Cambridge Assessment International Education

Cambridge O Level

CANDIDATE NAME

CENTRE NUMBER | | | | | |

CANDIDATE NUMBER | | | | |

STATISTICS **4040/23**

Paper 2 **October/November 2025**

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.

[Turn over

1 A museum curator collected data about the visitors to the museum.
The data were grouped into classes labelled 0–9, 10–19, 20–29 etc.

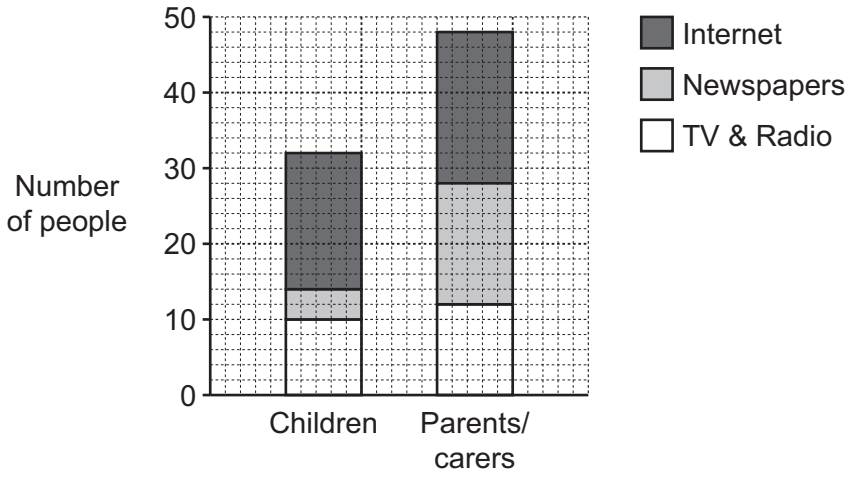
Complete the table for the '10–19' class for each of the three variables shown.

Variable	*Lower class boundary*	*Upper class boundary*	*Mid-point*	*Class width*
Number of visitors that arrive each hour				
Age, in completed years, of each visitor				
Length of time, to the nearest minute, each visitor spends in the museum				

[6]

2 Neo wants to compare where adults and children find out about the news. He asks the children in his class and their parents/carers. They must choose one of Internet, Newspapers or TV & Radio.

The sectional bar chart shows his results.



(a) Find

(i) the number of the children who said Internet,

..... [2]

(ii) the percentage of the parents/carers who said Newspapers.

..... [2]

(b) Neo concludes that adults are less likely than children to use TV & Radio to find out about the news.

(i) Make **two** comments about Neo's conclusion.

1 .....

.....

2 .....

.....

[2]

(ii) Name an alternative type of chart that would have been more convenient for Neo to use.

..... [1]

3 A taxi driver records the distances of the journeys for each of her first 100 customers.

Distance, x (km)	*Number of journeys*
$1 \leq x < 3$	14
$3 \leq x < 5$	23
$5 \leq x < 7.5$	25
$7.5 \leq x < 10$	18
$10 \leq x < 15$	10
$15 \leq x < 20$	8
$20 \leq x < 30$	2

(a) Use linear interpolation to calculate an estimate for the median of these distances.

..... [3]

For a journey, a customer is charged \$1.50 plus \$2 per km.
For example, a journey of 3.65 km would cost \$1.50 + 3.65 × \$2 = \$8.80 .

An estimate of the interquartile range of these distances is 5.35 km.

(b) Find an estimate for the interquartile range of the amounts charged for these first 100 journeys.

..... [1]

The taxi driver decides to reduce the rate charged for journeys longer than 18 km.

(c) Use linear interpolation to calculate an estimate for the number of these first 100 journeys that would have been charged at a reduced rate.

..... [3]

4 Two bags, one green and one red, each contain an unknown number of discs.
Each disc is numbered 1, 2, 3 or 4, and there is at least one disc with each number in each bag.
A disc is taken from each bag.

Here are some possible events.

A A disc with an **even** number is taken from the **green** bag.
B A disc numbered **3** is taken from the **green** bag.
C A disc with an **odd** number is taken from the **red** bag.
D A disc numbered **3** is taken from the **red** bag.

(a) Write down all the pairs of independent events and all the pairs of mutually exclusive events.

Independent events ……...

Mutually exclusive events ……
[3]

The probabilities of two of the events above are known.

$P(B) = 0.3$ $P(C) = 0.6$

(b) Find P(*B* and *C*) and P(*B* or *C*).

P(*B* and *C*) = ...

P(*B* or *C*) = ...
[4]

(c) Find the smallest possible total number of discs in the red bag.

..... [1]

5 The scores obtained by the competitors in a dancing competition are recorded.
To compare with competitors in another competition, these scores are scaled.

	Mean	Standard deviation
Recorded scores	23	6
Scaled scores	30	9

Asha receives a recorded score of 21.

(a) Find Asha's scaled score.

..... [2]

Pablo's score does not change when it is scaled.

(b) Find Pablo's recorded score.

..... [2]

Later, a different scaling was used with a scaled mean of 45.
With this scaling, a recorded score of 43 became a scaled score of 80.

(c) Find the standard deviation for these new scaled scores.

..... [2]

DO NOT WRITE IN THIS MARGIN
DO NOT WRITE IN THIS MARGIN
DO NOT WRITE IN THIS MARGIN
DO NOT WRITE IN THIS MARGIN

6 There are 360 vehicles for sale on a website. Lucy wants to check the service history of a simple random sample of the vehicles. She numbers the vehicles from 000 to 359.

(a) Use the random number table below to select a simple random sample of **size 4**.
Start on the left of the table.

24 80 74 15 39 48 12 83 47 28 51 36 05 26

..... [2]

She wants to do further checks on a sample of the 72 cars on the website.
Each car uses either Petrol or Diesel fuel and has either a Manual or Automatic gearbox.
The table shows the number of cars in each category.

	Manual	Automatic
Petrol	36	12
Diesel	24	0

She numbers the cars as shown.

	Manual	Automatic
Petrol	00–35	36–47
Diesel	48–71	

(b) She decides to take a sample of **size 6**, stratified by gearbox type.

(i) Use the random number table below to select the sample.
Start on the left of the table, and use every number if the gearbox type to which it relates has not yet been fully sampled.

49 08 42 39 94 08 53 61 27 32 65 23 06 50 64

..... [3]

(ii) Show whether or not your sample is representative of the Diesel cars.

[3]

7 The back-to-back stem-and-leaf diagram shows the percentages of populations that lived in urban areas in 11 southern African countries in the years 1960 and 2020.
For example, in one of these countries in 1960, 3% of the population lived in urban areas.

1960						2020			
7	5	4	4	3	**0**				
8	8	3	1	0	**1**	7			
					2	9			
					3	2	5	7	9
				7	**4**	5			
					5	2			
					6	7	7		
					7	1			

Key: 0 | **1** | 7 represents .....

.....

.....

(a) Complete the key for the back-to-back stem-and-leaf diagram.

[3]

(b) **(i)** Find the median and the range for 1960 and for 2020 and complete the table.

	1960	2020
Median		
Range		

[3]

(ii) For the data from 1960, explain whether or not you think the median is an appropriate measure of central tendency and the range is an appropriate measure of spread.

Median .....

.....

Range .....

.....

[2]

Anjali says, 'I can see from this back-to-back stem-and-leaf diagram that the percentage of the population in urban areas has increased in each of these 11 countries between 1960 and 2020.'

(c) State with a reason whether or not you think that she can use this back-to-back stem-and-leaf diagram to make this statement.

.....

..... [1]

Anjali wants to collect more data from three of these countries, chosen at random.

(d) Find the probability that exactly two of the three countries had more than half of their population living in urban areas in 2020.

..... [4]

8 Kopano manages a small hotel. His costs can be divided into 4 categories:
Buildings (rent, repairs etc.), Bills (electricity, internet, etc.), Food and Other.
Weights based on expenditure in 2022 and a price relative taking 2022 as the base year are shown in the table.

Category	*Weight*	*Price relative*	
		2022	2023
Buildings	12		
Bills	6		
Food	5		
Other	3		98.5

Between 2022 and 2023:

- the cost of Buildings increased from $350 per month to $371 per month
- the cost of Bills increased by 3.5%
- the cost of Food did not change.

(a) Complete the table with the remaining price relatives for 2022 and 2023.

[5]

(b) Explain what the price relative of 98.5 for Other tells you.

.....

.....

..... [2]

(c) Calculate a weighted aggregate cost index for 2023, taking 2022 as the base year.

..... [3]

The total expenditure for Kopano in 2022 was $9100.

(d) (i) Use the index calculated in **(c)** to estimate Kopano's expenditure in 2023.
Give your answer to the nearest $10.

..... [2]

(ii) This estimate might be very different from the true expenditure in 2023 if the weights have changed.

Give two possible reasons why the weights may have changed.

Reason 1 .....

.....

Reason 2 .....

.....
[2]

In 2024 the weights based on expenditure for the four categories had changed.

Category	*Weight*	
	2022	2024
Buildings	12	10
Bills	6	7
Food	5	5
Other	3	1

(e) Explain whether or not there has been a change in the relative weight of Food in 2024 compared to 2022.

.....

..... [1]

9 A game consists of throwing a ball at a target with two sections, labelled \$1 and \$2.
The probability of the ball hitting each section is shown in the table.

Section	\$1	\$2
Probability	0.25	0.15

(a) Find the probability that the ball misses the target.

..... [1]

A player throws a ball at the target.
If the ball hits the target, the player wins the amount shown on the section that the ball hits and the game ends.
If the ball misses the target, the player throws the ball one more time.
If the ball hits the target on the second throw, the player wins the amount shown on the section that the ball hits and the game ends.
If the ball misses the target twice, the player wins \$0 and the game ends.

(b) (i) Construct a probability distribution table for the amounts that a player can win.

[5]

(ii) State an assumption that you have made in giving your answer to **(b)(i)**.

.....

..... [1]

(iii) Find the amount that should be charged if this is to be a fair game.

..... [2]

In a second game, players are given two balls to throw at a target. The target has two sections labelled \$1 and \$2, and the ball cannot miss the target.
The probabilities of hitting each section with the first ball are shown in the table.

Section	\$1	\$2
Probability	0.65	0.35

If the player hits the \$1 section with the first ball, then the probabilities remain the same for the second ball.
If the player hits the \$2 section with the first ball, then the probability of hitting the \$2 section with the second ball increases to 0.75.
The player wins the total of the two amounts shown on the sections that the balls hit.

(c) (i) Find the expected amount that a player will win.

..... [5]

20 players each pay \$3 for two balls to throw at the target.

(ii) Find the expected profit or loss for the owner of the game.

..... [2]

10 Jamal collects data about the number of international flights that depart from Stats City airport each quarter.

Year and quarter	*Number of international flights*	*4-point moving average*	*Centred 4-point moving average*
2022 Q3	3460		
2022 Q4	$x =$		
		2461	
2023 Q1	1750		2463.5
		2466	
2023 Q2	2682		2472.5
		$y =$	
2023 Q3	3480		2480
		2481	
2023 Q4	2004		$z =$ .....
		2496	
2024 Q1	1758		2507.25
		2518.5	
2024 Q2	2742		2522
		2525.5	
2024 Q3	3570		
2024 Q4	2032		

(a) Use statistical language to describe fully the type of data that Jamal collects.

..... [2]

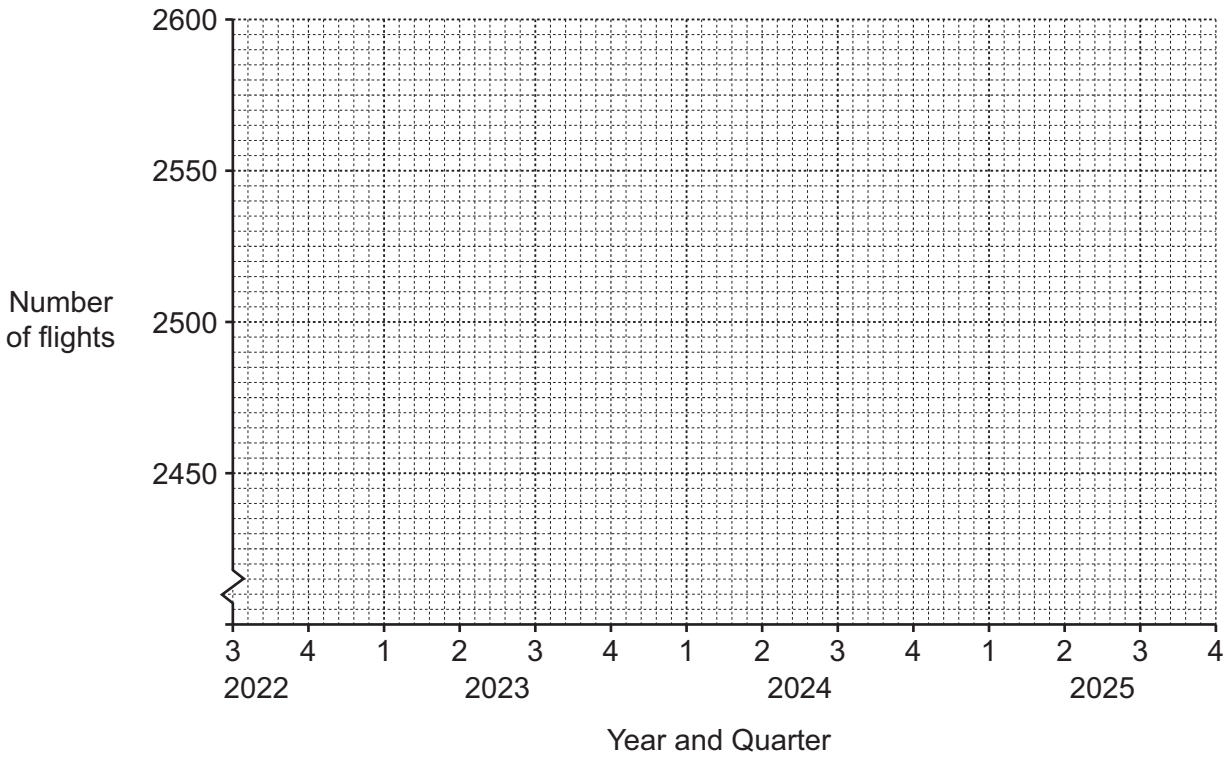
(b) Find the values of x, y and z and insert them in the table.

[3]

(c) Use appropriate values from the table to find an estimate for the seasonal component for quarter 1.

..... [3]

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



[3]

The seasonal component for international flights in quarter 2 is 215.

(e) Use your answer to **(d)** and this seasonal component to estimate the number of international flights from Stats City airport in quarter 2 of 2025.

..... [2]

[Question 10 continues on the next page]

Jamal also collects data about domestic flights from Stats City airport. The table shows the seasonal components for quarter 3.

	Seasonal component for Q3
International flights	1000
Domestic flights	−12

Jamal claims that this table shows that 'There are typically more international flights than domestic flights from Stats City airport in quarter 3.'

(f) Explain why this cannot be concluded from the data in this table.

.....

.....

..... [1]