



Cambridge O Level

CANDIDATE
NAME
CENTRE
NUMBER

--	--	--	--	--

CANDIDATE
NUMBER

--	--	--	--

STATISTICS

4040/22

Paper 2

October/November 2024

2 hours 15 minutes

You must answer on the question paper.

You will need: Calculator
Pair of compasses
Protractor

INSTRUCTIONS

- Answer **all** questions.
- Use a black or dark blue pen. You may use an HB pencil for any diagrams or graphs.
- Write your name, centre number and candidate number in the boxes at the top of the page.
- Write your answer to each question in the space provided.
- Do **not** use an erasable pen or correction fluid.
- Do **not** write on any bar codes.
- You should use a calculator where appropriate.
- You must show all necessary working clearly.

INFORMATION

- The total mark for this paper is 100.
- The number of marks for each question or part question is shown in brackets [].

This document has **16** pages.



1 A sports coach collected data about the people who attended one of her fitness sessions.

(a) In each case below, put a tick (✓) in the column with the correct description of the type of data she collected.

<i>Data</i>	<i>Not a variable</i>	<i>Qualitative variable</i>	<i>Discrete quantitative variable</i>	<i>Continuous quantitative variable</i>
The favourite sport of each person that attended				
The number of people who attended the session				
The distance each person travelled to get to the session				

[3]

At the start of the session the coach recorded the number of laps of a field each person ran in five minutes, **in completed laps**.

She grouped the numbers of laps into classes labelled 5–6, 7–8, 9–10 etc.

(b) State the lower and upper class boundaries of the 7–8 class.

Lower class boundary

Upper class boundary

[1]

At the end of the session the coach recorded the time each person took to complete a fitness circuit, **to the nearest minute**.

She grouped these times into classes labelled 5–6, 7–8, 9–10 etc.

(c) State the lower and upper class boundaries of the 7–8 class.

Lower class boundary

Upper class boundary

[1]





2 Three events A , B and C are such that

$$P(A) = 0.35 \quad P(B) = 0.2 \quad P(C) = 0.4 \quad P(A \cup C) = 0.52$$

(a) Given that A and B are mutually exclusive events, find $P(A \cap B)$ and $P(A \cup B)$.

$$P(A \cap B) = \dots\dots\dots$$

$$P(A \cup B) = \dots\dots\dots$$

[3]

(b) (i) Find $P(A \cap C)$.

$$\dots\dots\dots [2]$$

(ii) Show whether or not A and C are independent events.

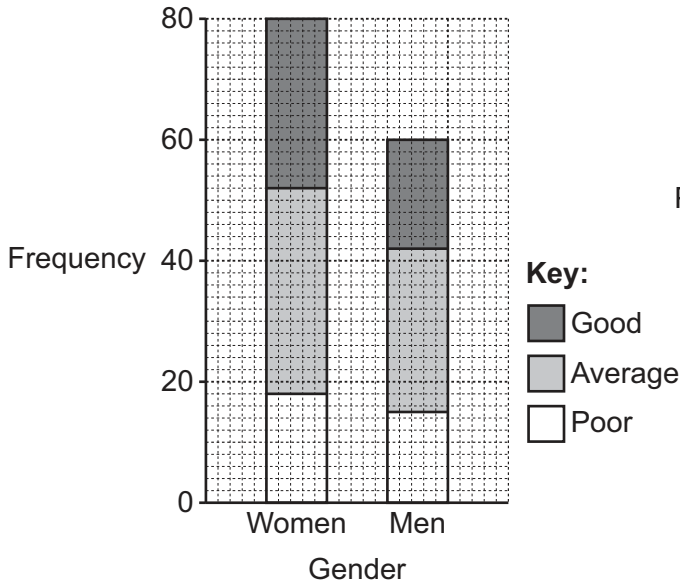
[3]



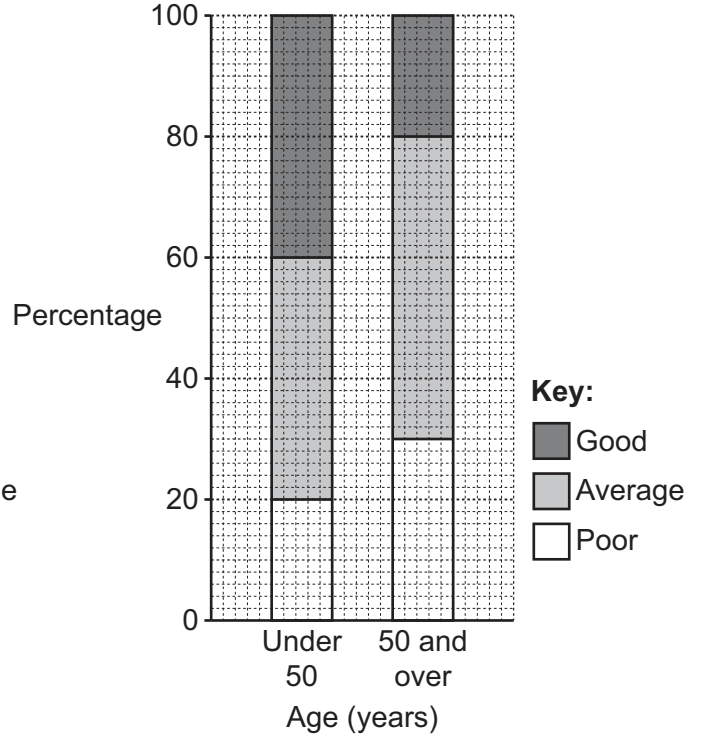


- 3 A sample of 140 women and men of various ages were each asked to rate a film as one of Good, Average or Poor.
Information about the results from these 140 people is shown below.

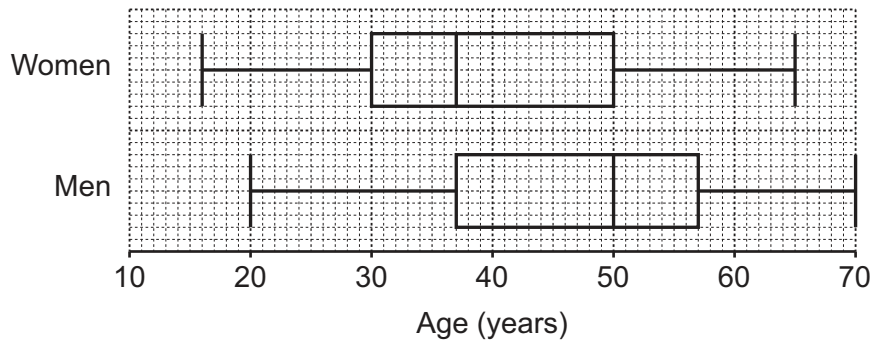
Sectional bar chart comparing film ratings by gender



Percentage sectional bar chart comparing film ratings by age category



Box-and-whisker diagram comparing ages by gender



- (a) Find the number of these people who said that the film was good.

..... [2]

- (b) Find the number of women aged under 30 in the sample.

..... [2]





- (c) Find the number of these people aged 50 and over who said that the film was average.

..... [3]

- 4 A photographer divides his total expenditure into four categories: Equipment, Travel, Advertising and Insurance. The table shows weights based on expenditure in 2022 and price relatives taking 2022 as the base year.

<i>Category</i>	<i>Weight</i>	<i>Price relative in 2023</i>
Equipment	10	94
Travel	5	103
Advertising	3	101
Insurance	2	

The weighted aggregate cost index for 2023 taking 2022 as the base year is 98.1 .

- (a) Find the price relative for insurance.

..... [4]

- (b) Explain what the figure 98.1 tells you.

.....

 [2]

The weighted aggregate cost index may not produce a reliable estimate for his expenditure in 2023 if the weights have changed.

- (c) Give **one** reason why the weights may have changed.

.....
 [1]





- 5 A bag contains five discs that are numbered 1 and one disc that is numbered 2.
Another bag contains three discs that are numbered 1 and one disc that is numbered 2.

In a game a disc is selected at random from each bag.
The two numbers obtained are added together, and the total gives the prize, in dollars, for playing the game.

- (a) Complete the probability distribution for the prizes.

<i>Prize (\$)</i>			
<i>Probability</i>			

[4]

Yuri is charged \$2.50 each time he plays the game.

- (b) Find the expected profit or loss for Yuri if he plays the game 15 times.

..... [4]





- 6 Amna is a sprinter. Her athletics club is holding a competition. Amna is trying to decide whether to enter the 100 m event or the 200 m event.

She finds the following summary statistics for the fastest times, in seconds, of all the sprinters in her club.

	100 m	200 m
Mean	12.1	24.3
Standard deviation	0.35	0.80

Her fastest time this year for the 100 m is 11.4 seconds and for the 200 m is 23.3 seconds.

- (a) By scaling each of her times to a mean of 0 seconds and a standard deviation of 1 second, suggest which event Amna should enter.

..... [4]

Amna's trainer says that to estimate an athlete's time in the 400 m you should double their time for 200 m and add 4 seconds.

- (b) Using this rule, estimate the mean and standard deviation of the fastest times of the sprinters in Amna's club for the 400 m.

Mean

Standard deviation [2]





- 7 The lengths of some sea cucumbers are measured in centimeters.
The frequency table and the cumulative frequency table give some information about the results.

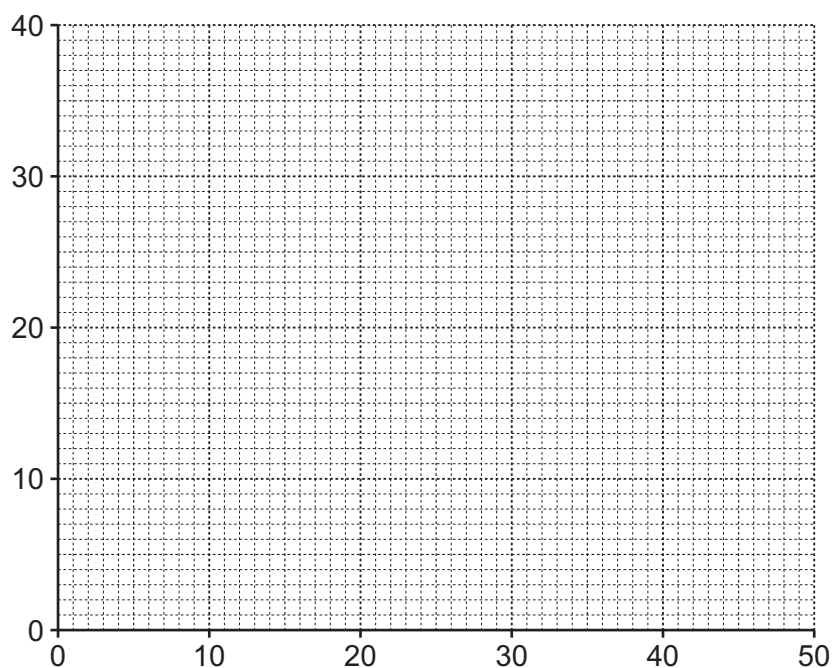
<i>Length, l (cm)</i>	<i>Frequency</i>
$0 \leq l < 8$	0
$8 \leq l < 16$	7
$16 \leq l < 24$	
$24 \leq l < 32$	
$32 \leq l < 40$	10
$40 \leq l < 48$	0

<i>Length, l (cm)</i>	<i>Cumulative frequency</i>
$l < 8$	0
$l < 16$	7
$l < 24$	32
$l < 32$	66
$l < 40$	
$l < 48$	

- (a) Complete the frequency table and the cumulative frequency table.

[3]

- (b) On the grid, draw a frequency polygon for the lengths of the sea cucumbers.
Add appropriate labels to the axes.



[4]





(c) Use linear interpolation to find an estimate for the median length of these sea cucumbers.

..... [4]

The width of each of these sea cucumbers is approximately half the length.

(d) Use linear interpolation to find an estimate for the number of these sea cucumbers that have a width of less than 11.5 cm.

..... [4]





- 8 The managers at a company decide to survey the employees about changing the company's working hours.

They decide to ask the open question, 'How would you like the working hours to change?'

- (a) Give one disadvantage of asking an open question.

.....
 [1]

Joseph thinks they should survey all of the employees.

- (b) Write down the name of the survey method that Joseph thinks they should use.

..... [1]

The other managers think they should survey a sample of the employees rather than all of them.

- (c) Give one advantage and one disadvantage of using a sample.

Advantage

.....

Disadvantage

..... [2]

The table shows the number of employees at the company in each age group and with each contract type.

		Contract type	
		Full-time	Part-time
Age group	Under 40	24	20
	40 and over	12	10

Adla, Fatima and Kami each want to take a sample of the 66 employees at the company, so they number the employees as shown below.

		Contract type	
		Full-time	Part-time
Age group	Under 40	00–23	36–55
	40 and over	24–35	56–65





Adla decides to take a systematic sample of **size 6**.

(d) Use the random number table below to help select her sample.

85 27 02 31 44 56 32 19 81 74 02 40 69

..... [3]

(e) Show whether or not Adla's sample is representative in terms of contract type.

[3]

Fatima decides to take a sample of **size 11** stratified by contract type and age group.

(f) Find the number of full-time employees aged under 40 that should be in her sample.

..... [1]

Kami decides to take a sample stratified by age group, rounding the sample size for each stratum to the nearest integer.

Kami wants to have 4 employees aged 40 and over in his sample.

(g) Find all the possible sizes of his total sample.

..... [3]





- 9 (a) Describe in which situations it is necessary to centre moving average values and explain why it is necessary in order to calculate seasonal components.

.....

.....

..... [2]

The table shows the quarterly profits made by a large company from the first quarter of 2021 to the second quarter of 2023.

<i>Year and quarter</i>	<i>Profit (\$100 000s)</i>	<i>4-point moving average (\$100 000s)</i>	<i>Centred 4-point moving average (\$100 000s)</i>
2021 Q1	4.2		
2021 Q2	5.7		
		4.5	
2021 Q3	3.6		
		4.4	
2021 Q4	4.5		
		4.2	
2022 Q1	3.8		
		4.15	
2022 Q2	4.9		
		4.05	
2022 Q3	3.4		
		3.95	
2022 Q4	4.1		
		3.8	
2023 Q1	3.4		
2023 Q2	4.3		

- (b) Complete the table with the centred 4-point moving average values.





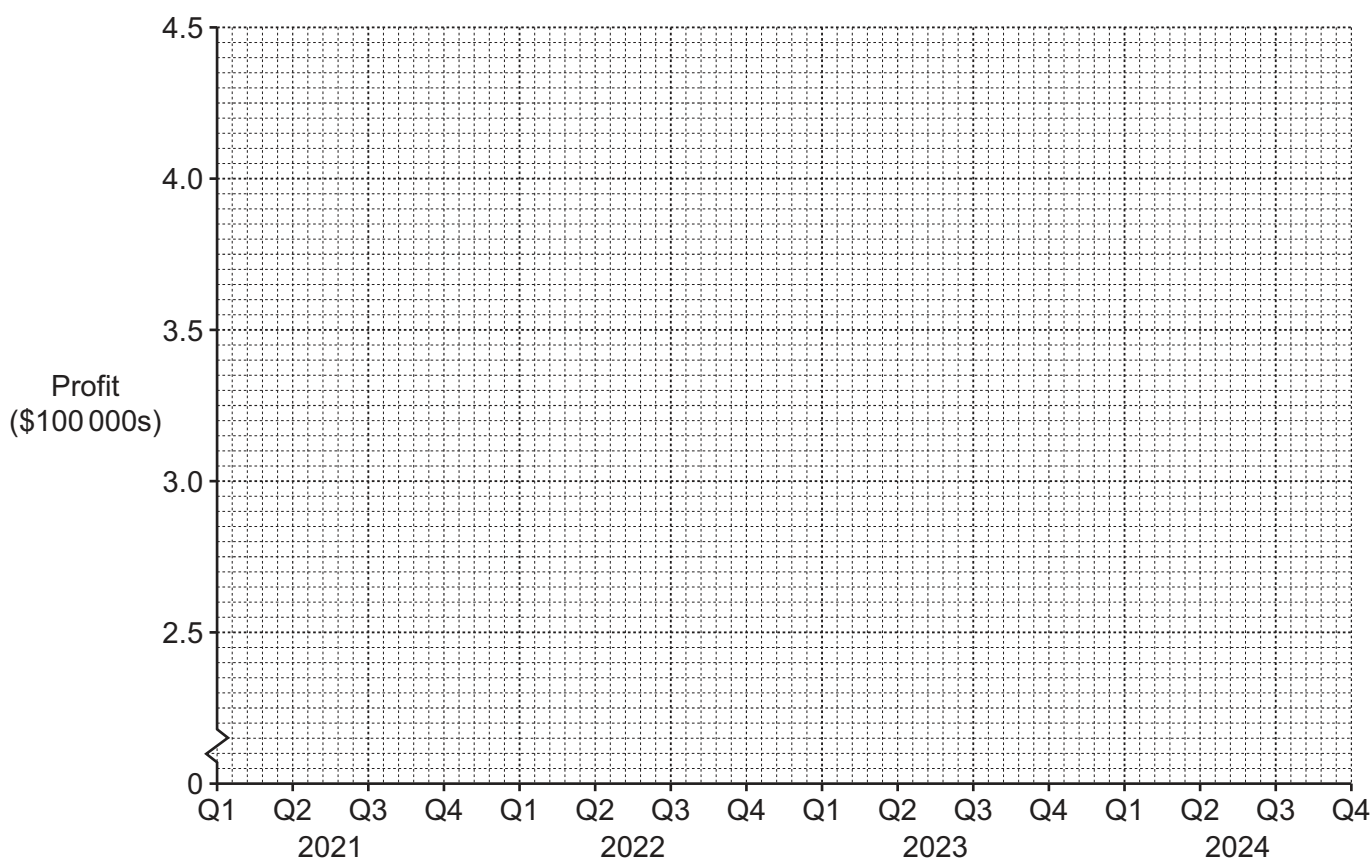
(c) Use appropriate values from the table to calculate the seasonal component for quarter 3.

..... [3]

(d) Explain what your answer to part (c) tells you about the profits in quarter 3.

.....
..... [1]

(e) Plot all the centred moving average values on the grid below and draw an appropriate trend line.



[3]

(f) Use your answers to parts (c) and (e) to estimate the profit made by the large company in quarter 3 of 2024.

..... [3]





- 10 Akal is studying the health of the birds in the nests in the eastern section of a forest. He counts the number of birds in each nest.

<i>Number of birds in each nest</i>	0	1	2	3
<i>Number of nests</i>	4	5	8	7

He numbers the nests and selects a nest at random.

- (a) Find the probability that he selects a nest containing 0 birds.

..... [1]

He decides instead to number the birds and select a bird at random.

- (b) (i) Find the total number of birds in these nests.

..... [2]

- (ii) Find the probability that he selects a bird from a nest containing 2 birds.

..... [2]

Zuri is studying the health of the birds in the nests in the western section of the forest. Some information about her results is shown in the table.

	Eastern section	Western section
Number of birds in nests		43
Number of nests		
Mean number of birds in nests		2.15
Range of the number of birds in nests		5

- (c) Complete the table for the eastern and western sections of the forest.

[4]



Zuri says that the number of birds in each nest is more varied in the western section of the forest than the eastern section.

- (d) Explain whether or not you think that the data in the table is sufficient to support Zuri's claim.

.....

.....

.....

..... [2]

- (e) Find the mean number of birds in each nest for the eastern and western sections of the forest combined.

..... [2]

- (f) For each situation in the table below, consider what the effect on the combined mean would be.

Write the appropriate letter in the table for each situation.

- A** The combined mean would increase.
B The combined mean would decrease.
C The combined mean would stay the same.
D There is not enough information to know the effect on the combined mean.

<i>Situation</i>	<i>Effect on the combined mean</i>
Another nest containing 0 birds is found in the eastern section of the forest	
Another nest containing 2 birds is found in the western section of the forest	
All the nests containing 0 birds are removed from the calculation	

[3]





Permission to reproduce items where third-party owned material protected by copyright is included has been sought and cleared where possible. Every reasonable effort has been made by the publisher (UCLES) to trace copyright holders, but if any items requiring clearance have unwittingly been included, the publisher will be pleased to make amends at the earliest possible opportunity.

To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced online in the Cambridge Assessment International Education Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download at www.cambridgeinternational.org after the live examination series.

Cambridge Assessment International Education is part of Cambridge Assessment. Cambridge Assessment is the brand name of the University of Cambridge Local Examinations Syndicate (UCLES), which is a department of the University of Cambridge.

