4.1 Statistics Toolkit

Mark Schemes
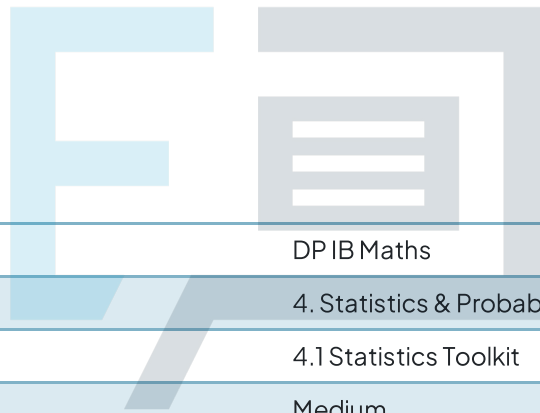
Course	DP IB Maths
Section	4. Statistics & Probability
Topic	4.1 Statistics Toolkit
Difficulty	Medium

To be used by all students preparing for DP IB Maths AI SL
Students of other boards may also find this useful

Exam Papers Practice

Question 1

a) Just take the first 6 weights:

$$\text{mean} = \frac{52 + 57 + 63 + 80 + 56 + 66}{6}$$

$$= \frac{187}{3} = 62.333333...$$

$$\boxed{\text{mean} = 62.3 \text{ kg} \ (1 \text{ d.p.})}$$

Note: There are other possible convenience samples! The one used here is just an especially obvious one.

52, 57, (63), 80, 56, (66), 101, 68, (55), 96, 70, (62), 66, 64, (99), 91, 55, (92)

b) $\frac{18}{6} = 3$ take every 3rd weight

$$\text{mean} = \frac{63 + 66 + 55 + 62 + 99 + 92}{6}$$

$$= \frac{437}{6} = 72.833333...$$

$$\boxed{\text{mean} = 72.8 \text{ kg} \ (1 \text{ d.p.})}$$

c) The sum of the 18 weights is 1293.
   So the true mean is:

$$\frac{1293}{18} = \frac{431}{6} = 71.8\dot{3} \text{ kg}$$

> The systematic sample mean (72.8)
> is a lot closer to the true mean here
> than the convenience sample mean
> (62.3) is.
>
> Because a systematic sample is
> generally more random than a
> convenience sample, it will often
> give more representative results.

This isn't guaranteed, though! Starting on
the first data value in (b) would give a
mean of 81, which is not much more accurate
than the value in (a). Starting on the
second value in (b) would give a mean
of 61.6, which is less accurate than
the value in (a).

Question 2   a) (i) Opportunity sampling

(ii)
> This survey will be quick, easy,
> and inexpensive to carry out.
>
> But it is unlikely to be
> representative of the shop's
> customers as a whole — it
> will only be surveying people
> who shop during lunchtime
> on a particular day of the
> week.

Exam Papers Practice

b) Quota sampling. For example, divide the shop's opening hours over a week into a number of 'blocks' and then survey a fixed number of people during each block to create the sample.

Question 3

(i) Kids small

size with highest frequency

(ii) The mode is a kids size, but the majority of masks sold (51 out of 84) are adult sizes.

(iii) Number sold during week:

29 + 4 + 8 + 24 + 15 + 4 = 84

84 ÷ 7 = 12 masks per day

Question 4

$$\text{standard deviation} = \sqrt{\frac{\Sigma x^2}{n} - \bar{x}^2}$$

$$\Sigma x = 696 \qquad \Sigma x^2 = 54998 \qquad n = 9$$

$$\sigma = \sqrt{\frac{54998}{9} - \left(\frac{696}{9}\right)^2} = 11.421227\ldots$$

$$\bar{x} = \frac{\Sigma x}{n} \quad (\text{mean})$$

$$\sigma = 11.4 \text{ cm } (3 \text{ s.f.})$$

Question 5

a) (i) median = 115 g

(ii) $Q_1 = 103$ g

(iii) $Q_3 = 120$ g

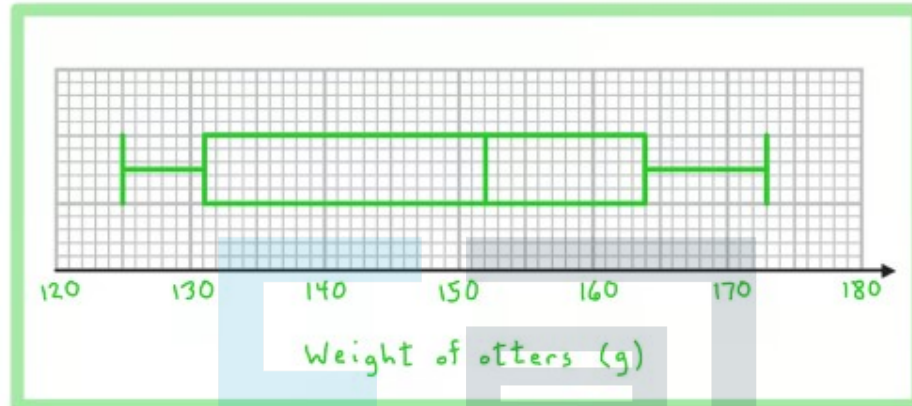$$Q_3 - Q_1 = 120 - 103 = 17$$

$$\text{IQR} = 17 \text{ g}$$

$$\underset{\underset{\text{smallest}}{\downarrow}}{}\qquad\underset{\underset{\text{range}}{\downarrow}}{}$$

b) largest weight = $125 + 48 = 173$ g

$$LQ = 164 - 33 = 131 \text{ g}$$

$$\underset{\underset{UQ}{\uparrow}}{}\qquad\underset{\underset{IQR}{\uparrow}}{}$$

Weight of otters (g)

Question 6

a) (i) Median is $\dfrac{20+1}{2} = 10.5^{th}$ value :

$$Q_2 = \dfrac{1.2 + 1.3}{2} = 1.25 \text{ m}$$

$$\dfrac{20}{4} = 5 \implies Q_1 \text{ is } 5.5^{th} \text{ value}$$

$$20 \times \dfrac{3}{4} = 15 \implies Q_3 \text{ is } 15.5^{th} \text{ value}$$

$$Q_1 = \dfrac{1.2 + 1.2}{2} = 1.2 \text{ m} \qquad Q_3 = \dfrac{1.4 + 1.4}{2} = 1.4 \text{ m}$$

(ii) $\boxed{IQR = 1.4 - 1.2 = 0.2 \text{ m}}$ $IQR = Q_3 - Q_1$

(iii) $1.2 - 1.5 \times 0.2 = 0.9$ lower boundary

$1.4 + 1.5 \times 0.2 = 1.7$ upper boundary

$\boxed{0.4 \text{ m is an outlier}}$

height (m)

b)

a) (i) Median is $\frac{20+1}{2} = 10.5^{th}$ value:

$$Q_2 = \frac{1.2 + 1.3}{2} = 1.25 \text{ m}$$

$\frac{20}{4} = 5 \implies Q_1$ is 5.5^{th} value

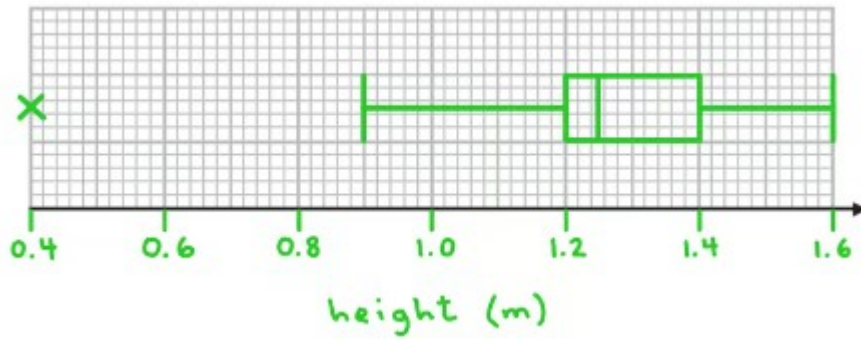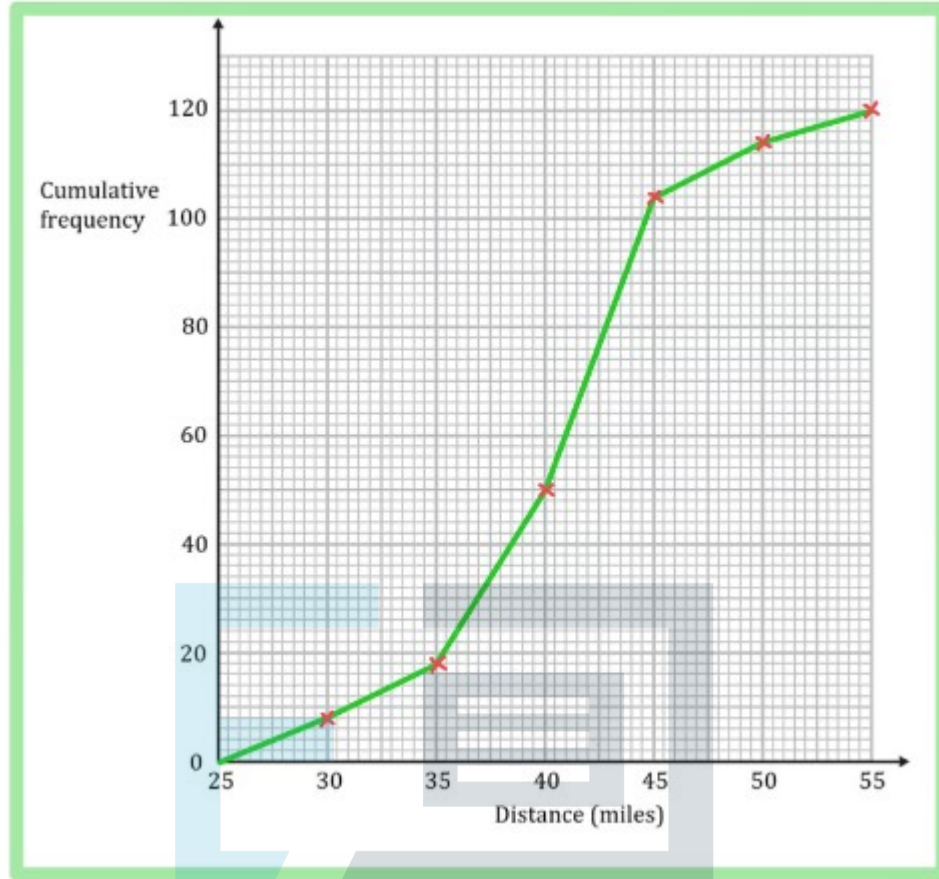$20 \times \frac{3}{4} = 15 \implies Q_3$ is 15.5^{th} value

$$Q_1 = \frac{1.2 + 1.2}{2} = 1.2 \text{ m} \qquad Q_3 = \frac{1.4 + 1.4}{2} = 1.4 \text{ m}$$

(ii) $\text{IQR} = 1.4 - 1.2 = 0.2 \text{ m}$ $\text{IQR} = Q_3 - Q_1$
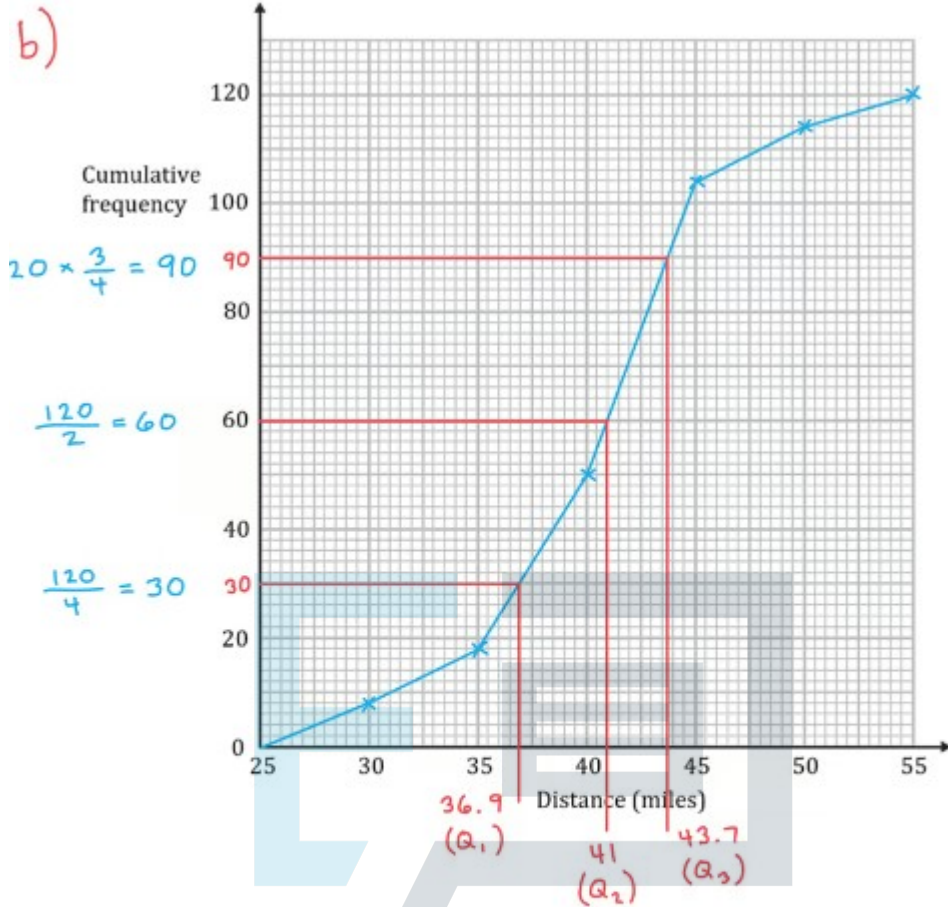
(iii) $1.2 - 1.5 \times 0.2 = 0.9$ lower boundary

$1.4 + 1.5 \times 0.2 = 1.7$ upper boundary

$$0.4 \text{ m is an outlier}$$

Question 7

b)

$20 \times \frac{3}{4} = 90$

$\frac{120}{2} = 60$

$\frac{120}{4} = 30$

36.9 (Q₁)

41 (Q₂)

43.7 (Q₃)

Cumulative frequency

Distance (miles)

median = 41 miles

IQR = 43.7 - 36.9 = 6.8 miles

Question 8

a)

$$\frac{80}{2} = 40$$
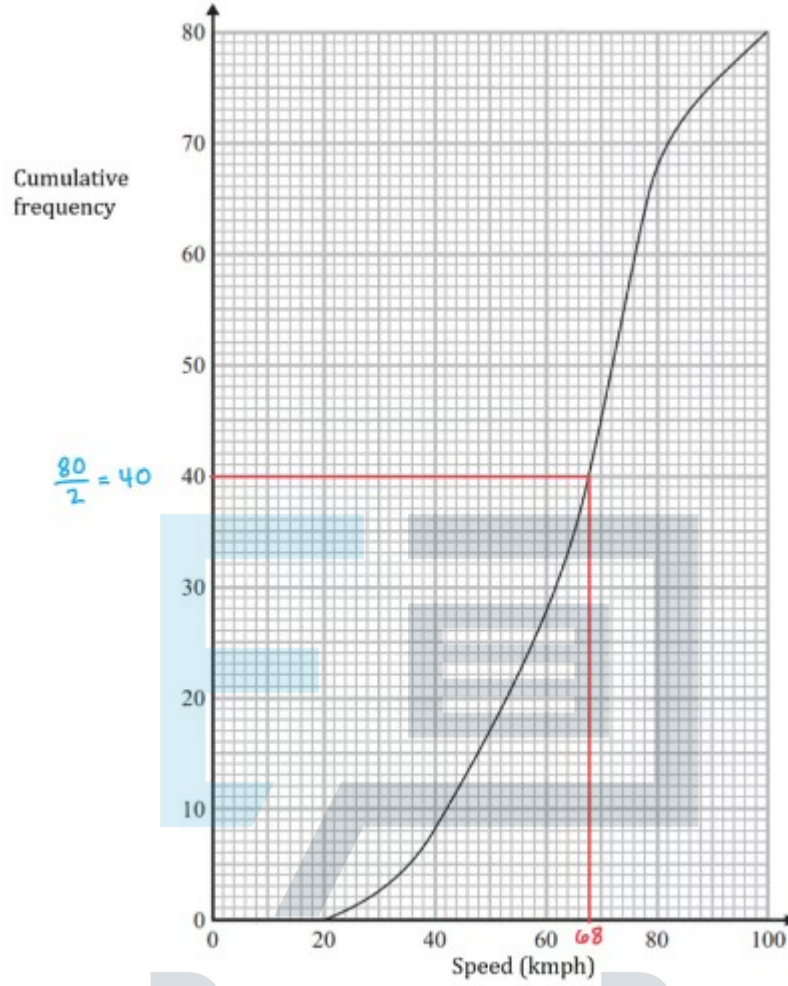
median = 68 kmph

b) 10% over 80 kmph = 88 kmph

74

Cumulative frequency

Estimate for number of people above 88 kmph
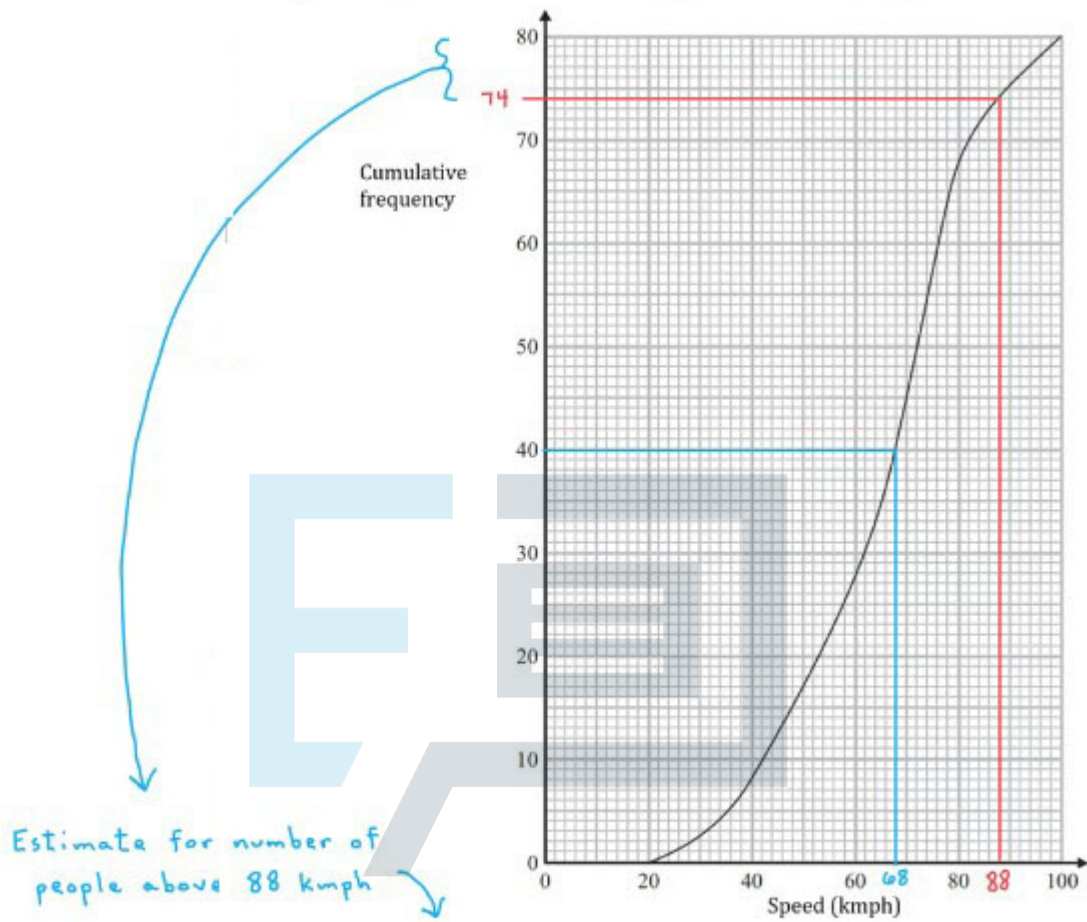
$80 - 74 = 6$

$\frac{6}{80} \times 100 = \boxed{7.5\%}$

Question 9   a)

$$\boxed{\text{Median} = 3.7 \text{ hours}}$$

b) $61 - 17 = 44$

44 students

c) $LQ = 2.9 \qquad UQ = 4.6$

$4.6 - 2.9 = 1.7 \qquad IQR = UQ - LQ$

IQR = 1.7 hours

d) $p = 48 - 9 = 39$

$q = 74 - 48 = 26$

p = 39   q = 26

$9 + 39 + 26 + 6 = 80 \checkmark$

e) $10\% \text{ of } 80 = 8$

$80 - 8 = 72$

d = 6.4

F) $4000 \times \dfrac{21}{80} = 1050$

Approximately 1050 students took less than 3 hours.