Basic stats 17.04 *[35 marks]*

Elsie, a librarian, wants to investigate the length of time, T minutes, that people spent in her library on a particular day.

1a. State whether the variable T is discrete or continuous. *[1 mark]*

Markscheme

continuous **A1**

[1 mark]

Elsie's data for 160 people who visited the library on that particular day is shown in the following table.

T (minutes)	$0 \le T < 20$	$20 \le T < 40$	$40 \le T < 60$	$60 \le T < 80$	$80 \le T < 100$
Frequency	50	62	k	14	8

1b. Find the value of k. *[2 marks]*

Markscheme

$160 - 50 - 62 - 14 - 8$ **(M1)**

$(k =)\ 26$ **A1**

[2 marks]

1c. Write down the modal class. *[1 mark]*

Markscheme

$20 \leq T < 40$ **A1**

[1 mark]

1d. Write down the mid-interval value for this class. *[1 mark]*

Markscheme

30 **A1**

[1 mark]

1e. Use Elsie's data to calculate an estimate of the mean time that people spent in the library. *[2 marks]*

Markscheme

33.5 minutes **A2**

Note: ***FT*** from their value of k and their mid-interval value. Follow through from part (c)(ii) but only if mid-interval value lies in their interval.

[2 marks]

1f. Using the table, write down the maximum possible number of people who spent 35 minutes or less in the library on that day. *[1 mark]*

Markscheme

112 **A1**

[1 mark]

Elsie assumes her data to be representative of future visitors to the library.

1g. Find the probability a visitor spends at least 60 minutes in the library. *[2 marks]*

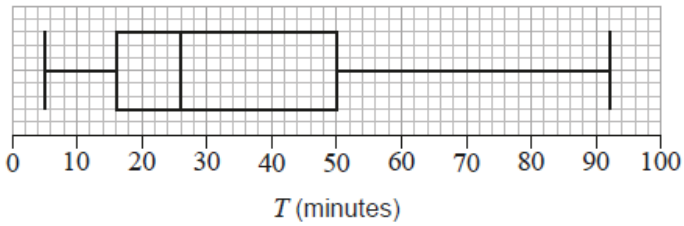
Markscheme

$\frac{22}{160}$ $\left[0.138,\ 0.1375,\ 13.75\%,\ \frac{11}{80}\right]$ ***A1A1***

Note: Award ***A1*** for correct numerator, ***A1*** for correct denominator.

[2 marks]

The following box and whisker diagram shows the times, in minutes, that the 160 visitors spent in the library.



1h. Write down the median time spent in the library. *[1 mark]*

Markscheme

26 minutes ***A1***

[1 mark]

1i. Find the interquartile range. *[2 marks]*

Markscheme

$50 - 16$ ***(M1)***

Note: Award ***M1*** for both correct quartiles seen.

34 minutes ***A1***

[2 marks]

1j. Hence show that the longest time that a person spent in the library is not an outlier. *[3 marks]*

Markscheme

correct substitution into outlier formula ***(M1)***

$50 + 1.5 \times 34$

$= 101$ ***A1***

$92 < 101$ **OR** highest value on diagram < 101 ***R1***

not an outlier ***AG***

Note: Award ***R1*** for their correct comparison. Follow through from their part (h). Award ***R0*** if their conclusion is "it is an outlier", this contradicts Elsie's belief.

[3 marks]

Elsie believes the box and whisker diagram indicates that the times spent in the library are not normally distributed.

1k. Identify one feature of the box and whisker diagram which might support Elsie's belief. *[1 mark]*

Markscheme

EITHER

the diagram is not symmetric or equivalent

e.g the median is not in the center of the box or
the lengths of the whiskers are (very) different or (positive or right) skew

OR

the mean and median are (very) different; ***A1***

[1 mark]

Hafizah harvested 49 mangoes from her farm. The weights of the mangoes, w, in grams, are shown in the following grouped frequency table.

Weight (g)	$100 \leq w < 200$	$200 \leq w < 300$	$300 \leq w < 400$	$400 \leq w < 500$	$500 \leq w < 600$
Frequency	4	7	14	16	8

2a. Write down the modal group for these data. *[1 mark]*

Markscheme

* This question is from an exam for a previous syllabus, and may contain minor differences in marking or structure.

$400 \leq w < 500$ ***(A1)*** ***(C1)***

Note: Accept alternative notation $[400,\ 500)$ or $[400,\ 500[$.
Do not accept "400-500".

[1 mark]

2b. Use your graphic display calculator to find an estimate of the standard deviation of the weights of mangoes from this harvest. *[2 marks]*

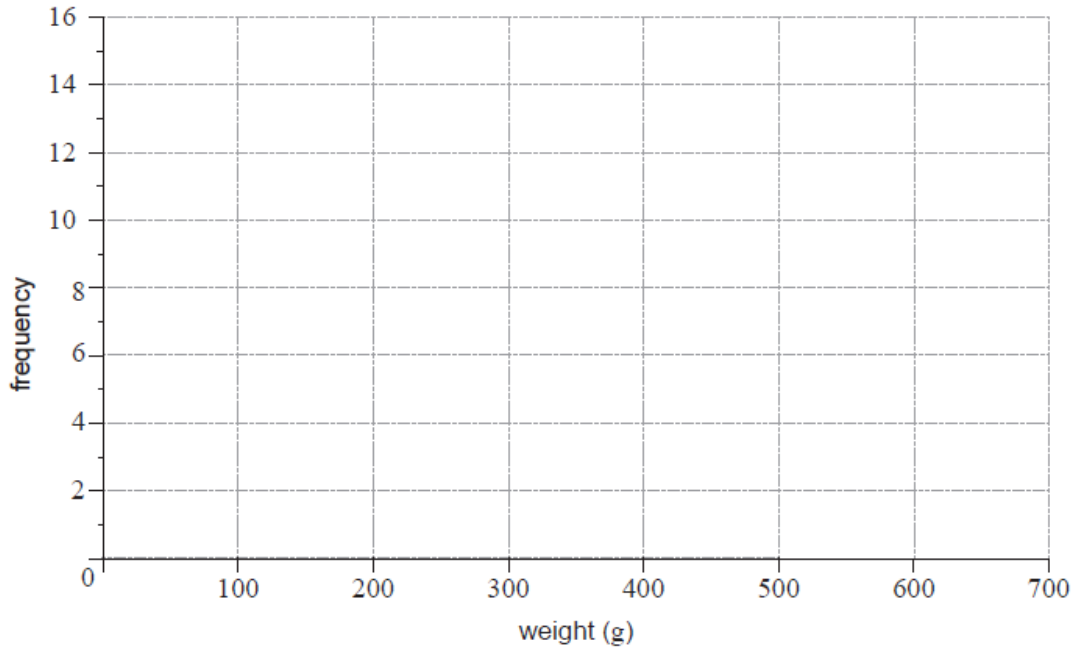
Markscheme

115 $(115.265\ldots\ (g))$ ***(A2)*** ***(C2)***

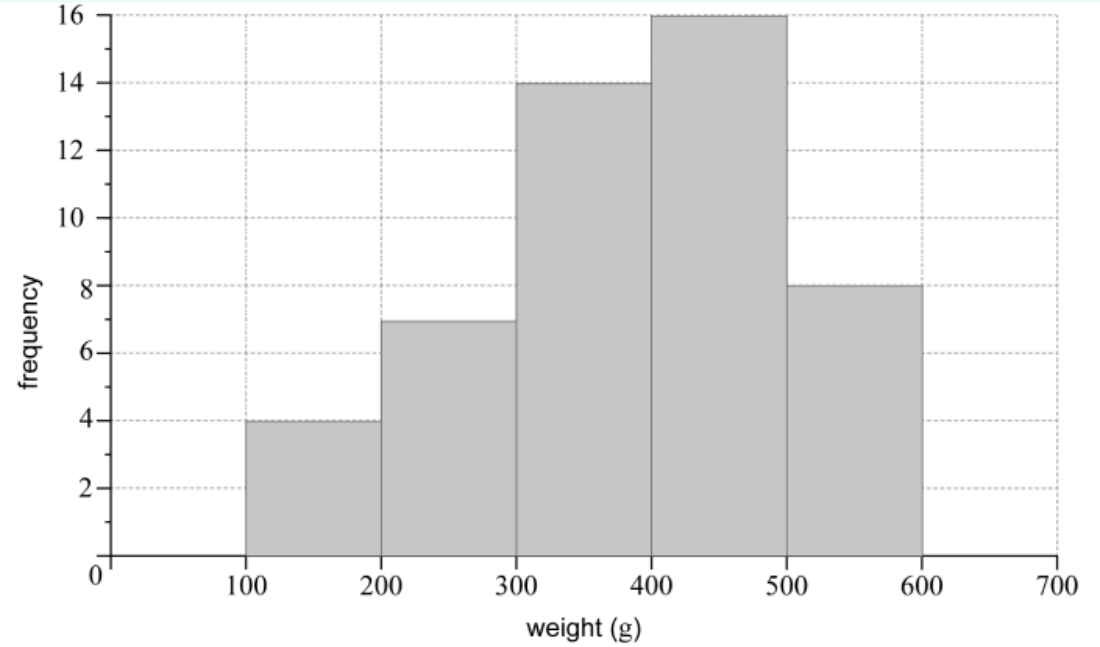
Note: Award ***(A1)(A0)*** for an answer of 116 $(116.459\ldots)$.

[2 marks]

2c. On the grid below, draw a histogram for the data in the table. *[3 marks]*



Markscheme



(A2)

(A1) ***(C3)***

Note: Award ***(A2)*** for all correct heights of bars or ***(A1)*** for three or four correct heights of bars.
Award ***(A1)*** for rectangular bars all with correct left and right end points ($100,\ 200,\ 300,\ 400,\ 500$ and 600) and for no gaps; the bars do **not** have to be shaded.
Award at most ***(A2)(A0)*** if a ruler is not used for all lines.

[3 marks]

Chicken eggs are classified by grade (4, 5, 6, 7 or 8), based on weight. A mixed carton contains 12 eggs and could include eggs from any grade. As part of the science project, Rocky buys 9 mixed cartons and sorts the eggs according to their weight.

Grade	Weight, w (grams)	Frequency
4	$40 \le w < 50$	3
5	$50 \le w < 60$	30
6	$60 \le w < 70$	45
7	$70 \le w < 80$	25
8	$80 \le w < 90$	5

3a. State whether the weight of the eggs is a continuous or discrete variable. *[1 mark]*

Markscheme

continuous *(A1)* *(C1)*

[1 mark]

3b. Write down the modal grade of the eggs. *[1 mark]*

Markscheme

6 *(A1)* *(C1)*

Note: Award *(A0)* for an answer of $60 \leq w < 70$.

[1 mark]

3c. Use your graphic display calculator to find an estimate for the standard deviation of the weight of the eggs. *[2 marks]*

Markscheme

8.97 $(8.97479\ldots)$ (g) *(A2)* *(C2)*

[2 marks]

3d. The mean weight of these eggs is 64.9 grams, correct to three significant figures. *[2 marks]*

Use the table and your answer to part (c) to find the **smallest possible** number of eggs that could be within one standard deviation of the mean.

Markscheme

$[55.9,\ 73.9]$ **OR** $55.9252\ldots \leq w \leq 73.8747\ldots$ ***(M1)***

Note: Award ***(M1)*** for correct endpoints seen. If the answer to part (c) is $14.1421\ldots$, award ***(M1)*** for endpoints of $[50.7578\ldots,\ 79.0421\ldots]$.

45 ***(A1)*****(ft)** ***(C2)***

Note: Follow through from their part (c). For a standard deviation between 0 and 5 inclusive, the ***FT*** answer is 0.

[2 marks]

A florist sells bouquets of roses. The florist recorded, in **Table 1**, the number of roses in each bouquet sold to customers.

Table 1

Number of roses in a bouquet (n)	2	3	4	5	6	7	8	9	10	11	12
Number of customers (f)	9	2	4	5	7	3	10	2	3	1	4

The roses can be arranged into bouquets of size small, medium or large. The data from **Table 1** has been organized into a cumulative frequency table, **Table 2**.

Table 2

Bouquet size	**Number of roses (n)**	**Frequency (f)**	**Cumulative frequency**
small	$2 \leq n \leq 4$	15	
medium	$5 \leq n \leq 8$	25	
large	$9 \leq n \leq 12$		

4a. Complete the cumulative frequency table. *[2 marks]*

Markscheme

* This question is from an exam for a previous syllabus, and may contain minor differences in marking or structure.

Bouquet size	Number of roses (n)	Frequency (f)	Cumulative frequency
small	$2 \le n \le 4$	15	15
medium	$5 \le n \le 8$	25	40
large	$9 \le n \le 12$	10	50

(A1)

(A1)**(ft)** ***(C2)***

Note: Award ***(A1)*** for 10; ***(A1)*****(ft)** for the last column all correct. Follow through from *their* 10 for *their* 50 in the last column.

[2 marks]

4b. Write down the probability that a bouquet of roses sold is **not** small. *[2 marks]*

Markscheme

$\frac{35}{50}\ \left(0.7,\ \frac{7}{10},\ 70\,\%\right)$ ***(A1)*****(ft)*****(A1)*****(ft)** ***(C2)***

Note: Award ***(A1)*****(ft)** for their numerator being 25 + *their* 10, and ***(A1)*****(ft)** for their denominator being *their* 50. Follow through from part (a).

[2 marks]

4c. A customer buys a large bouquet. *[2 marks]*

Find the probability that there are 12 roses in this bouquet.

Markscheme

$\frac{4}{10}\ \left(0.4,\ \frac{2}{5},\ 40\,\%\right)$ ***(A1)*****(A1)****(ft)** ***(C2)***

Note: Award ***(A1)*** for a numerator of 4 and ***(A1)*****(ft)** for *their* 10 as denominator. Follow through from part (a).

[2 marks]