

When there is a lot of data, you can organize it into groups in a grouped frequency table.

For Continuous data, you can draw a histogram. It is similar to a bar chart but it doesn't have gaps between the bars.

- Why are there no gaps in continuous data?
- **Only frequency histograms with equal class intervals will be examined.**
- You can use your GDC to draw histograms. *→ Show how to use - see calc directions*

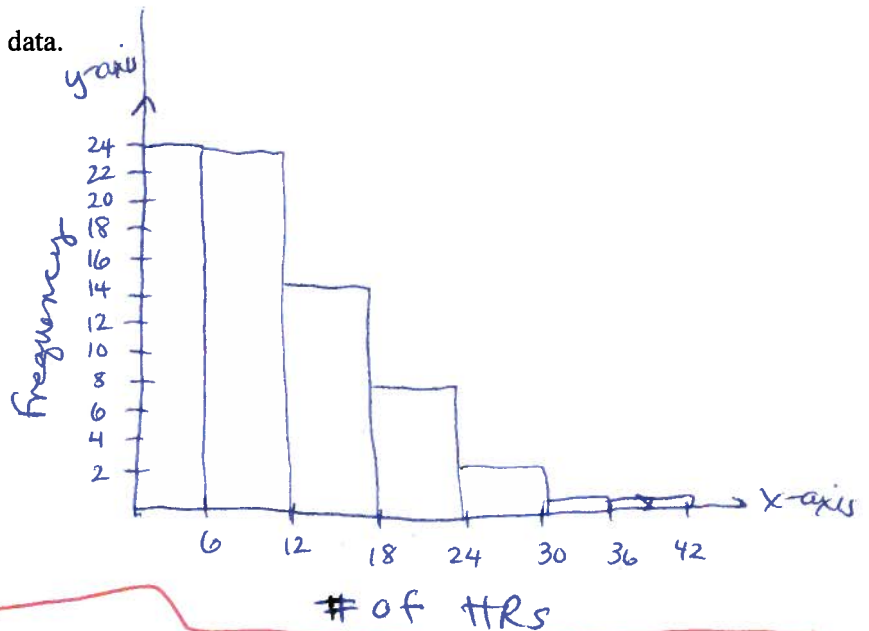
Example:

The homerun totals of 76 players (not including pitchers) (players played in a minimum of 40 games) of teams in the NL Central division in 2013 are:

8, 23, 7, 10, 12, 17, 6, 21, 13, 6, 6, 9, 11, 6, 1, 8, 9, 24, 18, 12, 19, 9, 21, 30, 2, 7, 0, 1, 0, 2, 18, 13, 10, 12, 12, 9, 24, 8, 13, 4, 1, 6, 4, 5, 11, 1, 4, 15, 7, 16, 5, 36, 12, 21, 5, 15, 8, 6, 3, 3, 1, 3, 12, 13, 11, 1, 9, 22, 7, 24, 5, 17, 2, 1, 0, 0

Draw a frequency table and histogram for the data.

<u># of HRs</u>	<u>Frequency</u>
$0 \leq HR < 6$	24
$6 \leq HR < 12$	24
$12 \leq HR < 18$	15
$18 \leq HR < 24$	8
$24 \leq HR < 30$	3
$30 \leq HR < 36$	1
$36 \leq HR < 42$	1



Exercise 8B

8.3 Measures of central tendency

The three most common measures of central tendency are

1. mean
2. median
3. mode

What is the mode?

the value that occurs most frequently in data

- ~~What is another name for the mode?~~
- Can there be more than one mode? If there are two modes, what is the set called?
yes *bimodal*
- When is there no mode?

When no numbers appear more than once

Example.

- a. Find the mode of 3, 4, 6, 6, 7, 7, 7, 8, 8, 9, 10, and 10.
- b. Given the frequency table find the modal class.

7

Weight	Frequency
$110 \leq w < 130$	11
$130 \leq w < 150$	6
$150 \leq w < 170$	12
$170 \leq w < 190$	7
$190 \leq w < 210$	3

modal class = $150 \leq w < 170$

Exercise 8C

THE MEAN!!!

The arithmetic mean is usually called the mean or the average.

The mean is the sum of the numbers divided by the number of numbers in a set of data.

$$\text{Mean} = \frac{\text{Sum of the data values}}{\text{Number of data values}}$$

The mean gives a single number that indicates a center of the data set.

- Usually not a member of the data set
- But a representative value
- The lower case Greek letter μ is the symbol for the population mean.

- o Population mean $\mu = \frac{\sum x}{N}$, where $\sum x$ is the sum of the data values and N is the number of data values in the population.

μ is pronounced 'mu', Σ (which tells us to find the sum here) pronounced 'sigma' and N is 'nu'.
FC ask Mrs. Haley to sing the Greek alphabet song!!!!

There is often a misunderstanding between the population mean and the sample mean. The populations mean uses Greek letters whereas the sample mean uses \bar{x} and n . Our course uses only the population mean.

4 mean examples (they aren't very nice)!!!

- a. Find the mean of 32, 43, 55, 30, 62, and 57. You can also do this on you GDC.

$$\frac{32+43+55+30+62+57}{6} = 46.5$$

- b. Find the mean of the sets below.

AP Score	Frequency
1	15
2	8
3	11
4	4
5	2

Weight	Frequency
$110 \leq w < 130$	11
$130 \leq w < 150$	6
$150 \leq w < 170$	12
$170 \leq w < 190$	7
$190 \leq w < 210$	3

midpt	Freq * midpt
120	1320
140	840
160	1920
180	1260
200	600

This is the formula as it appears in the IB Formula booklet:

$$\mu = \frac{\sum_{i=1}^n f_i x_i}{\sum_{i=1}^n f_i}$$

When the data is grouped, we can calculate the mean by assuming that all of the data values are equally spread around the midpoint.

WARNING: This method leads to small inaccuracies and that is why exam questions often say 'estimate the mean'. It does not mean 'guess' – it means work out, as in this example or with your GDC.

$$\mu = \frac{1(15) + 2(8) + 3(11) + 4(4) + 5(2)}{15 + 8 + 11 + 4 + 2}$$

$$= \frac{90}{40} = \frac{9}{4} = 2.25$$

$$\frac{\sum f \cdot m}{\sum f} = \frac{1320 + 840 + 1920 + 1260 + 600}{11 + 6 + 12 + 7 + 3}$$

$$= \frac{5940}{39} = 152.31$$

Last one: Kunal really likes to do well on tests. He has scores of 95, 89, 93, and 84. What score must Kunal need to get on the fifth test in order to get an A⁺ score of 93% (according to NAFC)?

$$93 = \frac{95 + 89 + 93 + 84 + x}{5}$$

$$465 = x + 361$$

$$x = 104\% \quad (\text{possible with extra credit?})$$

Exercise 8D

The median.

The median is the number in the middle when the numbers in a set of data are arranged in order of size. If the number of numbers in a data set is even, then the median is the mean of the two middle numbers.

Find the median of the 13, 11, 32, 18, 19, 20, 13, 15, 25, 29, 28, and 20.

11, 13, 13, 15, 18, 19, 20, 20, 25, 28, 29, 32

$$\text{med} = \frac{20 + 19}{2} = 19.5$$

If there are a lot of numbers and it is difficult to find the middle member we can use the formula

Median = $\left(\frac{n+1}{2}\right)$ th member, where n is the number of members in the set.

*Common error. This formula does not give the median. It gives the position of the median in the data set.

Exercise 8E

Summary of measures of central tendency

	Advantages	Disadvantages
<p>Mode The mode can be used for qualitative data or when asked to choose the most popular item.</p>	<ul style="list-style-type: none"> • Extreme values do not affect the mode. 	<ul style="list-style-type: none"> • Does not use all members of the data set. • Not necessarily unique – may be more than one answer. • When no values repeat in the data set, there is no mode. • When there is more than one mode, it is difficult to interpret and/or compare.
<p>Mean The mean describes the middle of a set of data.</p>	<ul style="list-style-type: none"> • Most popular measure in fields such as business, engineering and computer science. • Uses all members of the data set. • It is unique – there is only one answer. • Useful when comparing sets of data. 	<ul style="list-style-type: none"> • Affected by extreme values. <div style="border: 1px solid gray; padding: 5px; margin-top: 10px;"> <p>In the data set of salaries €15 000, €20 000, €22 000, €17 000, €75 000 how does the extreme value of €75 000 affect the mean?</p> </div>
<p>Median The median describes the middle of a set of data.</p>	<ul style="list-style-type: none"> • Extreme values do not affect the median as strongly as they do the mean. • Useful when comparing sets of data. • It is unique – there is only one answer. • As the median is the middle value, 50% of the data is either side of it. 	<ul style="list-style-type: none"> • Not as popular as mean. • Less used in further calculations.

Investigation
p8-267