Cambridge
O Level

Cambridge International Examinations
General Certificate of Education Ordinary Level

CANDIDATE NAME	
CENTRE NUMBER	
CANDIDATE NUMBER	

0123456789

STATISTICS **4040/02**

Paper 2 **For Examination from 2018**

SPECIMEN PAPER

2 hours 15 minutes

Candidates answer on the Question Paper.

Additional Materials: Pair of compasses
Protractor
Electronic calculator

READ THESE INSTRUCTIONS FIRST

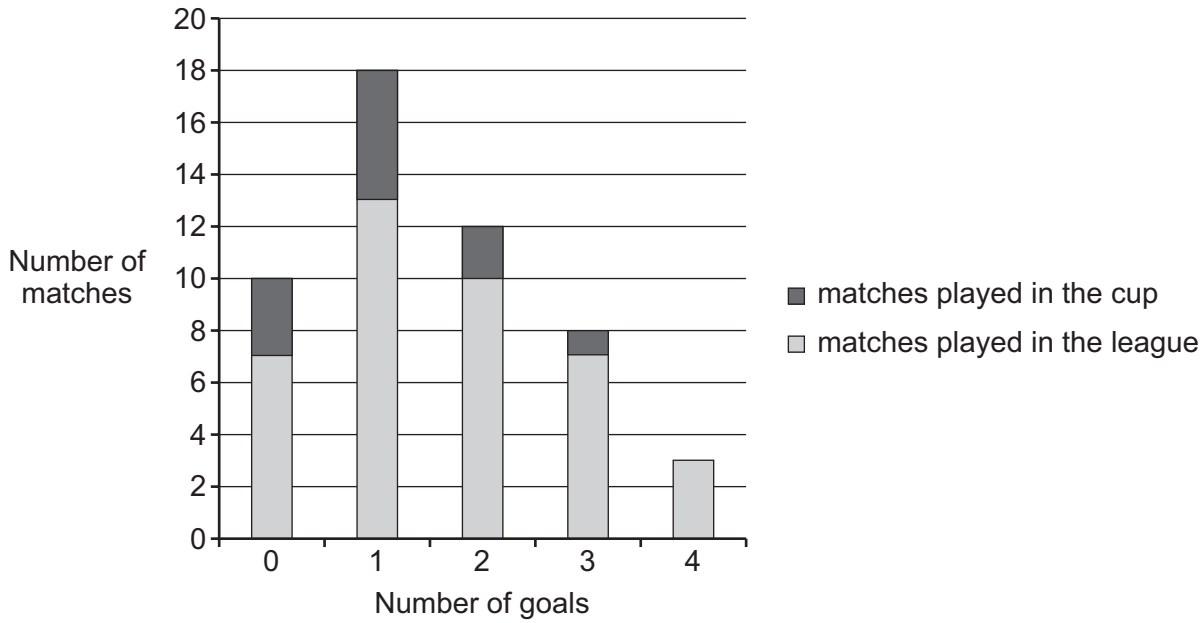
Write your Centre number, candidate number and name in the spaces at the top of this page.
Write in dark blue or black pen.
You may use an HB pencil for any diagrams or graphs.
Do not use staples, paper clips, glue or correction fluid.
DO **NOT** WRITE IN ANY BARCODES.

Answer **all** questions.
If working is needed for any question it must be shown below that question.
Essential working must be shown for full marks to be awarded.
Electronic calculators should be used.

At the end of the examination, fasten all your work securely together.
The number of marks is given in brackets [ ] at the end of each question or part question.
The total number of marks for this paper is 100.

This document consists of **17** printed pages and **1** blank page.

1 A football team used the diagram below to illustrate the number of goals it had scored per match in a season in both the league and cup competitions.



(a) State the full name given to this type of diagram.

.....[1]

(b) State whether the variable 'number of goals scored per match' is discrete or continuous and whether it is qualitative or quantitative.

.....

..... [1]

(c) Find the proportion of matches played in the cup in which the team scored 2 or more goals.

..... [2]

2 Events A, B, C and D are four of the possible outcomes of an experiment such that

$$P(A) = 0.15 , \quad P(B) = 0.2 , \quad P(C) = 0.4 \quad \text{and} \quad P(D) = 0.24 .$$

(a) If events A and B are independent, find

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

..... [2]

(ii) $P(A \cup B)$.

..... [2]

(b) If events C and D are mutually exclusive, find

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

..... [1]

(ii) $P(C \cup D)$.

..... [1]

3 **(a)** The annual salaries of the employees at a company have a mean of $\$m$ and a standard deviation of $\$s$, where $s \neq 0$.
A new employee arrives at the company and is paid an annual salary of $\$m$.
The mean and standard deviation of the salaries of the employees are now recalculated to include the salary of the new employee.

For each of the mean and the standard deviation, state whether it will increase, decrease, or stay the same when this new employee's salary is included.

Mean .....

Standard deviation ..... [2]

(b) At another company, at the end of 2013, the employees' annual salaries had a mean of \$12 000 and a standard deviation of \$1000.
During 2014, each of the employees' salaries increased by 5%. At the end of that year they each also received an annual bonus of \$200.

Calculate the mean and standard deviation of the annual incomes (salaries plus bonuses) of the employees at the end of 2014.

Mean \$ .....

Standard deviation \$ ..... [4]

4 Ariana and Bella are playing a game.
They each have 4 cards, which are numbered 1, 2, 3 and 4.
Each shuffles her own cards and turns one over at random.

(a) If the cards show the same number, Ariana wins and Bella must pay Ariana $3.
If the cards show different numbers, Bella wins and Ariana must pay Bella $1.

By finding the probabilities of Ariana and Bella winning, show whether or not the game is fair.

[3]

(b) In a second game the numbers shown on the cards are added together.
If the total is 4 or less, Ariana wins and Bella must pay Ariana $5.
If the total is 5 or more, Bella wins.

If the game is to be fair, how much should Ariana pay Bella if Bella wins?

$..... [3]

5 A small village has a population of 60 people aged 10 and over.
There is a proposal to change the timetable for the buses that pass through the village. A group of researchers wish to find out what the people of the village think about the proposed changes. Each researcher has a population list and thinks of a different way to select a sample.

(a) The first researcher plans to stand at the village bus stop at 7 am on a Monday morning and ask the first six people from the population who come to wait for a bus. Explain why this might not produce a reliable sample.

.....

.....

..... [2]

(b) A second researcher decides to take a simple random sample of size six, and numbers the population list from 00 to 59.

Use the random number table below to select a simple random sample of size six from the population of 60 people. Start at the beginning of the first row and work along the row, ensuring that no one is selected more than once.

RANDOM NUMBER TABLE

15 08 73 00 60 15 31 52 86 47 82 99 04 33

23 05 65 27 46 13 81 50 49 34 29 08 94 72

..... [2]

The table below shows the population, split into three different age groups.

	10–18 years	19–65 years	66 years and over	TOTAL
Number of people	20	30	10	60

(c) A third researcher decides to take a random sample of size six, stratified by age group.

(i) State how many people from each age group would be needed for such a sample.

10–18 years .....

19–65 years .....

66 years and over ..... [1]

(ii) Explain why a random sample, stratified by age group, might be a good idea in this situation.

.....

.....

..... [1]

6 The pupils in a class should arrive for registration at 9.00 am. On one particular day, 25 pupils were early, with a mean arrival time of 8.51 am. On the same day, 9 pupils were late, with a mean arrival time of 9.21 am, and 2 pupils arrived at 9.00 am exactly.

If x represents the number of minutes a pupil was late (a pupil who was early would have a negative value of x),

(a) find Σx, and hence find the mean arrival time for all 36 pupils.

$\Sigma x =$ .....

Mean = ... [3]

If $\Sigma x^2 = 5096$ for the 36 pupils,

(b) find the standard deviation of x, correct to one decimal place.

..... [3]

7 A class contains 15 girls and 15 boys whose heights have been measured to the nearest centimetre (cm).
The heights of the 15 girls are given below.

Heights of the girls in cm

171 152 145 154 169 150 152 149 161 158 157 148 155 164 163

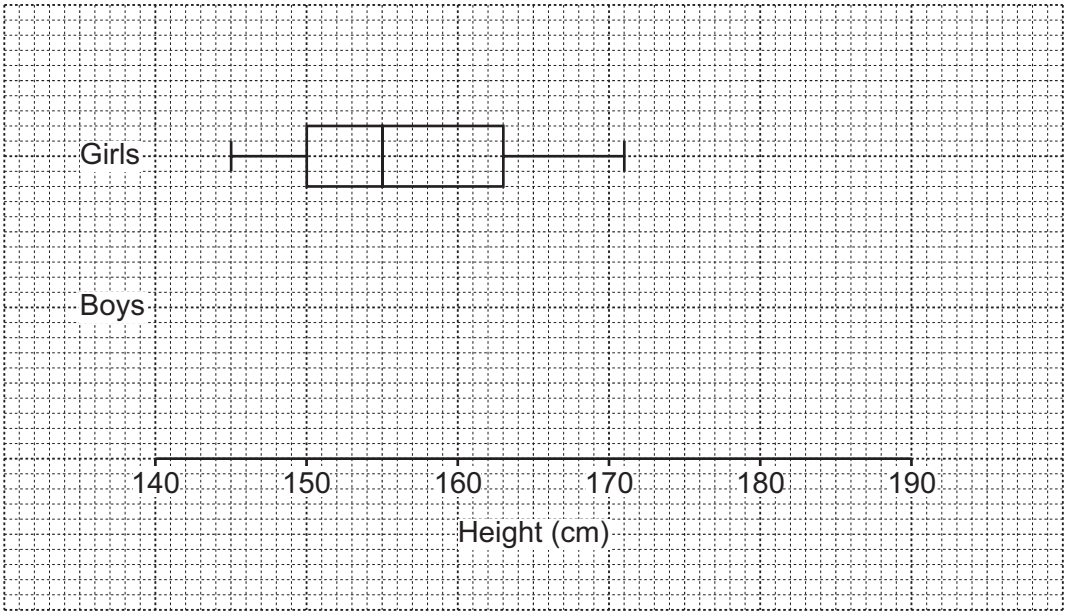
(a) Complete the ordered back-to-back stem-and-leaf diagram to include the data for the heights of the 15 girls.

Girls' heights		Boys' heights
	14	
	15	8 8 9
	16	4 4 6 8
	17	0 2 5 6 7 9
	18	1 4

Key: 0 | 15 | 8 represents 150 cm for girls and 158 cm for boys

[3]

(b) Complete the box-and-whisker diagram to display the data for the boys.



[3]

(c) Use the box-and-whisker diagram to make two comparisons between the heights of the boys and the heights of the girls.

.....

..... [2]

8 A fisherman recorded, in grams (g), to the nearest 100 grams, the masses of 100 fish he had caught in river *A*.

Mass of fish (grams)	*Number of fish*	*Cumulative frequency*
100–200	12	
300–400	31	
500–700	29	
800–1000	14	
1100–1400	8	
1500–2000	4	
2100–3000	2	

(a) State, with a reason, which of the mean or the median would be the more appropriate measure of central tendency for this set of data.

.....

..... [2]

(b) Find the cumulative frequencies and complete the table above. [1]

(c) State the class boundaries of the 300–400 class.

Lower class boundary .....

Upper class boundary ..... [1]

(d) Use linear interpolation to calculate an estimate of the interquartile range of the masses of the fish.

..... [6]

(e) The fisherman also recorded the masses of 100 fish caught in river *B* and found the interquartile range of the masses of these fish to be 352 g. Explain what this tells you about the masses of the fish caught in river *B* compared to those caught in river *A*.

.....

.....[1]

(f) Use linear interpolation to calculate an estimate of the percentage of fish in river *A* with a mass of less than 650 g.

..... [3]

9 A hospital records the number of patients admitted at two-monthly intervals over a period of two years. The results are shown in the table below, together with the 6-point moving average values for these data.

		Number of patients	*6-point total*	*6-point moving average value*	*Centred moving average value*
2012	Jan – Feb	241			
	Mar – Apr	208			
	May – Jun	x = ...			
			1272	212	
	Jul – Aug	185			
			1290	215	
	Sep – Oct	209			
			1290	215	
	Nov – Dec	261			
			1296	216	
2013	Jan – Feb	259			
			y = ...	z = ...	
	Mar – Apr	208			
			1323	220.5	
	May – Jun	174			
			1332	222	
	Jul – Aug	197			
	Sep – Oct	224			
	Nov – Dec	270			

(a) Calculate the values of x, y and z and insert them in the table.

[3]

(b) Explain clearly why it is necessary to centre the moving average values in this case.

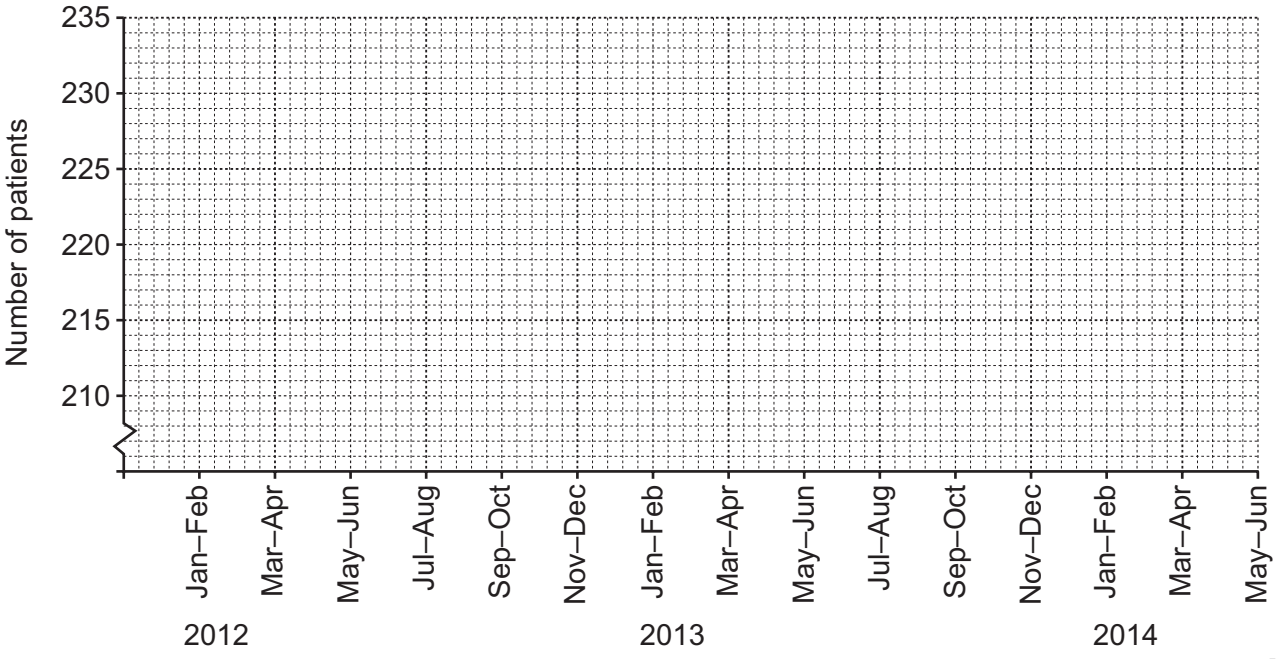
.....

..... [2]

(c) Calculate the centred moving average values and insert them in the appropriate places in the table.

[3]

(d) Plot the centred moving average values on the grid below and draw a trend line through the points.



[3]

(e) Explain what the trend line you have drawn tells you.

.....

..... [1]

The seasonal component for Mar–Apr is –11.25 .

(f) Estimate the number of patients admitted to the hospital during the period Mar–Apr 2014.

..... [2]

10 A hairdresser classifies the expenditure on her business into three categories: Rent, Equipment and Wages.

The cost of Rent has increased from $240 per month in 2012 to $256 per month in 2014.
The price relative of Equipment in 2014 is 110, taking 2012 as base year.
The hourly rate of the Wages of her employees has decreased by 2% between 2012 and 2014.

(a) (i) Calculate the price relative, to the nearest whole number, of Rent for 2014, taking 2012 as base year.

..... [2]

(ii) Explain what the price relative of 110 for Equipment indicates.

.....

.....

..... [3]

(iii) State the price relative of Wages for 2014, taking 2012 as base year.

..... [1]

The hairdresser wishes to calculate a weighted aggregate cost index, using weights calculated in 2012, for the three categories.

(b) (i) Briefly describe how these weights could be calculated.

.....

..... [1]

The weights in 2012 for Rent, Equipment and Wages were calculated as 7, 2 and 5 respectively.

(ii) Calculate, to the nearest integer, a weighted aggregate cost index for 2014, taking 2012 as base year.

..... [3]

(iii) Her total expenditure on the hairdressing business in 2012 came to $5760. Use your answer to part **(ii)** to estimate, to the nearest dollar, her total expenditure on the business in 2014.

..... [2]

(iv) Give two possible reasons why this estimate might be very inaccurate.

Reason 1 .....

.....

Reason 2 .....

..... [2]

11 The students at a college take one of three programmes of study: Physics, Chemistry and Mathematics (PCM) or Physics, Chemistry and Biology (PCB) or Economics, Geography and Mathematics (EGM). The numbers of students who study each programme are shown in the table below.

	PCM	PCB	EGM	TOTAL
Male	60	40	40	140
Female	40	90	30	160
TOTAL	100	130	70	300

(a) Find the probability that a student chosen at random

(i) is a male studying PCM,

..... [1]

(ii) is female,

..... [1]

(iii) is studying Physics as part of their programme,

..... [1]

(iv) is studying PCB, given that they are male.

..... [1]

(b) If two different students are chosen at random, find the probability that they are taking the same programme of study.

..... [3]

(c) If three different students are chosen at random, find the probability that they are each taking a different programme of study.

..... [3]

Students are required to buy textbooks for each subject that they study: one textbook for each of Physics, Chemistry and Biology and two textbooks for each of Mathematics, Economics and Geography.

(d) Find how many textbooks a student taking each programme of study must buy, and complete the table below.

Course	PCM	PCB	EGM
Number of textbooks			

[1]

(e) If one of the textbooks owned by a student at the college is selected at random, find the probability that it

(i) belongs to a student on the PCM programme,

..... [3]

(ii) is a Mathematics textbook.

..... [2]

BLANK PAGE