Multiple-choice section

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Question | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Answer | D | C | A | D | B | C | C | C | C |

Question 1 [2.6] [10A]

D

10A: mean =  10B: mean = 

Place scores in order to find quartiles:

10A: 55 56 67 67 78 78 83 88 88 98 IQR = 88 − 67 = 21

10B: 50 52 58 58 60 66 72 80 80 97 IQR = 80 − 58 = 22

Difference in means is 8.5.

Difference in IQR is 1.

Question 2 [2.3]

**C**

As the maximum temperature reached may not have occurred at the times when the measurements were recorded you cannot say with any certainty what that value would have been.

Question 3 [2.5]

**A**

QL = 1.5, median = 3 and QU = 5

Question 4 [2.2]

**D**

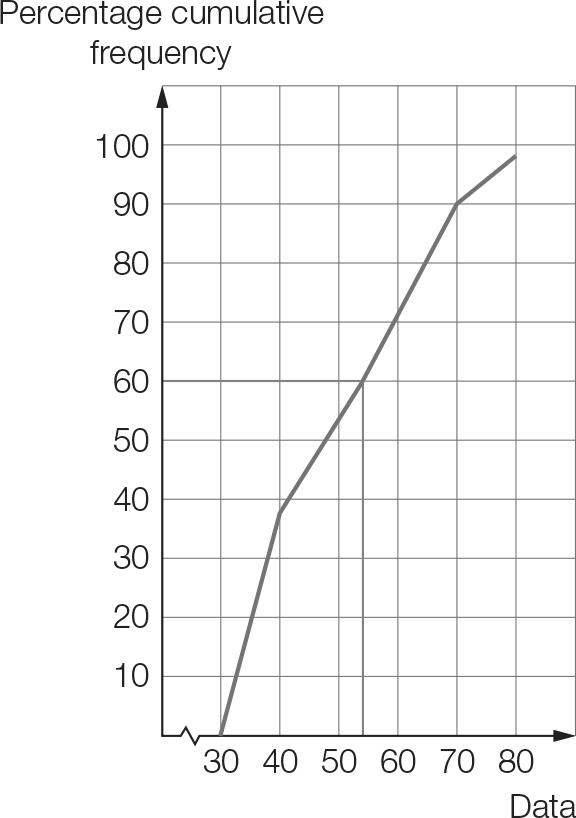
IQR = 25 − 17 = 8 1.5 × IQR = 12

QL − 1.5 × IQR = 17 − 12 = 5 QU + 1.5 × IQR = 25 + 12 = 37

Outliers are less than 5 or more than 37.

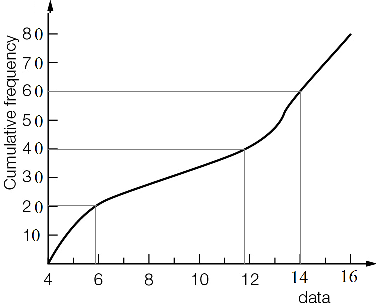
Question 5 [2.2]

**B**



Question 6 [2.1]

**C**



IQR = 14 − 6 = 8

Question 7 [2.2]

**C**

A: The range is 60 ∴ A is incorrect

B: IQR = 0 ∴ B is incorrect

C is correct

D: 25th percentile is 30 ∴ D is incorrect

Question 8 [2.2]

**C**

mean = 3.3, standard deviation = 2.9 (1 d.p.)

Question 9 [2.8] [10A]

**C**

Point A is too low.  
Point B is just too high.  
Point C seems very close.  
Point D is much too low.

Multiple-choice total marks: 9

Short answer section

Question 10 2 marks [2.1–2.7]

(a) A cumulative frequency curve can be used to find the five-number summary from a set of continuous data that is presented in grouped form.

(b) An outlier is a value that is significantly lower or higher than the majority of values in a data set.

Question 11 5 marks [2.1]

(a) To find a suitable set of numbers it is best to start with 15 dashes to represent each value required. Then fill in the highest and lowest values.  
5 \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ 16  
Now, the median is the eighth value but is not a whole number. This indicates we would be better to consider an even number of data values, say 16. To get a median of 10.5, you need two numbers that have a mean of 10.5, so add to 21. Use 10 and 11 as the 8th and 9th values.  
5 \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ 10 11 \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_ 16  
The lower quartile is between the 4th and 5th values, so these will be 7 and 8.  
The upper quartile is between the 12th and 13th values, so these will both be 15.  
5 \_\_\_\_ \_\_\_\_ 7 8 \_\_\_\_ \_\_\_\_ 10 11 \_\_\_\_ \_\_\_\_ 15 15 \_\_\_\_ \_\_\_\_ 16  
The remaining spaces can be filled in with any values that fit numerically with the pattern.  
As an example:  
5 5 6 7 8 9 10 10 11 12 14 15 15 15 16 16  
(The bold values could be different.)

(b) The following value is for the data set given above. Students’ answers will vary.  
Mean = 10.9 (1 d.p.)

Question 12 7 marks [2.1

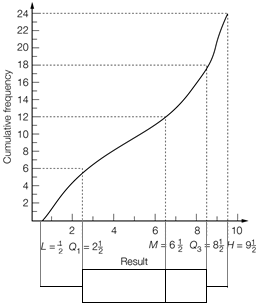
|  |  |  |  |
| --- | --- | --- | --- |
| x | f | x × f | Cumulative  frequency |
| 23 | 8 | 184 | 8 |
| 24 | 10 | 240 | 18 |
| 25 | 3 | 75 | 21 |
| 26 | 4 | 104 | 25 |
| 27 | 1 | 27 | 26 |
| Total | 26 | 630 |  |

(a) Mean =  ( 1 d.p.)

(b) QL is the 7th data value.  
The median is the midpoint of the 13th and 14th data value.  
QU is 20th data value.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lowest | QL | Median | QU | Highest |
| 23 | 23 | 24 | 25 | 27 |

Question 13 3 marks [2.2]



Question 14 5 marks [2.3]

(a) Set A: (i) median = 3 (ii) range = 6 − 1 = 5 (iii) IQR = 4.5 − 2 = 2.5  
Set B: (i) median = 2 (ii) range = 6 − 0 = 6 (iii) IQR = 4 − 0.5 = 3.5

(b) Set A values are higher on average. Set B data is more spread out than Set A. The highest value is in Set A and the lowest value is in Set B.

Question 15 3 marks [2.2]

Place the data in order:  
2.9 11.6 12.5 12.8 13.7 14.1 15.2 25.2

QL = 

QU =  = 14.65

IQR = 14.65 − 12.05 = 2.6

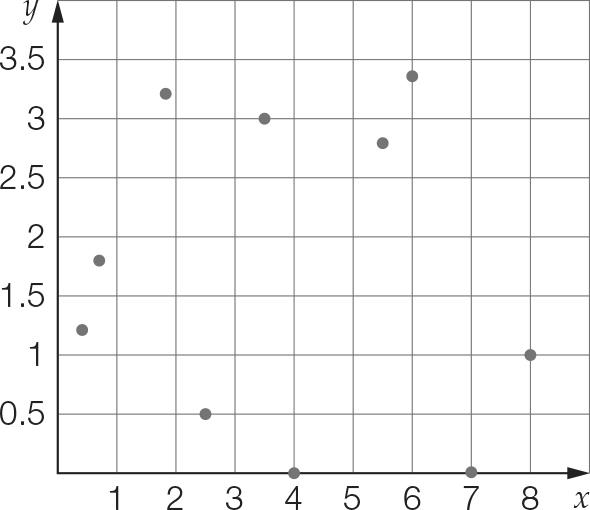
1.5 × IQR = 3.9

QL − 1.5 × IQR = 12.05 − 3.9 = 8.15

QU + 1.5 × IQR = 14.65 + 3.9 = 18.55

So, 2.9 and 25.2 would be considered outliers.

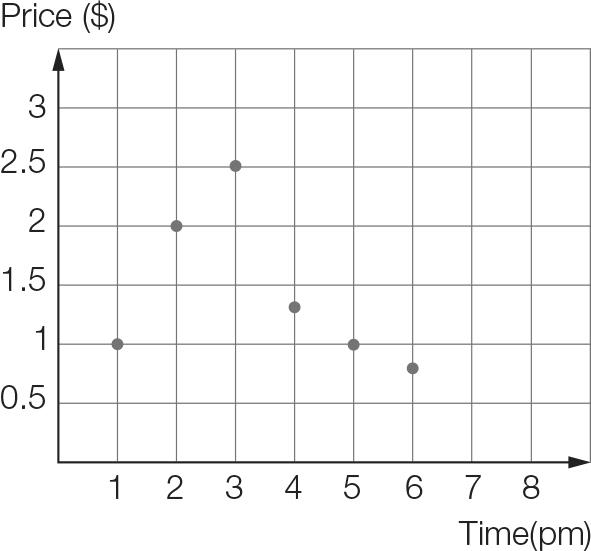
Question 16 4 marks [2.4]



There is no linear relationship evident between x and y.

Question 17 3 marks [2.5]

(a)



(b) At 1 pm the price started at $1 then rose significantly through to 3 pm, when the price had increased by a factor of 2.5. Then the price decreased over the next 3 hours, rapidly at first, to slightly below its initial level.

Question 18 9 marks [2.2]

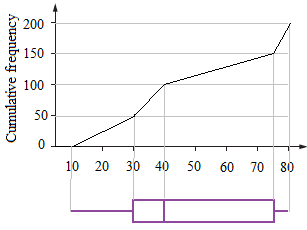
(a) Mean =  ≈ 1939  
Data in order:  
1866 1866 1866 1908 1931 1934 1934 1938 1942 1944 1944 1945 1945  
1947 1950 1953 1954 1954 1955 1955 1956 1961 1963 1963 1968 1974  
Median = 1946

(b) QL = 1938  
QU = 1956  
So the middle 50% of deaths occurred between 1938 and 1956.

(c) IQR = 1956 − 1938 = 18  
QL − 1.5 × IQR = 1938 − 1.5 × 18 = 1911  
QU + 1.5 × IQR = 1956 + 1.5 × 18 = 1983  
So, the three 1866 values and 1908 are outliers.

(d) Replacing all outlier values with 1931:  
Mean ≈ 1947  
Median = 1946  
The mean value increased by 8 years. The median is unchanged.

Question 19 4 marks [2.3]



Question 20 5 marks [2.7]

(a) Look at only the median price information. The median house price in Edgerton is approximately $460 000.

(b) The first section of the graph shows the average interest rates from 2012 to 2016. The second section of the graph is a prediction, based on the current economic status of the country.

(c) The percentage of first homebuyers increased sharply in 2008, which may indicate the presence of financial incentive.

(d) The graph ‘Home ownership: Edgerton’ shows actual data from the previous years, whereas the graph titled ‘Australian interest rates’ is forecasting interest rates.

(e) The interest rate is projected to decrease again, before increasing and remaining steady at approximately 3%. Lower interest rates means homebuyers, who take out a mortgage, pay less interest to a bank or other financial institute for borrowing money.

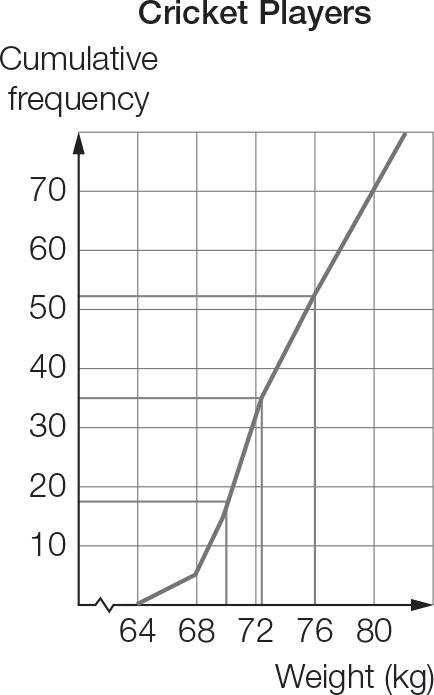
Short answer total marks: 50

Extended answer section

Question 21 10 marks [2.1]

|  |  |  |  |
| --- | --- | --- | --- |
| Weight (kg) | Frequency | Data  value | Cumulative  frequency |
| <64 | 0 | 64 | 0 |
| 64−<66 | 2 | 66 | 2 |
| 66−<68 | 4 | 68 | 6 |
| 68−<70 | 12 | 70 | 18 |
| 70−<72 | 15 | 72 | 33 |
| 72−<74 | 11 | 74 | 44 |
| 74−<76 | 10 | 76 | 54 |
| 76−<78 | 9 | 78 | 63 |
| 78−<80 | 7 | 80 | 70 |

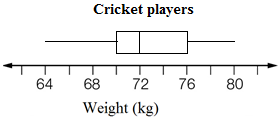
(a)



(b) As shown on the graph, the approximate values for the lower quartile is 70, the median is 72 and upper quartile is 76.

(c)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Min | QL | Median | QU | Max |
| 64 | 70 | 72 | 76 | 80 |



(d) The actual values could have occurred anywhere within the 2 kg class intervals for weights.  
So the minimum value of 64 kg would be in the interval 64–<66 kg.

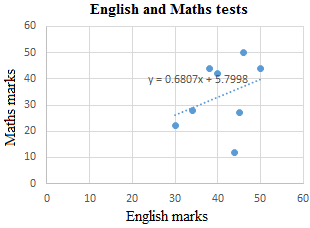
Question 22 10 marks [2.4, 2.6] [10A]

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| English | 45 | 46 | 30 | 34 | 50 | 44 | 38 | 40 |
| Maths | 27 | 50 | 22 | 28 | 44 | 12 | 44 | 42 |

Answers may vary, but should be close to the following.

(a) (i) y = 0.68x + 5.80

(ii)



(iii) Maths mark = English mark × 0.68 + 5.80

(iv) Maths mark = 36 × 0.68 + 5.80 = 30.28  
The predicted mark for Maths is 30.

(v) The Maths mark of 12 deviates most. This is the lowest point on the graph and clearly the furthest from the line measured vertically.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| English | 45 | 46 | 30 | 34 | 50 | 44 | 38 | 40 |
| Maths | 27 | 50 | 22 | 28 | 44 | 12 | 44 | 42 |
| Estimated Maths mark | 36 | 37 | 26 | 29 | 40 | 36 | 32 | 33 |
| Deviation | 9 | 13 | 4 | 1 | 4 | 24 | 12 | 9 |

(b) (i) y = 0.17x + 35.10

(ii) English mark = Maths mark × 0.17 + 35.10

(iii) English mark = 40 × 0.17 + 35.10 = 41.90  
The predicted mark for English is 42.

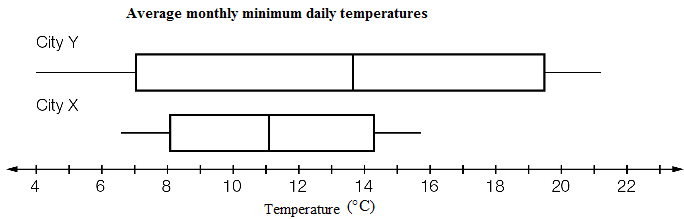
Question 23 15 marks [2.1, 2.3, 2.8]

Order the data:

City X: 6.8 7.4 7.5 8.6 9.4 10.4 11.6 12.3 14.3 14.3 15.5 15.7

City Y: 4 5.1 5.9 8.2 9.7 12.5 14.8 17.4 17.9 20.2 20.6 21.2

(a) City X: min = 6.8 °C, QL = 8.05 °C, median = 11 °C, QU = 14.3 °C, max = 15.7 °C  
City Y: min = 4 °C, QL = 7.05 °C, median = 13.65 °C, QU = 19.05 °C, max = 21.2 °C

(b) 

(c) The temperatures for City Y are, on average, higher and more widely spread than for City X. City Y has both the lowest temperature and the highest temperature.

[10A] (d) City X: mean = 11.2 °C, SD = 3.13 °C  
City Y: mean = 13.1 °C, SD = 6.13 °C

[10A] (e) The temperatures for City Y are, on average, higher and more widely spread than for City X.

Question 24 7 marks [2.1, 2.3]

(a)

|  |  |  |
| --- | --- | --- |
| x | f | x × f |
| 1 | 2 | 2 |
| 2 | 3 | 6 |
| 3 | 5 | 15 |
| 4 | 1 | 4 |
| 5 | 2 | 10 |
| 6 | 0 | 0 |
| 7 | 1 | 7 |
| 8 | 2 | 16 |
| 9 | 1 | 9 |
| 10 | 3 | 30 |
| Total | 20 | 99 |

Mean =  = 4.95

On average, the winning teams won 5 games.

(b) ‘More than half’ is 6–10 games.  
Number of teams  
= 0 + 1 + 2 + 1 + 3  
= 7 teams  
Proportion =  = 0.35  
35% of winning teams won more than half of their games.

(c)

|  |  |  |
| --- | --- | --- |
| x | f | x × f |
| 0 | 15 | 0 |
| 1 | 2 | 2 |
| 2 | 3 | 6 |
| 3 | 5 | 15 |
| 4 | 1 | 4 |
| 5 | 2 | 10 |
| 6 | 0 | 0 |
| 7 | 1 | 7 |
| 8 | 2 | 16 |
| 9 | 1 | 9 |
| 10 | 3 | 30 |
| Total | 35 | 99 |

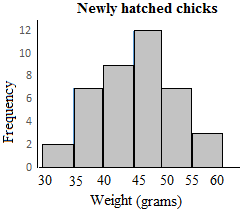
Mean =  = 2.828… ≈ 2.8  
On average, the teams won 2.8 games.

(d) Proportion =  =  or 0.2  
 or 20% of all teams won more than half of their games.

Question 25 18 marks [2.3, 2.8]

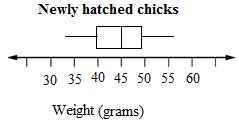
(a)

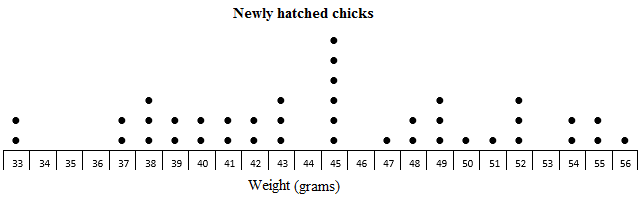
|  |  |
| --- | --- |
| Weight (kg) | Frequency |
| 30–<35 | 2 |
| 35–<40 | 7 |
| 40–<45 | 9 |
| 45–<50 | 12 |
| 50–<55 | 7 |
| 55–<60 | 3 |



(b) The data in order:  
33 33 37 37 38 38 38 39 39 40 | 40 41 41 42 42 43 43 43 45 45  
45 45 45 45 47 48 48 49 49 49 | 50 51 52 52 52 54 54 55 55 56

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Min | QL | Median | QU | Max |
| 33 | 40 | 45 | 49.5 | 56 |



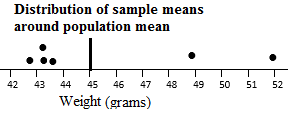
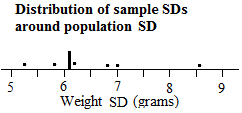
(c) 

(d) The dot plot has the raw data so you can see exactly how many of each weight. The weights range from 33 g to 56 g with six chicks at 45 g. The histogram shows a symmetrical distribution with a peak in the middle. Weights appear to range from 33 g to 60 g. The box plot shows the weights from 33 g to 56 g with an even spread overall.

[10A] (e) From a calculator (1 d.p.):  
Mean = 45.0 g, SD = 6.1 g

[10A] (f) (These are sample answers; students will almost certainly have different values.)  
14 24 9 11 28 → 52, 54, 49, 38, 52  
27 28 39 18 15 → 38, 52, 41, 33, 52  
4 16 15 24 6 → 54, 55, 52, 43, 56  
30 18 9 40 21 → 45, 33, 49, 49, 40  
2 19 33 4 11 → 45, 37, 38, 54, 40  
27 24 40 26 11 → 38, 43, 49, 50, 38  
Sample 1: 52, 54, 49, 38, 52  
Sample 2: 38, 52, 41, 33, 52  
Sample 3: 54, 55, 52, 43, 56  
Sample 4: 45, 33, 49, 49, 40  
Sample 5: 45, 37, 38, 54, 40  
Sample 6: 38, 43, 49, 50, 38

[10A] (g) From a calculator, using sample standard deviations:  
Sample 1: mean = 49, SD = 6.40  
Sample 2: mean = 43.2, SD = 8.53  
Sample 3: mean = 52, SD = 5.24  
Sample 4: mean = 43.2, SD = 6.80  
Sample 5: mean = 42.8, SD = 6.98  
Sample 6: mean = 43.6, SD = 5.77

[10A] (h)   
  


Extended answer results: \_\_\_ / 60

TOTAL test results: \_\_\_ / 119