

---

# Stats1 Chapter 2: Measures of Data

## Quartiles

# Quickfire Questions...

What position do we use for the median?

Lengths: 3cm, 5cm, 6cm, ...  
 $n = 11$

Median position: **6<sup>th</sup>**

Lengths: 4m, 8m, 12.4m, ...  
 $n = 24$

Median position: **12<sup>th</sup>/13<sup>th</sup>**

Age	Freq
$10 \leq a < 20$	12
$20 \leq a < 30$	5

Median position: **8.5**

Score	Freq
$150 \leq s < 200$	3
$200 \leq s < 400$	7

Median position: **5**

Ages: 5, 7, 7, 8, 9, 10, ...  
 $n = 60$

Median position: **30<sup>th</sup>/31<sup>st</sup>**

Score	Freq
$150 \leq s < 200$	15
$200 \leq s < 400$	6

Median position: **10.5**

Weights: 1.2kg, 3.3kg, ...  
 $n = 35$

Median position: **18<sup>th</sup>**

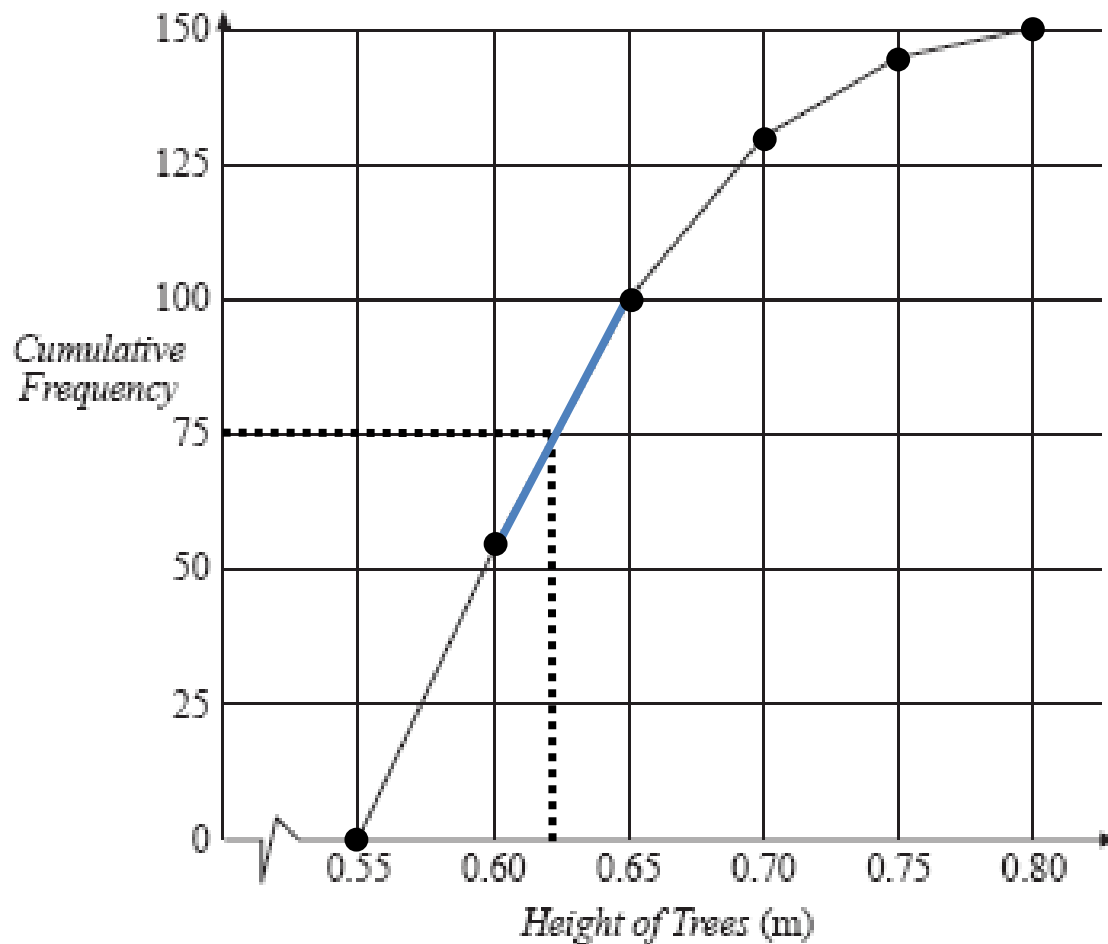
Volume (ml)	Freq
$0 \leq v < 100$	5
$100 \leq v < 200$	6
$200 \leq v < 300$	2

Median position: **6.5**

Weights: 4.4kg, 7.6kg, 7.7kg...  
 $n = 18$

Median position: **9<sup>th</sup>/10<sup>th</sup>**

# Linear Interpolation

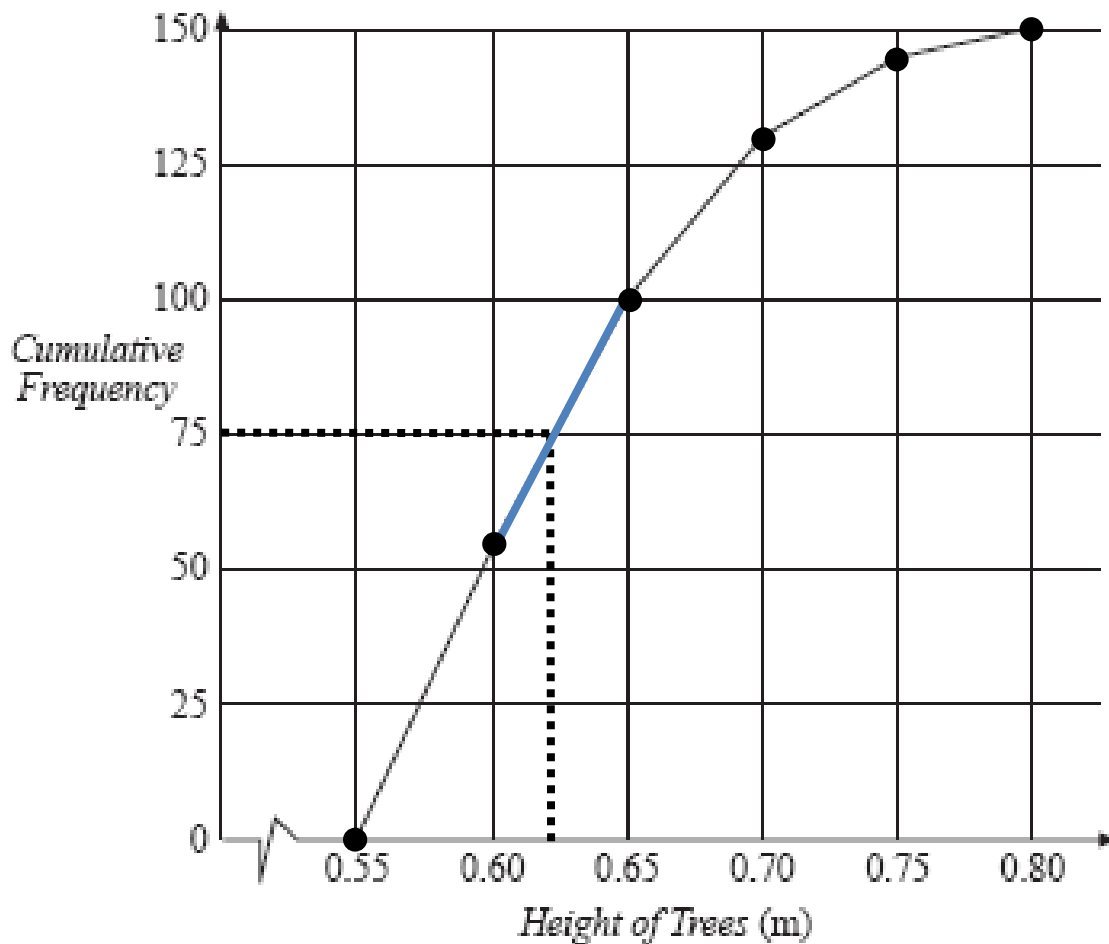


Height of tree (m)	Freq	C.F.
$0.55 \leq h < 0.6$	55	55
$0.6 \leq h < 0.65$	45	100
$0.65 \leq h < 0.7$	30	130
$0.7 \leq h < 0.75$	15	145
$0.75 \leq h < 0.8$	5	150

At GCSE we could find the median by drawing a suitable line on a cumulative frequency graph. How could we read off this value exactly using a suitable calculation?

**We could find the fraction of the way along the line segment using the frequencies, then go this same fraction along the class interval.**

# Linear Interpolation



Height of tree (m)	Freq	C.F.
$0.55 \leq h < 0.6$	55	55
$0.6 \leq h < 0.65$	45	100
$0.65 \leq h < 0.7$	30	130
$0.7 \leq h < 0.75$	15	145
$0.75 \leq h < 0.8$	5	150

The 75<sup>th</sup> item is within the  $0.6 \leq h < 0.65$  class interval because 75 is within the first 100 items but not the first 55.

Frequency up until this interval

55

Item number we're interested in.

75

Frequency by end of this interval

100

Height at start of interval.

0.6m

Med

0.65m

Height by end of interval.

# Linear Interpolation

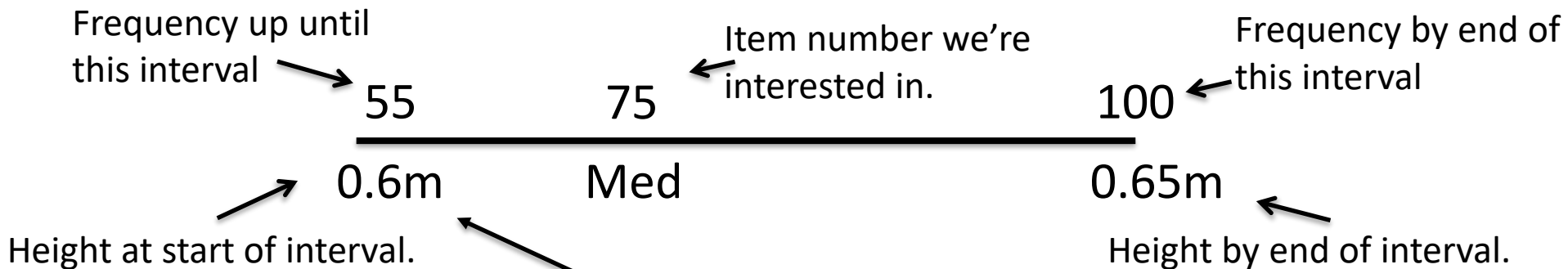
Height of tree (m)	Freq	C.F.
$0.55 \leq h < 0.6$	55	55
$0.6 \leq h < 0.65$	45	100
$0.65 \leq h < 0.7$	30	130
$0.7 \leq h < 0.75$	15	145
$0.75 \leq h < 0.8$	5	150

What fraction of the way across the interval are we?

$$\frac{20}{45}$$

Hence:

$$\text{Median} = 0.6 + \left( \frac{20}{45} \times 0.05 \right)$$



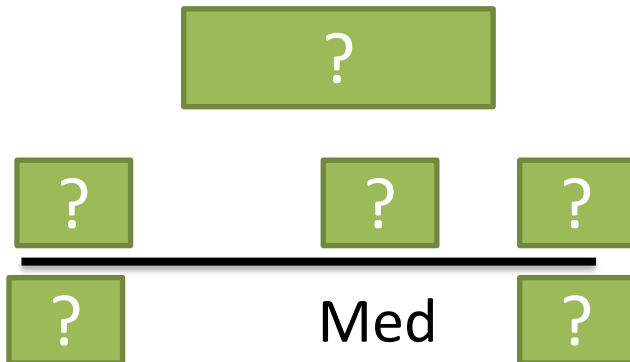
**Fro Tip:** I like to put the units to avoid getting frequencies confused with values of the variable.

**Fro Tip:** To quickly get frequency before and after, just look for the two cumulative frequencies that surround the item number.

# More Examples

Weight of cat (kg)	Freq	C.F.
$1.5 \leq w < 3$	10	10
$3 \leq w < 4$	8	18
$4 \leq w < 6$	14	32

Median class interval:



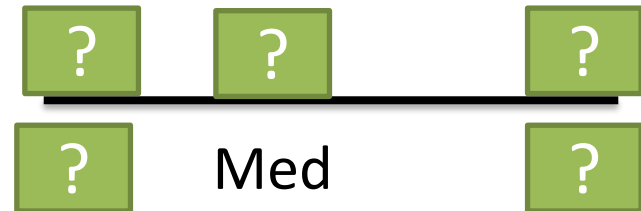
Fraction along interval:



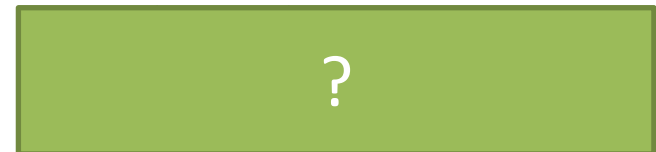
Median:



Time (s)	Freq	C.F.
$8 \leq t < 10$	4	4
$10 \leq t < 12$	3	7
$12 \leq t < 14$	13	20



Median:



# More Examples

Weight of cat (kg)	Freq	C.F.
$1.5 \leq w < 3$	10	10
$3 \leq w < 4$	8	18
$4 \leq w < 6$	14	32

Median class interval:

$$3 \leq w < 4$$

$$\begin{array}{ccc} 10 & 16 & 18 \\ \hline 3\text{kg} & \text{Med} & 4\text{kg} \end{array}$$

Fraction along interval:

$$\frac{6}{8}$$

Median:

$$3 + \left( \frac{6}{8} \times 1 \right) = 3.75\text{kg}$$

Time (s)	Freq	C.F.
$8 \leq t < 10$	4	4
$10 \leq t < 12$	3	7
$12 \leq t < 14$	13	20

$$\begin{array}{ccc} 7 & 10 & 20 \\ \hline 12\text{s} & \text{Med} & 14\text{s} \end{array}$$

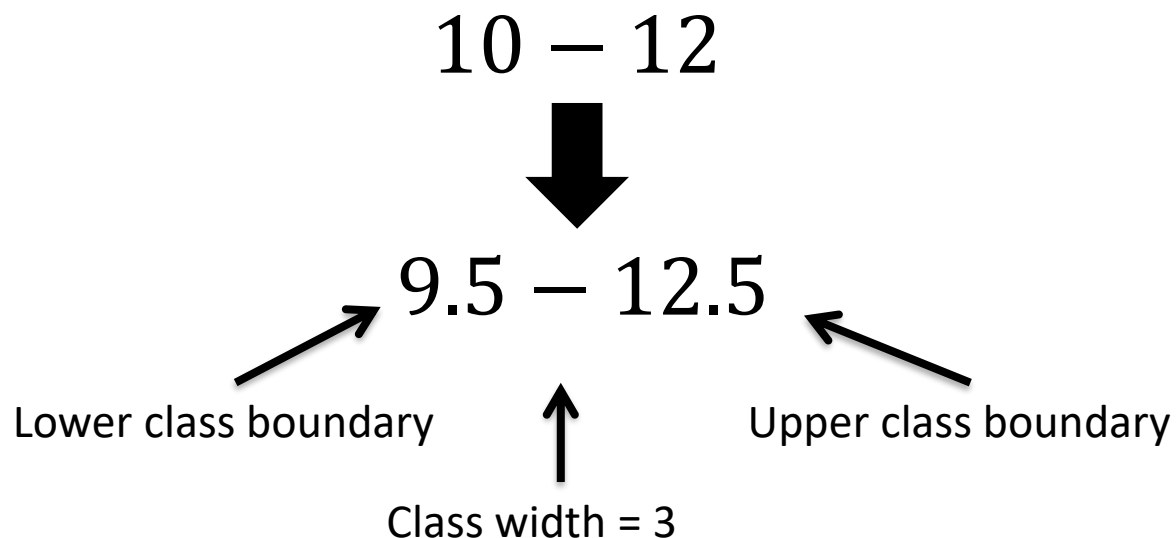
Median:

$$12 + \left( \frac{3}{13} \times 2 \right) = 12.5 \text{ (to 3sf)}$$

# What's different about the intervals here?

Weight of cat to nearest kg	Frequency
10 – 12	7
13 – 15	2
16 – 18	9
19 – 20	4

There are **GAPS** between intervals!  
What interval does this **actually** represent?





# Identify the class width

Distance $d$ travelled (in m)	...
$0 \leq d < 150$	
$150 \leq d < 200$	
<b><math>200 \leq d &lt; 210</math></b>	

Lower class boundary =

?

Class width =

?

Weight $w$ in kg	...
$10 - 20$	
$21 - 30$	
<b><math>31 - 40</math></b>	

Lower class boundary =

?

Class width =

?

Time $t$ taken (in seconds)	...
$0 - 3$	
<b><math>4 - 6</math></b>	
$7 - 11$	

Lower class boundary =

?

Class width =

?

Speed $s$ (in mph)	...
$10 \leq s < 20$	
$20 \leq s < 29$	
<b><math>29 \leq s &lt; 31</math></b>	

Lower class boundary =

?

Class width =

?

# Identify the class width

Distance $d$ travelled (in m)	...
$0 \leq d < 150$	
$150 \leq d < 200$	
<b><math>200 \leq d &lt; 210</math></b>	

Lower class boundary = 200

Class width = 10

Weight $w$ in kg	...
10 – 20	
21 – 30	
<b>31 – 40</b>	

Lower class boundary = 30.5

Class width = 10

Time $t$ taken (in seconds)	...
0 – 3	
<b>4 – 6</b>	
7 – 11	

Lower class boundary = 3.5

Class width = 3

Speed $s$ (in mph)	...
$10 \leq s < 20$	
$20 \leq s < 29$	
<b><math>29 \leq s &lt; 31</math></b>	

Lower class boundary = 29

Class width = 2

# Linear Interpolation with gaps

Edexcel S1 Jan 2007 Q4

Summarised below are the distances, to the nearest mile, travelled to work by a random sample of 120 commuters.

Distance (to the nearest mile)	Number of commuters	
0 – 9	10	10
10 – 19	19	29
20 – 29	43	72
30 – 39	25	97
40 – 49	8	105
50 – 59	6	111
60 – 69	5	116
70 – 79	3	119
80 – 89	1	120

For this distribution,

- (a) describe its shape, (1)  
(b) use linear interpolation to estimate its median. (2)



# Linear Interpolation with gaps

Edexcel S1 Jan 2007 Q4

Summarised below are the distances, to the nearest mile, travelled to work by a random sample of 120 commuters.

Distance (to the nearest mile)	Number of commuters	
0 – 9	10	10
10 – 19	19	29
20 – 29	43	72
30 – 39	25	97
40 – 49	8	105
50 – 59	6	111
60 – 69	5	116
70 – 79	3	119
80 – 89	1	120

For this distribution,

- (a) describe its shape, (1)  
 (b) use linear interpolation to estimate its median. (2)

$\frac{29}{19.5 \text{ miles}}$ 
 $\frac{60}{\text{Med}}$ 
 $\frac{72}{29.5 \text{ miles}}$

$$\text{Median} = 19.5 + \left( \frac{31}{43} \times 10 \right) = 26.7 \text{ kg}$$

# Test Your Understanding

Age of relic (years)	Frequency
0-1000	24
1001-1500	29
1501-1700	12
1701-2000	35

*Median* =

=

?

Shark length (cm)	Frequency
$40 \leq x < 100$	17
$100 \leq x < 300$	5
$300 \leq x < 600$	8
$600 \leq x < 1000$	10

*Median* =

=

?

# Test Your Understanding

Age of relic (years)	Frequency
0-1000	24
1001-1500	29
1501-1700	12
1701-2000	35

$$\begin{aligned} \text{Median} &= 1000.5 + \left( \frac{26}{29} \times 500 \right) \\ &= \mathbf{1448.7 \text{ years}} \end{aligned}$$

Shark length (cm)	Frequency
$40 \leq x < 100$	17
$100 \leq x < 300$	5
$300 \leq x < 600$	8
$600 \leq x < 1000$	10

$$\begin{aligned} \text{Median} &= 100 + \left( \frac{3}{5} \times 200 \right) \\ &= \mathbf{220cm} \end{aligned}$$

# Supplementary Exercise 1

1

Questions should be on a printed sheet...

The number of patients attending a hospital trauma clinic each day was recorded over several months, giving the data in the table below.

Number of patients	10 - 19	20 - 29	30 - 34	35 - 39	40 - 44	45 - 49	50 - 69
Frequency	2	18	24	30	27	14	5

Use linear interpolation to estimate the median of these data.

**Median** =

?

2

The ages of 300 houses in a village are recorded given the following table of results.

Age $a$ (years)	Number of houses
$0 \leq a < 20$	36
$20 \leq a < 40$	92
$40 \leq a < 60$	74
$60 \leq a < 100$	39
$100 \leq a < 200$	14
$200 \leq a < 300$	27
$300 \leq a < 500$	18

Use linear interpolation to estimate the median.

**Median** =

?

3

A cyber-café recorded how long each user stayed during one day giving the following results.

Length of stay (minutes)	Number of houses
$0 \leq l < 30$	15
$30 \leq l < 60$	31
$60 \leq l < 90$	32
$90 \leq l < 120$	23
$120 \leq l < 240$	17
$240 \leq l < 360$	2

Use linear interpolation to estimate the median of these data.

**Median** =

?

# Supplementary Exercise 1

1

Questions should be on a printed sheet...

The number of patients attending a hospital trauma clinic each day was recorded over several months, giving the data in the table below.

Number of patients	10 - 19	20 - 29	30 - 34	35 - 39	40 - 44	45 - 49	50 - 69
Frequency	2	18	24	30	27	14	5

Use linear interpolation to estimate the median of these data.

$$\text{Median} = 34.5 + \left( \frac{16}{30} \times 5 \right) = 37.2$$

2

The ages of 300 houses in a village are recorded given the following table of results.

Age $a$ (years)	Number of houses
$0 \leq a < 20$	36
$20 \leq a < 40$	92
$40 \leq a < 60$	74
$60 \leq a < 100$	39
$100 \leq a < 200$	14
$200 \leq a < 300$	27
$300 \leq a < 500$	18

Use linear interpolation to estimate the median.

$$\text{Median} = 40 + \left( \frac{22}{74} \times 20 \right) = 45.9$$

3

A cyber-café recorded how long each user stayed during one day giving the following results.

Length of stay (minutes)	Number of houses
$0 \leq l < 30$	15
$30 \leq l < 60$	31
$60 \leq l < 90$	32
$90 \leq l < 120$	23
$120 \leq l < 240$	17
$240 \leq l < 360$	2

Use linear interpolation to estimate the median of these data.

$$\text{Median} = 60 + \left( \frac{14}{32} \times 30 \right) = 73.125$$



# Supplementary Exercise 1

4

The following table summarises the times,  $t$  minutes to the nearest minute, recorded for a group of students to complete an exam.

Time (minutes) $t$	11 – 20	21 – 25	26 – 30	31 – 35	36 – 45	46 – 60
Number of students $f$	62	88	16	13	11	10

[You may use  $\sum ft^2 = 134281.25$ ]

(a) Estimate the mean of these data. **(5)**

(b) Use linear interpolation to estimate the value of the median. **(2)**

**Mean** =

?

**Median** =

?

# Supplementary Exercise 1

4

The following table summarises the times,  $t$  minutes to the nearest minute, recorded for a group of students to complete an exam.

Time (minutes) $t$	11 – 20	21 – 25	26 – 30	31 – 35	36 – 45	46 – 60
Number of students $f$	62	88	16	13	11	10

[You may use  $\sum ft^2 = 134281.25$ ]

(a) Estimate the mean of these data. **(5)**

(b) Use linear interpolation to estimate the value of the median. **(2)**

$$\text{Mean} = \frac{4837.5}{200} = 24.1875$$

$$\text{Median} = 20.5 + \left( \frac{38}{88} \times 5 \right) = 22.7$$

# Exercises 2.2

Pearson Statistics & Mechanics Year 1/AS

Pages 9-10

---

# Homework Exercise

- 1 The weekly wages (to the nearest £) of the production line workers in a small factory is shown in the table.

- Write down the modal class.
- Calculate an estimate of the mean wage.
- Write down the interval containing the median.

Weekly wage (£)	Frequency
175–225	4
226–300	8
301–350	18
351–400	28
401–500	7

- 2 The noise levels at 30 locations near an outdoor concert venue were measured to the nearest decibel. The data collected is shown in the grouped frequency table.

Noise (decibels)	65–69	70–74	75–79	80–84	85–89	90–94	95–99
Frequency	1	4	6	6	8	4	1

- Calculate an estimate of the mean noise level. (1 mark)
- Explain why your answer to part a is an estimate. (1 mark)

- 3 The table shows the daily mean temperature at Heathrow in October 1987 from the large data set.

© Crown Copyright Met Office

Temp (°C)	$6 \leq t < 8$	$8 \leq t < 10$	$10 \leq t < 12$	$12 \leq t < 14$	$14 \leq t < 16$	$16 \leq t < 18$
Frequency	3	7	9	7	3	2

- Write down the modal class. (1 mark)
  - Calculate an estimate for the mean daily mean temperature. (1 mark)
- 4 Two DIY shops (A and B) recorded the ages of their workers.

Age of worker	16–25	26–35	36–45	46–55	56–65	66–75
Frequency A	5	16	14	22	26	14
Frequency B	4	12	10	28	25	13

By comparing estimated means for each shop, determine which shop is better at employing older workers.

## Problem-solving

Since age is always rounded **down**, the class boundaries for the 16–25 group are 16 and 26. This means that the midpoint of the class is 21.

# Homework Answers

- 1 a £351 to £400      b £345      c £351 to £400
- 2 a 82.3 decibels  
b The mean is an estimate as we don't know the exact noise levels recorded.
- 3 a  $10 < t \leq 12$   
b  $11.4^{\circ}\text{C}$
- 4 Store B (mean 51 years) employs older workers than store A (mean 50 years).