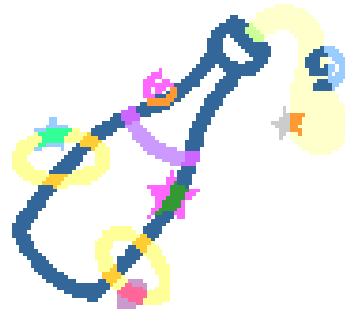# Data Handling – Collecting Data and Averages

Grades F to C

Two-Way Tables

Avoiding Bias

Designing Questionnaires

Types of Data and Sampling

Average and Range

Using Frequency Tables

Hyperlinks!

Stem-and-Leaf Diagrams

Calculations Using Stem-and-Leaf

Grouped Data

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

Two-way tables

Lesson Objective:

Can I read and complete a two-way table?

Grade F

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

What is a two-way table?
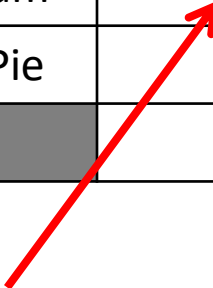
It is a method of showing data in a simple form.

The reason it's called a two-way table is because there are two headings.

Find the value that matches up with both conditions that you are looking for.

What a two-way table looks like

This is a two-way table about meal choices:

		First course			
		Chicken	Beef	Veggie	**Total**
Second course	Fruit Salad	4	2	3	**9**
	Ice Cream	6	1	4	**11**
	Apple Pie	3	0	2	**5**
Total		**13**	**3**	**9**	**25**

These 6 people chose chicken and ice cream.

A total of 9 people had the veggie option.

Answer these questions on the two-way table:

		Favourite Male Celebrity			
		Bruce Forsyth	Graham Norton	Nick Grimshaw	Total
Age Group	0 to 20	1	3	7	11
	21 to 40	3	5	2	10
	Over 40	7	6	1	14
	Total	11	14	10	35

1. How many people aged 21 to 40 chose Graham Norton?
2. How many people were aged Over 40?
3. Who did the youngest people like the most?

Answers: 5, 14, Nick Grimshaw
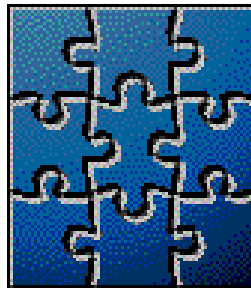
SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

Avoiding bias

Lesson Objective:

Can I spot and correct bias in questions?

Grade E

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

What is bias?

Bias is when you favour one particular answer or outcome.

Don't you think that Manchester United are the best football team?

Yes   Maybe   Definitely

This is a biased question.

Can you re-word these biased questions?

1. Is your favourite colour blue?

2. Do you agree that X Factor is rubbish?

3. What is your opinion of lasagne?

Lovely   Delicious   Beautiful

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

Designing a questionnaire

Lesson Objective:

Can I design an effective questionnaire?

Grade D

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

Remember:

Every question must have a time frame – something like "How many films do you watch *per week*?"

Give people options for each question so that you don't have too many possible answers. One of the options should be "Other" so everyone can answer.

Don't have groups that overlap – by this I mean, so that an answer could go in more than one group.

Comment on each question and improve it where necessary:

How much do you spend on clothes?

£0 to £10   £20 to £30   £30 to £40

I really like shopping in Brighton, don't you?

Yes   No

Which colour is your favourite?

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

Types of data and sampling

Lesson Objective:

Can I recognise different types of data and take a fair sample?

Grade D

?

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to chose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

Data's data, isn't it?

Not really, there are different types of data:

- **Qualitative** – words/names.
- **Quantitative** – numbers.
- **Primary** – data you collect yourself.
- **Secondary** – data someone else collected and you are using.
- **Discrete** – you can count it. eg. number of pets you own, number of red cars.
- **Continuous** – measured on a scale. eg. time, length.

Is the data below discrete or continuous?

1. The times of the songs on a CD.

   Continuous

2. The number of songs on a CD.

   Discrete

3. The number of brothers and sisters I have.

   Discrete

4. How tall my brothers and sisters are.

   Continuous

Samples of data

We use a sample instead of having to ask the whole population.

A sample saves time and money because they often give the same result as asking the whole population.

You must choose your sample fairly though.

How to take a fair sample:

The best/easiest ways are like this:

Put everyone's name in a hat and pick out the number of people you want.

Give the whole population a number.

Use the random number button on your calculator (marked "Ran#").

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

Average and range

Lesson Objective:

Can I calculate average and range?

Grade E

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

There are 3 types of average:

Mean

Add up the data and divide by how many bits there are.

Median

Put the data in order and pick the middle value.

Mode

Most common.

Range is not a type of average

The range is often used alongside average.

It is calculated by subtracting the smallest bit of data from the largest bit.

It is a measure of spread showing how close together the data is or how consistent it is.

A typical question:

Find the mean, median, mode and range of this set of test marks out of 10:

9, 6, 3, 1, 2, 5, 6, 7, 6, 5

Mean: $\dfrac{9+6+3+1+2+5+6+7+6+5}{10} = \dfrac{50}{10} = $ **5**

Median: 1, 2, 3, 5, 5, 6, 6, 6, 7, 9   **5.5**

Mode: **6**

Range: 9 – 1 = **8**

A couple of questions:

Find the mean, median, mode and range of these sets of numbers:

1. 5, 2, 7, 2, 8
Mean: 4.8; Median: 5; Mode: 2; Range: 6

2. 17, 21, 17, 26, 19, 20
Mean: 20; Median: 19.5; Mode: 17; Range: 9

3. 45, 38, 25, 38, 39, 42, 45, 37
Mean: 38.625; Median: 38.5; Mode: Bi-Modal (38 & 45); Range: 20

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

Using frequency tables

Lesson Objective:

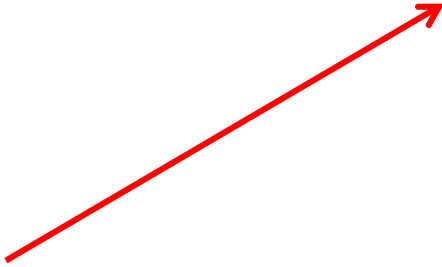*Can I calculate average and range from a frequency table?*

Grade D

SUCCESS CRITERIA: WHERE ARE WE NOW?

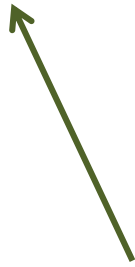| Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

What is a frequency table?

A frequency table shows what data has been collected.

People living in a house	1	2	3	4	5
Frequency	4	6	7	9	4

This means that there are 4 houses with 1 person living in it

This means that there are 7 houses with 3 people living in it

How do we calculate from frequency tables?

People living in a house	1	2	3	4	5
Frequency	4	6	7	9	4

Mean: $\dfrac{1\times4+2\times6+3\times7+4\times9+5\times4}{30} = \dfrac{93}{30} = $ **3.1**

Median: There are 30 bits of data so the value between the 15th and 16th values. That value lies in the "**3**" group.

Mode: The group with the largest frequency is **4**.

Range: The most people in a house is 5, the least is 1, so the range is **4**.

Have a go at this:

Find the mean, median, mode and range from this frequency table.

Number of pets	0	1	2	3	4	5
Frequency	5	8	4	3	2	1

Mean: 1.65
Median: 1
Mode: 1
Range: 5

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

Stem-and-leaf diagrams

Lesson Objective:

Can I draw a stem-and-leaf diagram from a set of data?

Grade D

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

What is a stem-and-leaf diagram?

They are way of organising data so that it's easier to work with.
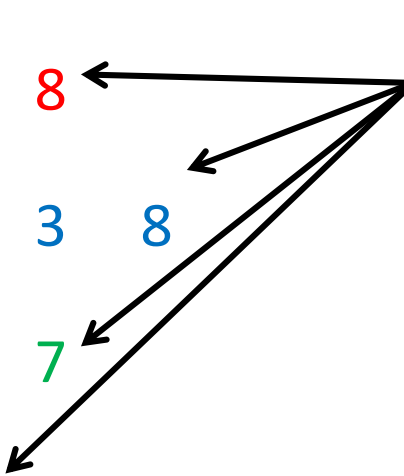
The stem is the larger place values.

The leaves are the smallest place value.

How does it work?

The speed of 15 cars on a residential road:

17, 23, 13, 28, 30, 41, 22, 23,

32, 18, 43, 37, 35, 15, 23

Stem	Leaves
1	3 5 7 8
2	2 3 3 3 8
3	0 2 5 7
4	1 3

The units must be in ascending order.
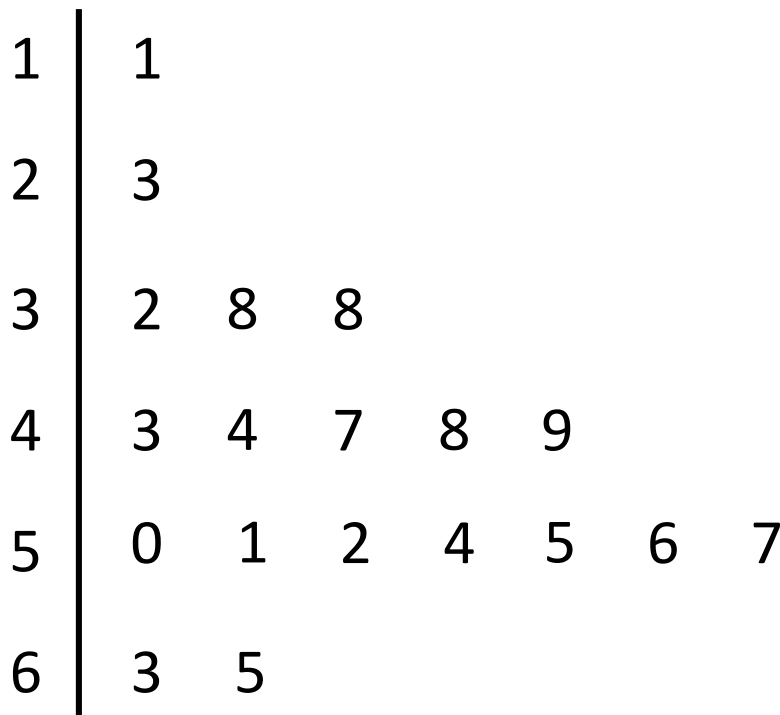
Don't forget the key!

Key:
1|3 means 13

Draw a stem-and-leaf diagram for this data below:

Goals scored by 19 football teams:

52, 48, 23, 47, 57, 38, 44, 63, 51, 50

56, 65, 43, 38, 54, 32, 11, 55, 49

1	1
2	3
3	2 8 8
4	3 4 7 8 9
5	0 1 2 4 5 6 7
6	3 5

Key:
2|3 means 23

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

Calculations using stem-and-leaf diagrams

Lesson Objective:

Can I use a stem-and-leaf diagram to find the median and inter-quartile range of a set of data?

Grade C

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

# What are stem-and-leaf diagrams used for?

They are mainly used for calculating the median and inter-quartile range (IQR).

The interquartile range finds the middle 50% of the data you have.

The point of the inter-quartile range is to eliminate extreme data – very low or very high results.

Calculating the median using the stem-and-leaf:

The speed of 15 cars on a residential road:

```
1 | 3 5 7 8
2 | 2 3 3 (3) 8
3 | 0 2 5 7
4 | 1 3
```
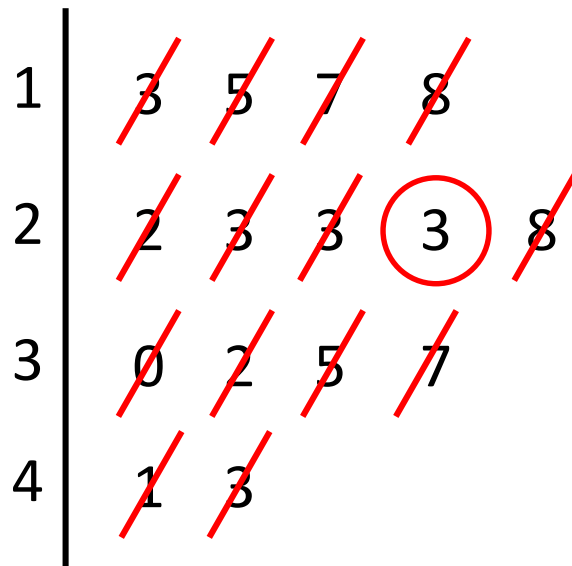
The median is the middle value after the data has been put in order.

The stem-and-leaf already has the data in order, so we just cross off from either end until the one or two values in the middle are left.
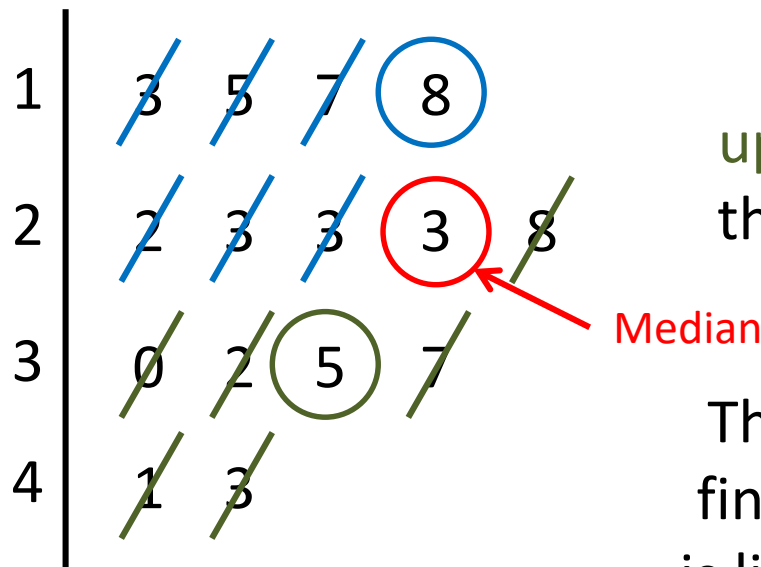
Key:
1|3 means 13

The median is **23**.

Calculating the inter-quartile range using the stem-and-leaf:

The speed of 15 cars on a residential road:

1	3 5 7 (8)
2	2 3 3 (3) 8
3	0 2 (5) 7
4	1 3

Median

Key:
1|3 means 13

There are two quartiles:
upper quartile (three quarters of the way) and lower quartile (one quarter of the way).

The easiest way to find them is to find the median first (the question is likely to have asked you to do this already). Then find what's in the middle of either side of the median.

The inter-quartile range is 35 – 18 = **17**.

You try this one:

Here are 23 test marks out of 50:

43, 22, 31, 27, 18, 13, 42, 31

15, 22, 48, 32, 26, 19, 22, 17

33, 36, 40, 20, 17, 24, 32

Draw a stem-and-leaf diagram for the test marks then find the median and inter-quartile range.

Answers:

```
1 | 3  5  7  7  8  9
2 | 0  2  2  2  4  6  7
3 | 1  1  2  2  3  6
4 | 0  2  3  8
```

Key:
3|1 means 31

Median: **26**

Lower Quartile: **19**

Upper Quartile: **33**

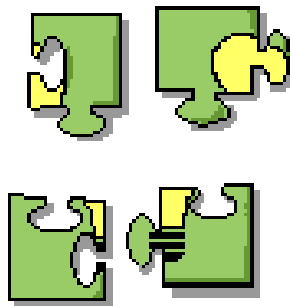Inter-quartile range: **14**

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

Grouped data

Lesson Objective:

Can I calculate an estimate of the mean from grouped data?

Grade C

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			

What is grouped data?

This is where data collected has a wide range, so it's easier to put each bit of data into a group.

From that we can find the "modal group", which is the group with the largest frequency.

Why do we estimate the mean?

We have to do this when data is in groups.

We don't know the exact values of each piece of data, just the group that it is in.

We have to assume that each bit of data is in the <u>middle</u> of each group

A typical question:

This is a table showing how long people take to get to work:

Time (t)	0≤t<10	10≤t<20	20≤t<30	30≤t<40	40≤t<50
Frequency	5	7	12	4	2
Mid-point	5	15	25	35	45

The modal group is 20≤t<30 as it has the largest frequency.

To estimate the mean we need to find the mid-point of each group, then multiply that by the frequency.

$$\frac{5 \times 5 + 7 \times 15 + 12 \times 25 + 4 \times 35 + 2 \times 45}{30} = \frac{660}{30} = \mathbf{\color{red}{22}}$$

Find the modal group and estimate the mean from this data on pocket money:

Money (m)	£0≤m<£5	£5≤m<£10	£10≤m<£15	£15≤m<£20	£20≤m<£25
Frequency	7	6	3	2	2

Modal group: **£0≤m<£5**

Estimate of the mean:

$$\frac{7 \times £2.50 + 6 \times £7.50 + 3 \times £12.50 + 2 \times £17.50 + 2 \times £22.50}{20}$$

Answer: £9

SUCCESS CRITERIA: WHERE ARE WE NOW?

Level	Learning outcomes:	R	A	G
F3	I can draw and use a two-way table.			
E3	I can understand and avoid bias in data.			
D3	I can design an effective questionnaire.			
D1	I can recognise types of data and explain how to choose a sample fairly.			
E2	I can calculate the mean, median, mode and range of a set of data.			
D2	I can calculate averages from a frequency table.			
D1	I can draw a stem-and-leaf diagram.			
C3	I can calculate average and range using a stem-and-leaf diagram.			
C2	I can calculate an estimated mean from grouped frequency.			