Cambridge O Level

CANDIDATE NAME	

CENTRE NUMBER		CANDIDATE NUMBER	
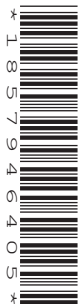
STATISTICS **4040/12**

Paper 1 **October/November 2019**

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 **18** printed pages and **2** blank pages.

Cambridge Assessment
International Education

[Turn over

1 A business organisation uses email for most correspondence, but also uses postal services for sending letters.

The pictogram below shows the number of letters sent during one particular month by each of three postal services: standard delivery, express delivery, and special delivery.

Standard delivery   ▣   ▣   ▣   ▣   ▣   ▣   ▣   ▣   ▣   ▣   ▣   ▣

Express delivery   ▣   ▣   ▣   ▣   ▣

Special delivery   ▣   ▣   ▣   ▣   ▣   ▣   ▣   ▣

▣ = 5 letters

For this month, for these postal services,

(a) state the number of letters sent by special delivery,

.... [1]

(b) state how many fewer letters were sent by express delivery than by standard delivery,

.... [1]

(c) calculate the percentage of letters sent by special delivery.

.... [2]

2 In a factory car park there are special covered parking spaces for use only by visitors to the factory.
The following data shows, for twelve consecutive working days, the number of these spaces occupied by vehicles at a certain time each day.

<div align="center">3   10   7   10   8   7   4   10   0   9   10   7</div>

For this data find

(a) the median,

.... [2]

(b) the mode.

.... [1]

At present there are 10 such parking spaces in the car park.

(c) State, giving a reason, whether or not you think the factory should provide more of these spaces for visitors, assuming that land is available for this purpose.

....

.... [1]

3 **(a)** Explain briefly what is meant by the statement that there is negative correlation between the variables in a bivariate distribution.

....

.... [1]

The variables in three bivariate distributions were measured and plotted. The following scatter diagrams were obtained.
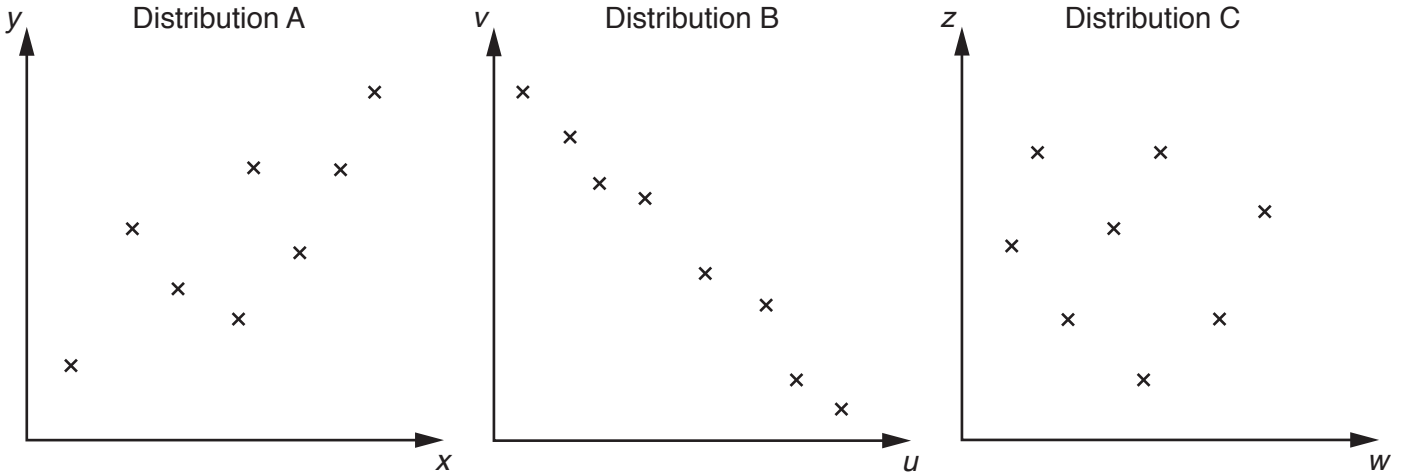
(b) Describe fully the correlation shown by the variables in each of these distributions.

Distribution A ....

Distribution B ....

Distribution C ....
[4]

4 Football matches may result in a win, draw or loss for a team. Matches are played either 'home' (on a team's own ground) or 'away' (on the ground of its opponent).

Last season, Mathtester United football team played 48 matches, of which 12 resulted in a draw, as shown in the following table.

	Win	Draw	Loss	TOTAL
Home				
Away				
TOTAL		12		48

The team played the same number of home matches as away matches.

(a) Use this information to insert two numbers into the table. [1]

Half of the home matches resulted in a win, and two-thirds of the away matches resulted in a loss.

(b) Use this information to insert two more numbers into the table. [1]

Of the matches that ended in a draw, there were two more home matches than away matches.

**(c)** Use this information to complete the table. [2]

The team earned 3 points for each win, 1 point for each draw, and 0 points for each loss.
At the start of the season the team's coach had aimed for a mean of at least 1.25 points per match as a minimum satisfactory level of performance.

(d) Determine whether or not she achieved her aim.

.... [2]

5 A job advertisement states that an applicant must have studied at least one of the subjects Architecture, Computing, and Management.
The diagram below represents the 52 applicants who meet this requirement.

Architecture

14

7 1
2

8 11 9

Computing   Management
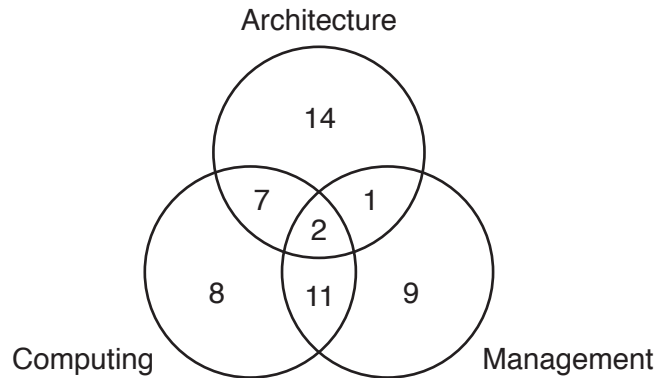
Use this information to find the number of applicants who have studied

(a) Architecture and Computing,

.... [1]

(b) Computing or Management or both.

.... [1]

The appointing committee studies the educational records of the 52 applicants and rejects 37 immediately.
The diagram below represents the subjects studied by the 37 rejected applicants.

11 5 6 7 8

Architecture   Computing   Management
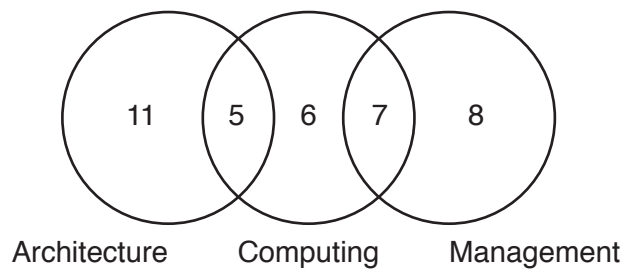
(c) Draw and label a Venn diagram to represent the subjects studied by the 15 applicants still under consideration for the job.

[3]

The applicants finally invited for interview have all studied Architecture and at least one other subject.

(d) Find the number of applicants invited for interview.

.... [1]

6 The headteacher of a school plans a survey to obtain the responses of pupils' parents to her proposal to start and finish daytime classes one hour earlier than at present. She considers the following survey methods.

A   Questionnaires will be posted to all the pupils' parents.

B   Telephone calls will be made to a randomly selected sample of the pupils' parents.

C   A questionnaire will be placed on the school's website inviting responses from the pupils' parents.

(a) Give one reason why

(i) method A might be preferred to method C,

....

.... [1]

(ii) method B might be preferred to method A,

....

.... [1]

(iii) method C might be preferred to method B.

....

.... [1]

(b) Design a closed question which could be used in the questionnaire in method A.

....

....

[2]

The headteacher plans to obtain also the responses of the pupils themselves, and the teachers, to her proposal.

(c) State, for which one of the groups, parents, pupils, or teachers, it will be easiest for her to obtain a census of the group's opinions. Explain your answer.

....

.... [1]

7 The table below gives information about the female population, and the number of births by age group, in a particular city for the year 2018.

Age group of females	Births	Population of females in age group
Under 25	340	6200
25–30	180	1300
31–40	220	2800
Over 40	90	5900

The fertility rate of a population is defined as the number of births per 1000 females.

(a) Calculate, to 1 decimal place, the crude fertility rate for the city in 2018.

.... [4]

The crude death rate for the city in 2018 was 8.7 per thousand of the population.

(b) Assuming that the city had equal numbers of males and females, estimate the increase in its population due to births and deaths combined in 2018.

.... [4]

   [Turn over

8 In a particular town, 140 households with internet access were questioned about the speed of their connections. The following table summarises the results obtained for download speeds, in megabits per second, Mbps.

Download speed (Mbps)	Number of households	Cumulative frequency
0–under 5	47	
5–under 10	17	
10–under 15	20	
15–under 20	42	
20–under 25	9	
25–under 30	3	
30–under 35	2	

(a) Complete the cumulative frequency column in the above table. [1]

(b) Plot the cumulative frequencies on the grid opposite, joining the points to form a cumulative frequency polygon. [3]
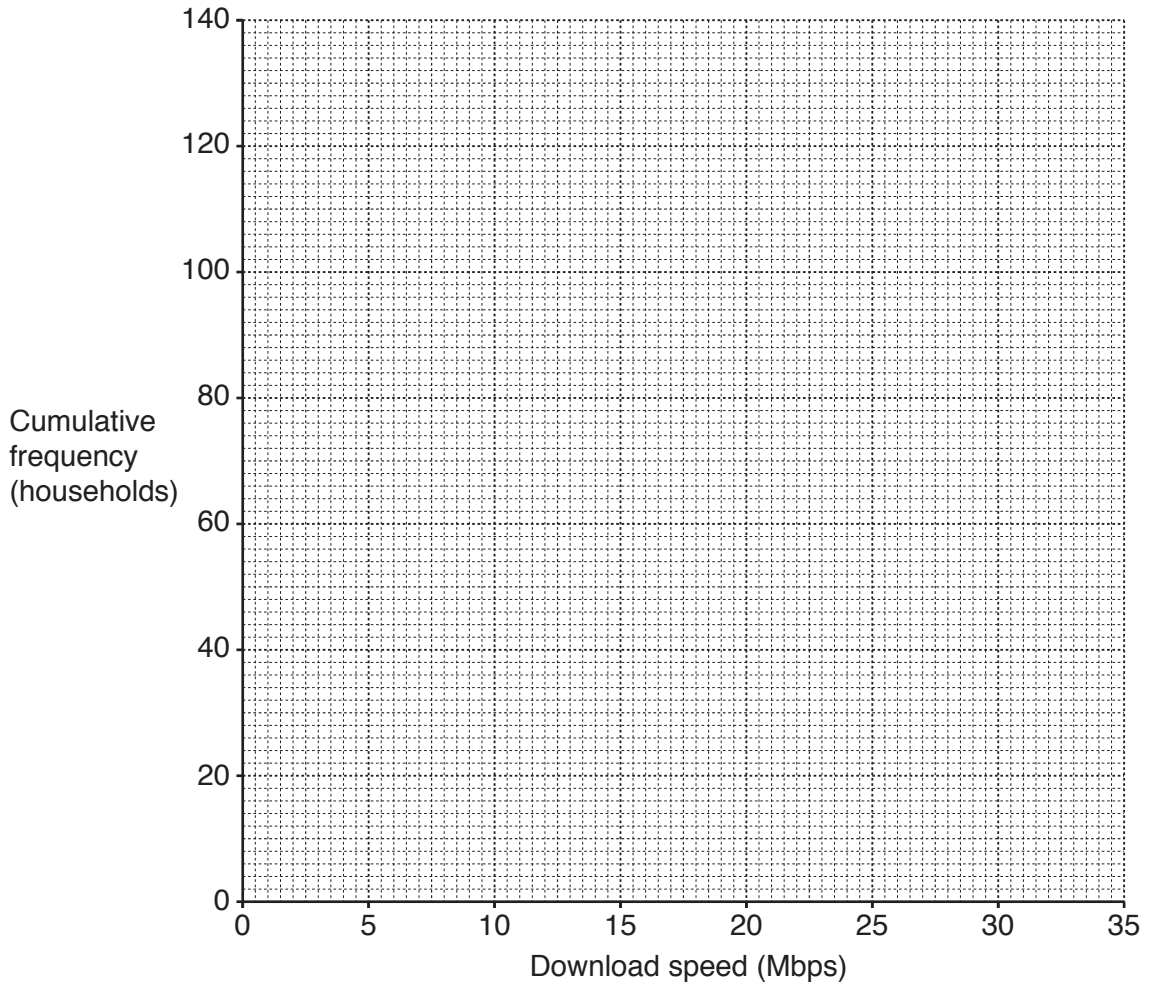
(c) Use your graph to estimate, for these speeds,

   (i) the median,

                            .... [1]

   (ii) the interquartile range, given that the lower quartile is 3.7 Mbps.

                            .... [3]

The government has two objectives for download speeds, for those with internet access.

Objective A: 60% of households should have at least 7 Mbps.
Objective B: 10% of households should have at least 22.5 Mbps.

(d) Determine, for this sample, whether or not these objectives are met.

.... [5]

(e) For estimating the measures in part **(c)**, explain why, in this case, you would expect the method of linear interpolation to give the same values as those obtained from the graph. (You are **not** required to recalculate the measures by this method.)

....

.... [1]

9 An office building contains offices of different sizes. The following histogram summarises the number of these offices, and their areas, in square metres.
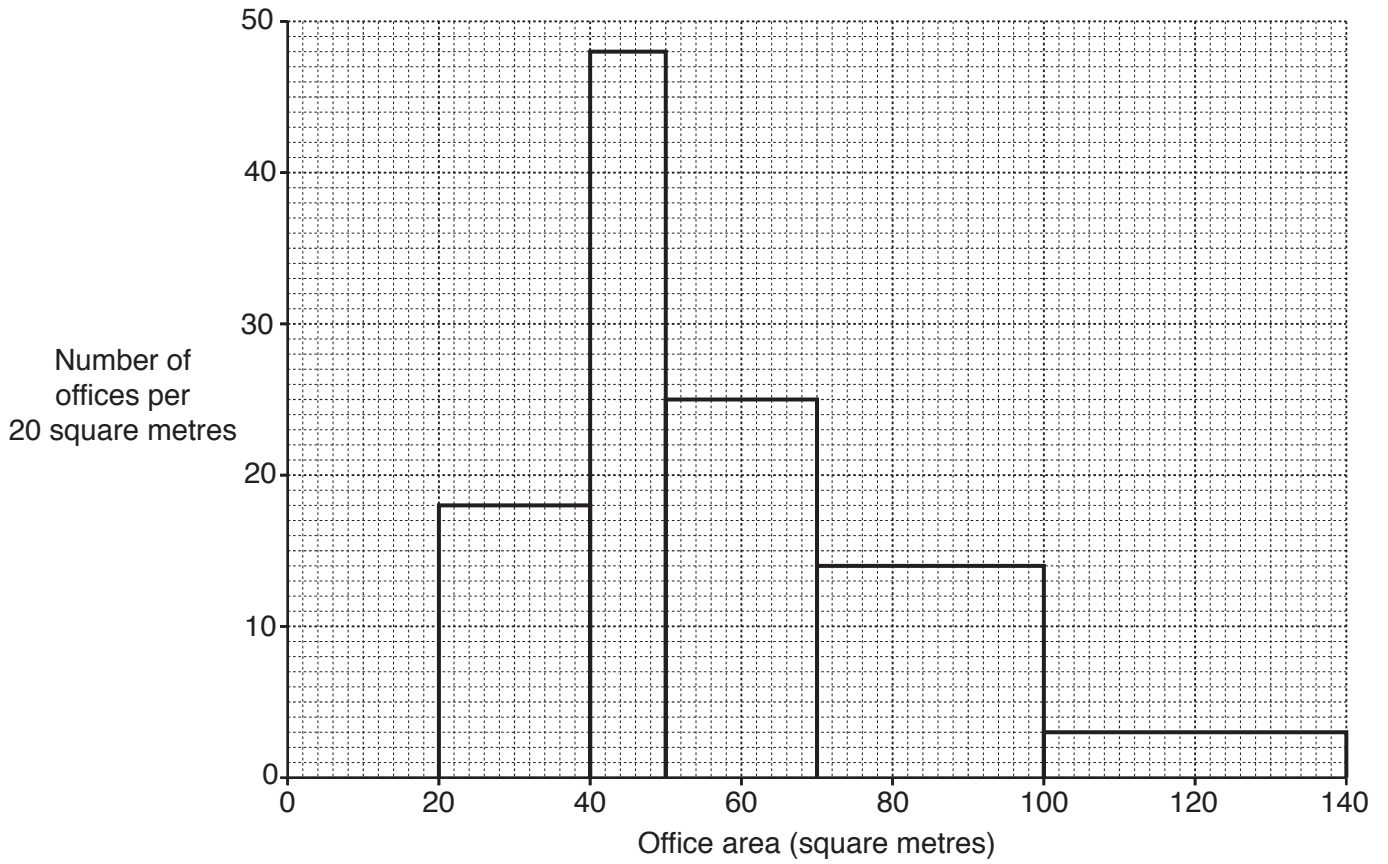
(a) Use the histogram to find the number of offices whose area, in square metres, is

(i) from 100 up to 140,

.... [2]

(ii) from 70 up to 100,

.... [1]

(iii) from 40 up to 50.

.... [1]

The building has two identical lifts (elevators). During office hours each lift is stationary at one of the floors (levels) of the building for 60% of the time, and is moving between floors for the remainder of the time.

Bernard, a statistics student, is studying the activity of the lifts. He assumes that the lifts operate independently of each other. He also assumes that the times when each of the lifts is moving are distributed randomly throughout office hours.

(b) Using Bernard's assumptions, find the probability that, at a randomly chosen moment during office hours,

    (i) both lifts are moving,

    .... [2]

    (ii) one lift is moving and the other is stationary.

    .... [3]

The building has five floors (including the ground floor), and each floor is served by both lifts. During office hours each lift is stationary at the ground floor for 24% of the time, and is stationary at each of the four upper floors for 9% of the time.

(c) Using Bernard's assumptions, find the probability that, at a randomly chosen moment during office hours,

    (i) the lifts are both stationary at the ground floor or both stationary at the second floor,

    .... [3]

    (ii) at least one lift is stationary at the top floor of the building when an office worker on that floor wishes to leave the building.

    .... [3]

10 Salman goes jogging to keep fit. On each occasion he records the distance jogged and the time taken. His results over a 14-day period are shown in the following table.

Distance, x (km)	3.6	5.4	6.9	5.4	6.9	3.6	6.9	6.9
Time, y (minutes)	27	33	46	37	42	23	48	40

(a) Plot these data on the grid below.
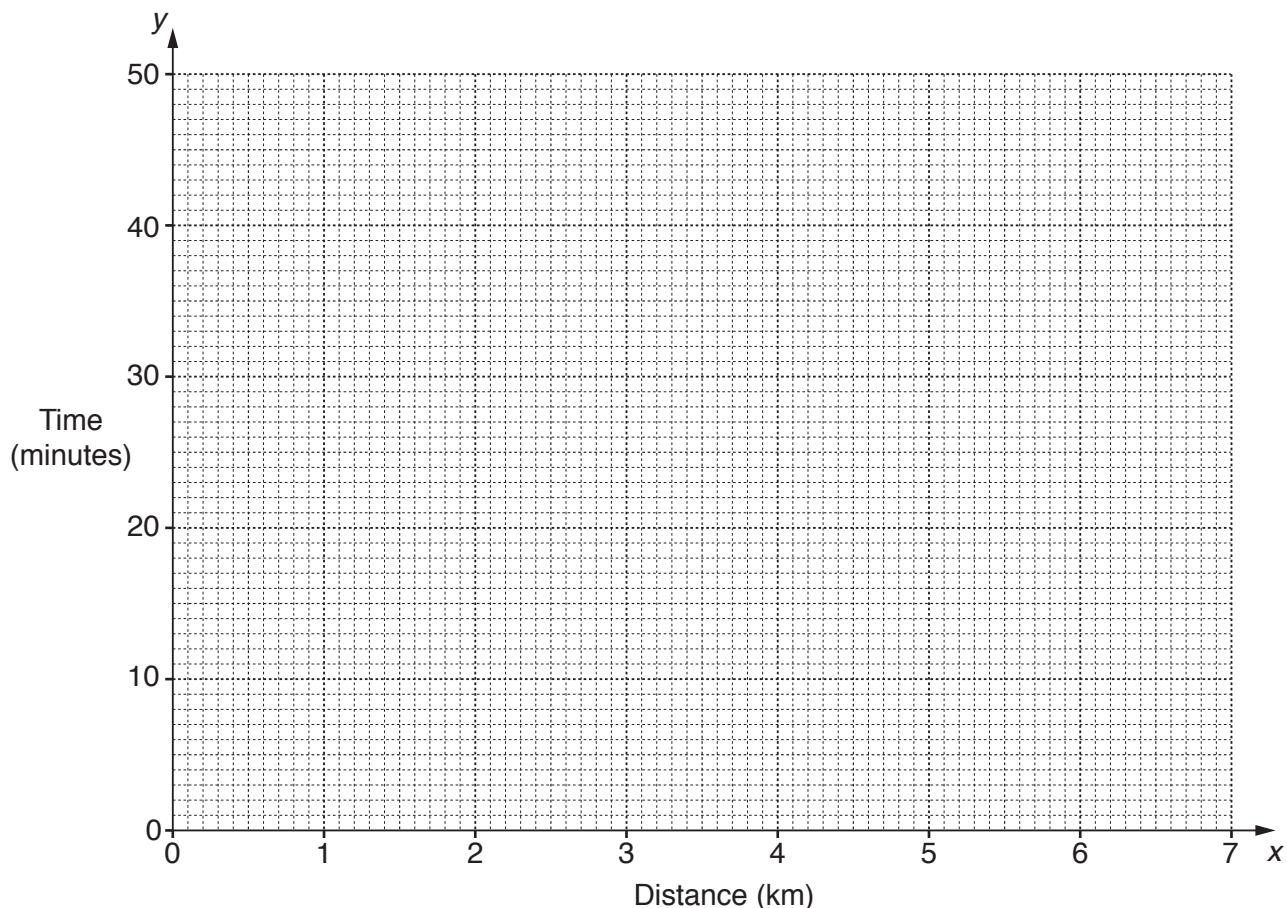
[2]

(b) How many different jogging routes does Salman appear to have used over this period? State any assumption made in giving your answer.

....

.... [1]

(c) On how many days did Salman appear **not** to go jogging over this period? State any assumption made in giving your answer.

....

.... [1]

The data have an overall mean of (5.7, 37) and an upper semi-average of (6.9, 44).

(d) Find the lower semi-average and plot this and the two given averages on your graph.

[3]

(e) Use your plotted averages to draw a line of best fit, and find its equation in the form $y = mx + c$.

.... [4]

Whilst jogging, Salman listens to songs stored on his portable music player. The songs have a mean length of 3.5 minutes.

(f) Use the equation you have found in part **(e)** to estimate the number of complete songs he will be able to hear whilst jogging a new route of distance 10.0 km.

.... [3]

(g) Give two reasons why your estimate in part **(f)** might be unreliable.

1 ....

....

2 ....

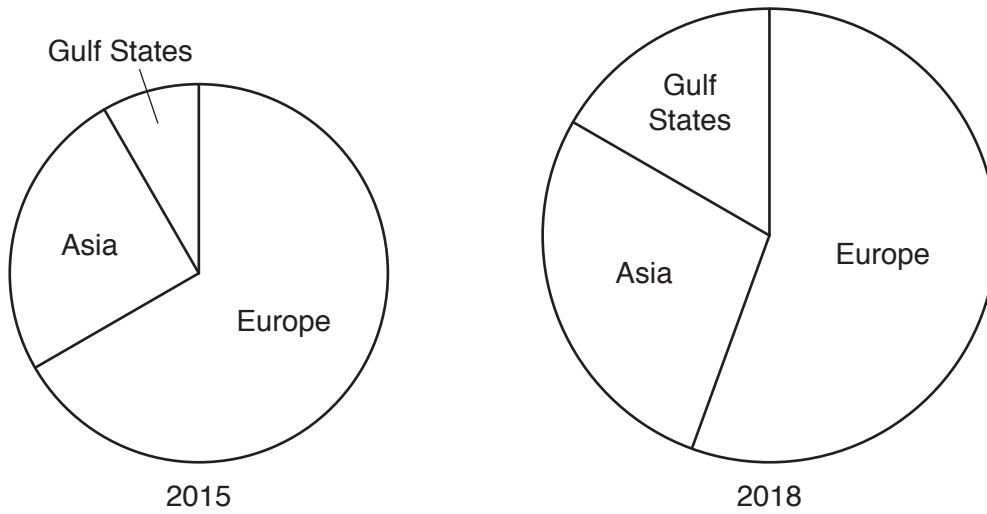... [2]

11 In the country of Blumeland the export of cut flowers is a major industry. The regions of the world to which it exported flowers in 2015 and 2018 are shown in the following pie charts, which are drawn to scale.

The chart for 2015 represents a quantity of 5400 tonnes of flowers.
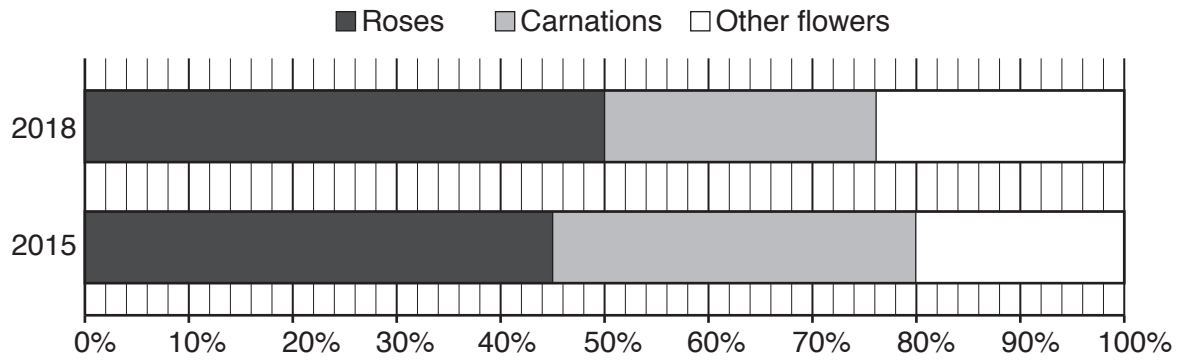
Find the quantity of flowers exported

(a) in 2015 to Asia,

.... tonnes [2]

(b) in 2018 to Europe.

.... tonnes [4]

Of the quantities of flowers exported in these two years, the percentages of different types of flower are shown in the following percentage bar chart.

■Roses ▨Carnations ☐Other flowers

(c) Find the change, from 2015 to 2018, in the quantity of carnations exported, stating also whether it was an increase or decrease.

.... tonnes

.... [4]

In 2015, of the quantity of flowers exported to the Gulf States, 68% were roses.

(d) Of all the roses exported in 2015, find the percentage which were exported to the Gulf States.

.... [4]

[Question 11 continues on the next page]

Cut flowers should have a long 'vase life' (the time for which a flower can be displayed before it dies). Measures for the vase life of samples of individual flowers of three different varieties of rose are shown below.

Rose Variety	Mean vase life (days)	Range of vase life (days)
A	10.7	2
B	11.1	4
C	10.9	3

(e) Judging on vase life alone, and assuming symmetrical distributions, from which one of these varieties would you choose to buy a bunch of flowers as a present for a relative? Justify your choice.

....

....

....

.... [2]

BLANK PAGE

20

BLANK PAGE