48	62	72	78	79

(d) On the grid above, use this data to draw a box plot to show the distribution of average life expectancy for males in these countries in the **Northern hemisphere**.

(3)

(e) Compare the distributions of the average life expectancy of males in the **Southern** hemisphere with the average life expectancy of males in the **Northern** hemisphere.

(2)

.....

.....

.....

.....

.....

(Total for Question 6 = 10 marks)

7 A supermarket manager wants to find out information about the vehicles going into the supermarket car park.

The vehicles are to be classified as Small cars, Large cars, Vans, or Other.

(a) Design a data collection sheet which could be used to record how many of each of these types of vehicles go into the car park.

(1)

The car park attendant collects information about cars entering the car park.

He stands by the entrance of the car park at 10am on one Tuesday for one hour.

(b) Give **one** reason why this survey method may not give representative results.

(1)

.....

.....

(Total for Question 7 = 2 marks)

 © Edexcel Limited 2008

8 The table gives information about the water resources and populations of seven countries in 2000.

Country	Renewable freshwater resource (m^3/yr per person)	Population (millions)	Shortfall in freshwater (km^3/yr)	Ocean coastline (approx.) km	Freshwater potential from wave powered desalination	
					km^3/yr	% of shortfall
Antigua	800	0.07	0.06	40	0.2	333
Barbados	307	0.27	0.38	40	0.2	53
Kenya	985	30.7	22	400	9.7	44
Morocco	971	29.9	21.8	1100	3.5	16
Oman	388	2.5	3.3	600	2.1	64
Somalia	1538	8.8	1.4	1600	8.4	600
South Africa	1154	43.3	23.6	700	7.1	30

(Source: Adapted from P A Davies School of Engineering University of Warwick)

(a) Write down the name of the country that had the biggest population.

(1)

.....

(b) Write down the name of the country that had the largest amount of renewable freshwater resource per person

(1)

.....

(c) Write down the names of the **two** countries that **could** supply all of their shortfall in freshwater from wave powered desalination.

(2)

..... and .....

(Total for Question 8 = 4 marks)

9 The time taken, in seconds, to answer each of 46 telephone calls to a call centre were recorded.

This information is summarised in the frequency table below.

Number of seconds	**Frequency**
$0 < t \leqslant 20$	7
$20 < t \leqslant 40$	17
$40 < t \leqslant 60$	14
$60 < t \leqslant 80$	5
$80 < t \leqslant 100$	2
$100 < t \leqslant 120$	1

(a) Complete the cumulative frequency table below for these data.

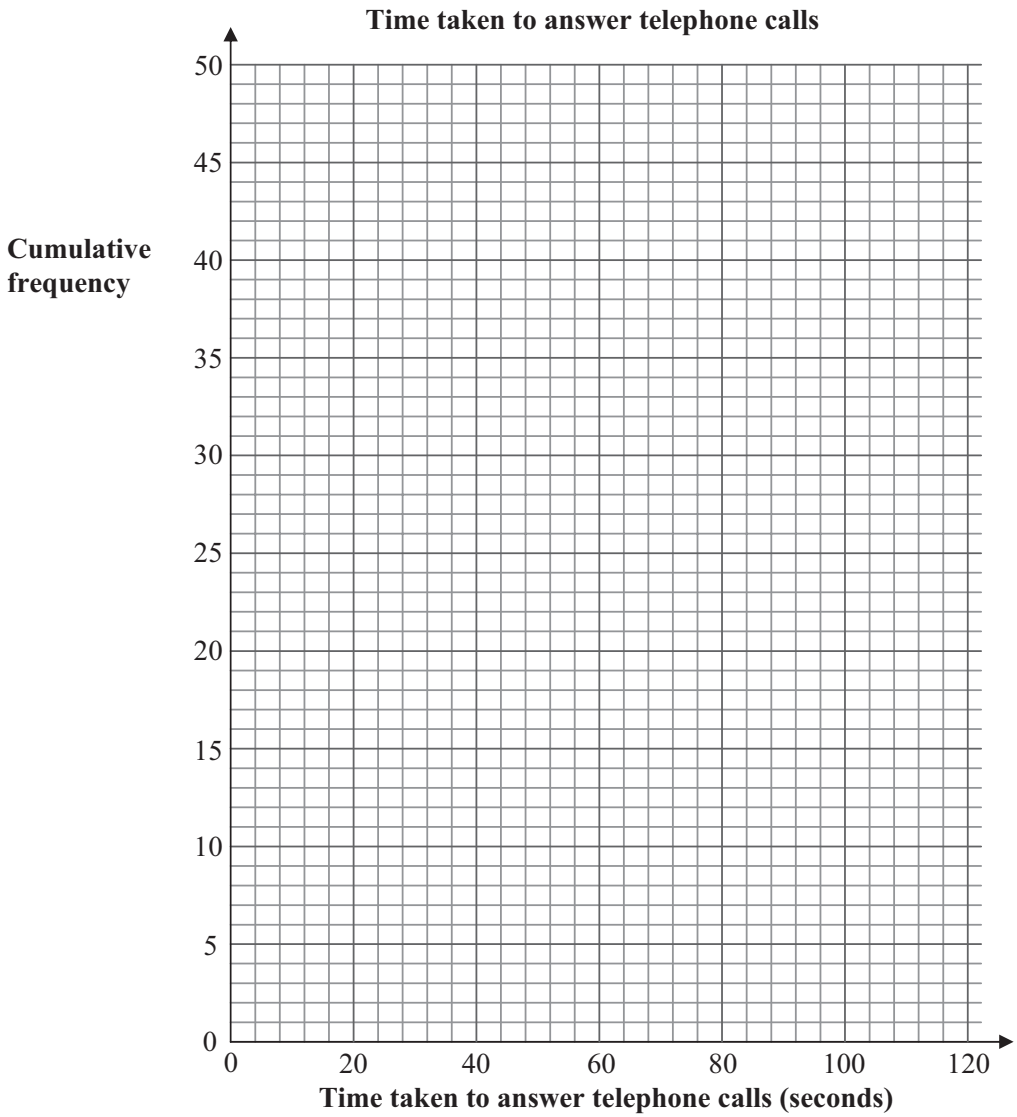
(2)

Number of seconds	**Cumulative frequency**
$0 < t \leqslant 20$	7
$0 < t \leqslant 40$	
$0 < t \leqslant 60$	
$0 < t \leqslant 80$	
$0 < t \leqslant 100$	
$0 < t \leqslant 120$	

 © Edexcel Limited 2008

(b) On the grid, draw a cumulative frequency diagram for these data.

(3)



(c) Use your cumulative frequency diagram to find an estimate for the median number of seconds taken to answer these telephone calls.

(2)

.....

The call centre aims to answer telephone calls within 30 seconds.

(d) Is the centre achieving this aim?
Give a reason for your answer.

(2)

.....

.....

.....

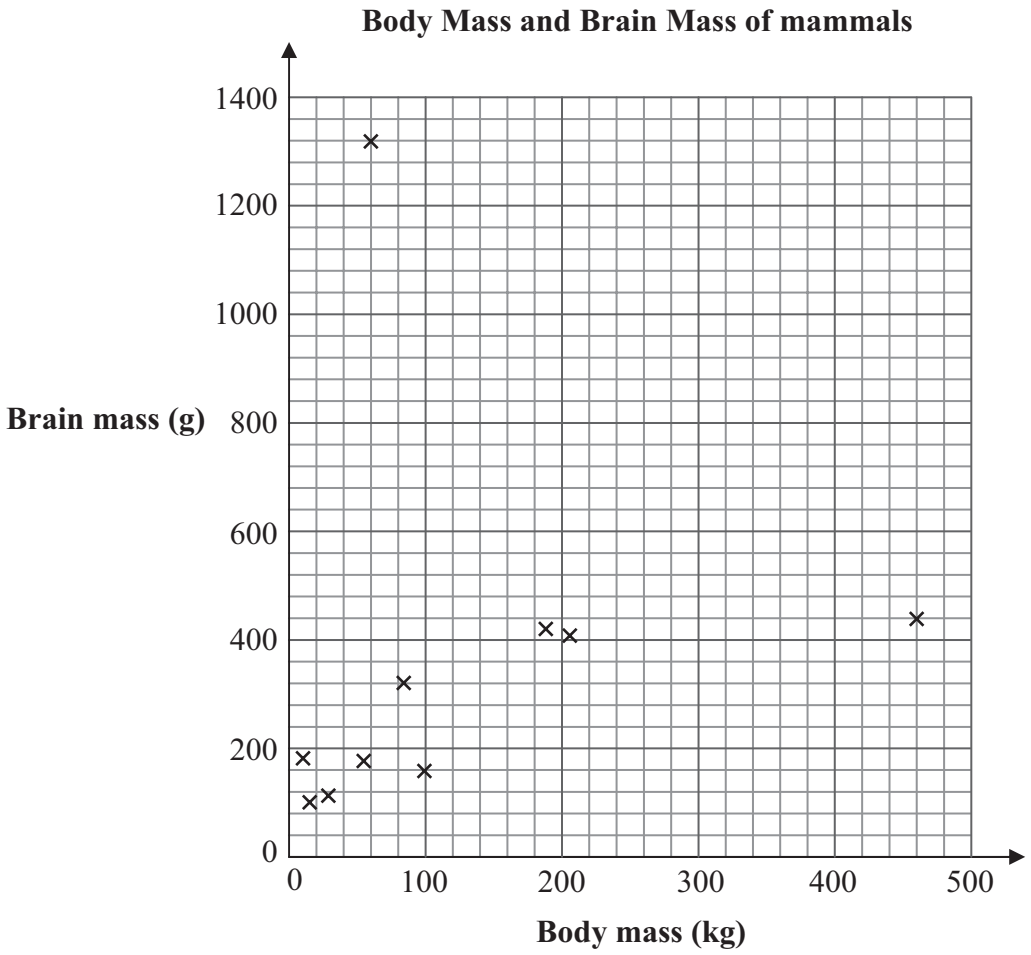
(Total for Question 9 = 9 marks)

10 The table gives information about the average body mass and average brain mass of ten mammals.

Large mammal	Average body mass (kg)	Average brain mass (g)
Baboon	10.55	179.5
Cow	460	440
Deer	14.83	98.2
Donkey	187.1	419
Goat	27.66	115
Gorilla	207	406
Jaguar	100	157
Human	62	1320
Seal	85	325
Sheep	55.5	175

(Source 'Sleep in Mammals: Ecological and Constitutional Correlates',1976)

These data are shown on the scatter diagram.



 © Edexcel Limited 2008

Horses have an average body mass of 520 kg and an average brain mass of 640 g.

(a) Plot this new information on the scatter diagram. **(1)**

(b) Write down the name of the mammal that does not fit the overall pattern. **(1)**

.....

(c) Ignoring the mammal in part (b),

(i) draw, by eye, a line of best fit on the diagram. **(1)**

(ii) For these data, describe and interpret the correlation. **(3)**

.....

.....

.....

..... kg

A mammal has an average body mass of 600 kg.

(d) (i) Estimate the average brain mass of the mammal. **(1)**

..... g

(ii) Give a statistical reason why your estimate might be unreliable. **(1)**

.....

.....

.....

(Total for Question 10 = 8 marks)

11 The local council is planning to build a new swimming pool.

The councillors want to get the views of the local people.

Councillor Green wants to take a census of the population of the town.

(a) (i) Give **one** advantage of taking a census. **(1)**

.....

(ii) Give **one** disadvantage of taking a census. **(1)**

.....

Councillor Smith suggests taking a sample from the people who attend the local sports centre.

(b) Explain why this would not be a representative sample. **(1)**

.....

.....

.....

.....

Councillor Singh suggests taking a simple random sample of 100 people.

(c) Describe, in detail, how the council could take a simple random sample. **(2)**

.....

.....

.....

.....

 © Edexcel Limited 2008

The council decided to use either **face-to-face interviews** or a **questionnaire**.

(d) Write down the advice you would give to the council, explaining:

- the advantages and disadvantages of these two methods of data collection
- the main features of a good question to use when collecting data.

(4)

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(Total for Question 11 = 9 marks)

 © Edexcel Limited 2008

12 In a series of air races, the pilots are awarded points based on their final position in each of ten races.

Paul is awarded 8, 9 or 10 points in each of his first nine races.

In his final race he makes a mistake and is awarded 0 points.

He wants to work out his average number of points.

(a) Write down **one** advantage to him of using the median. **(1)**

.....

.....

The range of Paul's points is 10

The range is not a fair measure of spread for his points.

(b) (i) Write down another measure of spread he could use. **(1)**

.....

(ii) Write down **one** advantage this new measure has that the range does not have. **(1)**

.....

(Total for Question 12 = 3 marks)

 © Edexcel Limited 2008

13 The cost of some groceries in June 2007 was £80
The cost of the same groceries in July 2007 was £81.60
June 2007 is the base month.

Month	**June 2007**	**July 2007**
Cost of groceries (£)	80	81.60
Index number	100	

(a) Work out the index number for the cost of the groceries in July 2007.

(2)

.....

The Index number for the cost of the same groceries in August 2007 is 103

(b) Work out the cost of the groceries in August 2007.

(2)

.....

The index number for September 2007 was 102.5

(c) Comment on how the price of the groceries changed between August 2007 and September 2007.

(1)

.....

.....

(Total for Question 13 = 5 marks)

14 Mary collects data on some people's arm measurements.

She measures the length from their shoulder to their fingertip.
She also measures the length from their elbow to their fingertip.

She suggests the following hypothesis,

$$\text{Arm length ratio} = \frac{\text{Length from shoulder to fingertip}}{\text{Length from elbow to fingertip}} = 1.6$$

Mary did this calculation for each of 15 males and 15 females.

Her results are summarised below.

Arm length ratio

	Lower quartile	**Median**	**Upper quartile**
Males	1.60	1.68	1.74
Females	1.58	1.63	1.72

(a) What conclusion could you draw about Mary's hypothesis from these data?
Give a reason for your answer.

(2)

.....

.....

.....

(b) What could Mary do to improve the quality of her investigation?

(1)

.....

.....

.....

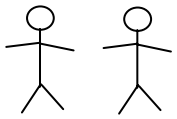
(Total for Question 14 = 3 marks)

TOTAL FOR PAPER = 80 MARKS

END

 © Edexcel Limited 2008

Foundation Tier

Question Number	Working	Answers	Notes/Comments	Mark
1(a)		Two completed stick men, i.e. 		1

Question Number	Working	Answers	Notes/Comments	Mark
1(b)		Saturday	B1	1

Question Number	Working	Answers	Notes/Comments	Mark
1(c)		90	B1	1

Question Number	Working	Answers	Notes/Comments	Mark
2(a)		59, 20, 49	**M1** Any one correct **A1** All correct in correct place on the table.	2

Question Number	Working	Answers	Notes/Comments	Mark
2(b)(i)		$\frac{30}{79}$	**B1** Accept 0.38 or 38%	1

Question Number	Working	Answers	Notes/Comments	Mark
2(b)(ii)		$\frac{12}{79}$	**B1** Accept 0.15 or 15%	1

Question Number	Working	Answers	Notes/Comments	Mark
3(a)		Correct Bars 2006 4 councillors 2007 3 councillors	 **B1** **B1**	2

Question Number	Working	Answers	Notes/Comments	Mark
3(b)		Liberal	B1	1

Question Number	Working	Answers	Notes/Comments	Mark
3(c)		Independent and Conservatives	**B1** Both needed in any order	1

Question Number	Working	Answers	Notes/Comments	Mark
3(d)	9 + 3	12	**M1** for method. Accept 10 + 4 for M1 **A1** CAO receives 2 marks	2

Question Number	Working	Answers	Notes/Comments	Mark
3(e)		**ANY TWO FROM:** • Liberals continued to have the most seats • Not much change • Independents had the least seats after both elections • Independents/ Conservatives both lost 1 seat • Labour and Liberals both won a seat	**B1 for 1st correct answer** **B1 for 2nd correct answer** Other equivalent wording acceptable Accept any other correct comment	2

Question Number	Working	Answers	Notes/Comments	Mark
4(a)(i)		201 180	**B1** **B1** Two correct moving averages: can be entered into incorrect place in table.	2

Question Number	Working	Answers	Notes/Comments	Mark
4(a)(ii)		Points plotted at (2007 1.5, 201) (2007 2.5, 180)	**B1** **B1**	2

 © Edexcel Limited 2008

Question Number	Working	Answers	Notes/Comments	Mark
4(b)		2	**B1** Accept 'quarter 2 of 2005'	1

Question Number	Working	Answers	Notes/Comments	Mark
4(c)		Reasonable trend line by eye through moving averages	**B1**	1

Question Number	Working	Answers	Notes/Comments	Mark
4(d)		Increasing or rising or equivalent	**B1**	1

Question Number	Working	Answers	Notes/Comments	Mark
5(a)(i)		B	**B1**	1

Question Number	Working	Answers	Notes/Comments	Mark
5(a)(ii)		A	**B1**	1

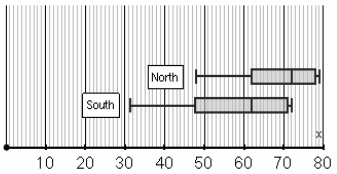
Question Number	Working	Answers	Notes/Comments	Mark
5(b)		Car or motorbike	**B1**	1

Question Number	Working	Answers	Notes/Comments	Mark
5(c)		**ANY TWO FROM:** A larger %/ proportion of those taking part in Sarah's survey Work from home **OR:** Use Public Transport The same %/ proportion Walk **OR:** Cycle A smaller %/ proportion of those taking part in Sarah's survey use a car or motorbike.	**B1** Answer without % or proportion, e.g. Reference to more gets B1 only. Reverses acceptable **B1**	2

Question Number	Working	Answers	Notes/Comments	Mark
6(a)		3 | 1 9 4 | 5 5 | 6 6 | 1 2 9 7 | 1 1 1 2	**B1** Effort to fill in numbers for first 3 stems **B1** For ordering leaves **B1** Completely correct diagram	**3**

Question Number	Working	Answers	Notes/Comments	Mark
6(b)		62	**B1**	**1**

Question Number	Working	Answers	Notes/Comments	Mark
6(c)		41	**B1**	**1**

Question Number	Working	Answers	Notes/Comments	Mark
6(d)			**B1** for Box with whiskers marked. **B1** for Median, ft **B1** quartiles and end points correct, ft	**3**

Question Number	Working	Answers	Notes/Comments	Mark
6(e)		**EITHER:** People in N appear to live longer than those in S **OR:** N larger Median than S **PLUS:** There is a greater range/ spread/IQR in the life expectancy of those who live in the S **OR:** Negative skew	 **B1** for a comparison concerning central tendency or comparative size. **B1** for a comparison of spread or a correct comment on skew Reverses acceptable.	**2**

 © Edexcel Limited 2008

Question Number	Working	Answers	Notes/Comments	Mark
7(a)		Tally or frequency chart	**B1** Chart with list of car categories and space for data collection, e.g. tally or frequency	**1**

Question Number	Working	Answers	Notes/Comments	Mark
7(b)		**ANY ONE OF:** It's only on one day in a week It's only for one hour It should be at different times of the day It's not a randomly selected time	**B1** Do not accept sampling from a different area as the question asks about the park. Other equivalent wording acceptable	**1**

Question Number	Working	Answers	Notes/Comments	Mark
8(a)		South Africa	**B1**	**1**

Question Number	Working	Answers	Notes/Comments	Mark
8(b)		Somalia	**B1**	**1**

Question Number	Working	Answers	Notes/Comments	Mark
8(c)		Antigua Somalia	**B1** **B1**	**2**

Question Number	Working	Answers	Notes/Comments	Mark
9(a)		Number of Seconds / Frequency ..20 / 7 ..40 / **24** ..60 / **38** ..80 / **43** ..100 / **45** ..120 / **46**	**B1** For at least 3 figures correct **B1** For all 5 numbers correct	**2**

Question Number	Working	Answers	Notes/Comments	Mark
9(b)		Points accurately plotted (within 1 small square) Line or curve through points including (0, 0)	**M1** **M1** **A1**	**3**

Question Number	Working	Answers	Notes/Comments	Mark
9(c)		38	**M1** line drawn on CF curve **A1** for 37.5 – 39.0 Correct answer only M1 A1	**2**

Question Number	Working	Answers	Notes/Comments	Mark
9(d)		No **PLUS:** A sensible reason such as: Only 16 calls were answered in less than 30 seconds. **OR:** Usually takes longer. **OR:** Less than half of the calls are answered in 30 seconds.	**B1** **B1**	**2**

Question Number	Working	Answers	Notes/Comments	Mark
10(a)		Cross at (520, 640)	**B1** Allow ½ square either way.	**1**

Question Number	Working	Answers	Notes/Comments	Mark
10(b)		Human	**B1** Correct answer only.	**1**

Question Number	Working	Answers	Notes/Comments	Mark
10(c)(i)		Line passing between points (0, 40) and (0, 160)at one end and between (500, 560) and (500, 680) at the other	**B1**	**1**

 © Edexcel Limited 2008

Question Number	Working	Answers	Notes/Comments	Mark
10(c)(ii)		**Positive** Correlation **PLUS:** As body mass increases so does brain mass. **OR:** As body mass decreases so does brain mass	**B1** Anything that suggests **positive** correlation. **B1** Second mark is for a correct in context answer. Wording may vary.	**2**

Question Number	Working	Answers	Notes/Comments	Mark
10(d)(i)		660 g to 800 g	**B1** Accept any figure in this range.	**1**

Question Number	Working	Answers	Notes/Comments	Mark
10(d)(ii)		**ANY ONE OF:** Extrapolation **OR:** A line drawn by eye is not reliable. oe	**B1** For any statement that suggests it would be outside the range of given data. **OR**: For any statement that suggests a poor line.	**1**

Question Number	Working	Answers	Notes/Comments	Mark
11(a)(i)		All the population give their views.	**B1** For anything that suggests everyone used	**1**

Question Number	Working	Answers	Notes/Comments	Mark
11(a)(ii)		**ANY ONE FROM:** Expensive **OR:** Time consuming **OR:** Difficult to do	**B1** The words used may vary but the meaning must be clear	**1**

Question Number	Working	Answers	Notes/Comments	Mark
11(b)		**ANY ONE FROM:** Biased **OR:** These people already take part in sports.	**B1** For anything that suggests this would be biased or unrepresentative of the population	1

Question Number	Working	Answers	Notes/Comments	Mark
11(c)		**EITHER:** Use a sampling frame **OR:** Use a spreadsheet of all the population. **PLUS:** Pick 100 using random sampling	**B1** A method for showing every member has an equal chance of being chosen is required. **B1** **Other acceptable answers** B1 Number the whole population B1 Use random generator/choose those assigned these numbers	2

 © Edexcel Limited 2008

Question Number	Working	Answers	Notes/Comments	Mark
11(d)		Answer should form two written paragraphs covering the points shown below. **1. Interview v face to face : any 3 of following** • An interviewer can explain questions but this is not possible with a questionnaire. • Responses easier with face to face. • Costs different. • Questionnaire less embarrassing. • Interviewer can influence answers. • Face to face could be more biased. **2. Suitable question: any 2 of following** Questions should be • clear, • brief, • free from bias and • not leading.	**B1 B1 B1** B1 for each point covered. The style of presentation should be appropriate. The information should be clear and coherent. Statistical vocabulary should be used where appropriate. If this is not the case then the mark for that point is not given. **B 1** The style of presentation should be appropriate. The information should be clear and coherent. Statistical vocabulary should be used where appropriate. If this is not the case then the mark is not given	**4**

Question Number	Working	Answers	Notes/Comments	Mark
12(a)		Unaffected by extremes	**B1** Accept: it could be easier to calculate.	1

Question Number	Working	Answers	Notes/Comments	Mark
12(b)(i) (ii)		**EITHER:** i) Interquartile range **PLUS** ii) Unaffected by extremes **OR** ii) Standard deviation **PLUS** ii) Uses all the data	**B1** **B1** These answers are linked so if i) is IQR then second mark only given if Unaffected by extremes written. Ditto other pair	1 1

Question Number	Working	Answers	Notes/Comments	Mark
13(a)	$\frac{81.60}{80}$ x 100	102	**M1** for $\frac{81.60}{80}$ or 1.02 **A1** for correct answer Correct answer no working M1A1	2

Question Number	Working	Answers	Notes/Comments	Mark
13(b)	80 + (3% of 80) **OR** 80 x $\frac{103}{100}$	£82.40	**M1** for working out 3% of 80 or seeing 2.4 **A1** for correct answer. Correct answer no working M1A1	2

 © Edexcel Limited 2008

Question Number	Working	Answers	Notes/Comments	Mark
13(c)		The price has gone down	**B1** Accept anything that suggests a decrease.	1

Question Number	Working	Answers	Notes/Comments	Mark
14(a)		**EITHER:** The hypothesis seems to be true. **PLUS** 50 % of the males and 50% of the females are within 0.14 of 1.6 **OR:** The hypothesis seems not true. **PLUS** 50% of the males are 1.68 or more and 50% of the females are 1.63 or more.	**B1** **B1** This is an open ended question. Accept either: The hypothesis is true/correct with a sensible reason **OR**: The hypothesis is untrue/incorrect with a sensible reason. The second mark is dependent on the first mark	2

Question Number	Working	Answers	Notes/Comments	Mark
14(b)		Take a larger sample	**B1** Accept anything that suggests more people are used. Accept as an alternative answer: The hypothesis would be better if it separated males and females. **PLUS:** The data suggests that analysing them separately would be better	1

 © Edexcel Limited 2008

Write your name here

Surname	Other names

Centre Number | Candidate Number

Edexcel GCSE

Statistics

Unit 1 – Paper 1H

Higher Tier

Sample Assessment Material

Time: 2 hours

Paper Reference

5ST1H/01

Ruler graduated in centimetres and millimetres, protractor, compasses, pen, HB pencil, eraser, electronic 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** the questions.
- Answer the questions in the spaces provided – *there may be more space than you need.*

Information

- The total mark for this paper is 100.
- The marks for **each** question are shown in brackets – *use this as a guide as to how much time to spend on each question.*
- Questions labelled with an **asterisk** (*) are ones where the quality of your written communication will be assessed. – *you should take particular care with your spelling, punctuation and grammar, as well as the clarity of expression, on these questions.*

Advice

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

Turn over ▶

N35542A

©2008 Edexcel Limited.

3/1

N 3 5 5 4 2 A 0 1 1 6

edexcel advancing learning, changing lives

GCSE Statistics

Higher Tier Formulae

You must not write on this page.

Anything you write on this page will gain NO credit.

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.

Variance $= \dfrac{\sum (x-\bar{x})^2}{n}$

Standard deviation (set of numbers) $\sqrt{\left[\dfrac{\sum x^2}{n} - \left(\dfrac{\sum x}{n}\right)^2\right]}$

or $\sqrt{\left[\dfrac{\sum (x-\bar{x})^2}{n}\right]}$

where $\bar{x}$ is the mean set of values.

Standard deviation (discrete frequency distribution) $\sqrt{\left[\dfrac{\sum fx^2}{\sum f} - \left(\dfrac{\sum fx}{\sum f}\right)^2\right]}$

or $\sqrt{\left[\dfrac{\sum f(x-\bar{x})^2}{\sum f}\right]}$

Spearman's Rank Correlation Coefficient $1 - \dfrac{6\sum d^2}{n(n^2-1)}$

 © Edexcel Limited 2008

Answer ALL the questions.

You must write down all stages of your working.

1 The table gives information about the water resources and populations of seven countries in 2000.

Country	Renewable freshwater resource (m^3/yr per person)	Population (millions)	Shortfall in freshwater (km^3/yr)	Ocean coastline (approx.) km	Freshwater potential from wave powered desalination	
					km^3/yr	% of shortfall
Antigua	800	0.07	0.06	40	0.2	333
Barbados	307	0.27	0.38	40	0.2	53
Kenya	985	30.7	22	400	9.7	44
Morocco	971	29.9	21.8	1100	3.5	16
Oman	388	2.5	3.3	600	2.1	64
Somalia	1538	8.8	1.4	1600	8.4	600
South Africa	1154	43.3	23.6	700	7.1	30

(Source: Adapted from P A Davies School of Engineering University of Warwick)

(a) Write down the name of the country that had the biggest population.

(1)

.....

(b) Write down the name of the country that had the largest amount of renewable freshwater resource per person

(1)

.....

(c) Write down the names of the **two** countries that **could** supply all of their shortfall in freshwater from wave powered desalination.

(2)

..... and .....

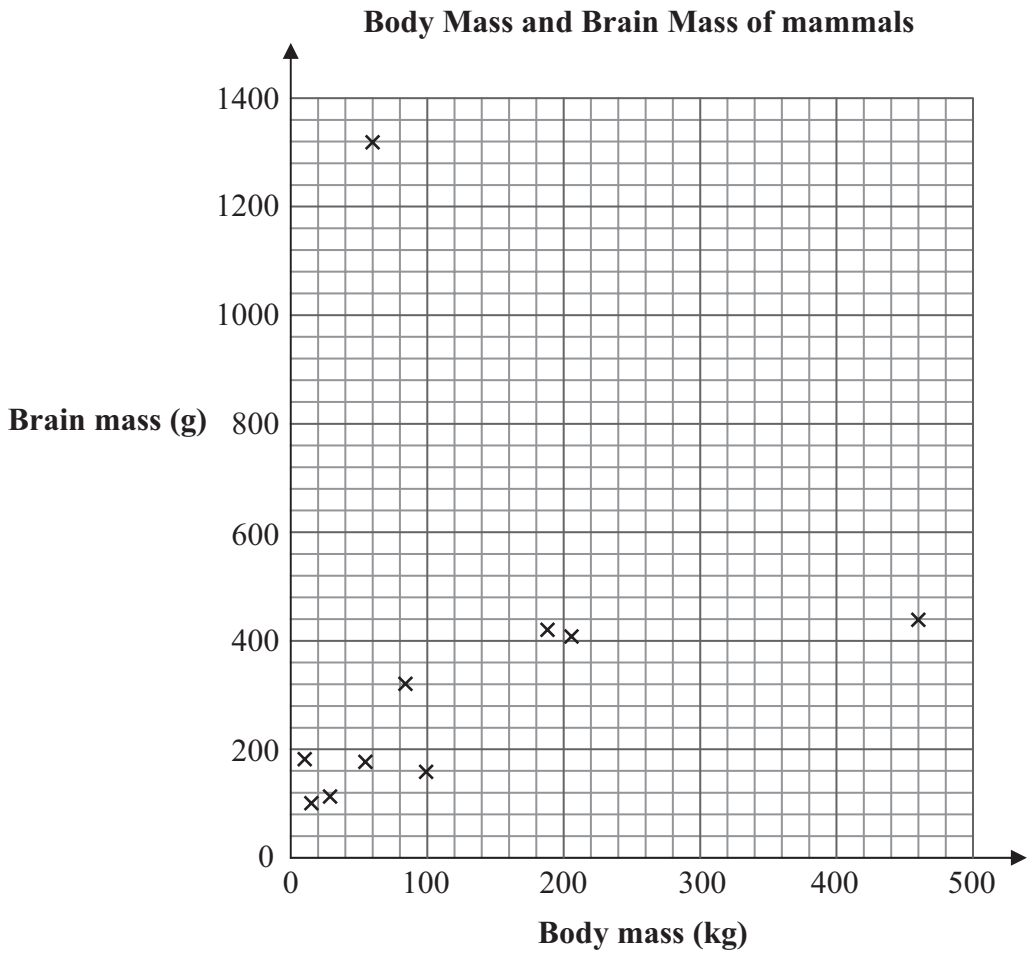
(Total for Question 1 = 4 marks)

2 The table gives information about the average body mass and average brain mass of ten mammals.

Large mammal	Average body mass (kg)	Average brain mass (g)
Baboon	10.55	179.5
Cow	460	440
Deer	14.83	98.2
Donkey	187.1	419
Goat	27.66	115
Gorilla	207	406
Jaguar	100	157
Human	62	1320
Seal	85	325
Sheep	55.5	175

(Source 'Sleep in Mammals: Ecological and Constitutional Correlates',1976)

These data are shown on the scatter diagram.



 © Edexcel Limited 2008

Horses have an average body mass of 520 kg and an average brain mass of 640 g.

(a) Plot this new information on the scatter diagram. **(1)**

(b) Write down the name of the mammal that does not fit the overall pattern. **(1)**

.....

(c) Ignoring the mammal in part (b),

(i) draw, by eye, a line of best fit on the diagram. **(1)**

(ii) For these data, describe and interpret the correlation. **(2)**

.....

.....

.....

A pig has an average brain mass of 400 g.

(d) Using your line of best fit estimate the average body mass of a pig. **(1)**

..... kg

A mammal has an average body mass of 600 kg.

(e) (i) Estimate the average brain mass of the mammal. **(1)**

..... g

(ii) Give a statistical reason why your estimate might be unreliable. **(1)**

.....

.....

.....

(Total for Question 2 = 8 marks)

3 The local council is planning to build a new swimming pool.

The councillors want to get the views of the local people.

Councillor Green wants to take a census of the population of the town.

(a) (i) Give **one** advantage of taking a census. **(1)**

.....

(ii) Give **one** disadvantage of taking a census. **(1)**

.....

Councillor Smith suggests taking a sample from the people who attend the local sports centre.

(b) Explain why this would not be a representative sample. **(1)**

.....

.....

.....

.....

Councillor Singh suggests taking a simple random sample of 100 people.

(c) Describe, in detail, how the council could take a simple random sample. **(2)**

.....

.....

.....

.....

 © Edexcel Limited 2008

The council decided to use either **face-to-face interviews** or a **questionnaire**.

*(d) Write down the advice you would give to the council, explaining:

- the advantages and disadvantages of these two methods of data collection
- the main features of a good question to use when collecting data.

(4)

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(Total for Question 3 = 9 marks)

4 In a series of air races, the pilots are awarded points based on their final position in each of ten races.

Paul is awarded 8, 9 or 10 points in each of his first nine races.

In his final race he makes a mistake and is awarded 0 points.

He wants to work out his average number of points.

(a) Write down **one** advantage to him of using the median. **(1)**

.....

.....

The range of Paul's points is 10

The range is not a fair measure of spread for his points.

(b) (i) Write down another measure of spread he could use. **(1)**

.....

(ii) Write down **one** advantage this new measure has that the range does not have. **(1)**

.....

(Total for Question 4 = 3 marks)

 © Edexcel Limited 2008

5 Some people were asked which one of three newspapers they prefer.

The results, by gender, are shown in the table.

	Males	**Females**	**Total**
The Sun	45	34	
The Daily Telegraph	11	10	
The Times	10	7	
Total			

(a) Complete the table. **(2)**

(b) Write down the **two** variables used in the table. **(2)**

.....

.....

One of these people is picked at random.

(c) (i) Write down the probability that this person is a female who prefers The Times. **(1)**

.....

(ii) Given that the person picked is a male, write down the probability that he prefers The Sun. **(2)**

.....

(Total for Question 5 = 7 marks)

6. The speeds of 100 cars passing a certain point on a motorway were recorded.
The table gives some information about these speeds.

Speed (miles per hour)	$40 < s \leqslant 60$	$60 < s \leqslant 65$	$65 < s \leqslant 70$	$70 < s \leqslant 75$	$75 < s \leqslant 80$	$80 < s \leqslant 100$
Frequency	15	12	17	20	17	19

The speed limit on this motorway is 70 miles per hour.

(a) Work out how many cars were exceeding the speed limit.

(2)

.....

(b) State whether there were more cars exceeding the speed limit or more cars within the speed limit. Give a reason for your answer.

(1)

.....

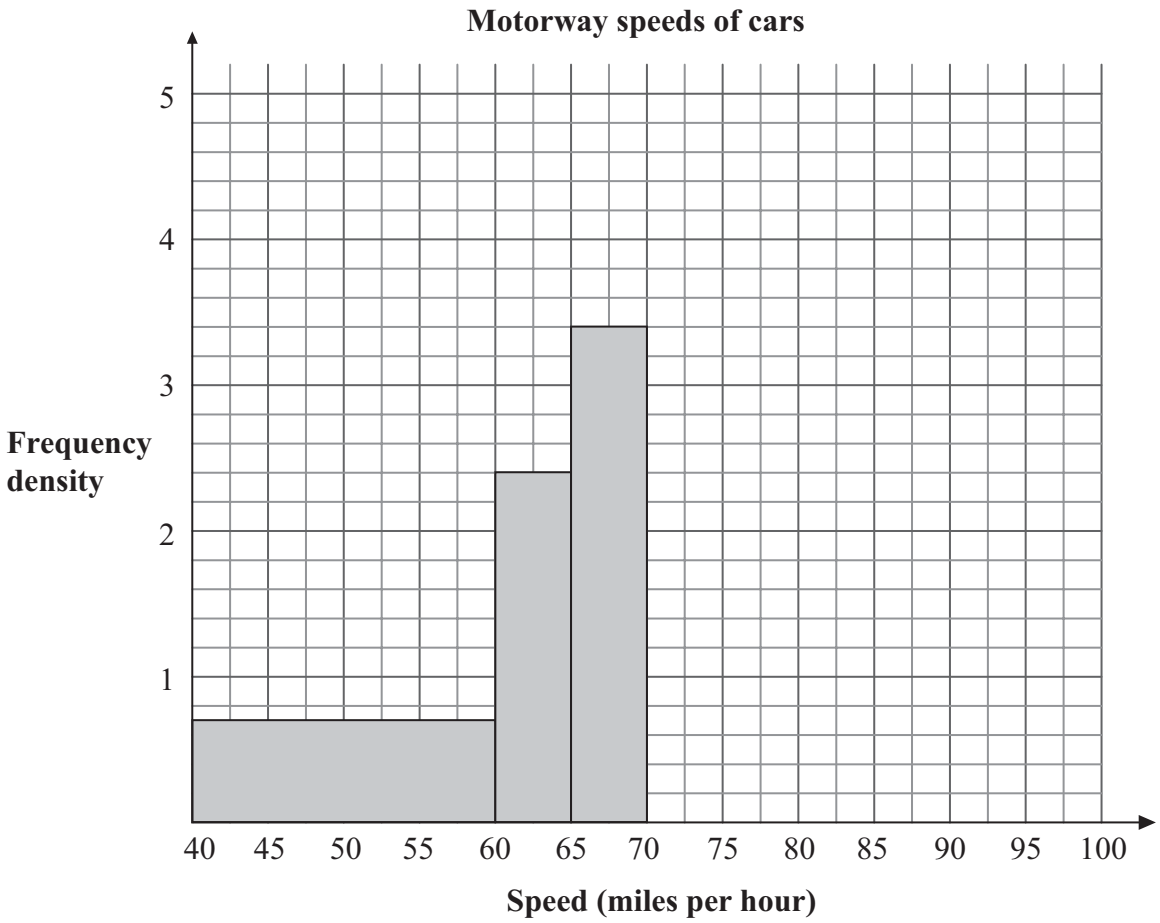
.....

.....

.....

 © Edexcel Limited 2008

Some of the data from the table are represented on the histogram.



(Source: Adapted from Dept for Transport)

(c) Use the table to complete the histogram.

(3)

(d) Estimate how many cars were travelling at between 60 and 68 miles per hour.
Show your working.

(3)

.....

(Total for Question 6 = 9 marks)

7 A council wants to find out views on recycling.

The council wants to know if there is a difference in the support for recycling between

males and females,

people aged 25 and under, and people aged over 25

(a) Write down **two** suitable hypotheses for this investigation. **(2)**

Hypothesis 1 .....

.....

Hypothesis 2 .....

.....

(b) Use the best word from the list to complete the sentences below.

qualitative **primary** **quantitative** **secondary** **ratio**

(i) Age group is ..... data. **(1)**

(ii) Gender is ..... data. **(1)**

Here are two sampling frames the council could use.

A A list of names in a telephone directory.

B A list of residents in one street.

(c) Suggest a possible disadvantage of using each of these sampling frames. **(2)**

Sampling frame **A** .....

.....

.....

Sampling frame **B** .....

.....

.....

 © Edexcel Limited 2008

The council will take a systematic sample of 1000 people from a sampling frame of 100 000 people.

(d) Describe how the council will do this.

(3)

.....

.....

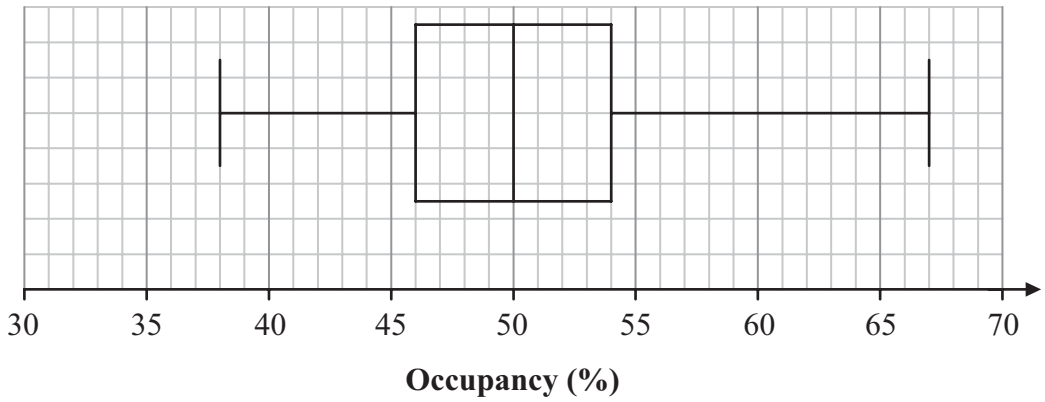
.....

(Total for Question 7 = 9 marks)

 © Edexcel Limited 2008

8 The box and whisker diagram gives information about the room occupancy rates (%) of 15 resorts in Yorkshire during 1999.

Room occupancy rates in Yorkshire in 1999 (%)



(Source: Yorkshire Tourist Board)

(a) Write down the median room occupancy rate. **(1)**

..... %

(b) Work out the interquartile range. **(2)**

..... %

(c) A normal distribution is suggested as a model for the room occupancy rates in Yorkshire in 1999. Give a reason which supports this choice. **(1)**

.....

.....

.....

 © Edexcel Limited 2008

During 2006 the room occupancy rates (%) for the same 15 resorts were

50 51 61 65 52 54 55 66 55 55 69 55 69 55 56

(Source: Yorkshire Tourist Board)

(d) For the 2006 room occupancy rates,

(i) find the median,

(2)

..... %

(ii) find the interquartile range,

(3)

..... %

(iii) describe the skew.

(1)

.....

(e) Compare the distribution of the 1999 room occupancy rates with the distribution of the 2006 room occupancy rates.

(2)

.....

.....

.....

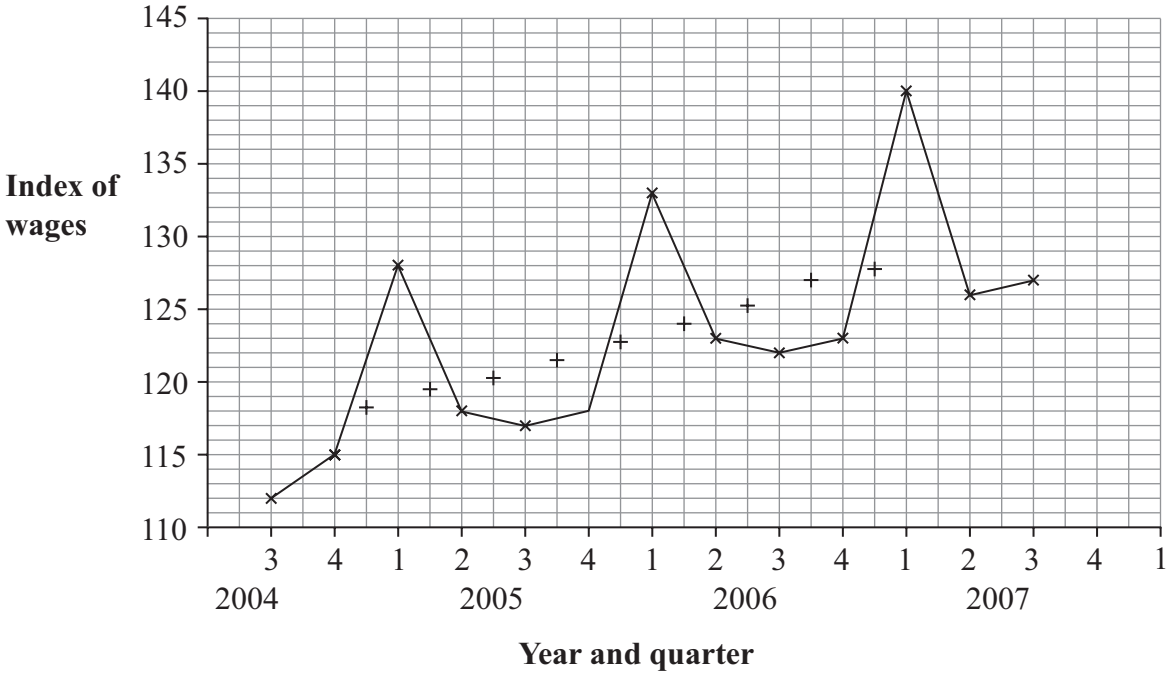
(Total for Question 8 = 12 marks)

9 The table shows the quarterly index of wages from the third quarter of 2004 to the third quarter of 2007.

The year 2000 is the base year for the index numbers.

	Quarter			
Year	**1**	**2**	**3**	**4**
2004	-	-	112	115
2005	128	118	117	118
2006	133	123	122	123
2007	140	126	127	-

(Source: statistics.gov.uk)



These data are plotted as a time series on the graph above.
The 4-point moving averages, except for the last one, are also plotted.

(a) Calculate the last 4-point moving average and plot it on the graph.

(2)

.....

(b) Draw a trend line on the graph.

(1)

 © Edexcel Limited 2008

(c) Write down the quarter in which wages were at their highest each year. **(1)**

.....

The average wage was £18 000 in 2000.

(d) Taking 2000 as the base year, use the information in the table to work out the average wage in quarter 1 of 2005. **(2)**

.....

(e) Work out the average seasonal effect for quarter 1.

Give your answer to the nearest whole number. Show your working. **(2)**

.....

(f) Predict the value of the index of wages for quarter 1 of 2008. **(3)**

.....

(Total for Question 9 = 11 marks)

10 JLD Engineering is supplied a part from two different companies.

70% of these parts are supplied by Company A.

30% of these parts are supplied by Company C.

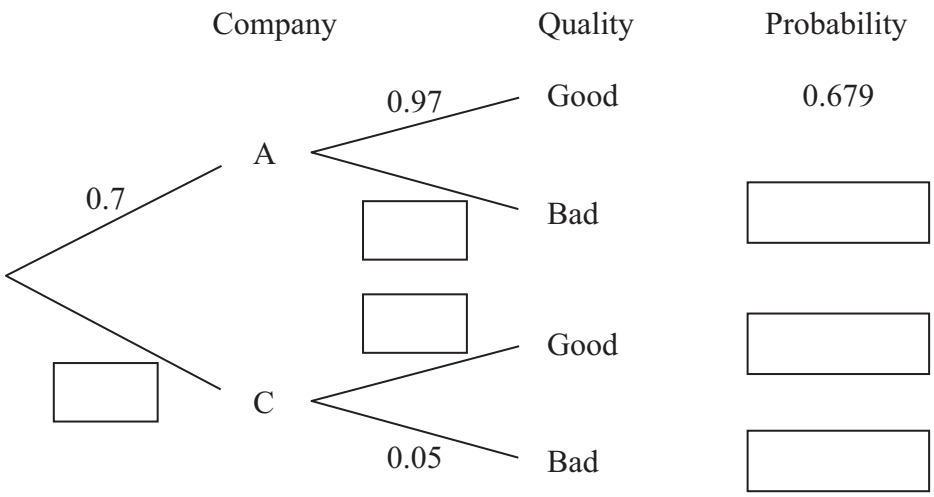
The quality of the parts supplied by Company A is 97% good and 3% bad.

The quality of the parts supplied by Company C is 95% good and 5% bad.

These data are represented on a probability tree diagram.

(a) Complete the tree diagram.

(2)



A part is selected at random.

(b) Work out the probability that the quality of the part is good.

.....

(2)

 © Edexcel Limited 2008

A part is selected at random.

(c) Given that the part is good, work out the probability that it was supplied by Company A.

(2)

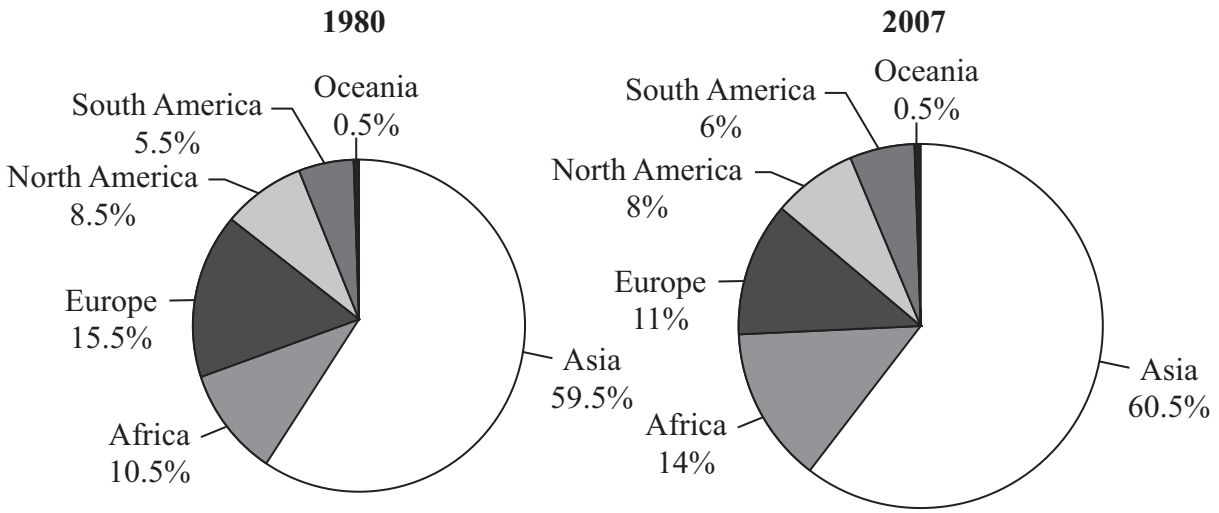
.....

(Total for Question 10 = 6 marks)

 © Edexcel Limited 2008

11 The comparative pie charts give information about the numbers of people living in six continents of the world.

The distribution of people in the world



(a) What has happened to the **total** number of people living in the world in 1980 compared to 2007? Explain how you can tell this from the pie charts.

(2)

.....

.....

.....

(b) Write down the name of the continent with the greatest increase in population.

(1)

.....

(c) Describe how the **total** number of people living in Oceania has changed in 1980 compared to 2007. Explain your answer.

(2)

.....

.....

.....

(Total for Question 11 = 5 marks)

 © Edexcel Limited 2008

12 The table gives information about the age, and the minimum stopping distance at 40 kilometres per hour, for each of 10 cars.

Car	Age of car (months)	Stopping distance (metres)				
A	9	28.4				
B	15	29.3				
C	24	37.6				
D	30	36.2				
E	38	36.5				
F	46	35.4				
G	53	36.3				
H	60	44.1				
I	64	44.8				
J	76	47.2				

(Source: www.bized.co.uk)

(a) Work out Spearman's rank correlation coefficient for these data.

You may use the blank columns in the table to help with your calculations.

(4)

.....

(b) Describe and interpret your answer to part (a).

(2)

.....

.....

.....

.....

(Total for Question 12 = 6 marks)

13 A company making cookers finds that the amount of sheet steel it uses each week is normally distributed with a mean of 2000 tonnes and a standard deviation of 200 tonnes.

(a) Use this information to predict the maximum amount of sheet steel the company is likely to use in any given week.
Explain your reasoning.

(2)

..... tonnes

.....

.....

.....

The manager of the company decides to make a regular order for the same amount of sheet steel to be delivered every Monday.

(b) Work out a suitable amount of sheet steel to be ordered.
Give a reason for your answer.

(2)

..... tonnes

.....

.....

.....

(c) Give **one** reason why the answer to part (a) might not be suitable as an answer to part (b).

(1)

.....

.....

.....

.....

(Total for Question 13 = 5 marks)

 © Edexcel Limited 2008

14 Packets of nuts conform to the European minimum weight standard (e).

The packets in a supermarket are marked as 250 g e. This means that, on average, 1 out of every 40 packets (2.5%) weighs less than 241 g.

Assume that the weights of the packets of nuts are normally distributed, with mean 250 g.

(a) Work out an estimate of the standard deviation of the weights of the packets.

(2)

..... g

The probabilities of buying 0, 1, 2, or 3 packets of nuts that do not conform to the European minimum weight standard can be modelled by a binomial distribution.

Greta buys three packets of these nuts.

(b) Work out the probability that she buys exactly two packets that do not conform to the European minimum weight standard.

(3)

You may use $(p + q)^3 = p^3 + 3p^2q + 3pq^2 + q^3$.

.....

(c) In order to use a binomial distribution as a model, an assumption is made about the probability that the weight is less than 241 g for each of the three packets. Write down this assumption.

(1)

.....

.....

(Total for Question 14 = 6 marks)

TOTAL FOR PAPER = 100 MARKS

END

BLANK PAGE

 © Edexcel Limited 2008

Higher Tier

Question Number	Working	Answers	Notes/Comments	Mark
1(a)		South Africa	**B1**	1

Question Number	Working	Answers	Notes/Comments	Mark
1(b)		Somalia	**B1**	1

Question Number	Working	Answers	Notes/Comments	Mark
1(c)		Antigua Somalia	**B1** **B1**	2

Question Number	Working	Answers	Notes/Comments	Mark
2(a)		Cross at (520, 640)	**B1** Allow ½ square either way.	1

Question Number	Working	Answers	Notes/Comments	Mark
2(b)		Human	**B1** Correct answer only.	1

Question Number	Working	Answers	Notes/Comments	Mark
2(c)(i)		Line passing between points (0, 40) and (0, 160)at one end and between (500, 540) and (500, 680) at the other	**B1**	1

Question Number	Working	Answers	Notes/Comments	Mark
2(c)(ii)		**Positive** Correlation **PLUS:** As body mass increases so does brain mass. **OR:** As body mass decreases so does brain mass	**B1** Anything that suggests **positive** correlation. **B1** Second mark is for a correct in context answer. Wording may vary.	2

Question Number	Working	Answers	Notes/Comments	Mark
2(d)		240 kg to 340 kg	**B1** Accept any figure in this range.	1

Question Number	Working	Answers	Notes/Comments	Mark
2(e)(i)		600g to 800 g	**B1** Accept any figure in this range.	1

Question Number	Working	Answers	Notes/Comments	Mark
2(e)(ii)		**ANY ONE OF:** Extrapolation **OR:** A line drawn by eye is not reliable. Oe	**B1** For any statement that suggests it would be outside the range of given data. **OR:** For any statement that suggests a poor line.	1

Question Number	Working	Answers	Notes/Comments	Mark
3(a)(i)		All the population give their views.	**B1** For anything that suggests everyone used	1

Question Number	Working	Answers	Notes/Comments	Mark
3(a)(ii)		**ANY ONE FROM:** Expensive **OR:** Time consuming **OR:** Difficult to do	**B1** The words used may vary but the meaning must be clear	1

Question Number	Working	Answers	Notes/Comments	Mark
3(b)		**ANY ONE FROM:** Biased **OR:** These people already take part in sports.	**B1** For anything that suggests this would be biased or unrepresentative of the population	1

 © Edexcel Limited 2008

Question Number	Working	Answers	Notes/Comments	Mark
3(c)		**EITHER:** Use a sampling frame and number each member **OR:** Use a spreadsheet of all the population. **PLUS:** Describe a method of randomly selecting 100 members	**B1** A method for showing every member has an equal chance of being chosen is required. **B1** **Other acceptable answers** B1 Number the whole population B1 Use random generator/choose those assigned these numbers	**2**

Question Number	Working	Answers	Notes/Comments	Mark
3(d) **QWC (i), (ii), (iii)**		Answer should form two written paragraphs covering the points shown below. **1. Interview v face to face : any 3 of following** • An interviewer can explain questions but this is not possible with a questionnaire. • Responses easier with face to face. • Costs different. • Questionnaire less embarrassing. • Interviewer can influence answers. • Face to face could be more biased. **2. Suitable question: any 2 of following** Questions should be • clear, • brief, • free from bias and • not leading.	**B1 B1 B1** B1 for each point covered. The style of presentation should be appropriate. The information should be clear and coherent. Statistical vocabulary should be used where appropriate. If this is not the case then the mark for that point is not given. **B 1** The style of presentation should be appropriate. The information should be clear and coherent. Statistical vocabulary should be used where appropriate. If this is not the case then the mark is not given.	**4**

 © Edexcel Limited 2008

Question Number	Working	Answers	Notes/Comments	Mark
4(a)		Unaffected by extremes	**B1**	1

Question Number	Working	Answers	Notes/Comments	Mark
4(b)(i)		**EITHER:** Interquartile range **OR:** Standard Deviation (σ)	**B1**	1

Question Number	Working	Answers	Notes/Comments	Mark
4(b)(ii)		If (i) IQR then (ii) must be 'Unaffected by extremes.' If (i) is SD then (ii) must be 'Uses all the data.'	**B1** The answer to (ii) is dependent on the answer to (i)	1

Question Number	Working	Answers	Notes/Comments	Mark
5(a)	<table><tr><td>M</td><td>F</td><td>T</td></tr><tr><td>45</td><td>34</td><td>**79**</td></tr><tr><td>11</td><td>10</td><td>**21**</td></tr><tr><td>10</td><td>7</td><td>**17**</td></tr><tr><td>**66**</td><td>**51**</td><td>**117**</td></tr></table>	66 & 51 & 117 79 & 21 & 17	**B1** **B1** Figures should be in positions shown in working column	2

Question Number	Working	Answers	Notes/Comments	Mark
5(b)		Gender **OR** Sex Paper **OR** Newspaper read	**B1 Do not allow** Male/Female. **B1 Do not allow** Name of paper	2

Question Number	Working	Answers	Notes/Comments	Mark
5(c)(i)		$\frac{7}{117}$ or 0.0598	**B1** Allow follow through.	1

Question Number	Working	Answers	Notes/Comments	Mark
5(c)(ii)		$\frac{45}{66}$ OR $\frac{15}{22}$ OR 0.682	**M1** for using 66 or their 66. ft **A1** for correct answer. ft Correct answer only M1A1 Follow through allowed for both marks	**2**

Question Number	Working	Answers	Notes/Comments	Mark
6(a)	20 + 17 + 19	56	**M1** for method – as in working column **A1** for correct answer 56 with no working gets M1 A1	**2**

Question Number	Working	Answers	Notes/Comments	Mark
6(b)		More cars exceeded the limit because 44 were under and 56 over	**B1** Idea of more plus reference to the figures required for 1 mark	**1**

Question Number	Working	Answers	Notes/Comments	Mark
6(c)	FD = $\frac{frequency}{classwidth}$	Bars: 70 – 75 fd 4 75 – 80 fd 3.4 80 – 100 fd 0.95	**M1** for effort at finding frequency density **A1** for 2 correct bars **A1** for third bar correct Histogram drawn correctly no working M1 A1 A1.	**3**

 © Edexcel Limited 2008

Question Number	Working	Answers	Notes/Comments	Mark
6(d)	(5 x 2.4) + (3 x 3.4) 12 + 10.2	22.2	**M1** 5 x 2.4 **M1** for 3 x 3.4 **A1** for correct answer Accept 22 instead of 22.2 22.2 or 22 with no working gets M1M1A1	**3**

Question Number	Working	Answers	Notes/Comments	Mark
7(a)		**EITHER:** Females are not so interested in recycling as males. **OR:** Males are not so interested in recycling as females. Oe ***PLUS***: **EITHER:** Over 25s are not so interested in recycling as those under 25. **OR:** Under 25s are not so interested in recycling as those over 25. oe	**B1** These could be expressed in different words. The reverses will be common. There must be one Hypothesis on age and one on gender. Accept: "the level of support is the same for both genders" as one answer **B1** Accept: "the level of support is the same for both age groups" as one answer	**2**

Question Number	Working	Answers	Notes/Comments	Mark
7(b)(i)		Quantitative	**B1**	**1**

Question Number	Working	Answers	Notes/Comments	Mark
7(b)(ii)		Qualitative	**B1**	**1**

Question Number	Working	Answers	Notes/Comments	Mark
7(c)		**A: ANY ONE OF:** Not everyone who recycles will have a telephone. **OR:** Many people who recycle will not have their name in the telephone book **OR:** The telephone book probably covers more than the council area **B: ANY ONE OF:** The street may not have a representative sample of people. **OR:** The street may be very short. **OR:** The street may be mainly shops	**B1** These could be expressed in different words. There may be other correct answers. Any sensible answer accepted **B1** These could be expressed in different words. Note: The answer 'Biased' is not enough. The reason why needs to be given.	**2**

Question Number	Working	Answers	Notes/Comments	Mark
7(d)		Numbered list of 100 000 people Find a random number between 1 and 100 Start at that number and take every 100th person.	**B1** Accept 'number the people' **B1** Accept use random tables or similar plus a reference to 1 to 100 **B1**	**3**

Question Number	Working	Answers	Notes/Comments	Mark
8(a)		50	**B1**	**1**

 © Edexcel Limited 2008

Question Number	Working	Answers	Notes/Comments	Mark
8(b)	54 – 46	8	**M1** for working (54 – 46) **A1** for correct answer Correct answer with no working M1 A1	**2**

Question Number	Working	Answers	Notes/Comments	Mark
8(c)		**EITHER:** The box plot is symmetrical about the median **OR:** $Q_3 - Q_2 = Q_2 - Q_1$	**B1** Accept 'quite symmetrical' Accept ' the median is in the middle of the distribution' Wording may vary.	**1**

Question Number	Working	Answers	Notes/Comments	Mark
8(d)(i)	50,51,52,54,55, 55, 55,55,55,56,61, 65,66,69,69	55	**M1** for ordering. **A1** for answer Correct answer with no working M1 A1	**2**

Question Number	Working	Answers	Notes/Comments	Mark
8(d)(ii)	Q_1 4th value = 54 Q_3 12th value = 65	11	**M1** for Q1 (54) **M1** for Q3 (65) **A1** for correct answer. Correct answer with no working M1 M1 A1 SC: If Order used in i) incorrect and used for ii) then the M1 M1 can be awarded on follow through.	**3**

Question Number	Working	Answers	Notes/Comments	Mark
8(d)(iii)		Positive skew	**B1** Do not accept anything other than positive or +ve	**1**

Question Number	Working	Answers	Notes/Comments	Mark
8(e)		2006 has the higher median. **OR:** 1999 has the lower median. **PLUS:** 2006 has the bigger IQR. **OR:** 1999 has the smaller IQR.	**B1** Accept equivalent wording **B1** Alternative acceptable: **PLUS** 2006 has the smaller range **OR** 1999 has the bigger range **B1** Alternative acceptable 2006 is symmetric **or** 1999 is positively skewed (Note: one mark is given for spread and one for median)	**2**

Question Number	Working	Answers	Notes/Comments	Mark
9(a)		129 Plot point on graph at (2007-1.5, 129)	**B1** for correct answer. **B1** Follow through for second B mark	**2**

Question Number	Working	Answers	Notes/Comments	Mark
9(b)		Reasonable trend line by eye through moving averages	**B1**	**1**

Question Number	Working	Answers	Notes/Comments	Mark
9(c)		Quarter 1	**B1** Accept 1	**1**

 © Edexcel Limited 2008

Question Number	Working	Answers	Notes/Comments	Mark
9(d)	$\frac{128}{100}$ x £18000) **OR:** (28% of £18000) + £18000	£23040	**M1** for working that shows in some way that the increase is 28% **A1** Correct answer with no working M1 A1	**2**

Question Number	Working	Answers	Notes/Comments	Mark
9(e)	(10 + 10 + 12) divide by 3	11 $10\frac{2}{3}$ **OR** $10.\dot{6}$	**M1** for reading 3 reasonable values from their line of best fit and dividing by 3 **A1** Follow through from their readings AWRT 11	**2**

Question Number	Working	Answers	Notes/Comments	Mark
9(f)	Trend from graph (133) Index is (133) + $10\frac{2}{3}$	$143\frac{2}{3}$ (Answer likely to range from $143\frac{1}{6}$ to $144\frac{1}{6}$ dependent on graph readings)	**M1** for reading the trend between 132.5 and 133.5 **M1** for their trend + variation as working. **A1** for answer. Accept their trend plus answer to (e) No Working but answer from $143\frac{1}{6}$ to $144\frac{1}{6}$ M1 M1 A1	**3**

Question Number	Working	Answers	Notes/Comments	Mark
10(a)		0.3, 0.03, 0.95 0.021, 0.285, 0.015	**B1** **B1** Follow through answers from tree diagram are acceptable for second B1	**2**

Question Number	Working	Answers	Notes/Comments	Mark
10(b)	0.679 + 0.285	0.964	**M1** for 0.679 + their total added. Ft from (a) **A1** for answer. Ft allowed Correct answer with no working M1 A1	**2**

Question Number	Working	Answers	Notes/Comments	Mark
10(c)	$\frac{0.679}{0.285+0.679}$	0.704 (anything which rounds to 0.7)	**M1** for the working fraction. 0.679 in correct places plus their value from (a) in correct place. **A1** for answer. Correct answer no working gets 2 marks This part allows follow through for both marks.	**2**

Question Number	Working	Answers	Notes/Comments	Mark
11(a)		They have increased. **PLUS:** 2006 circle has a larger area than the 1999 circle	**B1** Words such as risen/gone up may be used. **B1** A reference to the increase in circle **area** is required for second B1	**2**

 © Edexcel Limited 2008

Question Number	Working	Answers	Notes/Comments	Mark
11(b)		Asia	**B1**	**1**

Question Number	Working	Answers	Notes/Comments	Mark
11(c)		It has increased. **PLUS:** The percentages are the same but the numbers have risen because for 2007 it is 0.5% of a bigger number	**B1** **B1** Words may vary. Reference to the increase and reference to 0.5% of a bigger population is required for both marks.	**2**

Question Number	Working	Answers	Notes/Comments	Mark
12(a) **QWC** **(ii), (iii)**	Correct Ranking: 1,2,3,4,5,6,7,8,9, 10 1,2,7,4,6,3,5,8,9, 10 d^2 = 30 (M1) 1 - $\frac{6 \times 30}{10(100-1)}$	 awrt 0.82	**M1** for correct ranking Reverse ranking accepted. 10,9,8,7,6,5,4,3,2, 1 10,9,4,7,5,8,6,3,2, 1 for M1 **M1** for d^2 follow through from their ranking Incorrect ranking can still get this mark **M1** Correct working fraction only. (No ft) **A1** for correct answer. Correct answer no working 4 marks.	**4**

Question Number	Working	Answers	Notes/Comments	Mark
12(b)		Positive correlation **PLUS** The older the car the longer the stopping distance. oe	**B1** Follow through from (a) only if r between -1 and +1 **B1** Accept: The younger the car the shorter the stopping distance oe Ft again if (a) in limits above	**2**

Question Number	Working	Answers	Notes/Comments	Mark
13(a)		2600 **PLUS** **EITHER:** This is the mean +3 standard deviations. **OR:** Almost 100% of the time this will be enough	**B1** Accept: 2500 **B1** This is the mean + 2.5 standard deviations **Or** Almost 100% of the time this will be enough	**2**

 © Edexcel Limited 2008

Question Number	Working	Answers	Notes/Comments	Mark
13(b)		2000 tonnes **PLUS:** Over time the amount used would be the mean weekly total. or equivalent	**B1** **B1** **OR**: Accept answers in range 2000 – 2200 B1 **PLUS:** This would allow for the weeks when they go over the mean but in the long run the company would have spare sheet steel. or equivalent B1	**2**

Question Number	Working	Answers	Notes/Comments	Mark
13(c)		**EITHER:** Over time there would be an excess of 600 tonnes weekly. **OR**: This would be wasteful/economically unacceptable. Or equivalent	**B1** Wording may vary.	**1**

Question Number	Working	Answers	Notes/Comments	Mark
14(a)	$\frac{250-241}{2}$	4.5	**M1** for 250 - 241 divided by any whole number under 4 **A1** for answer Alternative acceptable: **M1** for $\frac{250-241}{1.96}$ **A1** for answer for 4.6	**2**

Question Number	Working	Answers	Notes/Comments	Mark
14(b) QWC (iii)	$q = \frac{39}{40}$ $p = \frac{1}{40}$ $3 \times (\frac{1}{40})^2 \times \frac{39}{40}$	0.00183	**M1** for $\frac{39}{40}$ in working **M1** for attempt at correct term **A1** for answer. Accept fractions or decimals.	**3**

Question Number	Working	Answers	Notes/Comments	Mark
14(c)		The probability of success is constant, oe	**B1** Could be expressed in different words	**1**

 © Edexcel Limited 2008

Edexcel GCSE

Statistics

Theme: Estimation
Student Booklet and Teachers' Notes

Sample Controlled Assessment Material

Paper Reference
5ST02/01

Turn over ▶

N35590A

©2008 Edexcel Limited.

N 3 5 5 9 0 A 0 1 1 6

edexcel
advancing learning, changing lives

Formulae sheets

Edexcel GCSE Statistics

Formulae Sheet

Foundation Tier

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.

 © Edexcel Limited 2008

Edexcel GCSE Statistics

Formulae Sheet

Higher Tier

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.

Variance $= \dfrac{\sum (x-\bar{x})^2}{n}$

Standard deviation (set of numbers) $\sqrt{\left[\dfrac{\sum x^2}{n} - \left(\dfrac{\sum x}{n}\right)^2\right]}$

or $\sqrt{\left[\dfrac{\sum (x-\bar{x})^2}{n}\right]}$

where $\bar{x}$ is the mean set of values.

Standard deviation (discrete frequency distribution) $\sqrt{\left[\dfrac{\sum fx^2}{\sum f} - \left(\dfrac{\sum fx}{\sum f}\right)^2\right]}$

or $\sqrt{\left[\dfrac{\sum f(x-\bar{x})^2}{\sum f}\right]}$

Spearman's Rank Correlation Coefficient $1 - \dfrac{6\sum d^2}{n(n^2-1)}$

 © Edexcel Limited 2008

THEME: Estimation

In life a valuable skill is the ability to estimate. For example, you may be asked to estimate how long a task is going to take or how much it is going to cost.

An estimate is an approximate idea of length, weight, time etc that is given without actually taking measurements, but is based upon your previous experience of such things.

All sorts of factors may affect our ability to estimate.

For example:

- **Age** of the person doing the estimation
- **Size** of the object being estimated
- **Orientation** of the object being estimated
- **Colour** of the object being estimated

In this task you will need to **collect your own data** so that you can investigate the process of estimating.

The project must be done in 3 stages.

The **first stage** must be done under formal supervision with your teacher present and you may **not** take this work out of the classroom until it has been completed. In this stage you will plan your investigation.

In the **second stage** you may work under informal supervision outside of the classroom. In this stage you collect your data, draw diagrams and do calculations.

In the **third stage** you must again work under formal supervision with your teacher present. You will write up your investigation. This will involve analysing and interpreting your data, drawing conclusions and evaluating your work.

Quality of written communication will be taken into account throughout the marking of your investigation. Quality of written communication includes clarity of expression, the structure and presentation of ideas and grammar, punctuation and spelling.

 © Edexcel Limited 2008

Stage 1: Planning

You will need to plan how you will carry out your investigation.

Think about estimating.
What factors affect our ability to estimate?
What are you going to investigate? Discuss this with your teacher.

Your teacher will give you some ideas but the highest marks are for originality – that means your own ideas.

Develop a question or hypothesis to investigate.

You can have just one question but it would be better if you considered related questions or hypotheses.

Remember

- The questions you decide to investigate will influence the sample you collect.
- You need to think about what you expect the answer to your questions or hypotheses to be and why you expect these answers.

Data collection

Think about the data you want to collect.

- Where will your data come from?
- How reliable is your source?
- How big a sample will you need? Discuss a suitable size with your teacher. The sample size will depend on the questions you are trying to answer but try to collect at least 30 pieces of data for each variable you are investigating.
- How are you going to collect the data? What sampling method is to be used? You can work with others to share the data collection but you must write down exactly what you did and what they were asked to do.
- How are you going to record the data you collect?
- What will you do about outliers or anomalies? (Pieces of data that appear to be out of place or incorrect).

Diagrams

- Which diagrams would give the most information about your questions and what are the reasons for your choice?
- You should always choose the most useful diagram for your question – there is no need to draw several diagrams to represent the same data.

Calculations

- What calculations will best summarise your data and answer the questions you have asked?
- Remember to consider outliers.
- You should **not** do calculations that do not answer your initial questions. You will get no extra credit and will be wasting your time.

Now write up your plan and give it in to your teacher.

Your plan should state the questions you are going to investigate and information on your plans for data collection, diagrams and calculations. You should write down reasons for your choices of all these.

Your teacher will discuss your plan with you. If you wish, at this point, you will be able to adapt your plan, taking into account your teacher's comments.

 © Edexcel Limited 2008

Stage 2: Collecting, processing and representing data

Now you have completed the first part of your work you need to do some work alone.

You need to:

- Collect your data as you have planned in stage 1, using any collection sheets you have designed, and keep a record of the raw data.
- Write down any problems that arose.
- Put your data into the form you need.
- Do your calculations and your diagrams neatly, by hand or using IT.
- Collect together everything you need to complete the writing of your report.

 © Edexcel Limited 2008

Stage 3: Interpreting and evaluating data

The final writing of the rest of your report should be done in the classroom.

You will have done most of your diagrams and calculations already, but now everything needs to be put together so that it makes a complete report.

You need to:

- Place your **Plan** at the beginning of your report.
- Under the title **Data Collection**, present the data you collected in tables.
- Write down any extra information about the data that was not in your plan. This might include problems that you found and how you resolved them.
- Under the title **Analysis**, present your diagrams and calculations in a form which helps you to interpret them.
- Draw everything together in a **Conclusion** that relates to the questions and hypotheses that you posed in your plan.
- **Evaluate** your work. Did you get the results you expected? Suggest improvements and discuss limitations.
- Create an **Appendix** for your raw data.
- Complete the Student Record Sheet and sign the authentication form and put them at the front of your report.
- Number the pages and fix your report together with a treasury tag or a piece of string.
- Hand in your report to your teacher.

 © Edexcel Limited 2008

Edexcel GCSE Statistics

Controlled Assessment

Estimation
Teachers' Notes

Estimation is a practical investigation suitable for GCSE Statistics students across both tiers of entry. The project must be undertaken by collecting primary data.

It is important that teachers advise students to include statistical techniques based on the content appropriate to the GCSE tier at which they are entered. The use of ICT to produce diagrams and to find statistical functions should be encouraged. The marks are awarded for the students' explanation of how they have used these tools and why they have used particular techniques.

The controlled assessment must be done in three stages.

1: Planning

Formal supervision (Classroom/IT suite)

10 marks
(1 – 2 hours)

This will take place under formal supervision. The students should be supervised at all times, but this is not examination conditions and students should be encouraged to discuss their planning with the teacher. Whilst word processors and spreadsheets may be useful NO INTERNET ACCESS should be available.

The teacher could conduct a whole class discussion of theme/s and possible data that might be collected. The class may discuss possible problems and how these could be overcome.

Suggestions as to how the task can be approached:

1. Discuss with students:
   - What sorts of things can be estimated?
   - How can we investigate people's ability to estimate?
   - What might affect a person's ability to estimate?
2. Discuss ideas for hypotheses or questions:
   - The hypotheses the students choose to investigate will determine the data to be collected
   - More able students will investigate the interrelationship of more than one variable, e.g. age and gender
   - Are older people more able to make accurate estimates?
   - Are males better at estimating than females?
   - Do physical factors such as size/orientation/colour affect our ability to make accurate estimates?

 © Edexcel Limited 2008

3. Discuss the issues relating to data collection:

- What is the population?
- What data is to be collected? (and how will it be recorded?)
- Which sampling method will be used? (and why?)
- How large should the sample be?
- Who will collect the data and how will bias be minimised?

Students may plan to share the collection of data in groups, but they will need to give an explicit account of how they will do this and the part they will play in their planning.

Whilst the use of techniques must be an integral part of the project, students require some guidance as to what is expected at each level. It should be pointed out that making a **choice** of techniques with reasons is essential if the student is to achieve marks for planning. Students should not attempt to demonstrate all known techniques. Some statistical techniques are more sophisticated than others. Students should be advised that the complexity of their investigation will be limited if the techniques they choose are not at a sufficiently complex level.

Students should write a plan and this should be reviewed and commented on before they are allowed to continue with their task. All feedback to the student should be written on the student record form.

Students should be advised to investigate questions or hypotheses that enable them to use appropriate diagrams/calculations at their ability level.

The guidance given by the teacher is essential at this stage to ensure each student is tackling a project which will enable them to demonstrate their statistical ability. Choosing too trivial a project will not allow a student to access the higher marks.

Here are some suggested questions and related possible analyses. There are many other possibilities. Questions should suit the student's ability level.

Are males better at estimating than females?
Collect data for males and females; draw comparative box plots; compare the distributions.

Are estimates normally distributed?
Collect data; draw a histogram; try to fit a normal distribution; discuss.

Are older people more able to make accurate estimates?
Collect data for older people and younger people; draw diagrams; determine spread; discuss.

Do people who under/over estimate small things also under/over estimate large things?
Collect data; draw a scatter diagram; calculate; discuss correlation.

Are people better at estimating discrete variables than they are at estimating continuous variables?
Collect data; draw suitable diagrams; calculate suitable measures of average and dispersion; compare and discuss.

Occasionally a very able student wants to go beyond the specification and this can be encouraged, however, the highest marks may be awarded for the creative application of techniques within the specification.

 © Edexcel Limited 2008

2: Collecting, processing and representing data

Informal supervision

Students may use the internet and bring in work from outside the classroom but the teacher will need to monitor the work in the classroom to ensure it relates to the initial plan.

The initial plan may be developed and adapted and this is to be encouraged but only following discussion with the teacher. This is to ensure that the project is the work of the individual student.

Collecting data
8 marks
(approx 2 – 3 hours)

Students **should** collect primary data for this task.
Data may be collected individually or in groups.
The method of collecting the data must be fully described and any problems should be fully discussed.

Processing, analysing and representing data
12 marks
(approx 2 – 4 hours)

The use of ICT should be encouraged, and there is no need to show evidence of calculations or hand drawn diagrams as these skills are assessed in the external assessments.

It is essential that sensible scales are used and labelled.

 © Edexcel Limited 2008

3: Interpret and evaluate

Formal supervision (Classroom/ICT suite)

10 marks
(up to 2 hours)

Students may will bring in work from outside the classroom but the teacher will need to monitor the work in the classroom to ensure it relates to the initial plan. The initial plan may have been developed and adapted and this is to be encouraged, but only following discussion with the teacher. This is to ensure that the project is the work of the individual student.

Students bring all the work they have compiled to the classroom and put together the whole report.

They will need to produce in final form:

- The written up hypothesis/es with their planned strategy.
- The data collection discussion describing exactly what they did.
- Raw data should be in an appendix with summary tables in the main body of the report.
- Discussion of problems and limitations of the data should be discussed.
- Reasons for choice should accompany the analysis. Diagrams and calculations should be interpreted.
- There should be an interpretation in the context of the whole investigation relating back to the original questions and hypotheses.
- Conclusions linking together the strands of their enquiry.
- Evaluation of the work discussing any limitations.
- A completed Student Record form should accompany the work and the student should sign the authentication form.
- Work can be handwritten or word processed.

When completed the work is handed in. This must be at the end of the highly controlled time – there must be no extension for finishing at home.

The teacher will need to complete the student record form, with their marks and any other information, which should be attached to the front of the project.

 © Edexcel Limited 2008

Quality of Written Communication (QWC)

Controlled assessments in GCSE Statistics provide opportunities across ability ranges to assess Quality of Written Communication (QWC).

Each stage of the controlled assessment offers opportunities where:

i) Learners ensure text is legible,
spelling, punctuation and grammar are accurate,
that meaning is clear

ii) learners select and use a form and style of writing appropriate
to purpose and to complex subject matter

iii) learners organise information clearly and coherently, using
specialist vocabulary when appropriate.

Throughout the controlled assessment there are opportunties to assess strand (i) of QWC, and to ensure learners are using clear and legible writing and checking their punctuation, grammar and spelling through their work.

Strand (ii) of QWC can be assessed differently through the stages. In Stages 2 and 3, for example, the learner may find it is more appropriate to use diagrammatic or tabular representations of information, with short sentances linking the different considerations within that stage, where in stages 1 and 4 a more extended writing structure would be appropriate.

Strand (iii) of QWC can be assessed throughout the stages of the controlled assessment, where learners can be expected to express themselves logically and clearly, using appropriate technical language and notation, and the overall investigation should be organised clearly and coherently.

Where students are required to provide clear aims, strategies, lines of enquiries, explanations, justifications or reasons for their work are all opportunities where QWC (ii) and QWC (iii) could be assessed.

In the assessment criteria, specific indicators where QWC (ii) and QWC (iii) can be assessed have been included. QWC (i) can be assessed throughout the controlled assessment.

 © Edexcel Limited 2008

BLANK PAGE

 © Edexcel Limited 2008

BLANK PAGE

 © Edexcel Limited 2008

BLANK PAGE

 © Edexcel Limited 2008

Edexcel GCSE

Statistics

Theme: Human Body
Student Booklet and Teachers' Notes

Sample Controlled Assessment Material

Paper Reference
5ST02/01

Turn over ▶

N35592A

©2008 Edexcel Limited.

N 3 5 5 9 2 A

edexcel
advancing learning, changing lives

Edexcel GCSE Statistics

Formulae Sheet

Foundation Tier

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.

 © Edexcel Limited 2008

Edexcel GCSE Statistics

Formulae Sheet

Higher Tier

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.

Variance $= \dfrac{\sum (x-\bar{x})^2}{n}$

Standard deviation (set of numbers) $\sqrt{\left[\dfrac{\sum x^2}{n} - \left(\dfrac{\sum x}{n}\right)^2\right]}$

or $\sqrt{\left[\dfrac{\sum (x-\bar{x})^2}{n}\right]}$

where $\bar{x}$ is the mean set of values.

Standard deviation (discrete frequency distribution) $\sqrt{\left[\dfrac{\sum fx^2}{\sum f} - \left(\dfrac{\sum fx}{\sum f}\right)^2\right]}$

or $\sqrt{\left[\dfrac{\sum f(x-\bar{x})^2}{\sum f}\right]}$

Spearman's Rank Correlation Coefficient $1 - \dfrac{6\sum d^2}{n(n^2-1)}$

 © Edexcel Limited 2008

THEME: Human body

When you look round a room full of people you will see that the human body has many shapes and sizes. We are very diverse animals. You can investigate the human body statistically in many different ways.

Some of the variables you could investigate are:

- eye/hair colours,
- handedness,
- lengths of arms/legs,
- height,
- weight,
- body mass index.

There are many questions that can be investigated statistically:

Do you expect there to be more people with brown eyes than any other colour?
Do you expect people with long arms to be tall?
Do you expect tall people to weigh more that short people?
Do you expect people to cluster around an average height or are heights distributed evenly?

There are many other questions you could ask and variables you could investigate.

The project must be done in three stages.

The **first stage** must be done under formal supervision with your teacher present and you will **not** be able to take this work out of the classroom until it has been completed. In this stage you will plan your investigation.

In the **second stage** you may work under informal supervision outside the classroom. In this stage you will collect your data, draw diagrams and do calculations.

In the **third stage** you must again work under formal supervision with your teacher present. You will write up your investigation. This will involve analysing and interpreting your data, drawing conclusions and evaluating your work.

Quality of written communication will be taken into account throughout the marking of your investigation. Quality of written communication includes clarity of expression, the structure and presentation of ideas and grammar, punctuation and spelling.

 © Edexcel Limited 2008

Stage 1: Planning

You will need to plan how you will carry out your investigation.

What are you going to investigate? Discuss this with your teacher.

Your teacher will give you some ideas but the highest marks are for originality – that means your own ideas.

Develop a question or hypothesis to investigate.

You can have just one question, but it would be better if you considered related questions or hypotheses.

Remember

- The questions you decide to investigate will influence the sample you collect.
- You need to think about what you expect the answers to your questions or hypotheses to be and why you expect these answers.

Data Collection

Think about the data you want to collect.

- Where will your data come from?
- How reliable is your source?
- How big a sample will you need? Discuss a suitable size with your teacher. The sample size will depend on the questions you are trying to answer but usually you should collect at least 30 pieces of data for each variable you are investigating.
- How are you going to collect the data? What sampling method is to be used? You can work with others to share the data collection but you must write down exactly what you did and what they were asked to do.
- How are you going to record the data you collect?
- What will you do about outliers or anomalies? (Pieces of data that appear to be out of place or incorrect).

Diagrams

- What are the best diagrams to draw to represent your data and what are the reasons for your choice?
- You should always choose the best diagram – there is no need to draw several diagrams to represent the same data.

Calculations

- What calculations will best summarise your data and answer the questions you have asked?
- Remember to consider outliers.
- You should **not** do calculations that do not answer your initial questions. You will get no extra credit and will be wasting your time.

Now write up your plan and give it in to your teacher.

Your plan should state the questions you are going to investigate and information on your plans for data collection, diagrams and calculations. You should write down reasons for your choices of all these.

Your teacher will discuss your plan with you. If you wish, at this point, you will be able to adapt your plan, taking into account your teacher's comments.

 © Edexcel Limited 2008

Stage 2: Collecting, processing and representing data

Now you have completed the first part of your work you need to do some work alone.

You need to:

- Collect your data as you have planned in stage 1, using any collection sheets you have designed, and keep a record of the raw data.
- Write down any problems that arose.
- Put your data into the form you need.
- Do your calculations and your diagrams neatly, by hand or using IT.
- Collect together everything you need to complete the writing of your report.

 © Edexcel Limited 2008

Stage 3: Interpreting and evaluating

The final writing of the rest of your report will be done in the classroom.

You will have completed most of your diagrams and calculations already, but now everything needs to be put together so that it makes a complete report.

You need to:

- Place your **Plan** at the beginning of your report.
- Under the title **Data Collection**, present the data you collected in tables.
- Write down any extra information about the data that was not in your plan. This might include problems that you found and how you resolved them.
- Under the title **Analysis**, present your diagrams and calculations in a form which helps you to interpret them.
- Draw everything together in a **Conclusion** that relates to the questions and hypotheses that you posed in your plan.
- **Evaluate** your work. Did you get the results you expected? Suggest improvements and discuss limitations.
- Create an **Appendix** for your raw data.
- Complete the Student Record Sheet and sign the authentication form and put them at the front of your report.
- Number the pages and fix your report together with a treasury tag or a piece of string.
- Hand in your report to your teacher.

 © Edexcel Limited 2008

Edexcel GCSE Statistics

Controlled Assessment

Human Body

Teachers' Notes

This task gives considerable scope for students of all abilities to produce projects. It is designed as a project that uses primary data collected in the student's school, college, academy or sports centre. Students will need guidance in choosing the questions they are going to pose. The questions will need to enable the students to use statistical techniques based on the specification content of the GCSE tier at which they are entered. They will also need advice about the collection of data.

The controlled assessment must be done in three stages.

1: Planning

Formal supervision (Classroom/IT suite)

10 marks
(1 – 2 hours)

This will take place under formal supervision. The students should be supervised at all times, but this is not examination conditions and students should be encouraged to discuss their planning with the teacher. Whilst word processors and spreadsheets may be useful NO INTERNET ACCESS should be available.

The teacher should discuss the task with the students.

Suggestions as to how the task can be approached can be discussed.

Discuss with the students:

- What sort of questions can be asked?
  Draw out of the students the possible questions they might consider. This will determine the sort of data they might collect.
- What data will be needed?
  This should include discussion of the population and how much data will be needed.
  (Generally at least 30 pieces required but at least 60 are needed for more complex work.)
- How could the data be collected?
  What data collection sheets will be needed?
  Students may collect data in groups but they will need to give an explicit account in their plan of how they will do this.
- What statistical techniques can be used?

 © Edexcel Limited 2008

Whilst the use of techniques must be an integral part of the project students require some guidance as to what is expected at each level. It should be pointed out that making a choice of techniques with reasons is essential if the student is to achieve marks for planning. Students should be encouraged to be selective. Some statistical techniques are more sophisticated than others. Students should choose techniques that suit their ability. The better students should be advised that the complexity of their investigation will be limited if their hypotheses and choice of techniques do not allow them to produce work of a reasonable degree of complexity.

Students should produce a written plan at the end of this highly controlled stage. This should be reviewed by the teacher and feedback should be given to the student before any further work is done. All feedback should be noted on the student's record form. At this point students can amend their plan in the light of the teacher's comments.

Teachers should give advice regarding the suitability of the student's choices in their feedback to the students at the end of stage 1.

Choosing too trivial a project will not allow a student to reach their potential.

Here are some suggested questions and related analyses.

- Students targeting lower grades (G – E) might consider collecting qualitative data such as eye colours or handedness. Questions such as 'What is the most common eye colour?' or 'What is the mean number of students with blue eyes in a class?' or 'Are there more males with blue eyes than females'. This would allow sampling from a single class (G/F) or sampling from several classes (F/E). Totals could be worked out and bar charts or pie charts could be drawn. There is scope for averages and simple comparisons.

- Students targeting middle grades (E – B) might look at associations between measurements. Such questions as 'Are boys taller than girls?' or 'Is there a relationship between heights and weights?' or 'Does your ability to pick up pegs depend upon your hand-span?'.

  This level of work could involve two body measurements or one body measurements allied with some other measure e.g. hand span and ability to pick up sticks. This should enable stratified sampling to be done.

  Any of the following could be used – scatter diagrams, cumulative frequencies, box plots, averages, quartiles, comparisons, Spearman's rank correlation etc

- Students targeting higher grades (B – A*) need to produce more individual work. They need to develop further some of the ideas for the mid-ability students to include such questions as 'Can height be modelled by a normal distribution?' or 'How does the spread of heights for girls compare with the spread of heights for boys?

  These students need to pose questions that give them the scope to cover the techniques suggested for middle ability student plus the use of statistical techniques such as histograms, standard deviation or normality.

 © Edexcel Limited 2008

2: Collecting, processing and representing data

Informal Supervision

Students in their own time, under informal supervision, need to complete:

Collecting data
(8 marks)
(approx 2 – 3 hours)

This should must have been planned in stage 1 and data collection sheets prepared. Students may want to collect in groups, but their contribution should be recorded in their plan. A group could visit a class and collect all the data needed for the various hypotheses to be tackled – heights, weights, eye colours, waist size, hair colours, experiment results. A spreadsheet could be created.

Processing, analysing and representing data
(12 marks)
(approx 2 – 4 hours)

This means that the students need to draw relevant diagrams and do relevant calculations. The use of ICT is encouraged but not essential. There is no need to show evidence of calculations or hand drawn diagrams as these skills are assessed in the external assessments. It is important that ICT is used sensibly and diagrams have sensible scales and labels.

 © Edexcel Limited 2008

3: Interpret and evaluate

Formal supervision (Classroom/ICT suite)

10 marks
(up to 2 hours)

Students may bring in work from outside the classroom but the teacher will need to monitor the work in the classroom to ensure it relates to the initial plan. The initial plan may have been developed and adapted and this is to be encouraged, but only following discussion with the teacher. This is to ensure that the project is the work of the individual student.

Students must bring all the work they have compiled to the classroom and put together the whole report.

They will need to produce in final form:

- The written up hypothesis/es with their planned strategy.
- The data collection discussion describing exactly what they did.
- Raw data should be in an appendix with summary tables in the main body of the report.
- Discussion of problems and limitations of the data should be discussed.
- Reasons for choice should accompany the analysis. Diagrams and calculations should be interpreted.
- There should be an interpretation in the context of the whole investigation relating back to the original questions and hypotheses.
- Conclusions linking together the strands of their enquiry.
- Evaluation of the work discussing any limitations.
- A completed Student Record form should accompany the work and the student should sign the authentication form.
- Work can be handwritten or word processed.

When completed the work is handed in. This must be at the end of the highly controlled time – there must be no extension for finishing at home.

The teacher will need to complete the student record form, with their marks and any other information, which should be attached to the front of the project.

 © Edexcel Limited 2008

Quality of Written Communication (QWC)

Controlled assessments in GCSE Statistics provide opportunities across ability ranges to assess Quality of Written Communication (QWC).

Each stage of the controlled assessment offers opportunities where:

i) Learners ensure text is legible,spelling, punctuation and grammar are accurate, that meaning is clear

ii) learners select and use a form and style of writing appropriate to purpose and to complex subject matter

iii) learners organise information clearly and coherently, using specialist vocabulary when appropriate.

Throughout the controlled assessment there are opportunties to assess strand (i) of QWC, and to ensure learners are using clear and legible writing and checking their punctuation, grammar and spelling through their work.

Strand (ii) of QWC can be assessed differently through the stages. In Stages 2 and 3, for example, the learner may find it is more appropriate to use diagrammatic or tabular representations of information, with short sentances linking the different considerations within that stage, where in stages 1 and 4 a more extended writing structure would be appropriate.

Strand (iii) of QWC can be assessed throughout the stages of the controlled assessment, where learners can be expected to express themselves logically and clearly, using appropriate technical language and notation, and the overall investigation should be organised clearly and coherently.

Where students are required to provide clear aims, strategies, lines of enquiries, explanations, justifications or reasons for their work are all opportunities where QWC (ii) and QWC (iii) could be assessed.

In the assessment criteria, specific indicators where QWC (ii) and QWC (iii) can be assessed have been included. QWC (i) can be assessed throughout the controlled assessment.

 © Edexcel Limited 2008

BLANK PAGE

 © Edexcel Limited 2008

BLANK PAGE

 © Edexcel Limited 2008

BLANK PAGE

 © Edexcel Limited 2008

Assessment Criteria for Controlled Assessment

Introduction

Controlled assessment is marked on a common mark scale across both tiers of entry. The maximum mark is 40, which we then convert to a mark out of 25 by a direct scaling factor for each tier of entry.

Each piece of work must be assessed under the following strand headings with the mark for each strand recorded on the *Student Record Form* in the specification.

The assessment criteria are sub-divided into four strands, these being:

1: Planning

2a Data collection

2b: Processing, analysing and representing data

3: Interpretation and discussion of results.

Strands 1 and 3 will be on a 10-mark scale, Strand 2a will be on an 8-mark scale and Strand 2b will be on a 12-mark scale.

The mark awarded in each strand must reflect the degree of difficulty and sophistication of the line of enquiry.

 © Edexcel Limited 2008

Quality of Written Communication (QWC)

Controlled assessments for the Edexcel GCSE in Statistics provide opportunities across ability ranges to assess Quality of Written Communication (QWC).

Each stage of the controlled assessment offers opportunities where:

i) student ensures text is legible
spelling, punctuation and grammar are accurate
that meaning is clear

ii) students select and use a form and style of writing appropriate to purpose and to complex subject matter

iii) students organise information clearly and coherently, using specialist vocabulary when appropriate.

Throughout the controlled assessment there are opportunties to assess strand (i) of QWC, and to ensure students are using clear and legible writing and checking their punctuation, grammar and spelling through their work.

Strand (ii) of QWC can be assessed differently through the stages. In Stages 2 and 3, for example, the student may find it is more appropriate to use diagrammatic or tabular representations of information, with short sentances linking the different observations and findings. In stages 1 and 4 a more extended writing structure would be appropriate.

Strand (iii) of QWC can be assessed throughout the stages of the controlled assessment, where students can be expected to express themselves logically and clearly, using appropriate technical language and notation, and the overall investigation should be organised clearly and coherently.

Areas where students are required to provide clear aims, strategies, lines of enquiries, explanations, justifications or reasons for their work are all opportunities where QWC (ii) and QWC (iii) could be assessed.

In the assessment criteria, specific indicators where QWC (ii) and QWC (iii) can be assessed have been included. QWC (i) can be assessed throughout the controlled assessment.

 © Edexcel Limited 2008

Planning (10 marks)

Mark	Performance descriptor
0	The student provides no evidence of an implicit plan to process or display some data.
1	The student provides evidence of an implicit plan to process or display some data.
2 **QWC (iii)**	The student gives a clear aim to process or display some data.
3	The student gives a simple aim and provides a strategy to use a simple statistical technique to process or display data.
4 **QWC (ii)**	The student chooses a simple aim and provides a strategy to use a simple statistical technique (diagram **or** calculation) to make a comparison.
5	The student chooses a simple aim and provides a strategy to use simple statistical techniques (diagrams **and** calculations) to make a comparison.
6 **QWC (ii) and QWC (iii)**	The student chooses a **more complex line of enquiry** to use statistical techniques to make a comparison. They give a clear aim and sensible **reasons** for the diagrams and calculations they will use.
7	The student chooses a more complex line of enquiry to use statistical techniques to make a comparison. They give a clear aim and justify which diagrams and calculations they use. They **identify** potential problems with the data (eg anomalies, different sized populations, scales etc).
8 **QWC (ii) and QWC (iii)**	The student plans to test hypotheses, which have been carefully specified in clear statistical terms. They give a clear aim and justify all of the diagrams and calculations they use, ensuring that diagrams are drawn so that **comparisons** can be made. They **plan** how they will deal with any potential problems with the data.
9 **QWC (ii) and QWC (iii)**	The student plans to test hypotheses, which have been carefully specified in clear statistical terms. They should consider a **number of interrelated variables** and justify their plan to use a number of different techniques. They must **plan and justify** how they will deal with any potential problems with the data.
10 **QWC (ii) and QWC (iii)**	The student plans to test a hypothesis, which has been carefully specified in clear statistical terms. They must **foresee possible problems**, which might arise and **justify their methods** for dealing with these. They should consider a number of interrelated variables and plan to use a number of different advanced techniques.

 © Edexcel Limited 2008

Collecting data (8 marks)

Mark	Performance descriptor
0	The student does not use any data.
1	The student uses some data.
2	The student **collects** some data (at least 10 items).
3 **QWC (ii)**	The student collects some data, **indicates its source and how it was collected**. The data should be shown in some way. They may use the whole population but should indicate that they are doing so. (The word census is not required at this level.)
4	The student uses a recognised sampling method and gives a brief account of how the data was collected and its source. The student collects sufficient data in two or more data sets to make comparisons. If a census is used reasons for this must be given.
5 **QWC (iii)**	The student uses a recognised sampling method and gives a detailed account of how they collected their data. They discuss the type(s) of data which may be discrete, continuous, qualitative or quantitative. Any anomalies in the data collected should be identified as they occur.
6	The student gives a detailed account of the **sampling mechanism** for their data collection and justifies the size of the sample. Any anomalies should be identified as they occur and a decision made, with **reasons**, as to whether they should be included or omitted.
7	The student justifies their choice of a particular sampling technique. Limits for outliers, set at the planning stage, should have been used. Problems in data collection which were identified at the planning stage (for example, different sized populations or samples, missing data) have been acted upon.
8 **QWC (ii) and QWC (iii)**	Reliability of the data source should be discussed with reference to source, collection, strategy and the proportion of anomalies found. Bias, how it may arise and what is being done to avoid it should be discussed. All of the techniques used for sampling and dealing with problems must be justified.

 © Edexcel Limited 2008

Processing, analysing and representing data (12 marks)

Mark	Performance descriptor
0	The student does not attempt to draw a simple diagram or perform a calculation.
1	The student attempts to draw a simple diagram or perform a calculation.
2	The student produces a simple diagram (correct labels and scales) or calculation successfully.
3	The student produces a simple correct statistical diagram or calculation.
4	The student produces simple correct statistical diagrams and calculations. These may be simply to display or summarise the data.
5 **QWC (iii)**	The student provides a diagram **or** calculation to make a simple comparison following on from their planning.
6	The student uses diagrams **and** calculations to make simple comparisons following on from their planning. At least one of the statistical techniques should be more complex than the techniques used for mark 4. The diagrams must be correct with scales and labels.
7	The student uses diagrams and calculations to make a comparison, at least one of which must be more complex.
8	The student use both diagrams and calculations which must be more complex to make comparisons.
9 **QWC (ii)**	The student justifies their use of diagrams and calculations, having ensured that diagrams enable comparisons to be made. They draw a series of diagrams and perform calculations to explore one or more variables without making connections between the variables.
10 **QWC (ii) and QWC (iii)**	The student uses diagrams and calculations to test a complex hypothesis. They draw a series of diagrams and perform calculations to explore one or more variables without making connections between the variables. The work is accurate with few errors. There is little irrelevant work present and outliers are considered if they occur.
11	The student should consider a number of interrelated variables and use a number of different techniques to explore possible connections or effects. They draw a series of diagrams and perform calculations to explore one or more variables making connections between **all** the variables. At least one of the techniques they use must be complex.
12 **QWC (ii) and QWC (iii)**	The student should consider a number of inter-related variables and plan to use a number of different techniques beyond those associated with mark 11. They link diagrams and calculations to explore possible connections, distributions or effects. The student must deal with the problems they foresaw in their plan and justify their approach.

 © Edexcel Limited 2008

Interpreting and evaluating data (10 marks)

Marks	Performance indicator
0	The student makes no comments about the data.
1	The student makes a comment about the data. For example: I collected 10 pieces of data.
2	The student makes a comment to draw a conclusion about the data. For example: The largest is...
3	The student makes a simple statistical comment about the diagram or calculation. For example: The mode is...
4	The student interprets a diagram or calculation using a simple techniques, to make a simple statistical comparison. For example: The bar charts show that the most popular drink is X. Drink Y is the least popular.
5 **QWC (iii)**	The student interprets a diagram **and** calculation to make a simple statistical comparison. In the case of multiple conclusions at least one but not all need to be correct.
6 **QWC (ii)**	The student **summarises their results** and comment upon their work. They make some simple written comparisons.
7	The student summarises and makes detailed explanations of their results with correct interpretations of statistical techniques. They correctly interpret their data and make comparisons.
8 **QWC (ii) and QWC (iii)**	The student summarises and makes detailed explanations of their results with correct interpretations of statistical techniques. They correctly interpret their data making in-depth comparisons and commenting on the effect of anomalies in their data. They evaluate their sampling or strategy.
9 **QWC (ii) and QWC (iii)**	The student summarises their strategy, discussing the interrelationships between the variables, interpreting their results and evaluating their planning. They relate summary statistics to confirm or refute their hypothesis. All techniques, some of which must be complex, must be used and commented upon.
10 **QWC (ii) and QWC (iii)**	The student summarises and evaluates as above to use a number of different techniques, at least one of which must be at more complex than those for mark 9. All commentaries should be correct and concise. Any limitations are discussed and quantified.

 © Edexcel Limited 2008