Cambridge IGCSE™

CANDIDATE NAME	

CENTRE NUMBER		CANDIDATE NUMBER	

STATISTICS **0479/01**

Paper 1 **For examination from 2027**

SPECIMEN PAPER **2 hours 15 minutes**

You must answer on the question paper.

You will need: Calculator   Ruler
              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. Do **not** use correction fluid or tape.
- 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 [ ].

                                                         [Turn over

This document has **16** pages.

v1

1 The list shows six terms used in statistical analysis.

classification
interpolation
population
correlation
expectation
representation

In each statement, write one of these terms to complete the statement.

When a census is conducted, information is collected on every item in the ....

A scatter diagram will display the amount of .... between the variables
in a bivariate distribution.

To calculate the median from a grouped frequency distribution table, the method of linear

.... can be used.

A bar chart is preferred to a histogram for the pictorial .... of discrete
quantitative data.

[4]

2 A manufacturer designs shirts made from various fabrics. The diagram shows the number of designs made using one or more of the fabrics cotton, polyester and rayon.
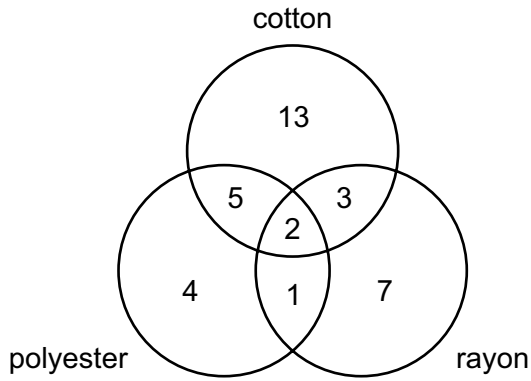
(a) Interpret the value 2 in the diagram.

....

.... [1]

(b) Find the number of designs made using:

  (i) rayon

.... [1]

  (ii) cotton or polyester or both

.... [1]

  (iii) exactly two of these fabrics.

.... [1]

(c) Find how many more designs are made using cotton than are made using polyester.

.... [1]

(d) Of the designs made using polyester, find the percentage that also use rayon.

.... [1]

3 A perfume company creates a new product.
The company wants 20% of the volume of the new product to be aromatic oils.
The company wants any variation in the volume of aromatic oils to be as small as possible.

They test five 100-millilitre bottles of the new product.
The exact amounts of aromatic oils they contain, in millilitres, are:

20.1   19.8   19.7   21.3   20.8

For the volume of aromatic oils in a test sample of five bottles, the company rules that:

* the mean should be larger than 19.5 millilitres and smaller than 20.5 millilitres
* the standard deviation should **not** be larger than 0.62 millilitres.

(a) Determine whether or not this sample satisfies the company's ruling.

.... [5]

(b) Explain whether or not the company should also rule that the standard deviation of a test sample should **not be smaller** than a certain amount.

....

.... [1]

4 A sociologist surveys the amount of time teenagers spend each day looking at screens (television, computers, smartphones etc.). The histogram summarises her results for the recorded 'daily screen time' of a sample of teenagers.
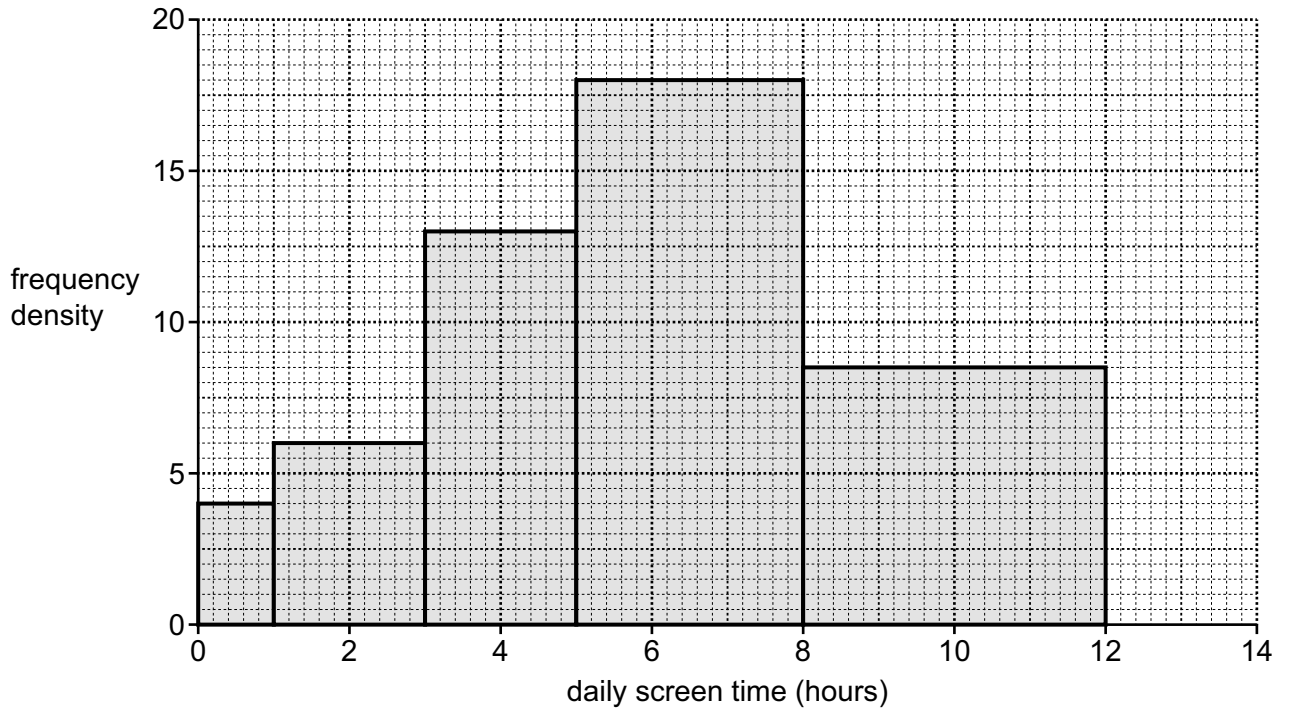
(a) Use the histogram to estimate the number of teenagers whose daily screen time is:

(i) up to 3 hours

.... [2]

(ii) from 6 hours up to 12 hours.

.... [3]

(b) The teenagers recorded their own estimates of daily screen time for answers to the survey. The sociologist thinks the teenagers have underestimated the values by one hour.

Assume that the sociologist is correct.
Estimate the number of teenagers whose daily screen time is from 3 hours up to 6 hours.

.... [2]

5 The table shows the number of university applications made by 39 pupils at a school.

number of university applications, *x*	0	1	2	3	4	5 or more
number of pupils, *f*	7	5	6	8	4	9

For this distribution:

(a) name a measure of dispersion which **cannot** be found exactly

.... [1]

(b) name a measure of dispersion which **can** be found exactly

.... [1]

(c) name a measure of central tendency which **can** be found exactly and find the value of the measure.

measure ....

value ....

[3]

(d) Three pupils were absent when the data was collected. They made no university applications. This extra data is now included.

(i) For the 42 pupils, find the value of the measure of central tendency you have named in **5(c)**.

.... [2]

(ii) For the 42 pupils, there is a measure of central tendency which can now be found that could **not** be found for the original 39 pupils.

Name this measure of central tendency and state the value of the measure.

measure ....

value ....

[1]

6 In 3 x 3 basketball there are three players on each of two teams.
The manager of a 3 x 3 basketball team has a squad of 12 players.
The manager categorises the players into three types: there are 4 strong players, 5 creative players and 3 fast players.

The manager selects three players at random from the squad for a match.

Find the probability of selecting:

(a) three creative players

.... [2]

(b) at least one fast player

.... [2]

(c) three different types of player.

.... [3]

(d) For a practice match, the manager selects three players at random from the squad to form a red team. The manager then selects three players at random from the remaining members of the squad to form a blue team.

Find the probability that the red team contains exactly one strong player and the blue team contains no strong players.

.... [4]

7 Cheetahs are the fastest land animals. Scientists fitted a wild cheetah with an electronic tracking collar to measure its hunting speeds.
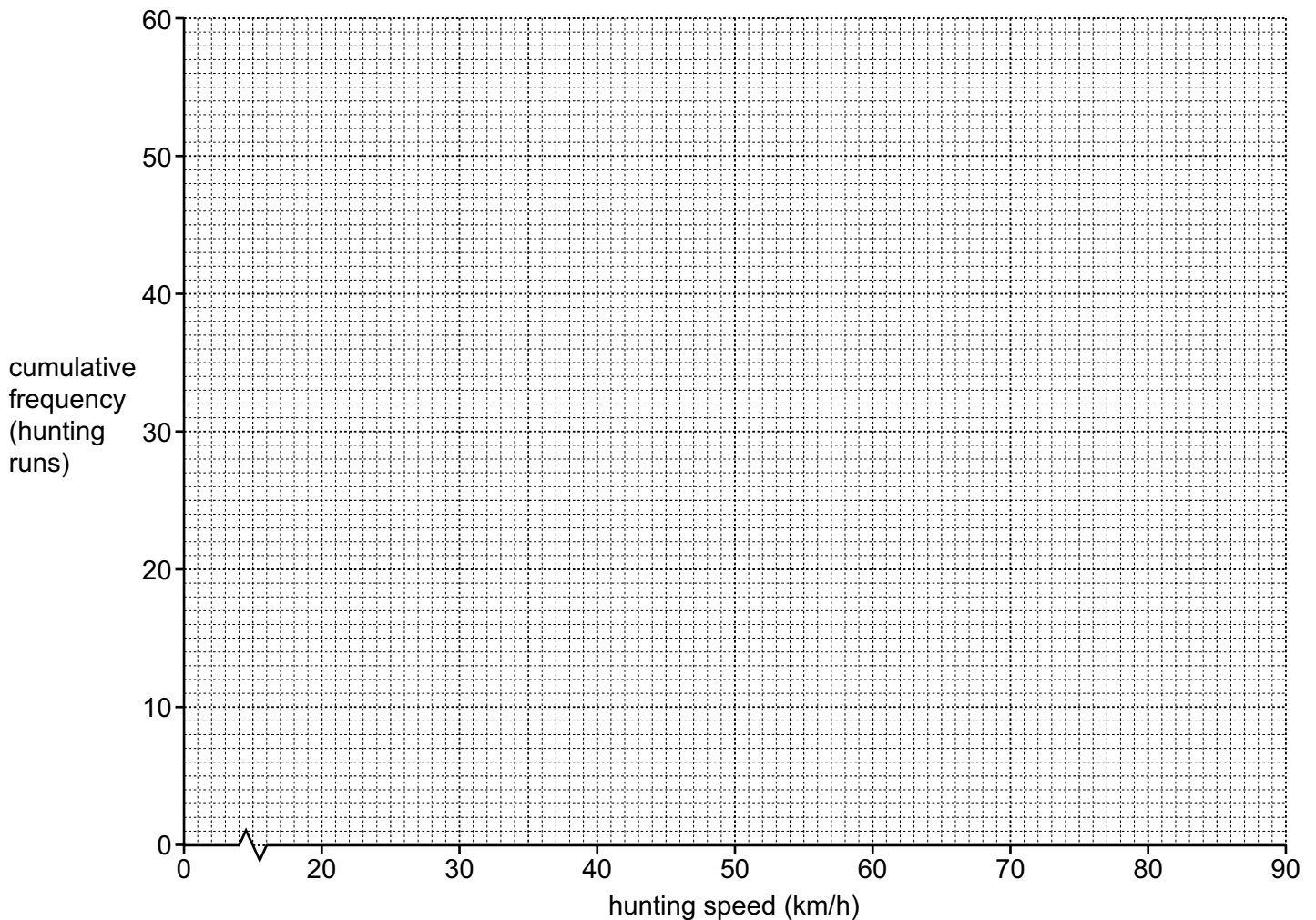
The table summarises the results for 56 hunting runs.

hunting speed (km/h)	number of hunting runs	cumulative frequency
20–under 30	2	
30–under 40	8	
40–under 50	20	
50–under 60	15	
60–under 70	7	
70–under 80	3	
80–under 90	1	

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

(b) Plot the cumulative frequencies on the grid and join the points with a smooth curve.

[3]

(c) For these hunting speeds, use the graph to estimate:

    (i) the median

    .... [1]

    (ii) the upper quartile

    .... [2]

    (iii) the value of p if the pth percentile is 44 km/h.

    .... [2]

(d) A cheetah's hunting speed depends on the speed which the prey can maintain when trying to escape. Cheetahs often chase gazelles that can maintain speeds of 50 km/h.

Use the graph to estimate the lower quartile of the cheetah's hunting speed for speeds of 50 km/h or more.

    .... [2]

(e) Of the cheetah's 56 hunting runs, 17 ended in success. The remainder ended in failure.

Estimate the probability that the cheetah will have only one success in its next two hunting runs.

    .... [2]

8 A farmer collects data from eight orchards where apricots are produced. The following data were collected in one growing season.

orchard	A	B	C	D	E	F	G	H
number of apricot trees, _x_	30	60	50	35	20	70	60	40
mass of apricots produced, _y_ (thousands of kg)	3.5	1.5	4.5	3.0	2.0	6.0	5.5	1.0

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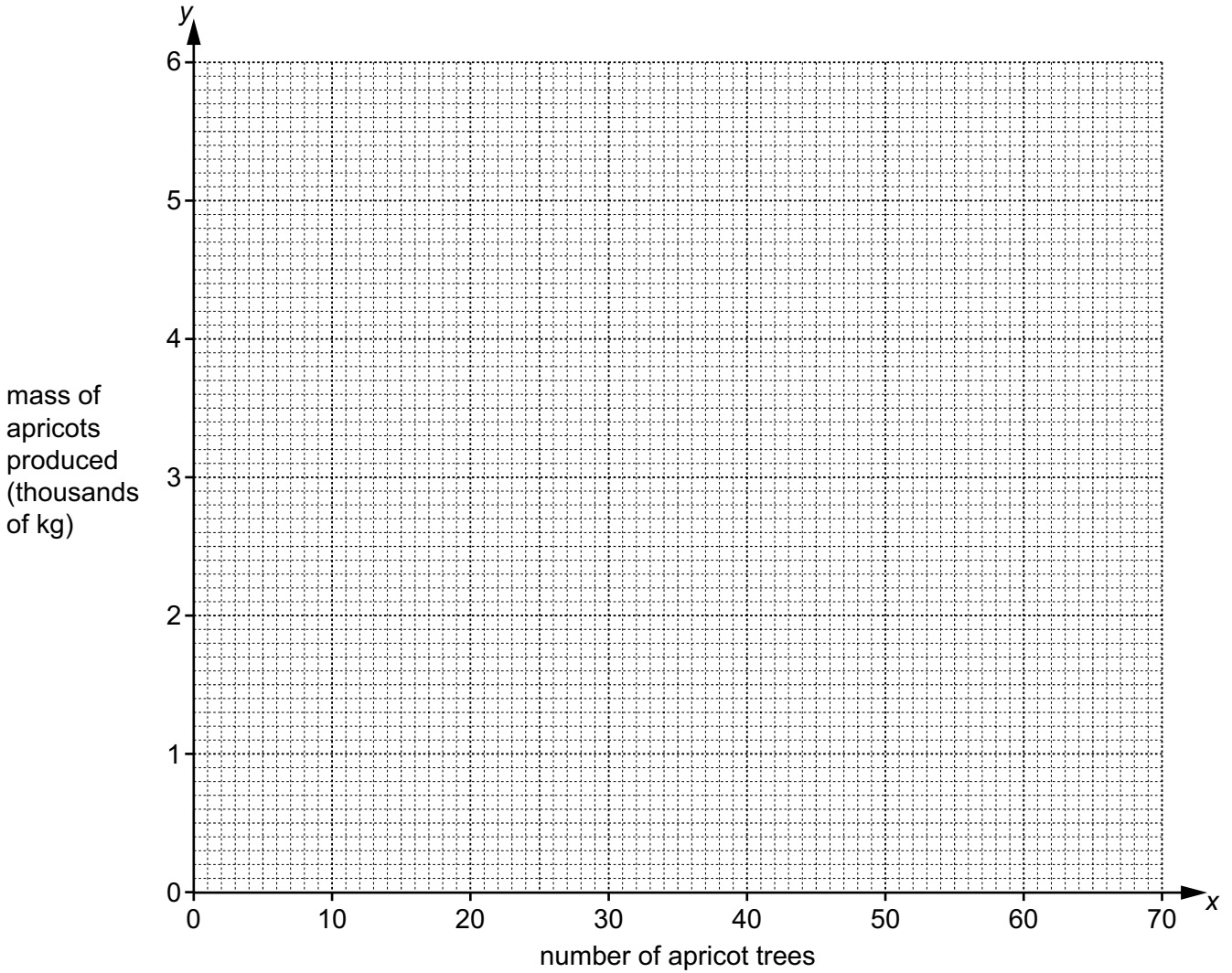
[2]

(b) The data have an overall mean of (45.625, 3.375) and an upper semi-average of (60, 4.375).

　　(i) Find the lower semi-average.

....　[2]

　　　　0479/01/SP/27

(ii) Without plotting the averages, and without drawing the line, find the equation of the line of best fit in the form *y* = *mx* + *c*.

.... [3]

(c) Orchards B and H contain only young trees. The remaining orchards contain only mature trees.

(i) Identify a difference between young and mature trees that is shown by your answer to **8(a)**.

....

.... [1]

(ii) Ignore the orchards B and H. Draw a line of best fit by eye for the remaining six points on the grid. [1]

(iii) Use the line you have drawn in **8(c)(ii)** to estimate the mass of apricots produced in one growing season from an orchard of 40 mature trees.

.... kg [2]

(d) Fresh apricots are often dried before selling. In the drying process about 80% of the mass is lost.

The farmer wants one orchard to produce 1000 kg of dried apricots in one growing season. Estimate the number of mature apricot trees the orchard needs to have.

.... [3]

9 In this question, all injury rates are expressed as injuries per thousand workers.

Workplace injury rates can be linked to job tenure. Job tenure is the length of time an employee has worked for an employer.
The table shows the workplace injury rates for workers in three areas of employment in a city in 2023. The table also shows the standard population of workers in all areas of employment in the city.

job tenure group (years)	job tenure group injury rate			standard population of workers (%)
	manufacturing	construction	services	
under 1	50.3	88.0	55.9	10
1–under 2	41.7	42.5	38.0	15
2–under 5	34.6	35.7	27.7	50
5 or more	29.3	32.4	20.1	25

(a) Explain how the information shows that the correlation between job tenure and injury rate was negative for all these areas of employment.

....

.... [1]

(b) Without calculation, explain how the information shows that the standardised injury rate for construction was the highest of the standardised injury rates for the three areas of employment.

....

.... [1]

(c) (i) Calculate the standardised injury rate for manufacturing and the standardised injury rate for services.

manufacturing ....

services ....

[5]

(ii) State which of these two areas of employment seems to have the safer working environment.
Give a reason for your answer.

.... [1]

(d) The table gives information on the number of workers in manufacturing and in construction in the city in 2023.

job tenure group (years)	number of workers in job tenure group	
	manufacturing	construction
under 1	800	700
1–under 2	1100	900
2–under 5	4300	2600
5 or more	2500	1200

(i) Find which of these two areas of employment had the greater number of injuries and state by how many.

.... [5]

(ii) Calculate the crude injury rate for construction.

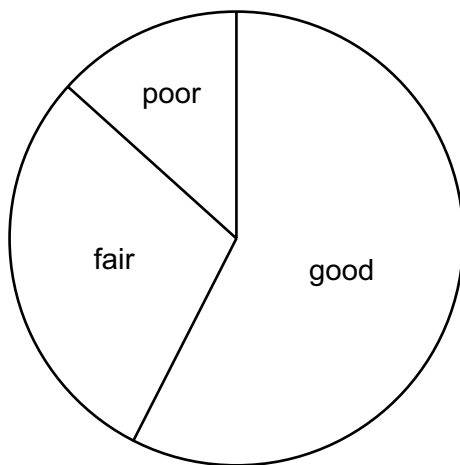.. [2]

10 Two tourist attractions in a city are the aquarium and the museum. A researcher from the tourist office surveyed samples of visitors to these attractions and recorded visitors' opinions of them. Visitors were asked to choose between 'good', 'fair' and 'poor'. The table shows the results.

opinion	number of visitors	
	aquarium	museum
good	69	85
fair	35	71
poor	16	24
total	120	180

The opinions are to be represented in two comparative pie charts. One pie chart is for the aquarium and the other for the museum. The pie chart for the aquarium has already been drawn.

(a) Draw the pie chart for the museum.

opinions of visitors to the aquarium   **opinions of visitors to the museum**

[6]

(b) Write down three observations that can be made by comparing the two pie charts that could **not easily** be seen in the original data.

1 ....

2 ....

3 ....

[3]

(c) The researcher recorded the ages of the visitors surveyed. The table summarises the ages of visitors to each attraction whose opinion had been recorded as 'good'.

age (years)	number of visitors	
	aquarium	museum
0–under 20	13	9
20–under 40	26	23
40–under 60	21	34
60–under 80	9	19

Calculate estimates of the mean ages of these visitors to each attraction.
Write your answers correct to 3 significant figures.

Write down a conclusion which may be made from your answers.

aquarium ....

museum ....

conclusion ....

....

[5]

(d) To conduct the survey, the researcher interviewed every tenth visitor as they left each attraction over a fixed period of time.

Name this type of sampling.

.... [1]

(e) In conducting a survey, open or closed questions may be used.

State which of these types of question were used when recording the opinions of visitors. Explain your answer.

....

.... [1]