Cambridge Assessment International Education

Cambridge O Level

CANDIDATE NAME

CENTRE NUMBER

CANDIDATE NUMBER

4187645222

STATISTICS **4040/22**

Paper 2 **October/November 2021**

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.

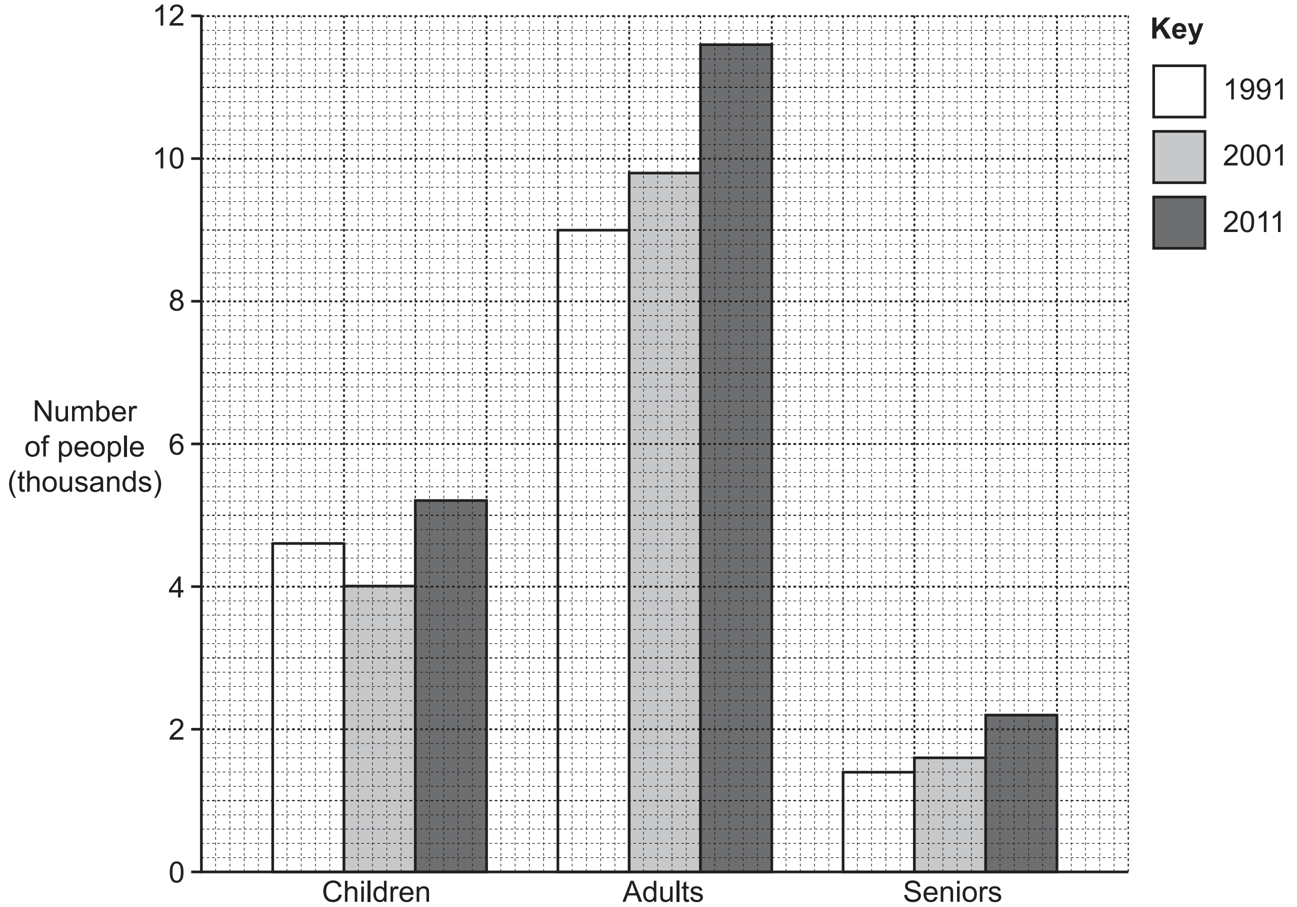
INFORMATION

- 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. Any blank pages are indicated.

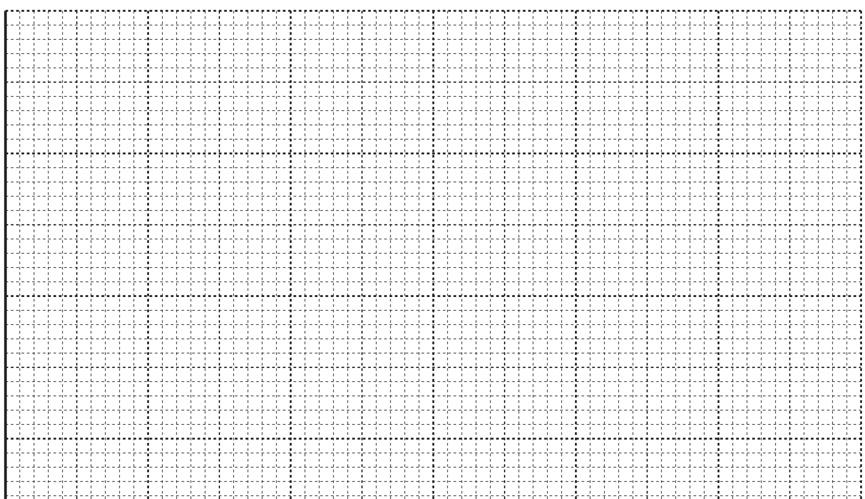
DC (CE/CGW) 206410/1

1 The population of a town is counted every ten years. The population is divided into three age groups: children, adults and seniors.



Azeeb claims that 'there has been a steady increase in the overall population of this town in the period 1991–2011'.

(a) Using the information in the chart above, draw a sectional bar chart on the grid below to help show whether or not Azeeb's claim is correct.



[4]

(b) Comment on whether or not you think that Azeeb's claim has been shown to be correct.

....

....

.... [2]

2 A teacher wants to find out what the students in his class think about a new textbook. He gives a sample of the 30 students in his class a questionnaire about the textbook.

(a) Define the population in this situation.

.... [1]

The students are each given a number from 01 to 30.

(b) Use the random number table below, starting at the beginning of the table, to select a simple random sample of **size 6**.

47 21 15 74 21 84 09 10 28 53 02 68 27 36

.... [2]

(c) Use the random number table below to help select a systematic sample of **size 6**.

52 04 91 23 86 30 19 47 07 72 00 24 56 08

.... [3]

3 A discrete variable V, with frequency distribution below, has mean 2.96 and standard deviation 0.82 .

V	1	2	3	4
Frequency	p	q	r	s

(a) Find the mean and standard deviation for each of the discrete variables W and X, with frequency distributions below.

W	10	11	12	13
Frequency	p	q	r	s

Mean ....

Standard deviation ....

X	10	20	30	40
Frequency	p	q	r	s

Mean ....

Standard deviation ....
[4]

The continuous variable Y, with frequency distribution below, represents measurements, in centimetres, to the nearest centimetre.

Y	10–14	15–19	20–24	25–29
Frequency	p	q	r	s

(b) For the 15–19 class, find the lower class boundary, the upper class boundary, the mid-point and the class interval.

Lower class boundary ....

Upper class boundary ....

Mid-point ....

Class interval ....
[2]

(c) Find an estimate for the mean and standard deviation of the continuous variable Y.

Mean ....

Standard deviation ....
[2]

4 Two friends Raashida and Takala count the number of characters in the last 12 text messages that each of them sent.
Here are the results.

Messages sent by Raashida						*Messages sent by Takala*					
18	29	14	34	23	11	33	29	12	28	1	65
4	22	23	18	13	10	2	9	17	13	9	40
Mean = 18.25 characters						Mean = 21.5 characters					

(a) Draw a back-to-back stem-and-leaf diagram to help compare the numbers of characters in the text messages of the two friends.

[5]

Takala compares the means and says that her messages are longer than Raashida's.

(b) Using your stem-and-leaf diagram, comment on whether or not you think Takala is correct.

....

....

.... [2]

5 A class takes a test and the marks, x, are such that

$$\text{Mean of } x = 66 \qquad \Sigma x = 2046 \qquad \Sigma x^2 = 138\,776$$

(a) Show that the standard deviation of the marks for the class is 11.0, correct to three significant figures.

[4]

The marks are to be scaled so that they have a mean of 60 and a standard deviation of 15.

(b) Find the mark that will remain unchanged by the scaling process.

.... [2]

6 A team plays 80% of its matches wearing its red kit.

The team may win, draw or lose a match with the following probabilities, depending on whether they are wearing their red kit or not.

	Wearing red kit	Not wearing red kit
Win	50%	40%
Draw	30%	25%
Lose	20%	35%

The team will score 3 points if they win, 1 point if they draw, and 0 points if they lose.

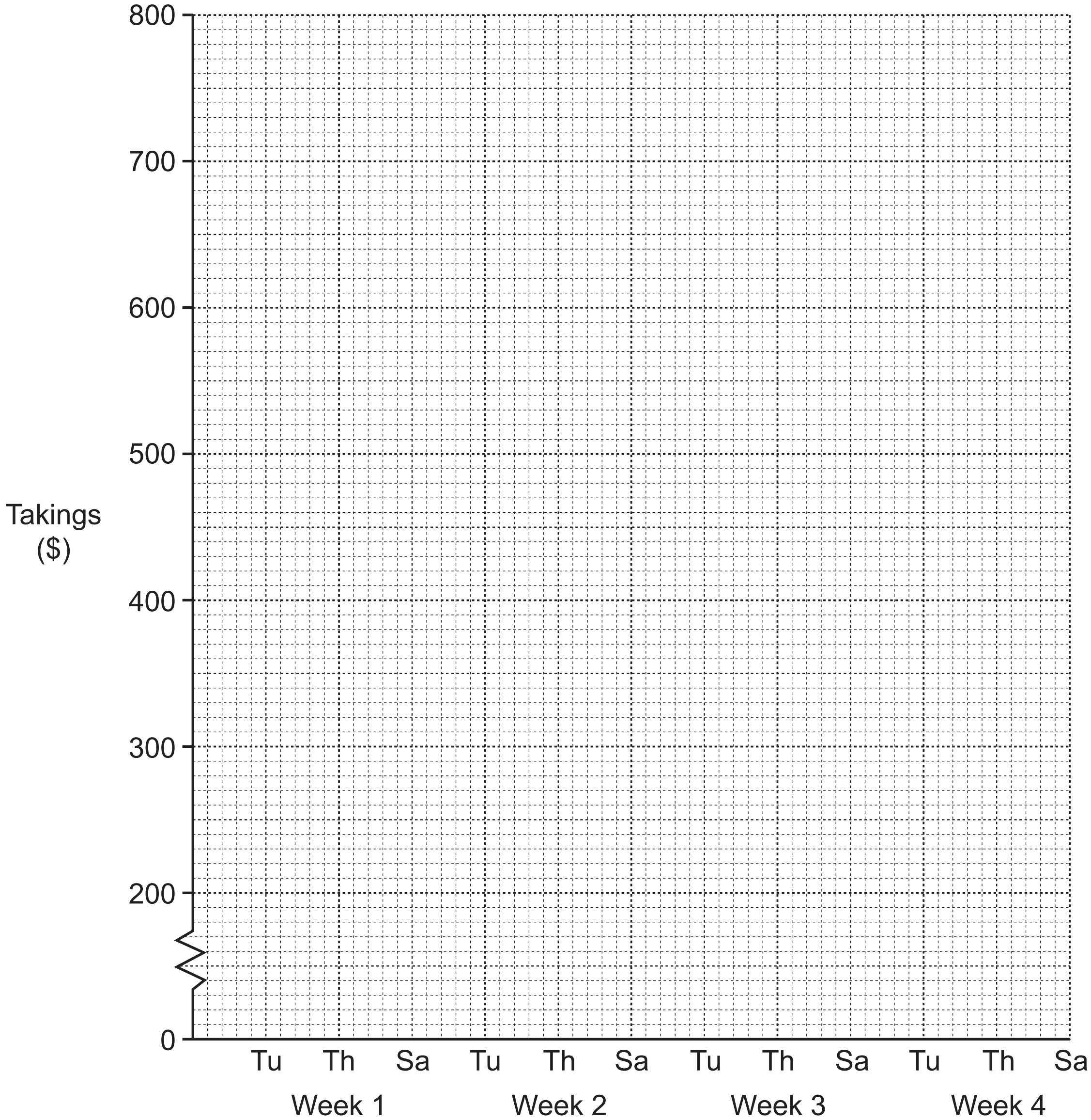
Calculate the expected number of points scored per match for this team.

.... [5]

7 A market stall holder sells clothes on three days each week, Tuesday (Tu), Thursday (Th) and Saturday (Sa).
His takings, in dollars ($), over a three-week period are shown below.

	Tuesday	Thursday	Saturday
Week 1	320	351	721
Week 2	298	343	709
Week 3	308	312	691

(a) Plot a time series graph for this information on the grid below.



[3]

(b) Calculate all the 3-point moving average values and insert them in the appropriate parts of the table below.

	Tuesday	Thursday	Saturday
Week 1			
Week 2			
Week 3			

[3]

(c) Using the original data and your moving average values find an estimate for the seasonal component for Thursday.

.... [3]

(d) Plot the moving average values on the graph on page 8, and draw an appropriate trend line. [3]

(e) Use your trend line and answer to part **(c)** to estimate the takings on the Thursday of Week 4.

.... [2]

(f) State an assumption that you have made in reaching your answer to part **(e)**.

....

.... [1]

8 Data has been collected from 80 cars.

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

Data	*Qualitative*	*Discrete quantitative*	*Continuous quantitative*
The length of each car			
The manufacturer of each car			
The number of airbags in each car			
The weight of each car			

[3]

The fuel consumption, *c*, measured in litres per 100 kilometres (l/100 km), was found for each of these 80 cars, and is summarised in the table below.

Fuel consumption, c, (l/100 km)	*Cumulative frequency*
$c < 7.0$	7
$c < 10.0$	22
$c < 10.5$	45
$c < 11.0$	60
$c < 12.0$	73
$c < 14.0$	80

(b) Use linear interpolation to find an estimate for the percentage of these cars that have a fuel consumption of more than 9 l/100 km.

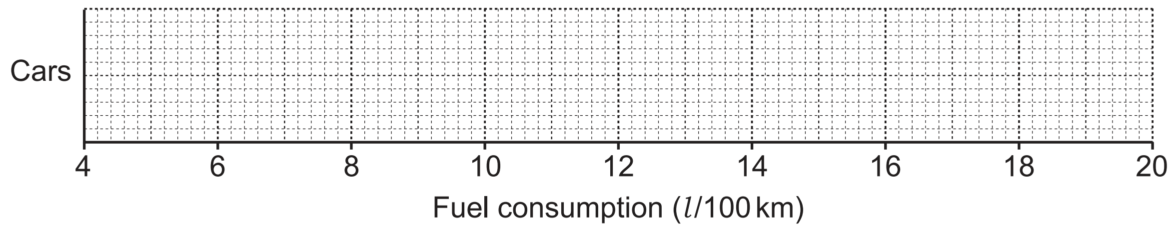
.... [4]

(c) Use linear interpolation to find an estimate for the median of the fuel consumptions of these 80 cars.

.... [4]

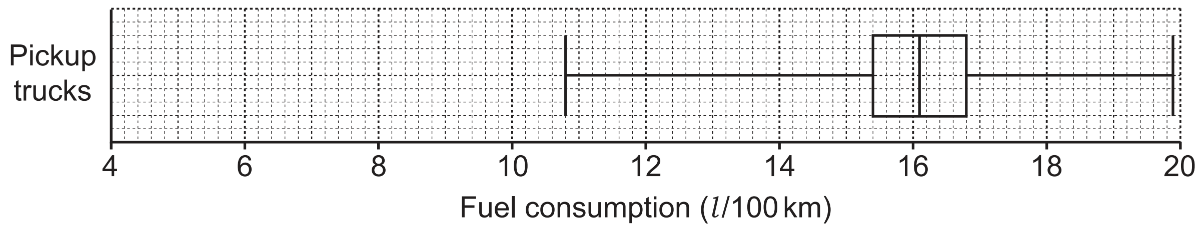
For these 80 cars, the lowest fuel consumption was 4.6 l/100 km, the range was 9.1 l/100 km, the upper quartile was 11.0 l/100 km and the interquartile range was 1.4 l/100 km.

(d) Using your answer to part **(c)** and the information above, draw a box-and-whisker diagram for the cars, on the grid below.



[3]

The fuel consumptions of 80 pickup trucks were also found and are summarised in the box-and-whisker diagram below.



(e) Using the box-and-whisker diagrams, make two comparisons between the fuel consumptions for the cars and the fuel consumptions for the pickup trucks.

1 ....

....

2 ....

....

[2]

9 A cinema manager divides her total expenditure into four categories: Licencing, Wages, Electricity and Other.
The expenditure for each category in 2016 is shown below.

Licencing	\$22 000
Wages	\$36 000
Electricity	\$ 8 000
Other	\$10 000

Some of the price relatives for 2019 and 2020, taking 2016 as the base year, are shown in the table below.

	2019	2020
Licencing	107	113
Wages	111	108
Electricity	100	....
Other	102	....

(a) Explain what the 100 in the table tells you.

....

....

.... [2]

To find an estimate for the expenditure in 2019, the manager does the following calculation:

$$\text{Mean price relative for 2019} = \frac{107 + 111 + 100 + 102}{4} = 105$$

$$\text{Estimate for expenditure in 2019} = \$76\,000 \times \frac{105}{100} = \$79\,800$$

(b) Explain why this is **not** likely to be a good estimate for the total expenditure in 2019.

....

....

.... [2]

The cost of Electricity increased from \$0.12 per unit in 2016 to \$0.15 per unit in 2020.
Costs in the Other category decreased by 1% between 2016 and 2020.

(c) Use this information to complete the table above part **(a)**.

[3]

(d) **(i)** Using expenditure in 2016 for weights, calculate, correct to one decimal place, a weighted aggregate cost index for 2020, taking 2016 as the base year.

.... [3]

(ii) Hence find an estimate for the total expenditure in 2020.

.... [3]

Later, an audit of the expenditure in 2020 showed that an estimate calculated using the method in part **(d)** was inaccurate.

The manager considered some possible explanations.

A The weights had changed between 2016 and 2020.
B The average wage had increased by 8% between 2016 and 2020.
C The number of employees had increased between 2016 and 2020.
D The cost of licencing had increased by 13% between 2016 and 2020.

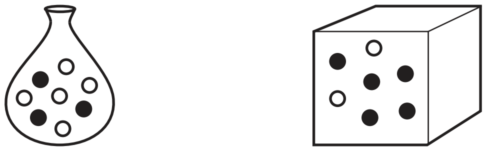
(e) State, with a reason, which **two** of the above are **not** possible explanations for the inaccurate estimate.

....

....

.... [2]

10 There are two containers, a bag and a box.



The bag contains 3 black counters and 5 white counters.
The box contains 5 black counters and 2 white counters.

(a) A counter is selected at random from each container.

Some of the possible outcomes are listed below.

A A black counter is selected from the bag.
B A white counter is selected from the box.
C Both counters selected are black.
D Both counters selected are white.

From this list, state all the possible pairs of mutually exclusive events and all the possible pairs of independent events.

Mutually exclusive events ....

Independent events ....
[3]

(b) Find

(i) $P(A \cap D)$,

.... [1]

(ii) $P(A \cap B)$,

.... [2]

(iii) $P(B \cup C)$.

.... [3]

The counters are returned to their original containers, as shown on page 14.

(c) Three counters are selected at random and removed from the bag.

Find the probability that exactly one of the selected counters is white.

.... [3]

These counters are now returned to the bag, as shown on page 14.

(d) Three counters are selected at random from the bag and placed in the box. Then two counters are selected at random from the box and placed in the bag.

Find the probability that, at the end of this process, the bag and the box each contain counters of just one colour.
Give your answer as a fraction in its lowest terms.

.... [4]

BLANK PAGE