EXERCISE 20F.2

1 **a** $s_n \approx 6.77$ kg $\therefore \sigma \approx 6.77$ kg **b** $\mu \approx 93.8$ kg

2 **a** $\overline{x} \approx 77.5$ g, $s_n \approx 7.44$ g **b** $\mu \approx 77.5$ g, $\sigma \approx 7.44$ g

EXERCISE 20F.3

1 **a** $\overline{x} \approx 1.72$ children, $s_n \approx 1.67$ children
b $\mu \approx 1.72$ children, $\sigma \approx 1.67$ children

2 **a** $\overline{x} \approx 14.5$ years, $s_n \approx 1.75$ years
b $\mu \approx 14.5$ years, $\sigma \approx 1.75$ years

3 **a** $\overline{x} \approx 37.3$ toothpicks, $s_n \approx 1.45$ toothpicks
b $\mu \approx 37.3$ toothpicks, $\sigma \approx 1.45$ toothpicks

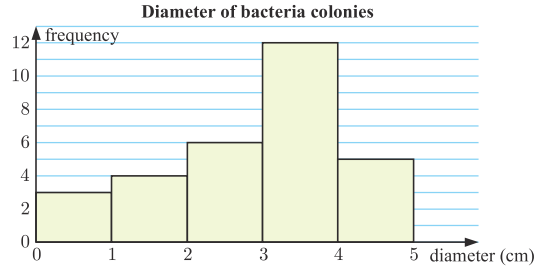
4 **a** $\overline{x} \approx 48.3$ cm, $s_n \approx 2.66$ cm
b $\mu \approx 48.3$ cm, $\sigma \approx 2.66$ cm

5 **a** $\overline{x} \approx \$390.30$, $s_n \approx \$15.87$
b $\mu \approx \$390.30$, $\sigma \approx \$15.87$

REVIEW SET 20A

1 **a** **i** 3.15 cm **ii** 4.5 cm

b



c The distribution is slightly negatively skewed.

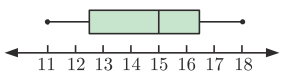
2 $a = 8$, $b = 6$

3 **a**

Distribution	*Girls*	*Boys*
shape	pos. skewed	approx. symm.
median	36 s	34.5 s
mean	36 s	34.45 s
modal class	34.5 - 35.5 s	34.5 - 35.5 s

b The girls' distribution is positively skewed and the boys' distribution is approximately symmetrical. The median and mean swim times for boys are both about 1.5 seconds lower than for girls. Despite this, the distributions have the same modal class because of the skewness in the girls' distribution. The analysis supports the conjecture that boys generally swim faster than girls with less spread of times.

4



5 **a** 58.5 s **b** 6 s **c** 53 s

6 **a** $p = 12$, $m = 6$ **c** $\frac{254}{30} = \frac{127}{15}$

b

Measure	*Value*
mode	9
median	9
range	4

7 **a** 88 students **b** $m \approx 24$

c

Time t (min)	$5 \leqslant t < 10$	$10 \leqslant t < 15$	$15 \leqslant t < 20$
Frequency	20	40	48

Time t (min)	$20 \leqslant t < 25$	$25 \leqslant t < 30$	$30 \leqslant t < 35$
Frequency	42	28	17

Time t (min)	$35 \leqslant t < 40$
Frequency	5

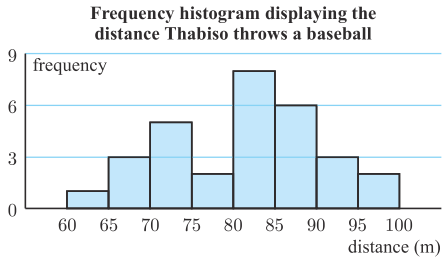
REVIEW SET 20B

1 **a** highest = 97.5 m, lowest = 64.6 m
b use groups $60 \leqslant d < 65$, $65 \leqslant d < 70$, ...

c **Distances thrown by Thabiso**

Distance (m)	*Tally*	*Frequency*
$60 \leqslant d < 65$	|	1
$65 \leqslant d < 70$	|||	3
$70 \leqslant d < 75$	𝍸	5
$75 \leqslant d < 80$	||	2
$80 \leqslant d < 85$	𝍸 |||	8
$85 \leqslant d < 90$	𝍸 |	6
$90 \leqslant d < 95$	|||	3
$95 \leqslant d < 100$	||	2
	Total	30

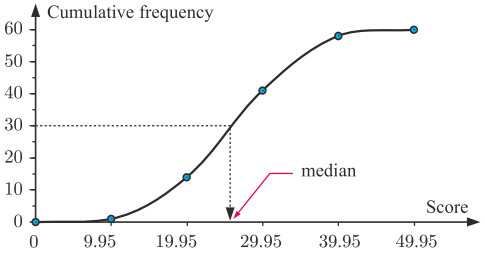
d



e **i** ≈ 81.1 m **ii** ≈ 83.1 m

2 **b** $k + 3$

3 **a**



b ≈ 25.9 **c** ≈ 12.0 **d** $\overline{x} \approx 26.0$, $s \approx 8.31$

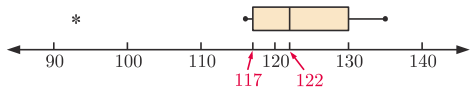
4 **a** $\overline{x} \approx 33.6$ L, $s \approx 7.63$ L **b** $\mu \approx 33.6$ L, $\sigma \approx 7.63$ L

5 **a** **i** 101.5 **ii** 98 **iii** 105.5 **b** 7.5
c $\overline{x} = 100.2$, $s \approx 7.59$

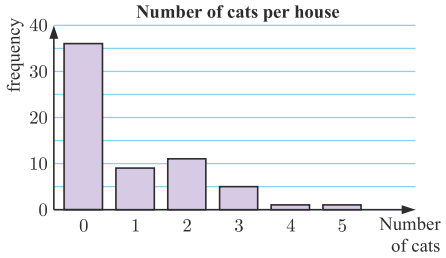
6 **a** $\overline{x} \approx 49.6$, $s \approx 1.60$
b Does not justify claim. Need a larger sample.

7 **a** $\sigma \approx 11.65$ **b** $Q_1 = 117$, $Q_3 = 130$ **c** yes, 93

d



8 **a**



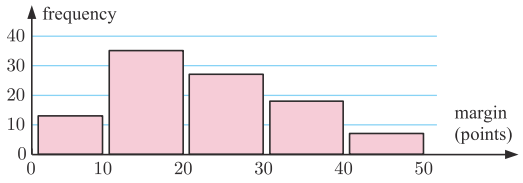
b positively skewed **c** **i** 0 **ii** 0.87 **iii** 0

d The mean, as it suggests that some people have cats. (The mode and median are both 0.)

REVIEW SET 20C

1 **a** discrete

b



c No, as we do not know each individual data value, only the intervals they fall in.

2 **a** $x = 7$ **b** $s^2 \approx 10.2$ **3** ≈ 414 customers

4 **a**

	A	B
Min	11	11.2
Q_1	11.6	12
Median	12	12.6
Q_3	12.6	13.2
Max	13	13.8

b

		A	B
i	Range	2	2.6
ii	IQR	1	1.2

c **i** We know the members of squad A generally ran faster because their median time is lower.

ii We know the times in squad B are more varied because their range and IQR are higher.

5 **a** $\overline{x} = €103.51$, $s \approx €19.40$ **b** $\mu = €103.51$, $\sigma \approx €19.40$

6 **a** No, extreme values have less effect on the standard deviation of a larger population.

b **i** mean **ii** standard deviation

c A low standard deviation means that the weight of biscuits in each packet is, on average, close to 250 g.

7 **a** 120 students **b** 65 marks **c** 54 and 75

d 21 marks **e** 73% of them **f** 82 marks

EXERCISE 21A

1 **a** A scatter diagram consists of points plotted on a set of axes where the independent variable is placed on the horizontal axis and the dependent variable on the vertical axis.

b Correlation refers to the relationship or association between two variables.

c Positive correlation describes the relationship when increasing the independent variable generally results in the dependent variable increasing.

d Negative correlation describes the relationship when increasing the independent variable generally results in the dependent variable decreasing.

e An outlier is a data point that does not fit the general trend of the data and is isolated from the main body of data.

2 **a** **i** no correlation **ii** zero
iii non-linear **iv** no outliers

b **i** positive correlation **ii** weak
iii roughly linear **iv** no outliers

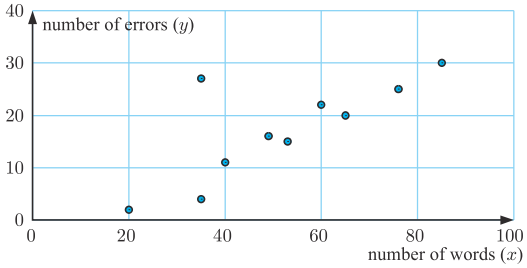
c **i** negative correlation **ii** moderate
iii non-linear **iv** one outlier

d **i** positive correlation **ii** moderate
iii linear **iv** no outliers

e **i** negative correlation **ii** strong
iii linear **iv** one outlier

f **i** positive correlation **ii** moderate
iii non-linear **iv** no outliers

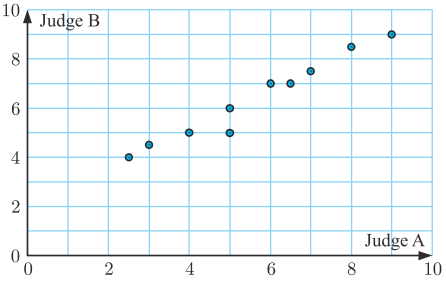
3 **a**



b **i** C **ii** G **iii** I

c positive, strong correlation **d** It is approximately linear.

4 **a**



b There appears to be a **strong**, **positive** correlation between Judge A's scores and Judge B's scores. This means that as Judge A's scores increase, Judge B's scores **increase**.

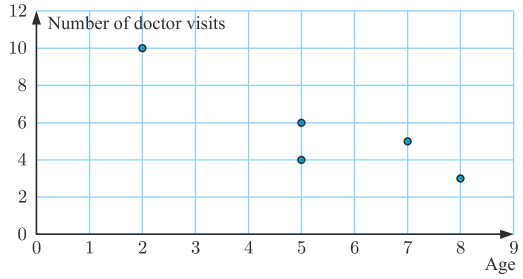
c No

EXERCISE 21B

1 **a** **B** **b** **A** **c** **D** **d** **C** **e** **E**

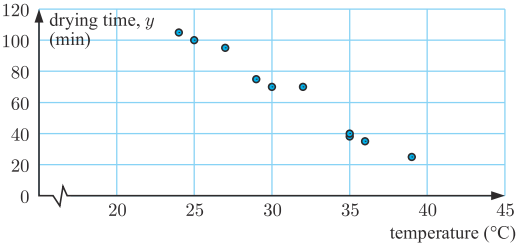
2 **a** $r = 1$ **b** $r = -1$ **c** $r = 0$

3 **a**



b $r \approx -0.892$ **c** strong, negative correlation

4 **a**



b $r \approx -0.987$ **c** very strong, negative correlation