

--	--	--	--	--

--	--	--	--

Statistics

Unit 1
Higher Tier



GST12

[GST12] Assessment

TIME

2 hours.

Assessment Level of Control:

Tick the relevant box (✓)

Controlled Conditions	
Other	

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 ten** 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	
Total Marks	

HIGHER TIER FORMULA SHEET

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

Spearman's Rank Correlation Coefficient

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

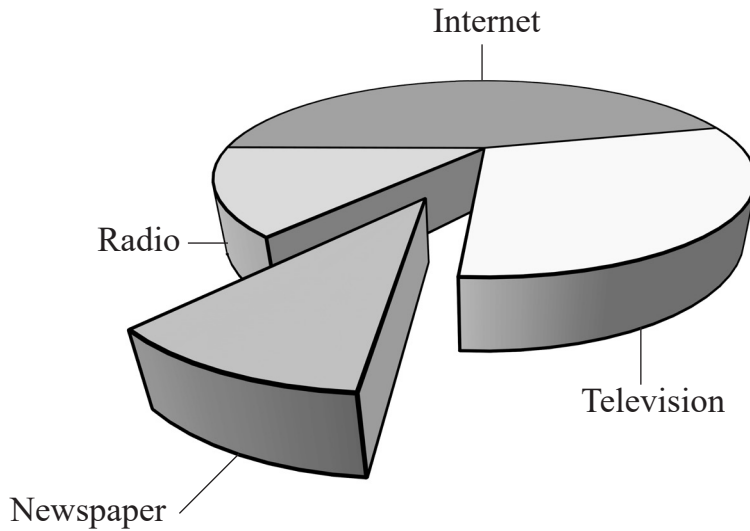
Answer **all** questions

Examiner Only

Marks Remark

- 1 In a survey, some people were asked to name their main source of local news.

The results are summarised in the pie chart below.



Give two reasons why the pie chart might be considered misleading.

- 1. _____
- _____
- _____ [1]

- 2. _____
- _____
- _____ [1]

Examiner Only	
Marks	Remark

2 Paul wants to survey pupils at his school about proposed changes to the canteen menu.

He plans to collect primary data.

(a) Explain what is meant by primary data.

_____ [1]

(b) State one advantage and one disadvantage of using primary data.

Advantage

_____ [1]

Disadvantage

_____ [1]

(c) The survey will include the following question:

“What do you think of the proposed changes to the canteen menu?”

Tick below:

Strongly Agree **Agree** **No opinion** **Disagree**

Give one reason why the response section is not suitable.

_____ [1]

For his survey, Paul chooses a sample of ten pupils from Year 8

(d) Give two reasons why this sample may give biased results.

1. _____
_____ [1]

2. _____
_____ [1]

Examiner Only	
Marks	Remark

3 A manufacturing company uses a machine which is programmed to fill bags with flour.

Every hour, a sample of five bags of flour is taken and the mean weight of the five bags of flour is calculated.

(a) Explain why the mean weight is calculated and not the median.

[2]

(b) Give one consequence to the company of overfilling the bags with flour.

[1]

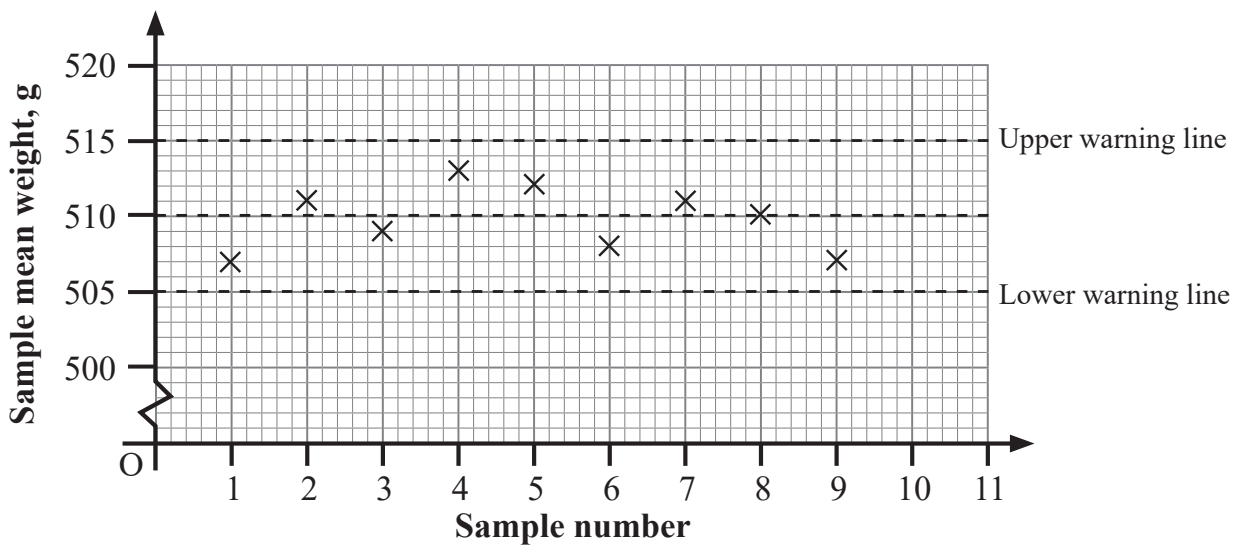
(c) Give one consequence to the company of underfilling the bags with flour.

[1]

Examiner Only	
Marks	Remark

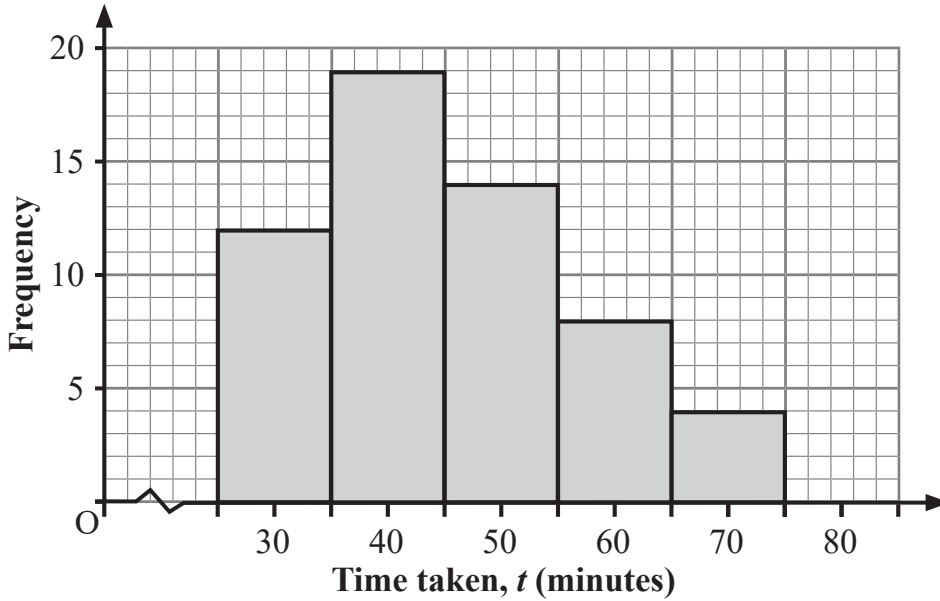
Evelyn is responsible for monitoring the mean weight of each sample over time.

She has plotted the first nine sample means on a control chart.



4 A sports coach recorded the times taken by a group of runners to complete a race.

He presented his results in the graph below.



(a) What is the name of this type of graph?

_____ [1]

(b) What type of skewness is shown in the graph?

_____ [1]

(c) Explain why it is not possible to use the graph to find the range of times taken by the runners to complete the race.

 _____ [1]

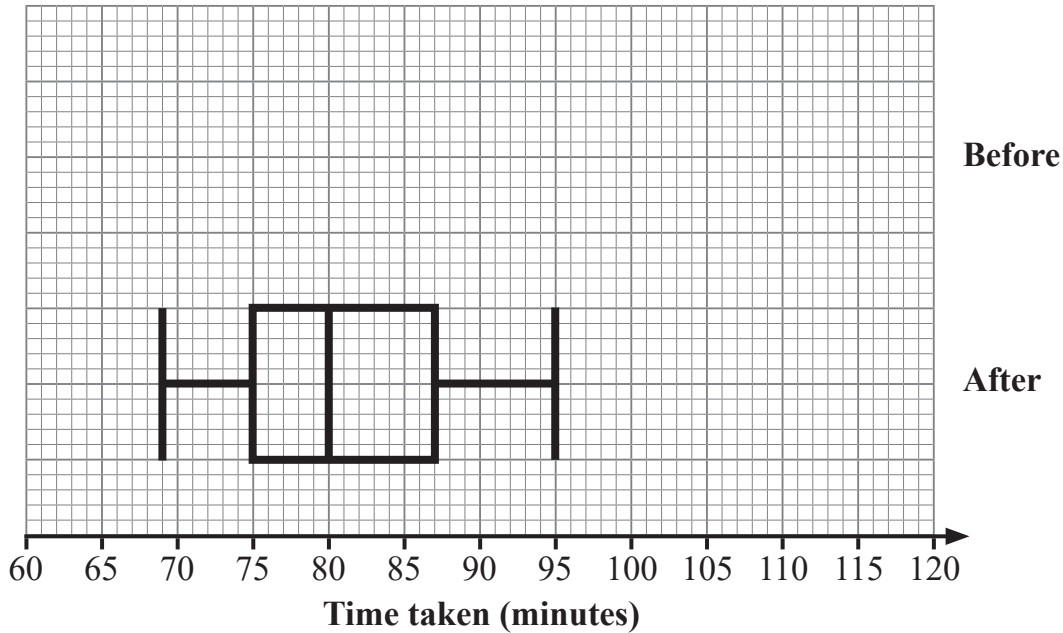
(d) How many runners completed the race?

Answer _____ [2]

Examiner Only	
Marks	Remark

The 15 workers went on a training course.

After their return, they each assembled another wardrobe and their times were recorded again. The results are summarised in the box plot below.



(c) On the same diagram, draw a box plot to show the times taken by the workers to assemble a wardrobe before they completed the training course. [2]

(d) Comment on the impact of the training course on the times taken by the workers to assemble wardrobes.

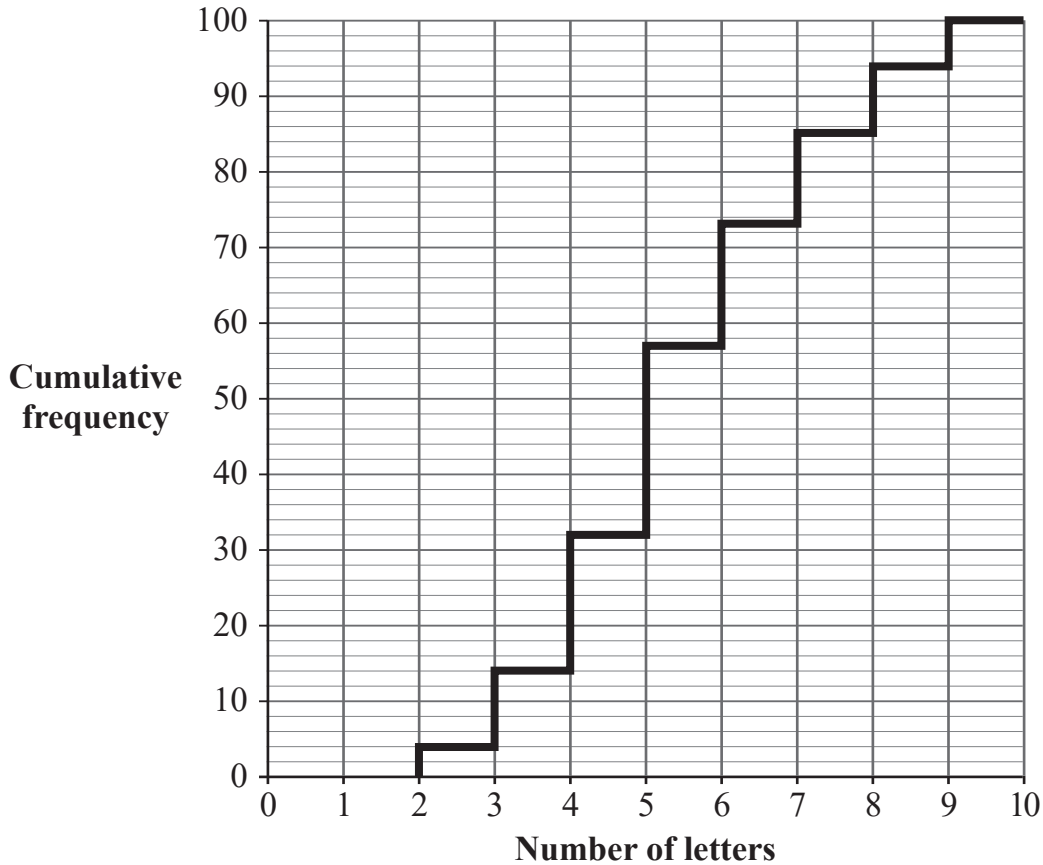
[4]

Examiner Only	
Marks	Remark

- 6 Joseph wishes to investigate the number of letters in boys' names compared to the number of letters in girls' names.

He asked 100 boys in Year 12 how many letters their first name contained.

Joseph's results are displayed in the stepped cumulative frequency diagram below.



- (a) Explain why the cumulative frequency diagram is stepped.

_____ [1]

- (b) Find the median.

Answer _____ letters [2]

Examiner Only	
Marks	Remark

(c) (i) Use the diagram to find the 90th percentile.

Answer _____ [2]

(ii) Give an interpretation, in context, of this value.

_____ [2]

Joseph repeated the survey for 100 girls from Year 12

The median for the girls' first names was 6 letters.

(d) Using your answer from (b), compare the median for the boys and the girls.

_____ [2]

The 10th percentile for girls' names was 3 letters.

The 90th percentile for girls' names was 7 letters.

(e) (i) Use these values to calculate an interpercentile range for girls' names.

Answer _____ letters [2]

(ii) Complete the sentence below.

Among the girls, _____ % had names containing
between _____ and _____ letters. [2]

Examiner Only	
Marks	Remark

7 Anne is a manager in a large bank which employs 720 full-time and part-time staff across three of its departments as shown in the table below.

	Legal	Sales	Finance
Full-time	104	280	128
Part-time	56	112	40

Anne wants to know if the full-time and part-time staff in each of these departments of the bank are satisfied in their jobs.

She decides to conduct a survey using a stratified random sample of 90 staff.

(a) Give a reason why Anne decided to use stratification in her sample.

_____ [1]

(b) Work out the number of part-time staff from the sales department Anne needs to survey.

Answer _____ [3]

Examiner Only	
Marks	Remark

- (c) Suggest one extra category other than those shown in the table which Anne could use to stratify the staff across these three departments.

_____ [1]

To improve job satisfaction, Anne decides to run a free prize draw. Every member of the three departments in the table is entered into the draw and a name is selected at random.

- (d) Work out the probability that the prize is won by a full-time member of staff.

Answer _____ [2]

Examiner Only	
Marks	Remark

8 Kevin wants to estimate the total number of eels in a lough.

He captures 45 eels, tags them and releases them unharmed.

One week later, Kevin captures 85 eels.

In this sample, six eels had a tag.

(a) Use this information to calculate an estimate of the number of eels in the lough.

Answer _____ [4]

(b) Write down two assumptions Kevin made to calculate this estimate.

1. _____

_____ [1]

2. _____

_____ [1]

Examiner Only	
Marks	Remark

Kevin measured the lengths, in centimetres, of the 45 eels in his original sample.

His results are shown in the table below.

Length (cm)	Frequency	
$65 < x \leq 70$	4	
$70 < x \leq 75$	5	
$75 < x \leq 80$	9	
$80 < x \leq 85$	16	
$85 < x \leq 90$	11	

(c) What type of data is the length of the eels?

Circle your answer.

Discrete **Continuous** **Categorical** **Bivariate**

[1]

(d) Calculate an estimate of the median length of the eels in Kevin's sample.

Answer _____ cm [4]

(e) Explain why your calculation in part (d) is an estimate.

_____ [1]

Examiner Only	
Marks	Remark

- 9 In a factory, two machines are used to produce metal rods of different weights.

For both machines, the weights of the rods are normally distributed.

The mean and standard deviation for the weights of the rods produced by each machine are shown in the table below.

Machine	Mean (g)	Standard deviation (g)
A	24.8	0.8
B	28.2	1.2

- (a) (i) Write down the percentage of rods produced by Machine A which are expected to have weights between 23.2 g and 26.4 g.

Answer _____ % [1]

- (ii) Explain how you obtained your answer to part (a)(i).

_____ [1]

A rod produced by Machine A was chosen at random. The standardised score for the weight of this rod was calculated to be 0.25

Peter says this rod must have weighed more than 24.8 g.

- (b) Explain why Peter is correct.

_____ [2]

Examiner Only

Marks Remark

The standardised score for a rod produced by Machine B was -0.1

(c) Calculate the weight of this rod.

Answer _____ g [3]

A rod produced by Machine A weighs 27.4 g.

(d) Show that this weight is an outlier for Machine A.

[3]

A rod weighing 26.4 g was found on the factory floor.

Peter calculated the standardised score for this rod using the mean and standard deviation for each machine as follows:

Machine A 2.0

Machine B -1.5

(e) Which machine was the rod more likely to have come from?

Machine A Machine B

Give a reason for your answer.

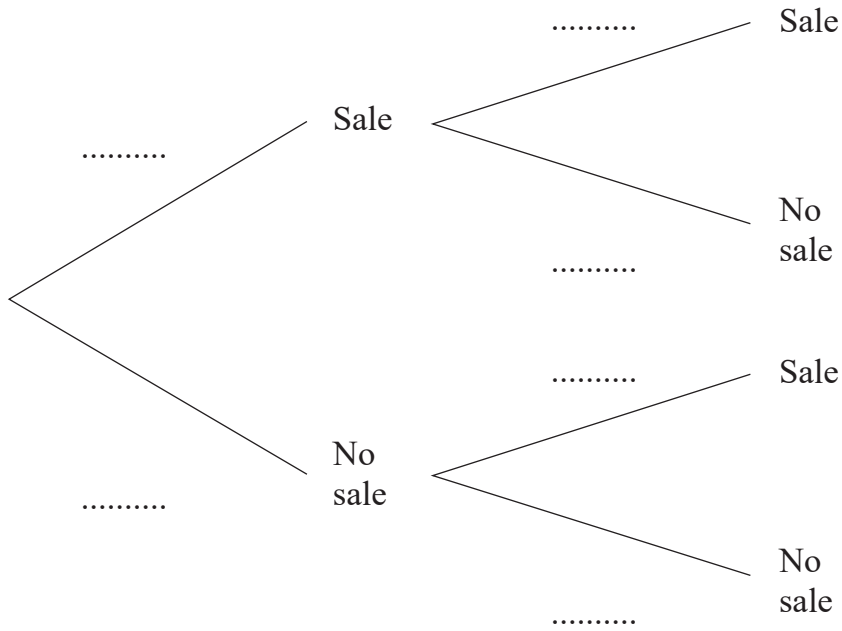
[2]

Examiner Only	
Marks	Remark

10 Cara works as a telephone sales representative.

The probability that Cara makes a sale when a customer calls is 0.62

(a) Complete the probability tree diagram for two consecutive calls.



[3]

(b) Calculate the probability that Cara makes just one sale from two consecutive calls.

Answer _____ [3]

Examiner Only	
Marks	Remark

Cara receives four calls during one hour.

She decides to use the binomial distribution as a model for the number of sales she makes from these four calls.

- (c) Explain why the binomial distribution is a suitable model for Cara to use.

[2]

- (d) Calculate the probability that Cara will make more than two sales during a one hour period.

You may use $(p + q)^4 = p^4 + 4p^3q + 6p^2q^2 + 4pq^3 + q^4$

Answer _____ [3]

- (e) Work out the most likely number of sales Cara will make during one hour.

You may use $(p + q)^4 = p^4 + 4p^3q + 6p^2q^2 + 4pq^3 + q^4$

Answer _____ [3]

THIS IS THE END OF THE QUESTION PAPER

Examiner Only

Marks Remark

Permission to reproduce all copyright material has been applied for.
In some cases, efforts to contact copyright holders may have been unsuccessful and CCEA
will be happy to rectify any omissions of acknowledgement in future if notified.