New Specification

Rewarding Learning

Centre Number

Candidate Number

General Certificate of Secondary Education
2018

Statistics

Unit 1
Higher Tier

GST12

[GST12]
THURSDAY 21 JUNE, MORNING

TIME

2 hours.

INSTRUCTIONS TO CANDIDATES

Write your Centre Number and Candidate Number in the spaces provided at the top of this page.
Write your answers in the spaces provided in this question paper.
Answer **all thirteen** questions.
Any working should be clearly shown in the spaces provided since marks may be awarded for partially correct solutions.
You **may** use a calculator for this paper.

INFORMATION FOR CANDIDATES

The total mark for this paper is 100.
Figures in brackets printed down the right-hand side of pages indicate the marks awarded to each question or part question.
You should have a calculator, ruler, compasses and protractor.
The formula sheet is on page 2.

For Examiner's use only	
Question Number	Marks
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
Total Marks	

HIGHER TIER FORMULAE SHEET

$$\text{Standard deviation} = \sqrt{\frac{\sum fx^2}{n} - \left(\frac{\sum fx}{n}\right)^2}$$

Spearman's Rank Correlation Coefficient

$$r_s = 1 - \left(\frac{6\sum d^2}{n(n^2 - 1)}\right)$$

BLANK PAGE

(Questions start overleaf)

Answer **all** questions

1 The table below shows the number of each type of traffic offence recorded by the police in Northern Ireland each month in 2016

Offence	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
Breach of signs and signals	46	37	30	30	39	14	13	29	26	43	41	39	**387**
Careless driving	486	551	527	405	386	250	94	79	95	129	140	103	**3245**
Construction and use	411	357	315	206	165	174	98	78	70	106	99	66	**2145**
Dangerous driving	7	25	30	16	6	7	7	2	0	4	1	4	**109**
Lighting	58	56	60	29	19	20	47	26	26	52	40	49	**482**
Miscellaneous	68	78	105	82	51	23	13	14	10	12	14	14	**484**
Parking	145	87	122	57	57	45	35	45	62	152	181	70	**1058**
Pedal cyclists	2	0	1	2	4	3	1	0	1	2	1	1	**18**
Pedestrian	12	8	9	3	5	7	2	1	3	5	3	4	**62**
Seatbelt	54	121	111	68	88	98	84	50	75	83	67	67	**966**
Speeding	505	680	569	498	472	492	524	411	386	513	462	326	**5838**
Tachograph	0	1	0	0	0	0	1	0	0	2	1	1	**6**
Using a mobile phone	423	467	537	468	403	428	396	352	413	379	465	275	**5006**
No insurance	76	81	98	60	58	69	84	72	56	74	76	67	**871**
No vehicle test certificate	50	64	68	54	22	36	71	38	41	44	56	67	**611**
No driving licence	6	9	5	4	1	2	0	0	0	0	0	0	**27**
TOTAL	**2349**	**2622**	**2587**	**1982**	**1776**	**1668**	**1470**	**1197**	**1264**	**1600**	**1647**	**1153**	**21 315**

(a) Write down the total number of traffic offences committed in 2016

Answer _____ [1]

(b) How many Seatbelt or Speeding offences were committed in August?

Answer _____ [1]

(c) A newspaper printed the following headline:

"21 000 bad drivers in Northern Ireland"

Give one reason why this headline could be misleading.

_____

_____

_____ [1]

Examiner Only	
Marks	Remark

Examiner Only	
Marks	Remark

2 Sophie is designing a questionnaire to survey her class about how they travel to school.

(a) Write down an example of a closed question Sophie could use.

_____

_____

_____

_____ [2]

(b) Give one advantage and one disadvantage of using closed questions.

Advantage _____

_____

_____ [1]

Disadvantage _____

_____

_____ [1]

While carrying out her survey, Sophie discovers that she keeps having to explain what one of the questions on the questionnaire means.

(c) What should Sophie have done with her questionnaire before using it for her survey? Explain why she should have done this.

_____

_____

_____

_____

_____ [2]

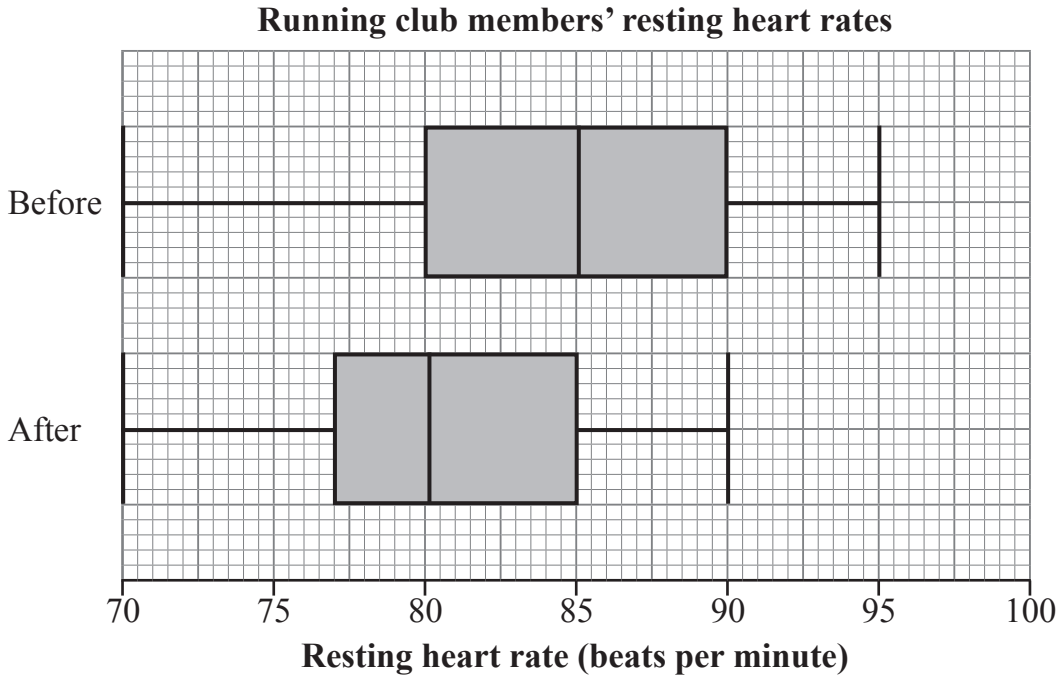
Examiner Only	
Marks	Remark

3 During the summer, a running club organises a boot camp where members can try to improve their general fitness.

A group of 60 members took part in the boot camp. The instructor measured their resting heart rates, in beats per minute, before the boot camp and found that they were, in general, too high.

He then measured their resting heart rates again after the boot camp.

The results for both measurements are shown in the following box plots.



Using the information in the box plots, describe the effect of the boot camp on the resting heart rates of the group.

_____

_____

_____

_____

_____ [3]

Examiner Only	
Marks	Remark

4 A mail order company operates a customer service helpline.

On its website the company claims that calls are answered, on average, within 20 seconds.

One day, the manager records the length of time taken to answer all calls to the helpline.

The results are given in the table below.

Time, t (seconds)	**Frequency, f**		
$0 < t \leqslant 10$	343		
$10 < t \leqslant 20$	214		
$20 < t \leqslant 30$	197		
$30 < t \leqslant 40$	145		
$40 < t \leqslant 50$	72		
$50 < t \leqslant 60$	29		
TOTAL	**1000**		

(a) Give one advantage and one disadvantage of using a grouped frequency table for this data.

Advantage _____

_____

_____ [1]

Disadvantage _____

_____

_____ [1]

(b) Calculate an estimate of the mean time taken to answer a call.

You may use the blank columns in the table opposite to help you with your working.

Answer _____ seconds [4]

(c) Using your answer to part **(b)**, explain whether or not the claim made by the company is justified.

_____

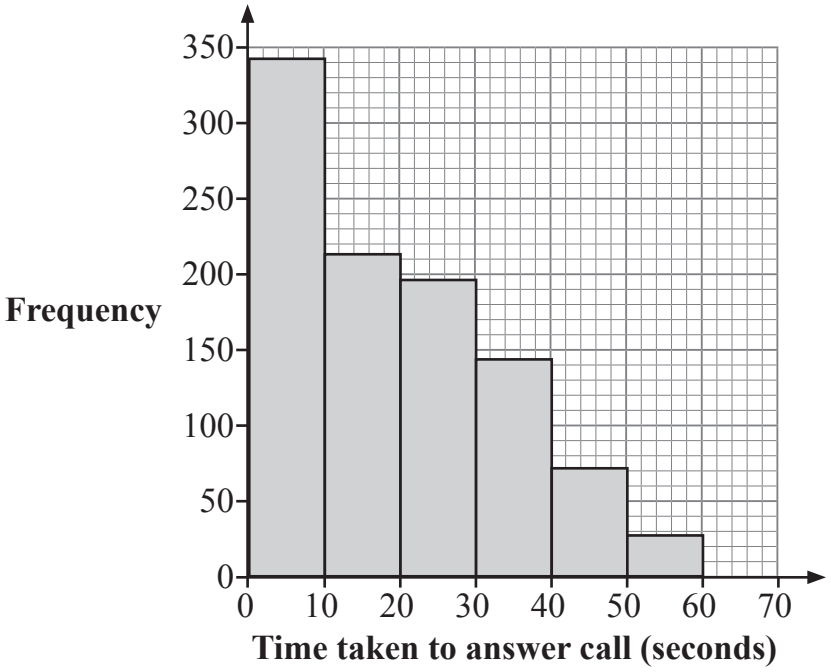
_____

_____

_____ [2]

Examiner Only	
Marks	Remark

A histogram showing the distribution of the times taken to answer the calls is shown below.



(d) (i) Do you think that this data could be modelled by a normal distribution?

Tick the correct box.

Yes ☐ **No** ☐ [1]

(ii) Explain your answer.

_____

_____

_____

_____ [1]

Examiner Only

Marks	Remark

Examiner Only	
Marks	Remark

5 A local council is planning to close the library in a small town.

The council decides to investigate the opinions of the people who live in the town about their plan to close the library.

(a) What is the population for this investigation?

_____

_____ [1]

One of the questions on the questionnaire is:

How many times have you used the local library recently?

1–5 ☐ **6–10** ☐ **More than 10** ☐

(b) Give **two** reasons why this is not a suitable question.

Reason 1 _____

_____

_____ [1]

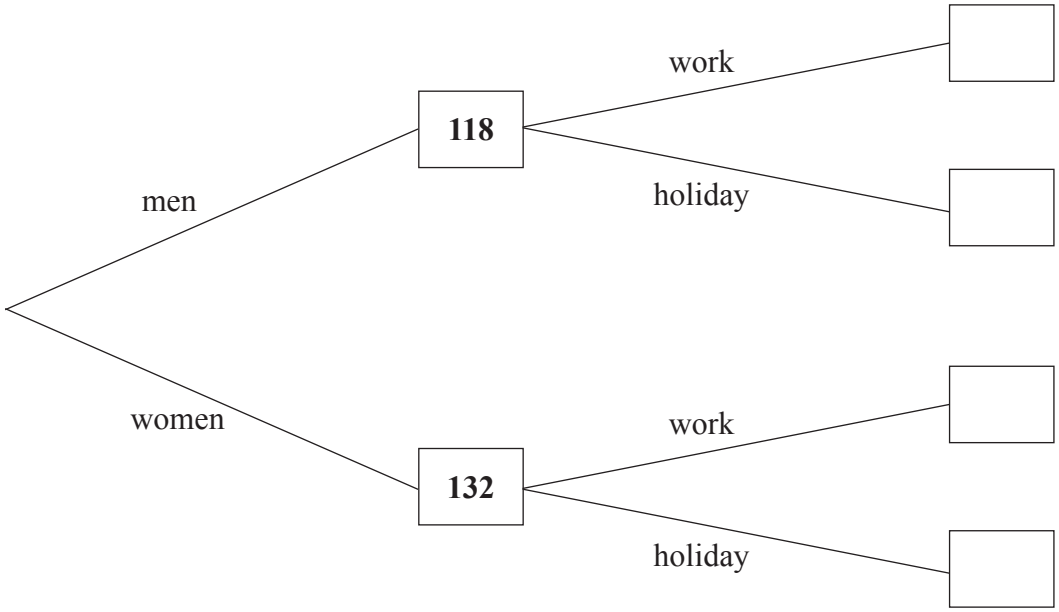
Reason 2 _____

_____

_____ [1]

Examiner Only	
Marks	Remark

6 In a survey, all passengers in an airport lounge said they were either travelling for work or on holiday. The partially completed frequency tree shows some of the information collected.



(a) 71 men and 64 women were travelling for work.

Complete the frequency tree. [2]

(b) A female passenger is chosen at random.

What is the probability that she is travelling on holiday?

Answer _____ [2]

(c) One of the passengers travelling on holiday is chosen at random.

What is the probability that the passenger is male?

Answer _____ [2]

Examiner Only	
Marks	Remark

Examiner Only	
Marks	Remark

7 Eight friends are talking about the number of times they had been to the cinema in the past year.

Their results are as follows.

3 8 9 5 12 4 2 1

(a) Calculate the mean of these numbers.

Answer _____ [2]

(b) Lisa says that the mode of these numbers is 12

Explain why Lisa is **not** correct.

_____

_____

_____

_____ [1]

(c) One of the group remembered that she had been to the cinema 7 times and not 9 times.

What effect will this have on the value of the mean?

Tick the correct box.

Increase ☐ **Decrease** ☐ **No change** ☐ [1]

(d) Jenny thinks that the number of times a person goes to the cinema might be related to how far they live from it.

Write down a hypothesis Jenny could use.

Hypothesis

_____

_____ [1]

Jenny plans to ask some people about:

- the distance they live from the nearest cinema; and
- the number of times they have been to the cinema in the past year.

(e) What kind of data is the number of visits to the cinema?

Circle **two** words from the list below that describe the data.

Qualitative	**Discrete**	**Continuous**	
Bivariate	**Quantitative**	**Categorical**	[2]

After collecting her data, Jenny calculated the product moment correlation coefficient to be –0.769

(f) What conclusion could Jenny draw from this value?

_____

_____

_____

_____ [2]

Examiner Only

Marks	Remark

8 A manufacturing process produces tins of baked beans.

The label on the tin states that there is 415 g of beans in the tin.

(a) Give one reason why it would not be sensible to check the weight of beans in every tin.

_____

_____

_____ [1]

At regular intervals, samples of tins are taken and the mean weight of beans in each sample is plotted on a control chart.

(b) Explain briefly why control charts are useful in a production process.

_____

_____

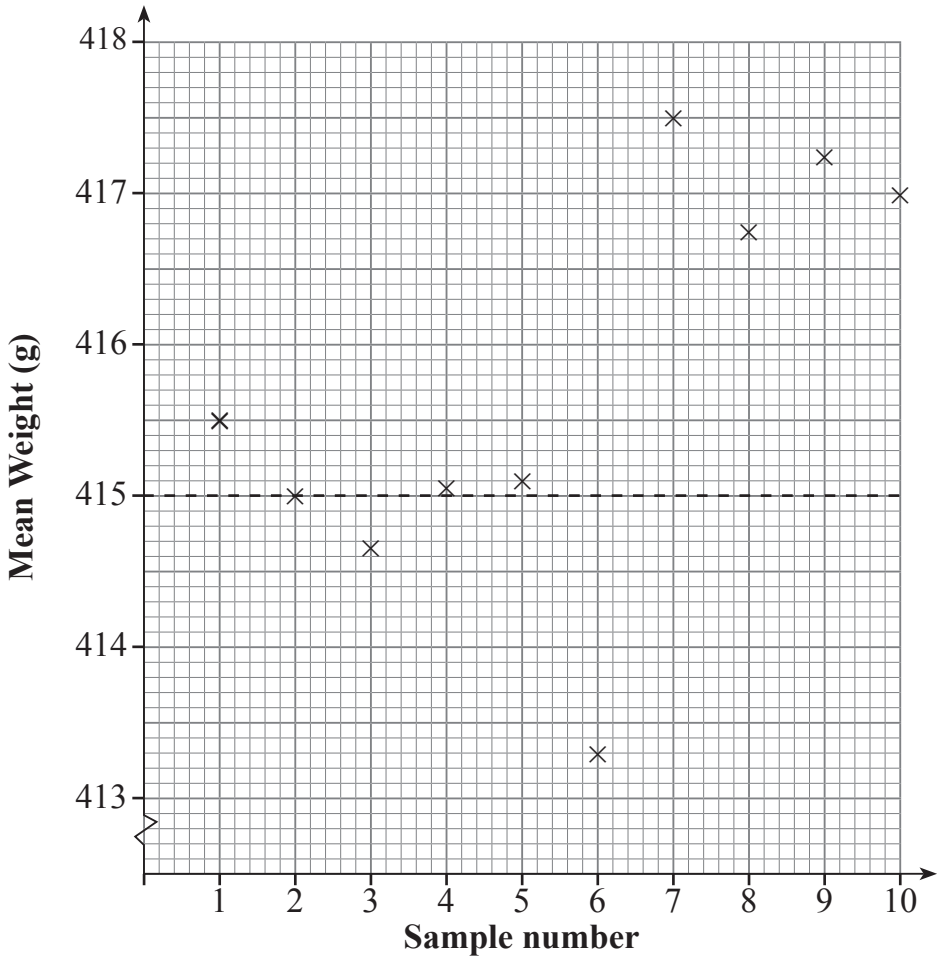
_____ [1]

The production manager has set the lower and upper warning limits as (415 ± 1.2) g and the lower and upper action limits as (415 ± 1.8) g.

The control chart for one shift's production is shown at the top of the next page.

(c) On the control chart, draw and label the warning lines and action lines. [3]

Examiner Only	
Marks	Remark



(d) During the shift, the production manager went home sick so the control chart went unchecked.

Explain, with reasons, how the production manager would have responded to the chart if he had been at work.

_____

_____

_____

_____

_____

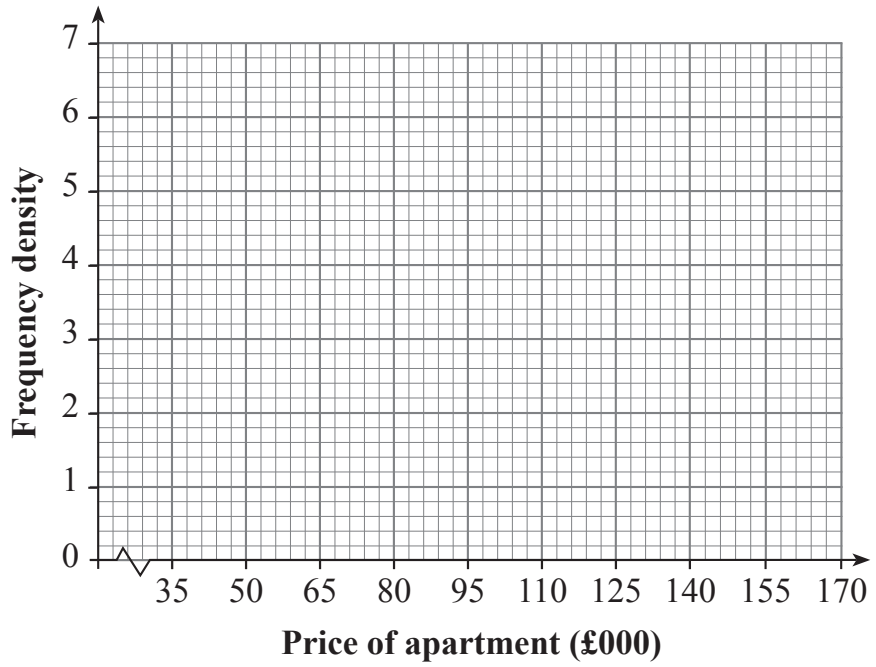
_____ [4]

Examiner Only	
Marks	Remark

9 An estate agent is doing some research on property values in a city. The prices in thousands of pounds (£000) of a sample of apartments which were sold during one year are given in the grouped frequency table below.

Price (£000)	No. of Apartments
$35 \leqslant p < 65$	42
$65 \leqslant p < 95$	114
$95 \leqslant p < 110$	96
$110 \leqslant p < 125$	81
$125 \leqslant p \leqslant 155$	18

(a) Use the information in the table to draw a histogram for the prices of the apartments.



[4]

Examiner Only

Marks	Remark

(b) Victor says that the table indicates that the most expensive apartment in the city cost £155 000

Give two reasons why Victor may not be correct.

Reason 1

_____

_____

_____ [1]

Reason 2

_____

_____

_____ [1]

Examiner Only	
Marks	Remark

10 The table below shows the number of complaints received by a company each quarter over a three-year period.

Year	Quarter	No. of complaints	4-point moving averages
2012	Q1	79	
	Q2	102	103
	Q3	120	105
	Q4	111	106
2013	Q1	87	108.5
	Q2	106	109.5
	Q3	130	112
	Q4	115	
2014	Q1	97	
	Q2	120	
	Q3	134	
	Q4	125	

(a) Explain briefly why **4-point** moving averages are used for this data.

_____

_____

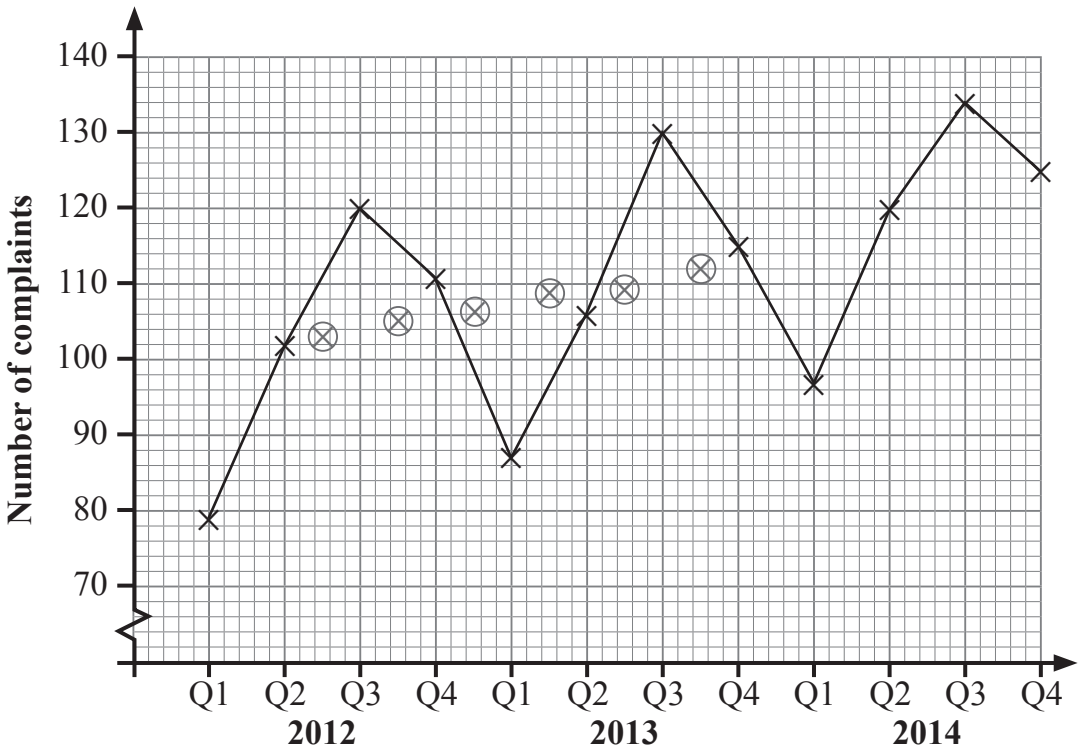
_____ [1]

(b) Complete the table above. [3]

Examiner Only

Marks	Remark

The data from the table is plotted as a time series graph below.



(c) Plot the remaining 4-point moving averages and draw a trend line. [2]

(d) Describe the trend in the number of complaints received by the company over the three-year period.

_____

_____ [1]

(e) Using your trend line, estimate the number of complaints the company could expect to receive during the first quarter of 2015

Answer _____ [3]

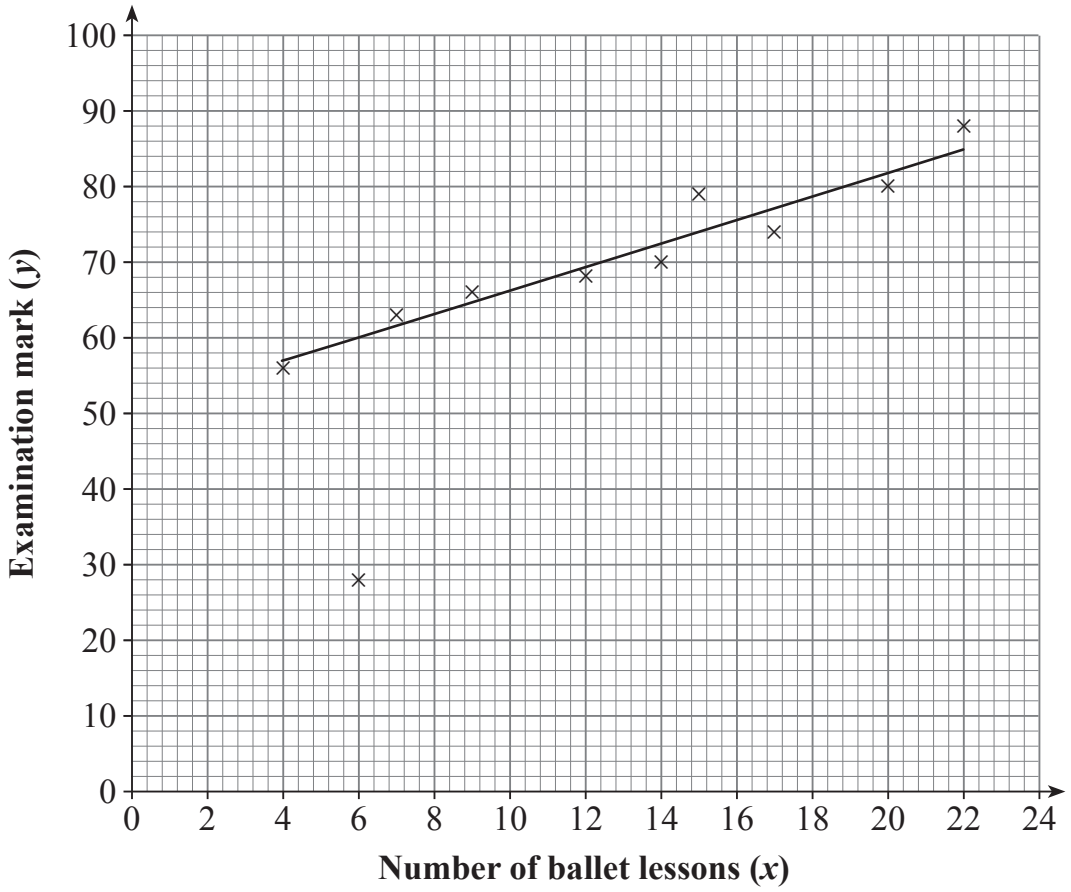
Examiner Only

Marks	Remark

Examiner Only	
Marks	Remark

11 Grace carries out a survey among ten of her friends about the number of ballet lessons they took and the mark they got in their recent ballet examination.

Her results are given in the scatter diagram below.



(a) One of the results is an outlier.

Put a circle around the point on the graph representing this result. [1]

The outlier is now removed.

Grace draws a line of best fit and calculates its equation to be $y = 50.6 + 1.58x$

(b) Give an interpretation of the number 50.6 in this equation.

_____

_____

_____

_____ [2]

Grace's friend Anna was unable to attend the examination but she took 11 lessons.

(c) Use the equation of the line of best fit to show that an estimate of Anna's examination result is 68

[2]

Examiner Only	
Marks	Remark

During the summer holidays, Anna took part in a ballet competition.

The eight dancers were ranked by two judges. The ranks are as follows:

Dancer	Judge 1	Judge 2		
A	1	2		
B	2	1		
C	3	3		
D	4	4		
E	5	6		
F	6	5		
G	7	7		
H	8	8		

(d) Calculate Spearman's rank correlation coefficient for this data, giving your answer correct to three decimal places.

Answer _____ [3]

(e) Anna thinks one of the judges is not judging consistently.

Using your result from **(d)**, comment on Anna's claim.

_____

_____

_____ [2]

Examiner Only

Marks	Remark

BLANK PAGE

(Questions continue overleaf)

12 The table below shows the price of a jar of coffee on 1st January over four years.

Year	**2014**	**2015**	**2016**	**2017**
Price	£3.09		£3.22	£3.17
Chain base index number	100	101.0	103.2	

(a) Calculate the price of a jar of coffee on 1st January 2015

Answer £ _____ [2]

(b) Calculate the chain base index number for 2017

Answer _____ [3]

Examiner Only

Marks	Remark

(c) Explain the meaning of the value of the chain base index number for 2017 which you calculated in part **(b)**.

Examiner Only	
Marks	Remark

_____

_____

_____

_____ [2]

The company which produces the coffee has three departments: production, administration and management.

The percentage of the total wages bill for each of the departments is as follows:

60% production

25% administration

15% management

Taking 2015 as the base year, the index numbers for the company's total wages bill for each department in 2016 are as follows:

Production	Administration	Management
106	103	107

(d) In 2015, the total wages bill for the company was £340 540

Calculate the wages bill for the **production** department in 2016

Answer _____ [4]

Examiner Only	
Marks	Remark

(e) Calculate the weighted index number for the company's total wages bill for 2016

Answer _____ [4]

Examiner Only	
Marks	Remark

13 For his Geography project, Simon is comparing the lengths of stones found at two different locations along a river.

At the first location, the mean length of a stone is 33.4 mm and the standard deviation is 1.7 mm.

Simon selects a stone at random and measures its length to be 35 mm.

(a) Calculate the standardised score for the length of this stone.

Answer _____ [3]

At the second location, the mean length of a stone is 31.6 mm and the standard deviation is 1.1 mm.

Simon selects a stone at random from this location and finds that the standardised score for the length of the stone is –1.455

(b) Find the actual length of this stone.

Answer _____ mm [2]

Examiner Only	
Marks	Remark

Upon returning to school, Simon accidentally mixed up some of the stones.

He selected one at random and measured its length to be 32 mm.

(c) Using standardised scores, work out which location the stone is more likely to have come from.

_____ [4]

(d) What assumption must be made about the distributions of the lengths of the stones at each location in order to use standardised scores?

_____

_____ [1]

THIS IS THE END OF THE QUESTION PAPER

Examiner Only	
Marks	Remark

11633.02 **R**