Representations of Data - Edexcel Past Exam Questions

1. The number of caravans on Seaview caravan site on each night in August last year is summarised as follows: the least number of caravans was 10. The maximum number of caravans on this site was 64. The three quartiles for this site was 33, 41 and 52 respectively

During a month, the least number of caravans on Northcliffe caravan site was 31.
The maximum number of caravans on this site on any night that month was 72.
The three quartiles for this site were 38, 45 and 52 respectively.

(*a*) On graph paper and using the same scale, draw box plots to represent the data for both caravan sites. You may assume that there are no outliers. **(6)**

(*b*) Compare and contrast these two box plots. **(3)**

(*c*) Give an interpretation to the upper quartiles of these two distributions. **(2)**

Jan 05 Q2*(edited)*

---

2. Aeroplanes fly from City *A* to City *B*. Over a long period of time the number of minutes delay in take-off from City *A* was recorded. The minimum delay was 5 minutes and the maximum delay was 63 minutes. A quarter of all delays were at most 12 minutes, half were at most 17 minutes and 75% were at most 28 minutes. Only one of the delays was longer than 45 minutes.

An outlier is an observation that falls either 1.5 × (interquartile range) above the upper quartile or 1.5 × (interquartile range) below the lower quartile.

(*a*) On graph paper, draw a box plot to represent these data. **(7)**

(*b*) Suggest how the distribution might be interpreted by a passenger who frequently flies from City *A* to City *B*. **(1)**

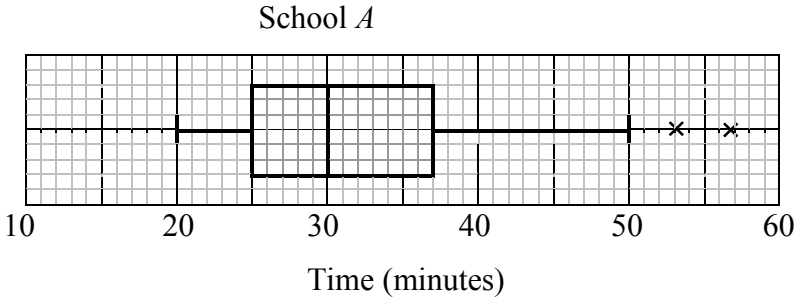
June 05 Q4*(edited)*

---

3. (*a*) Describe the main features and uses of a box plot. **(3)**

Children from schools *A* and *B* took part in a fun run for charity. The times, to the nearest minute, taken by the children from school *A* are summarised in Figure 1.

Figure 1



(*b*) (i) Write down the time by which 75% of the children in school *A* had completed the run.

(ii) State the name given to this value. **(2)**

(*c*) Explain what you understand by the two crosses (×) on Figure 1. **(2)**

For school *B* the least time taken by any of the children was 25 minutes and the longest time was 55 minutes. The three quartiles were 30, 37 and 50 respectively.

(*d*) On graph paper, draw a box plot to represent the data from school *B*. **(4)**

(*e*) Compare and contrast these two box plots. **(4)**

June 06 Q1

4. A teacher recorded, to the nearest hour, the time spent watching television during a particular week by each child in a random sample. The times were summarised in a grouped frequency table and represented by a histogram.

One of the classes in the grouped frequency distribution was 20–29 and its associated frequency was 9. On the histogram the height of the rectangle representing that class was 3.6 cm and the width was 2 cm.

(*a*) Give a reason to support the use of a histogram to represent these data. **(1)**

(*b*) Write down the underlying feature associated with each of the bars in a histogram. **(1)**

(*c*) Show that on this histogram each child was represented by 0.8 cm^2. **(3)**

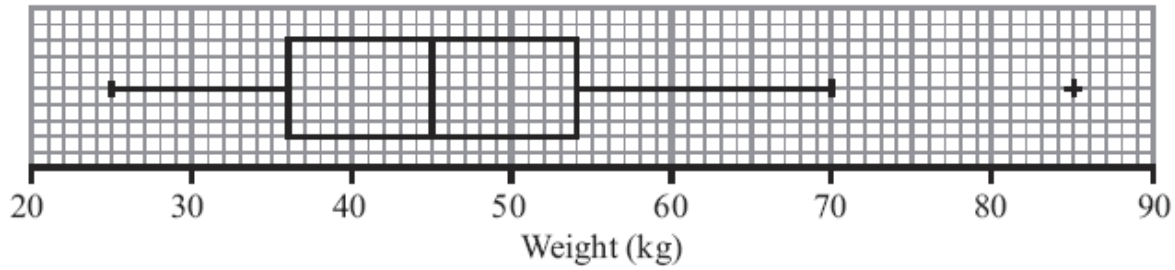
The total area under the histogram was 24 cm^2.

(*d*) Find the total number of children in the group. **(2)**

Jan 07 Q5

---

5. The box plot shows a summary of the weights of the luggage, in kg, for each musician in an orchestra on an overseas tour.



The airline's recommended weight limit for each musician's luggage was 45 kg.

Given that none of the musician's luggage weighed exactly 45 kg,

(*a*) state the proportion of the musicians whose luggage was below the recommended weight limit. **(1)**

A quarter of the musicians had to pay a charge for taking heavy luggage.

(*b*) State the smallest weight for which the charge was made. **(1)**

(*c*) Explain what you understand by the + on the box plot in Figure 1, and suggest an instrument that the owner of this luggage might play. **(2)**

June 07 Q2*(edited)*

6.

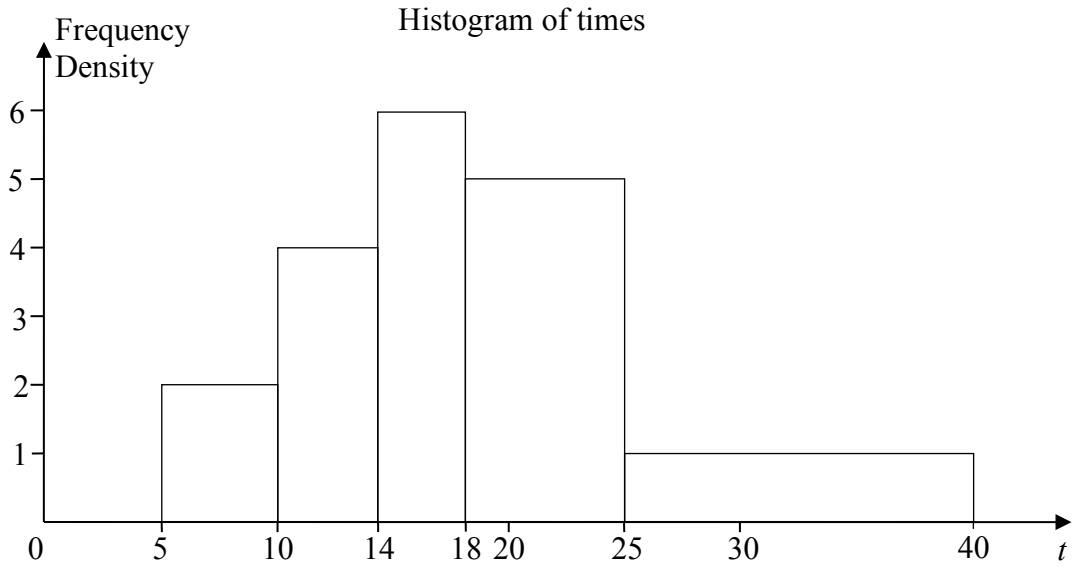


Figure 2

Figure 2 shows a histogram for the variable t which represents the time taken, in minutes, by a group of people to swim 500 m.

(a) Copy and complete the frequency table for t.

t	5 – 10	10 – 14	14 – 18	18 – 25	25 – 40
Frequency	10	16	24		

(2)

(b) Estimate the number of people who took longer than 20 minutes to swim 500 m.

(2)

(c) Find an estimate of the mean time taken.

(4)

(d) Find an estimate for the standard deviation of t.

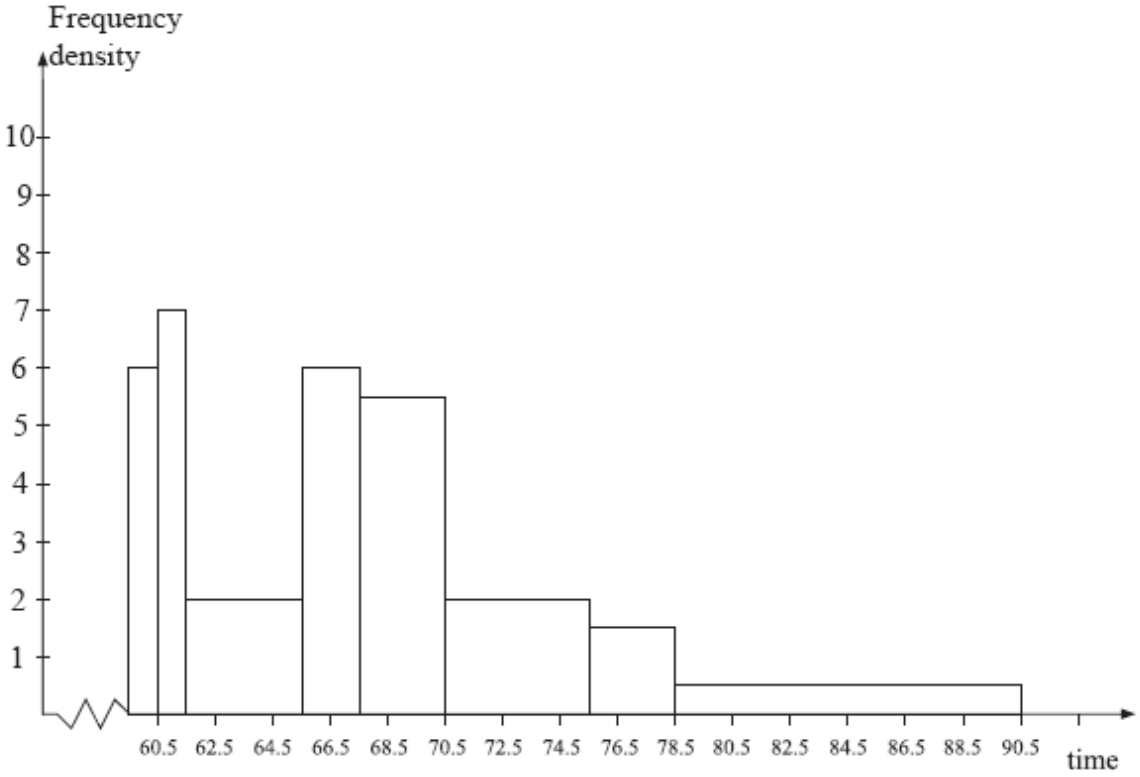
(3)

(e) Find the median and quartiles for t.

(4)

June 07 Q5*(edited)*

7. The histogram shows the time taken, to the nearest minute, for 140 runners to complete a fun run.



Use the histogram to calculate the number of runners who took between 78.5 and 90.5 minutes to complete the fun run. **(5)**

Jan 08 Q3

---

8. In a study of how students use their mobile telephones, the phone usage of a random sample of 11 students was examined for a particular week.

The total length of calls, y minutes, for the 11 students were

17, 23, 35, 36, 51, 53, 54, 55, 60, 77, 110

(a) Find the median and quartiles for these data. **(3)**

A value that is greater than $Q_3 + 1.5 \times (Q_3 - Q_1)$ or smaller than $Q_1 - 1.5 \times (Q_3 - Q_1)$ is defined as an outlier.

(b) Show that 110 is the only outlier. **(2)**

(c) Draw a box plot for these data indicating clearly the position of the outlier. **(3)**

Jan 09 Q4*(edited)*

---

9. In a shopping survey a random sample of 104 teenagers were asked how many hours, to the nearest hour, they spent shopping in the last month. The results are summarised in the table below.

Number of hours	Mid-point	Frequency
0 – 5	2.75	20
6 – 7	6.5	16
8 – 10	9	18
11 – 15	13	25
16 – 25	20.5	15
26 – 50	38	10

A histogram was drawn and the group (8 – 10) hours was represented by a rectangle that was 1.5 cm wide and 3 cm high.

(*a*) Calculate the width and height of the rectangle representing the group (16 – 25) hours.

(3)

(*b*) Use linear interpolation to estimate the median and interquartile range.

(5)

(*c*) Estimate the mean and standard deviation of the number of hours spent shopping.

(4)

(*d*) State, giving a reason, which average and measure of dispersion you would recommend to use to summarise these data.

(2)

Jan 09 Q5*(edited)*

10. The variable x was measured to the nearest whole number. Forty observations are given in the table below.

x	10 – 15	16 – 18	19 –
Frequency	15	9	16

A histogram was drawn and the bar representing the 10 – 15 class has a width of 2 cm and a height of 5 cm. For the 16 – 18 class find

(a) the width, **(1)**

(b) the height **(2)**
of the bar representing this class.

June 09 Q3

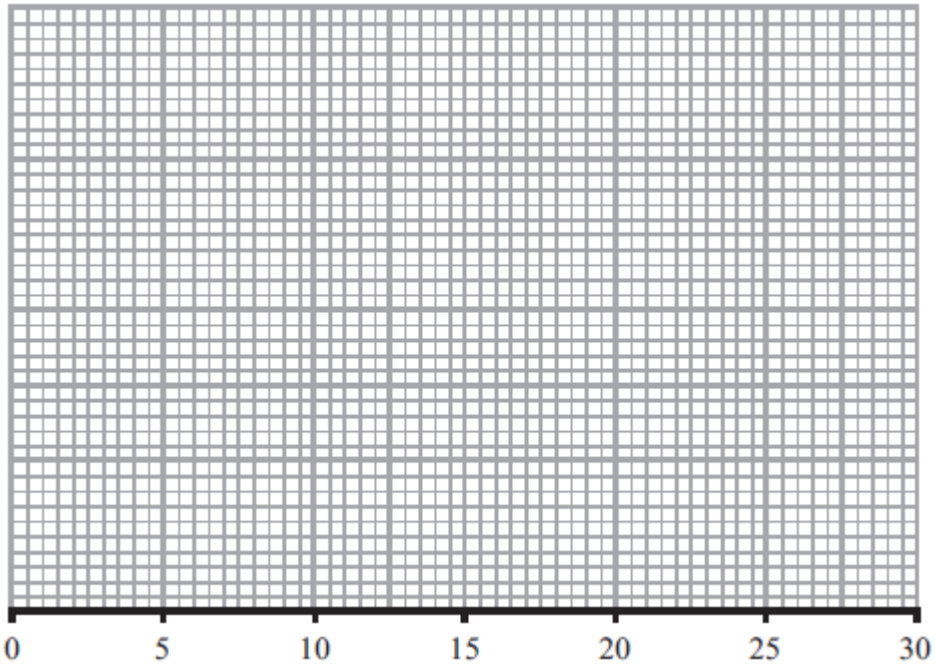
11. Over a long period of time a small company recorded the amount it received in sales per month. The results are summarised below.

	Amount received in sales (£1000s)
Two lowest values	3, 4
Lower quartile	7
Median	12
Upper quartile	14
Two highest values	20, 25

An outlier is an observation that falls
either 1.5 × interquartile range above the upper quartile
or 1.5 × interquartile range below the lower quartile.

(*a*) On the graph paper below, draw a box plot to represent these data, indicating clearly any outliers.

(5)



(*b*) The company claims that for 75 % of the months, the amount received per month is greater than £10 000. Comment on this claim, giving a reason for your answer.

(2)

Jan 11 Q3*(edited)*