Year 7 Knowledge Organiser – Statistics

Objectives

Interpret and construct tables, charts and diagrams, and know their appropriate uses

Interpret, analyse and compare the distributions of data sets through median, mean, mode and range.

Key Vocabulary

Average - a number expressing the central or typical value in a set of data

Spread - the measure of how far the numbers in a data set are away from the mean or the median

Data - facts and statistics collected for reference or analysis

Approximate - an estimation of a number or rounding a number to its nearest place value

Discrete data - data that can only take certain values e.g. shoe size

Continuous data - Data that can take any value (within a range)

Distribution - shows the possible values for a variable and how often they occur

Frequency - the number of times the data value occurs

Averages and Range...

Mean – Find the total value of the data and divide this by the number of pieces of data in the set. Round accordingly.

Mode – The piece of data that appears the most frequently

Median – Order your data and find the value in the middle. If there are two values in the middle, find the mean of these.

Range – The difference between the lowest and highest pieces of data.

> "Hey diddle diddle,
> the median's the middle!
> You add and divide for the mean.
> The mode is the one that appears the most,
> and the range is the difference in-between!"

Ungrouped Data

The number of times an event happened

The table shows the number of siblings students have. The answers were

3, 1, 2, 2, 0, 3, 4, 1, 1, 2, 0, 2

2 people had 0 siblings. This means there are 0 siblings to be counted here

Number of siblings	Frequency
0	2	→ 0
1	3	→ 3
2	4	2 + 2 + 2 + 2 OR 2 x 4 = 8
3	2	3 + 3 OR 3 x 2 = 6
4	1	4

2 people have 3 siblings so there are 6 siblings in total

Best represented by discrete data (Not always a number)

OVERALL there are
0 + 3 + 8 + 6 + 4
Siblings = 21 siblings

Grouped Data

If we have a large spread of data it is better to group it. This is so it is easier to look for a trend. Form groups of equal size to make comparison more valid and spread the groups out from the smallest to the largest value.

Discrete Data The groups do not overlap

Cost of TV (£)	Tally	Frequency
101 - 150	THL II	7
151 - 200	THL THL I	11
201 - 250	THL	5
251 - 300	III	3

We do not know the exact value of each item in a group – so an estimate would be used to calculate the overall total (Midpoint)

Continuous Data To make sure all values are included inequalities represent the subgroups

x Weight(g)	Frequency
$40 < x \le 50$	1
$50 < x \le 60$	3
$60 < x \le 70$	5

eg this group includes every weight bigger that 60kg, up to and including 70kg

Pictograms, bar and line charts

Represents quantitative data

Pictogram

Language	
French	
Spanish	
German	

= 4 people

- Need to remember a key
- Visually able to identify mode

Bar Chart

How Y8 travel to school

- Gaps between the bars
- Clearly labelled axes
- Scale for the axes
- Title for the bar chart
- Discrete Data

Line Chart

- Gaps between the lines
- Clearly labelled axes
- Scale for the axes
- Discrete Data

Multiple Bar chart

Compares multiple groups of data

- Clearly labelled axes
- Scale for axes
- Comparable data bars drawn next to each other

Average temperatures — Key: London, Moscow, Seoul

Key/ Colour code for separate groups of information

Gap between different categories of data

Draw and interpret Pie Charts

Remember a circle has 360°

Type of pet	Dog	Cat	Hamster
Frequency	32	25	3

There were 60 people asked in this survey (Total frequency)

$\frac{32}{60}$ "32 out of 60 people had a dog"

This fraction of the 360 degrees represents dogs

$\frac{32}{60} \times 360 = 192°$

Multiple method
As 60 goes into 360 – 6 times
Each frequency can be multiplied by 6 to find the degrees (proportion of 360)

Use a protractor to draw
This is 192°

Represents quantitative, discrete data

Draw and interpret line graphs

- Commonly used to show changing over time
- The points are the recorded information and the lines join the points

Line graphs do not need to start from 0

More than one piece of data can be plotted on the same graph to compare data

Playground Temperature

It is possible to make estimates from the line e.g. temperature at 9.30am is 5°C