

Andy is investigating the lateness of the trains that arrive at Manchester Piccadilly.

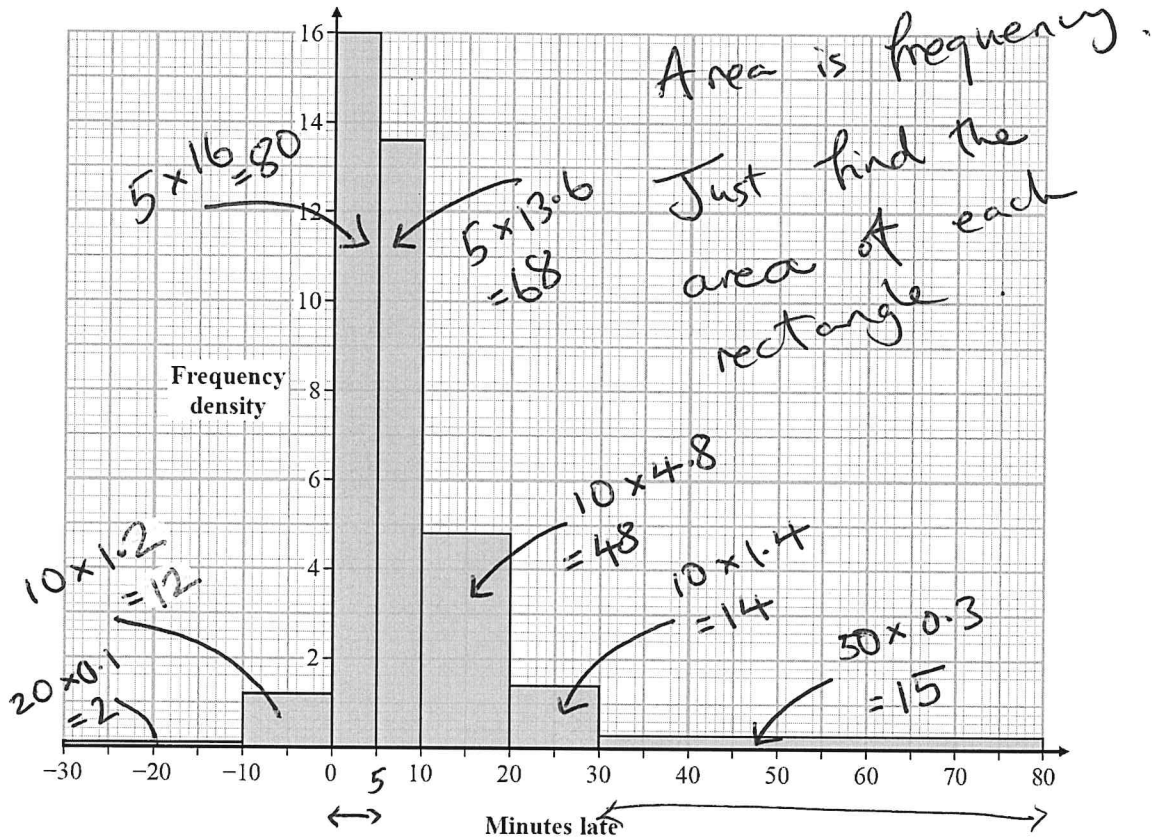
(a) Explain why it is sensible for Andy to collect secondary data for his investigation.

The train company already have collected it.

Q1 It would take huge amount of time and effort to collect primary data on this (1)

Andy found out data on the lateness of each of 238 trains arriving at Manchester Piccadilly from London on Monday to Friday for one week.

The histogram below was drawn using Andy's data.



Histograms

A negative time for a train means that the train arrived early at Manchester Piccadilly.

In the week, 14 trains arrived early.

(b) Find how many of the 238 trains did **not** arrive more than 10 minutes late at Manchester Piccadilly.

$$14 + 80 + 68 = 162$$

162

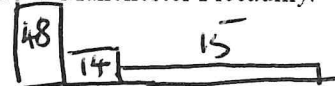
(3)

For a long distance journey, such as London to Manchester Piccadilly, a train is only considered to be late when it arrives more than 10 minutes late.

Andy says that more than 1 in 3 trains from London arrive late at Manchester Piccadilly.

77 out of 238 are late (c) Determine whether or not Andy is correct.

77/238 = 0.323, 32.3% You should comment on the reliability of your conclusion.



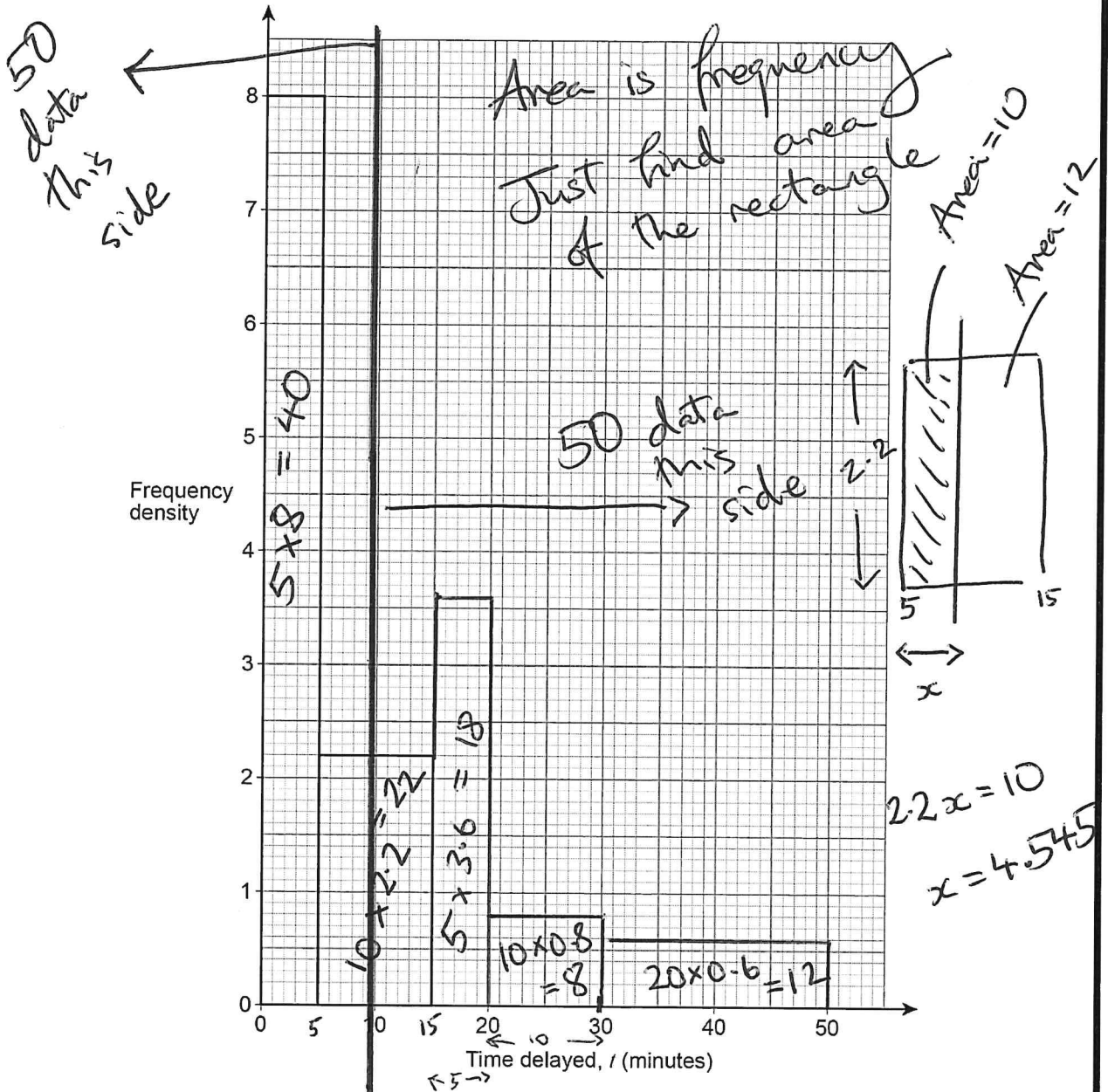
so Andy is almost precisely correct

The table shows, for one train company, some data about the amount of time by which 100 trains were delayed.

Q2

Time delayed, t (minutes)	Frequency
$0 < t \leq 5$	40
$5 < t \leq 15$	22
$15 < t \leq 20$	18
$20 < t \leq 30$	8
$30 < t \leq 50$	12

Here is a partly completed histogram for the data in the table. 100



By using the histogram or the table, calculate an estimate for the median value.

[3 marks]

Median is 50th value.

This is in the 2nd bar. 10 data values into 2nd bar

$5 + 4.545$ Answer 9.55 minutes

The frequency table below shows the masses, in kilograms, of all suitcases weighed at an airport security section before one flight.

Q3

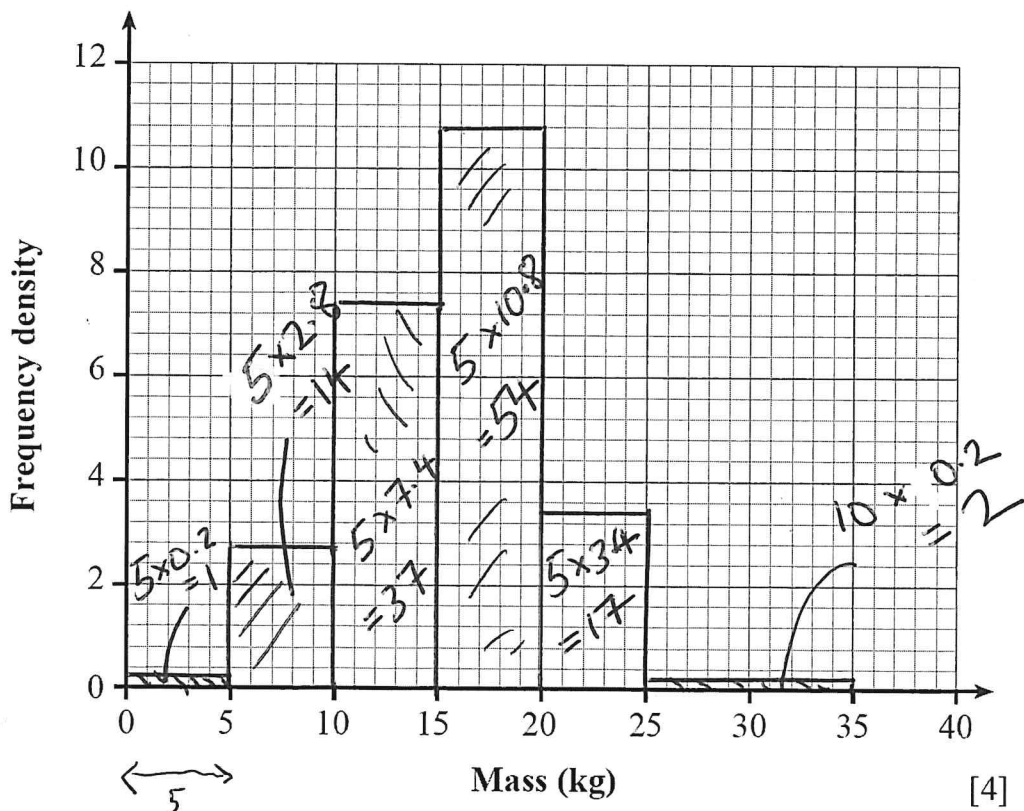
Mass, m (kg)	Frequency	freq. density		
$0 \leq m < 5$	1	0.2		
$5 \leq m < 10$	14	2.8		
$10 \leq m < 15$	37	7.4		
$15 \leq m < 20$	54	10.8		
$20 \leq m < 25$	17	3.4		
$25 \leq m < 35$	2	0.2		

- (a) Give one advantage and one disadvantage of using a grouped frequency table.

Advantage Easier to do calculations [1]

Disadvantage Loss of original data [1]

- (e) Draw a histogram to display the data from the frequency table.

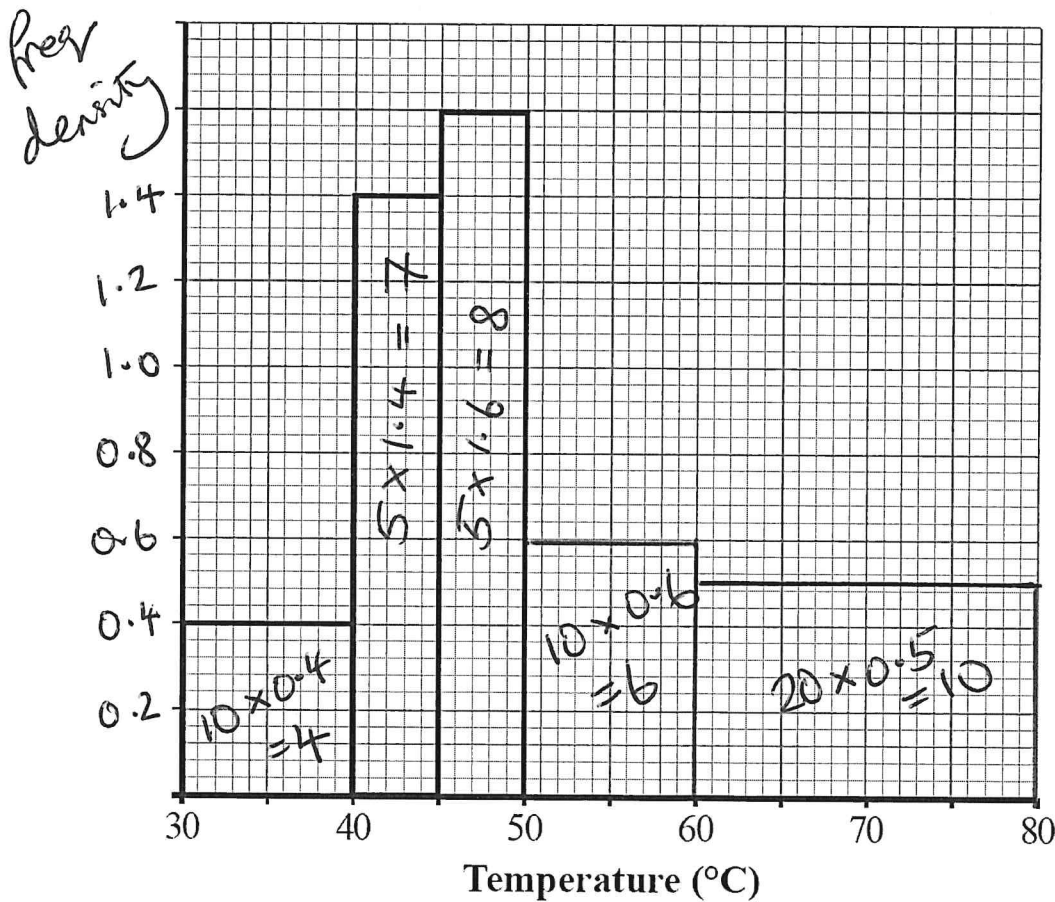


Sam records the typical operating temperature (in °C) of different types of engines as shown in the table below.

Q4

Temperature (°C)	Frequency	class width	freq density
$30 < x \leq 40$	4	10	0.4
$40 < x \leq 45$	7	5	1.4
$45 < x \leq 50$	8	5	1.6
$50 < x \leq 60$	6	10	0.6
$60 < x \leq 80$	10	20	0.5

Use the data in the table to complete the histogram below.



Just check by finding the areas of the rectangles

[3]

The grouped frequency table below shows the distribution of ages of a sample of people who live in Lisburn.

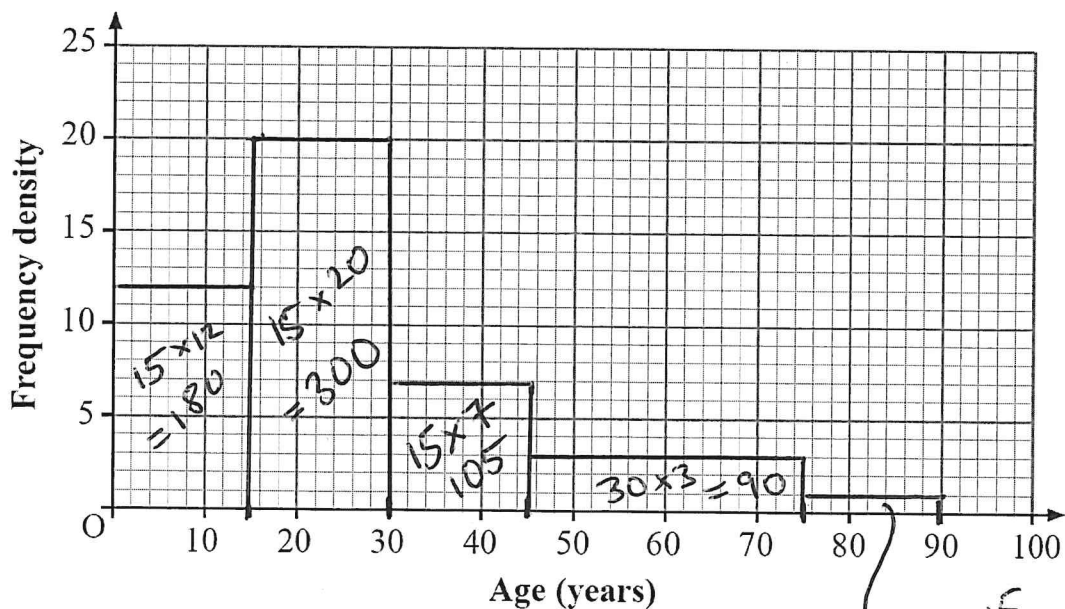
Q5

Age, A , (years)	Frequency	class width	freq. density
$0 \leq A \leq 14$	180	15	12
$15 \leq A \leq 29$	300	15	20
$30 \leq A \leq 44$	105	15	7
$45 \leq A \leq 74$	90	30	3
$75 \leq A \leq 89$	15	15	1
$A \geq 90$	0		

- (a) Explain why the width of the $15 \leq A \leq 29$ class is 15

Age $15 \leq A \leq 29$. This includes everyone up to 29 years 364 days. You don't round to nearest year. You only get into [2] the next group when it is your 30th birthday

- (b) On the grid below, draw a histogram to illustrate the distribution of ages of the sample of people who live in Lisburn.



$15 \times 1 = 15$ [4]