4.1 Statistics Toolkit

Question Paper

Course	DP IB Maths
Section	4. Statistics & Probability
Topic	4.1 Statistics Toolkit
Difficulty	Hard

Time allowed: 100

Score: /80

Percentage: /100

Question 1a

The following cumulative frequency curve shows the distance travelled, in kilometres, to work by 160 people in Cape Town, South Africa, during 2021.
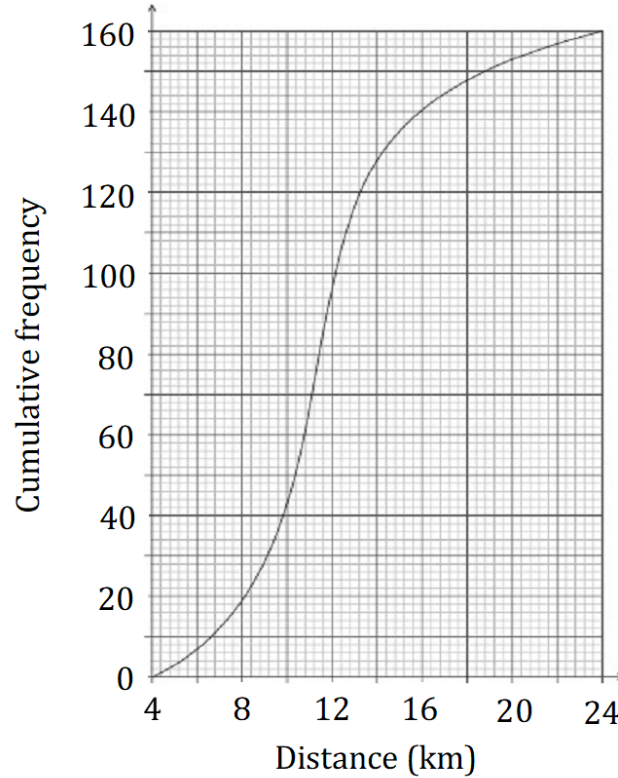
(a) Rounding your answer to the nearest half kilometre, use the graph to find the

    (i)   median distance.

    (ii)   lower quartile.

    (iii) upper quartile.

[3 marks]

Question 1b

(b) Draw a box-and-whisker diagram to represent this sample.

[3 marks]

Question 1c

(c) Using your answers from part (a), calculate the maximum distance that can be travelled by someone and still not be considered an outlier.

[2 marks]

Question 2a

The following cumulative frequency curve shows the amount of water, in litres, that 60 people drink over a month.

Amount of water, in litres
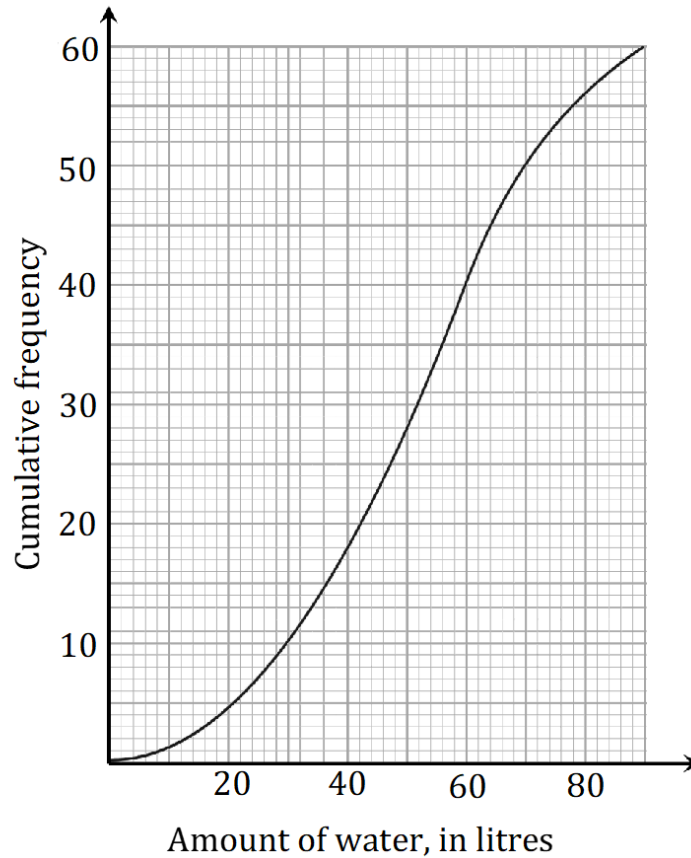
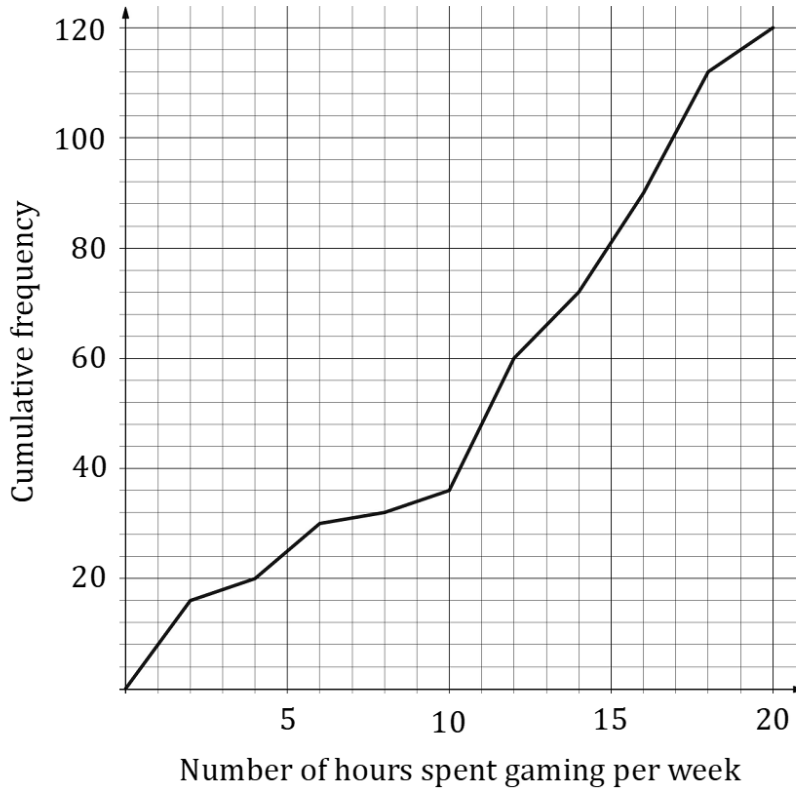(a) State whether the data is discrete or continuous.

[1 mark]

Question 2b

(b) Complete the following frequency table.

Litres of water, L	$0 \leq L \leq 30$	$30 < L \leq 50$	$50 < L \leq 70$	$70 < L \leq 90$
Frequency				
Cumulative Frequency				

[6 marks]

Question 3a

The following cumulative frequency curve shows the number of hours spent gaming per week by 120 high school students.

(a) Find the

(i)   median.

(ii)   interquartile range.

[3 marks]

Question 3b

(b) Calculate the percentage of students that spent less than 17 hours gaming per week.
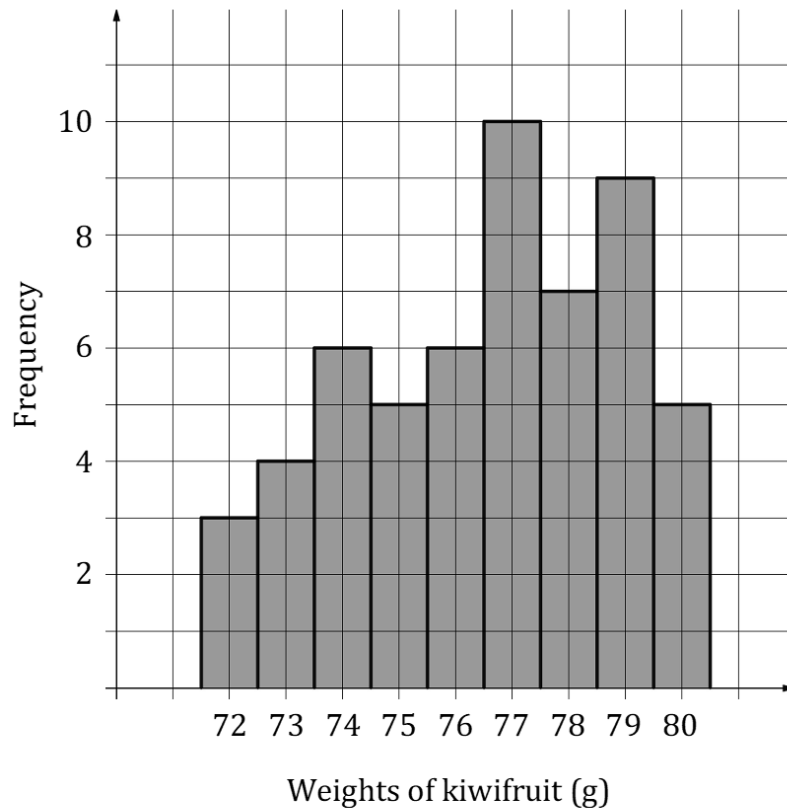
[2 marks]

Question 3c

The 120 students were chosen randomly by sampling 60 senior students and 60 junior students. The school has the same number of senior and junior students.

(c) Write down the sampling method used.

[1 mark]

Question 4a

The histogram below shows the weights of kiwifruit, each measured to the nearest gram.

(a) Write down the modal weight of the kiwifruits.

[1 mark]

Question 4b

(b) Find the median weight of the kiwifruits.

[2 marks]

Question 4c

(c) Write down two inequalities that represent the weight, w, of a kiwifruit that is considered an outlier.

[4 marks]

Question 5a

A group of people who use a gym participated in a research survey and the ages of the participants were recorded in the following table:

Age, in years (a)	$15 \leq a < 18$	$18 \leq a < 30$	$30 \leq a < 50$	$50 \leq a < 65$	$65 \leq a < 80$
Frequency	4	x	34	28	14

It is known that $34 < x < 40$.

(a) Write down

  (i)   the modal class.

  (ii) the mid interval value of the modal class.

[2 marks]

Question 5b

(b) Determine the class in which the lower quartile lies.

[2 marks]

Question 5c

(c) Calculate the mean age of participants between the ages of 30 and 80.

[2 marks]

Question 5d

The participants in this survey were chosen by selecting every person who entered the gym who was not wearing headphones.

(d) Write down the type of sampling method used.

[1 mark]

Question 6a

The table below shows the distribution of deliveries made by a group of food delivery drivers on a working day in Berlin.

Deliveries	6	7	8	9	10	11
Frequency	15	18	22	41	12	5

(a) Find

    (i)   the mean number of deliveries made.

    (ii)   the standard deviation.

    (iii) the median.

    (iv) the interquartile range.

[4 marks]

Question 6b

(b) Determine if a delivery driver who made 4 deliveries would be considered an outlier.

[2 marks]

Question 6c

The delivery drivers were selected for the survey by ordering their names alphabetically, then selecting every 20th number.

(c) Identify the sampling technique used in the sampling method.

[1 mark]

Question 7a

The following table shows the number of passes made by 11 players on a rugby team:

Player	1	2	3	4	5	6	7	8	9	10	11
Number of passes	12	18	22	41	9	18	22	28	30	21	18

(a) Write down

   (i)   the mean.

   (ii)   the median.

   (iii) the mode.

[3 marks]

Question 7b

(b) Find the interquartile range.

[3 marks]

Question 7c

(c) Determine if any of the players would be considered an outlier, and if so, state which player(s).

[3 marks]

Question 8a

The table below shows the number of TVs school students have at their home:

Number of TVs	0	1	2	3	4	5
Frequency	12	42	56	42	30	15

(a) Find

(i)   the mean.

(ii)   the standard deviation.

(iii) the median.

(iv) the interquartile range.

[4 marks]

Question 8b

(b) Determine if a student who has seven TVs would be considered an outlier.
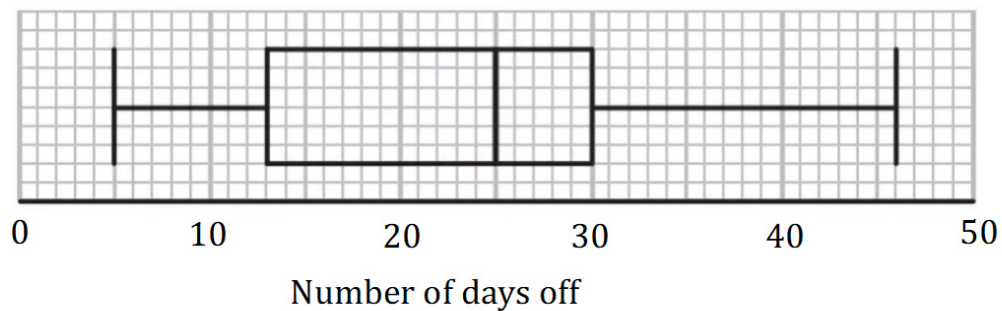
[2 marks]

Question 8c

The students were selected for the survey by randomly selecting student ID numbers, using a random number generator.

(c) Identify the sampling technique used in the sampling method.

[1 mark]

Question 9a

The number of days off taken by employees in a company during a two-year period was recorded. The data was organised into a box and whisker diagram as shown below:

Number of days off

(a) For this data, write down

    (i)   the maximum number of days off by an employee during the two years.

    (ii)   the median.

    (iii) the interquartile range.

[4 marks]

Question 9b

Steven claims that this box and whisker diagram can be used to infer that the percentage of employees who took between 14 and 30 days off is greater than the percentage of employees who took 25 days or more off.

(b) State whether Steven is correct. Justify your answer.

[2 marks]

Question 10a

The table below shows the points scored per game from two basketball players, Karo and Anna, across 9 games:

Anna	22	25	27	22	21	20	31	29	28
Karo	17	12	8	6	19	18	20	19	96

(a) State a statistical measure that would be helpful for a coach who wants to measure the consistency of players scoring performances.

[1 mark]

Question 10b

(b) For both players, find

    (i)   the mean.

    (ii)   the standard deviation.

    (iii) the median.

    (iv) the range.

    (v)   the interquartile range.

[6 marks]

Question 10c

(c) Determine whether the mean or median is a better representation of Karo's scoring ability. Justify your answer.

[1 mark]

Question 11a

Ashika runs a maths revision course three months before the final exam. In order to determine the current levels of the students Ashika gives them a practice exam the day before the course starts. On the first day of the course Ashika gives the students another test which has similar questions to the first test.

(a) State whether Ashika is testing the reliability or validity of her method for determining the current levels of the students. State the name of the test she uses.

[2 marks]

Question 11b

Ashika wants to determine whether the course improves the students' levels so she asks the following question to each of the students after the revision course:

"Do you think that this course will help you to achieve at least a level 5 in your maths exam?"

(b) Explain whether this question has content validity. Give two reasons for your answer.

[3 marks]

Question 11c

(c) Explain whether asking the question anonymously would increase its validity.

[2 marks]

Question 11d

After the final exam, Ashika analyses how well the levels from the test at the start of her revision course predict the levels in the final exam.

(d) State the name of the validity test that Ashika is using.

[1 mark]