

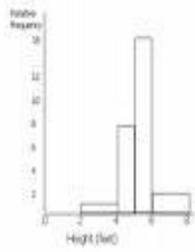
<p>When working with statistical diagrams, make sure you look at the scale carefully</p>	<p>Range = largest data value – smallest value</p>
<p>Grouping data is good because it makes the data easier to calculate with and interpret.</p> <p>However, by grouping we lose the original data so our calculations are estimates.</p>	<p>Mode = most common</p> <p>Modal class = group with highest frequency</p>
<p>Median = middle data value when the data is in numerical order</p>	<p>Mean = sum of all data divided by how many pieces of data there are</p>
<p>Interquartile range = upper quartile – lower quartile</p>	<p>The interquartile range is the range of the middle 50% of the data. It is the length of a box on a box plot</p> <p>Interquartile range calculations tend not to include anomalies; for this reason interquartile range is more accurate than range.</p>

<p>When comparing distributions, refer to</p> <p>An average – such as median or mean</p> <p>e.g. if boys have higher median exam marks than girls they did better on average</p>	<p>When comparing distributions, refer to</p> <p>The spread of the data; the range or interquartile range</p> <p>Large range or interquartile range implies less consistency</p>
--	--

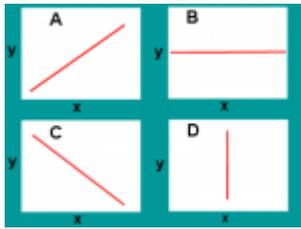
<p>Stem and leaf diagram</p> <p>Leaves must be in order</p> <p>There must be a key</p>	<p>Mean from a frequency table DIVIDE BY TOTAL FREQUENCY</p> <table border="1"> <thead> <tr> <th>SCORE</th> <th>FREQUENCY</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>3</td> </tr> <tr> <td>1</td> <td>2</td> </tr> <tr> <td>2</td> <td>3</td> </tr> <tr> <td>3</td> <td>2</td> </tr> </tbody> </table> <p>Mean score = $((0 \times 3) + (1 \times 2) + (2 \times 3) + (3 \times 2))/10 = 14/10 = 1.4$</p>	SCORE	FREQUENCY	0	3	1	2	2	3	3	2
SCORE	FREQUENCY										
0	3										
1	2										
2	3										
3	2										

<p>To work out a 3 – point moving average, work out the average of the first three points, move along one and then work out the average of the next three e.g.</p> <table border="1"> <tr> <td>5</td><td>7</td><td>6</td><td>2</td><td>4</td><td>6</td><td>8</td> </tr> <tr> <td></td><td>6</td><td>5</td><td>4</td><td>4</td><td>6</td><td></td> </tr> </table> <p>The moving average will be written in the middle of the points it has used as above.</p>	5	7	6	2	4	6	8		6	5	4	4	6		<p>On a moving averages graph, the trend line should be a line of best fit of the moving averages.</p>
5	7	6	2	4	6	8									
	6	5	4	4	6										

<p>Mean from a grouped frequency table USE MIDPOINTS</p> <p>Total the midpoint x frequency ; divide by total frequency</p> <table border="1"> <thead> <tr> <th>GROUP</th> <th>FREQUENCY</th> <th>MIDPOINT</th> <th>MIDPOINT x FREQUENCY</th> </tr> </thead> <tbody> <tr> <td>0-2</td> <td>5</td> <td>1</td> <td>5</td> </tr> <tr> <td>3-5</td> <td>4</td> <td>4</td> <td>16</td> </tr> <tr> <td>6-10</td> <td>3</td> <td>8</td> <td>24</td> </tr> <tr> <td>11-15</td> <td>3</td> <td>13</td> <td>39</td> </tr> <tr> <td>TOTAL</td> <td>15</td> <td></td> <td>84</td> </tr> </tbody> </table> <p>MEAN = 84/15 = 5.6</p>	GROUP	FREQUENCY	MIDPOINT	MIDPOINT x FREQUENCY	0-2	5	1	5	3-5	4	4	16	6-10	3	8	24	11-15	3	13	39	TOTAL	15		84	<p>RELATIVE FREQUENCY is a probability found from experiment.</p> <p>If the relative frequency of a 5 on a biased dice is 0.2 then after</p> <p>10 throws you would expect $10 \times 0.2 = 2$ fives</p> <p>50 throws you would expect $50 \times 0.2 = 10$ fives</p> <p>100 throws you would expect $100 \times 0.2 = 20$ fives</p>
GROUP	FREQUENCY	MIDPOINT	MIDPOINT x FREQUENCY																						
0-2	5	1	5																						
3-5	4	4	16																						
6-10	3	8	24																						
11-15	3	13	39																						
TOTAL	15		84																						

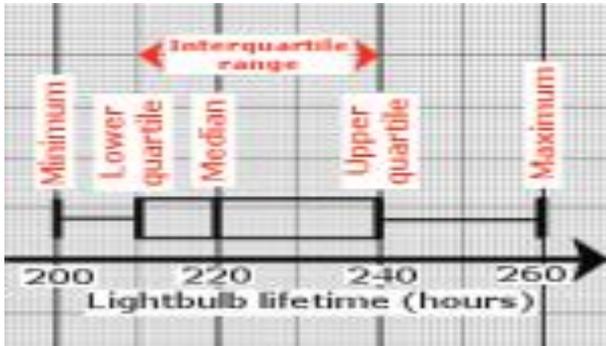
<p style="text-align: center;">GOOD SURVEY</p>	<p style="text-align: center;">BAD SURVEY</p>
<ol style="list-style-type: none"> 1. Keep it simple 2. Use tick boxes 3. Make sure responses cover all possibilities 	<ol style="list-style-type: none"> 1. Do not ask leading questions; Don't influence people's decisions 2. Do not ask personal questions 3. Do not include any overlapping responses
<p>To find median on a cumulative frequency, draw a line at halfway up the cumulative frequency and see where it meets the curve.</p> <p>Lower quartile- same but $\frac{1}{4}$ of way up</p> <p>Upper quartile- same but $\frac{3}{4}$ of way up</p>	<p>On a histogram plot frequency density on y axis where</p> $\text{Frequency density} = \frac{\text{Frequency}}{\text{Class width}}$
<p>The y axis will always have something relating to frequency on it ie</p> <p>Frequency</p> <p>Cumulative frequency</p> <p>Frequency density</p>	<p>Cumulative frequency means running total.</p>
<p>In a histogram, area under bars = frequency.</p>  <p>Total area = total frequency</p> <p>Work out areas by doing width x height.</p>	<p>A line of best fit on a scatter diagram must have about the same number of points above and below the line.</p> <p>It must be a straight line, but it doesn't have to go through the origin.</p>

- A POSITIVE CORRELATION
- B NO CORRELATION
- C NEGATIVE CORRELATION
- D NO CORRELATION



GROUP	FREQUENCY
$0 \leq x < 2$	9
$2 \leq x < 4$	3
$4 \leq x < 6$	4
$6 \leq x < 8$	4
TOTAL	20

The modal class interval is the one with highest frequency: $0 \leq x < 2$



Lower quartile = 210 = value $\frac{1}{4}$ of way into data.
 Upper quartile = 240 = value $\frac{3}{4}$ of way into data.
 Median = $\frac{1}{2}$ way into data.
 Interquartile range = upper quartile - lower quartile = 30

In a stem and leaf diagram you must put LEAVES in order and give a key.

In this example 1|9 means 19.

The **mode** is 48,

the **range** is highest - lowest = $62 - 19 = 43$

the **median** (middle value) is halfway between 42 and 43 = 42.5.

To find median cross off values from start and finish until you reach the middle.

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

Random sampling

All members of the population must have the same chance of being chosen

This can happen by picking names from a hat or by using a list of random numbers.

Stratified sampling is when the population is divided into categories and a number of each category is surveyed in the **same proportion** as the population

The sample is chosen **randomly**

REMEMBER WHOLE NUMBERS

A sample should consider things like the different genders, ages and cultures appropriately.

On a scatter diagram, always draw a line of best fit!

<p>When plotting a frequency polygon use MID-POINTS</p>	<p>When plotting a cumulative frequency curve use END-POINTS and join with a SMOOTH CURVE.</p>
<p>Systematic sampling is sampling where every 10th or 20th item is surveyed after population displayed in a list with some given order</p>	<p>To be sure that a dice is biased, you must roll the dice enough to be sure that one number is rolled more often than the others.</p>
<p>Using moving averages to find a predicted value</p> <ol style="list-style-type: none"> 1. Work out the next supposed moving average using a line of best fit 2. See what the next value would have to be to achieve this moving average! 	<p>To write a suitable question, think</p> <p>Are all responses covered? Do you need to give a time frame? Make sure responses don't overlap</p>
<p>Label all axes on graphs</p>	<p>Use a ruler to draw straight lines</p>

<p>Pie charts</p> <p>To find an angle, divide by total frequency and multiply by 360 If easy frequency e.g. 90, multiply by 4 to get angle 60 multiply by 6 to get angle</p>	<p>When asked to comment on a trend in moving averages, try to comment on a general trend ie general increase or decrease</p>																																																	
<p>A sample space diagram shows all possible outcomes, e.g. adding the scores of 2 dice</p> <table border="1" data-bbox="68 535 774 804"> <tr> <td>+</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> <td>6</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> <td>6</td> <td>7</td> </tr> <tr> <td>2</td> <td>3</td> <td>4</td> <td>5</td> <td>6</td> <td>7</td> <td>8</td> </tr> <tr> <td>3</td> <td>4</td> <td>5</td> <td>6</td> <td>7</td> <td>8</td> <td>9</td> </tr> <tr> <td>4</td> <td>5</td> <td>6</td> <td>7</td> <td>8</td> <td>9</td> <td>10</td> </tr> <tr> <td>5</td> <td>6</td> <td>7</td> <td>8</td> <td>9</td> <td>10</td> <td>11</td> </tr> <tr> <td>6</td> <td>7</td> <td>8</td> <td>9</td> <td>10</td> <td>11</td> <td>12</td> </tr> </table> <p>Use this to find probs e.g. $p(5) = 3/36 = 1/12$</p>	+	1	2	3	4	5	6	1	2	3	4	5	6	7	2	3	4	5	6	7	8	3	4	5	6	7	8	9	4	5	6	7	8	9	10	5	6	7	8	9	10	11	6	7	8	9	10	11	12	<p>If events are independent they don't affect each other</p> <p>To find probability of both happening, multiply their probabilities together!!</p>
+	1	2	3	4	5	6																																												
1	2	3	4	5	6	7																																												
2	3	4	5	6	7	8																																												
3	4	5	6	7	8	9																																												
4	5	6	7	8	9	10																																												
5	6	7	8	9	10	11																																												
6	7	8	9	10	11	12																																												
<p>In probability, if a question asks for who has the most accurate results then it is always the person who has performed more trials</p>	<p>AT LEAST or NONE implies the "1-" rule. For example at least one rainy day means everything but no rainy days so we can do</p> <p>$P(\text{at least one rainy day}) = 1 - P(\text{no rainy days})$</p>																																																	
<p>In tree diagrams, Multiply along, add between</p> <p>In tree diagrams each branch adds up to 1</p> <p>In tree diagrams final probabilities (the ones you have multiplied to get) add up to 1</p>	<p>Mutually exclusive events are events which can't happen together</p> <p>EXAMPLE: You can't get a 1 on a dice and an even number at the same time!</p> <p>$P(1 \text{ and even}) = 0$</p>																																																	

The probabilities of all possible mutually exclusive outcomes of an event add to 1.

Given a group $0 \leq x < 2$

0 is contained within the group as X is greater than or equal to zero
2 is NOT contained as it is less than zero

To find median using a histogram, work out AREA, as this will give you the frequency.

e.g. area = 60, total frequency = 60 so median lies at 30th value.

See where this 30th value lies.

Upper quartile = 45th value, lower quartile = 15th value

When working out probabilities for tree diagrams, be careful to check whether the probabilities change or not!!

For example, if you are talking about sweets, you will not return the sweet to the bag!!

2 WAY TABLES

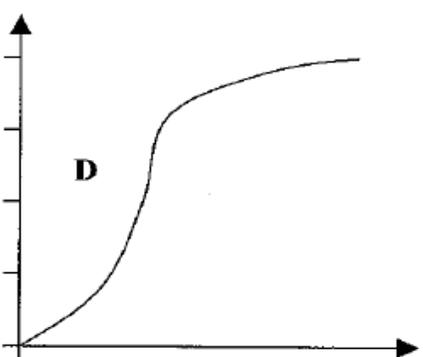
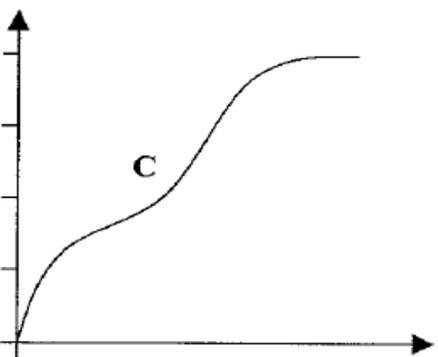
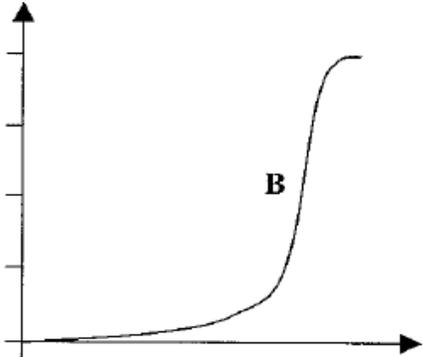
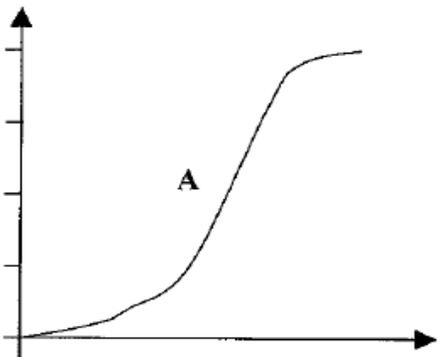
One variable shown by the rows of the table

One variable shown by the columns

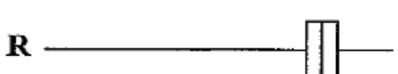
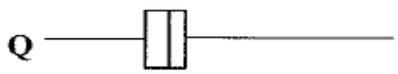
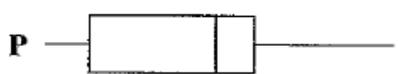
	First class	Business class	Economy
American	6	8	51
British	3	5	73
French	0	4	34
German	1	3	12

If the question asks you for a probability of **both** or **one of** or **at least one of** two things happening, you must draw a TREE DIAGRAM

20. Here are four cumulative frequency diagrams.



Here are four box plots.



For each box plot, write down the letter of the appropriate cumulative frequency diagram.