

## Topic: Summarising Data

Topic/Skill	Definition/Tips	Example																				
1. Types of Data	<p><b>Qualitative Data</b> – non-numerical data</p> <p><b>Quantitative Data</b> – numerical data</p> <p><b>Continuous Data</b> – data that can take <b>any numerical value</b> within a given range.</p> <p><b>Discrete Data</b> – data that can take <b>only specific values</b> within a given range.</p>	<p>Qualitative Data – eye colour, gender etc.</p> <p>Continuous Data – weight, voltage etc.</p> <p>Discrete Data – number of children, shoe size etc.</p>																				
2. Grouped Data	<p>Data that has been <b>bundled in to categories</b>.</p> <p>Seen in grouped frequency tables, histograms, cumulative frequency etc.</p>	<table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>Foot length, <math>l</math>, (cm)</th> <th>Number of children</th> </tr> </thead> <tbody> <tr> <td><math>10 \leq l &lt; 12</math></td> <td>5</td> </tr> <tr> <td><math>12 \leq l &lt; 17</math></td> <td>53</td> </tr> </tbody> </table>	Foot length, $l$ , (cm)	Number of children	$10 \leq l < 12$	5	$12 \leq l < 17$	53														
Foot length, $l$ , (cm)	Number of children																					
$10 \leq l < 12$	5																					
$12 \leq l < 17$	53																					
3. Primary /Secondary Data	<p><b>Primary Data</b> – <b>collected yourself</b> for a specific purpose.</p> <p><b>Secondary Data</b> – <b>collected by someone else</b> for another purpose.</p>	<p>Primary Data – data collected by a student for their own research project.</p> <p>Secondary Data – Census data used to analyse link between education and earnings.</p>																				
4. Mean	<p><b>Add</b> up the values and <b>divide</b> by how many values there are.</p>	<p>The mean of 3, 4, 7, 6, 0, 4, 6 is</p> $\frac{3 + 4 + 7 + 6 + 0 + 4 + 6}{7} = 5$																				
5. Mean from a Table	<ol style="list-style-type: none"> <li>Find the midpoints (if necessary)</li> <li>Multiply Frequency by values or midpoints</li> <li>Add up these values</li> <li>Divide this total by the Total Frequency</li> </ol> <p>If <b>grouped</b> data is used, the answer will be an <b>estimate</b>.</p>	<table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>Height in cm</th> <th>Frequency</th> <th>Midpoint</th> <th>F × M</th> </tr> </thead> <tbody> <tr> <td><math>0 &lt; h \leq 10</math></td> <td>8</td> <td>5</td> <td><math>8 \times 5 = 40</math></td> </tr> <tr> <td><math>10 &lt; h \leq 30</math></td> <td>10</td> <td>20</td> <td><math>10 \times 20 = 200</math></td> </tr> <tr> <td><math>30 &lt; h \leq 40</math></td> <td>6</td> <td>35</td> <td><math>6 \times 35 = 210</math></td> </tr> <tr> <td>Total</td> <td>24</td> <td>Ignore!</td> <td>450</td> </tr> </tbody> </table> <p><b>Estimated Mean</b> height: <math>450 \div 24 = 18.75\text{cm}</math></p>	Height in cm	Frequency	Midpoint	F × M	$0 < h \leq 10$	8	5	$8 \times 5 = 40$	$10 < h \leq 30$	10	20	$10 \times 20 = 200$	$30 < h \leq 40$	6	35	$6 \times 35 = 210$	Total	24	Ignore!	450
Height in cm	Frequency	Midpoint	F × M																			
$0 < h \leq 10$	8	5	$8 \times 5 = 40$																			
$10 < h \leq 30$	10	20	$10 \times 20 = 200$																			
$30 < h \leq 40$	6	35	$6 \times 35 = 210$																			
Total	24	Ignore!	450																			
6. Median Value	<p>The <b>middle</b> value.</p> <p>Put the data in order and find the middle one.</p> <p>If there are <b>two middle values</b>, find the number half way between them by <b>adding them together and dividing by 2</b>.</p>	<p>Find the median of: 4, 5, 2, 3, 6, 7, 6</p> <p>Ordered: 2, 3, 4, <b>5</b>, 6, 6, 7</p> <p>Median = 5</p>																				
7. Median from a Table	<p>Use the formula <math>\frac{(n+1)}{2}</math> to find the position of the median.</p> <p><math>n</math> is the total frequency.</p>	<p>If the total frequency is 15, the median will be the <math>\left(\frac{15+1}{2}\right) = 8\text{th}</math> position</p>																				
8. Mode /Modal Value	<p><b>Most</b> frequent/common.</p> <p>Can have more than one mode (called bi-modal or multi-modal) or no mode (if all values appear once)</p>	<p>Find the mode: 4, 5, 2, 3, 6, 4, 7, 8, 4</p> <p>Mode = 4</p>																				
9. Range	<p><b>Highest value subtract the Smallest value</b></p>	<p>Find the range: 3, 31, 26, 102, 37, 97.</p> <p>Range = <math>102 - 3 = 99</math></p>																				

	Range is a 'measure of spread'. The smaller the range the more <u>consistent</u> the data.	
10. Outlier	A value that ' <b>lies outside</b> ' most of the other values in a set of data. An outlier is <b>much smaller or much larger</b> than the other values in a set of data.	
11. Lower Quartile	<b>Divides the bottom half</b> of the data into <b>two halves</b> .  $LQ = Q_1 = \frac{(n+1)}{4} \text{th value}$	Find the lower quartile of: 2, <u>3</u> , 4, 5, 6, 6, 7  $Q_1 = \frac{(7+1)}{4} = 2nd \text{ value} \rightarrow 3$
12. Lower Quartile	<b>Divides the top half</b> of the data into <b>two halves</b> .  $UQ = Q_3 = \frac{3(n+1)}{4} \text{th value}$	Find the upper quartile of: 2, 3, 4, 5, 6, <u>6</u> , 7  $Q_3 = \frac{3(7+1)}{4} = 6th \text{ value} \rightarrow 6$
13. Interquartile Range	The <b>difference</b> between the <b>upper quartile and lower quartile</b> .  $IQR = Q_3 - Q_1$  The <b>smaller</b> the <b>interquartile range</b> , the <b>more consistent</b> the data.	Find the IQR of: 2, 3, 4, 5, 6, 6, 7  $IQR = Q_3 - Q_1 = 6 - 3 = 3$