



Unit 6 – Collecting and Representing Data part 1

What do I need to be able to do?

- By the end of this unit you should be able to:
- Construct and interpret frequency tables and polygon. two-way tables, line, bar, & pie charts
 - Find and interpret averages from a list and a table
 - Construct and interpret time series graphs, stem and leaf diagrams and scatter graphs

Keywords

- Population:** the whole group that is being studied
Sample: a selection taken from the population that will let you find out information about the larger group
Representative: a sample group that accurately represents the population
Random sample: a group completely chosen by chance. No predictability to who it will include.
Bias: a built-in error that makes all values wrong by a certain amount
Primary data: data collected from an original source for a purpose.
Secondary data: data taken from an external location. Not collected directly.
Outlier: a value that stands apart from the data set

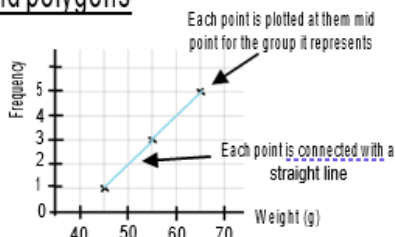
Frequency tables and polygons

x Weight(g)	Frequency
40 < x ≤ 50	1
50 < x ≤ 60	3
60 < x ≤ 70	5

We do not know from grouped data where each value is placed so have to use an estimate for calculations

MID POINTS

Mid-points are used as estimated values for grouped data. The middle of each group



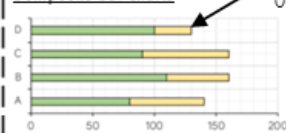
The data about weight starts at 40. So the axis can start at 40

Mid-point
Startpoint + Endpoint
2

Bar and line charts

Compare the bars green compared to yellow. The size of each bar is the frequency. Overall total easily comparable

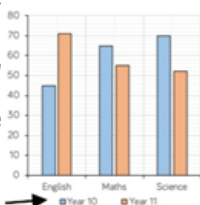
Composite bar charts



Dual bar charts

Bars are compared side by side
Easier to compare subgroups

Categories clearly indicated



Two way tables

60 people visited the zoo one Saturday morning. 26 of them were adults. 13 of the adult's favourite animal was an elephant. 24 of the children's favourite animal was an elephant.

Extract information to input to the two-way table.

	Adult	Child	Total
Elephant	13	24	37
Other	13	10	23
Total	26	34	60

Subgroups each have their own heading

Needs subgroup totals

Overall total

Draw and interpret Pie Charts

Type of pet	Dog	Cat	Hamster
Frequency	32	25	3

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

32
60
"32 out of 60 people had a dog"

This fraction of the 360 degrees represents dogs
 $\frac{32}{60} \times 360 = 192^\circ$



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°

Comparing Pie Charts:
You NEED the overall frequency to make any comparisons

Averages from lists

The Mean

A measure of average to find the central tendency... a typical value that represents the data

24, 8, 4, 11, 8,

Find the sum of the data (add the values)

55

Divide the overall total by how many pieces of data you have

$55 \div 5$

Mean = 11

The Mode (The modal value)

This is the number OR the item that occurs the most (it does not have to be numerical)

24, 8, 4, 11, 8,

This can still be easier if the data is ordered first

Mode = 8

The Median

The value in the center (in the middle) of the data

Put the data in order 4, 8, 8, 11, 24

Find the value in the middle 4, 8, 8, 11, 24

Median = 8

NOTE: If there is no single middle value find the mean of the two numbers left

For Grouped Data

The modal group – which group has the highest frequency.

Averages from a table

Non-grouped data

Number of Siblings	0	1	2
Frequency	6	8	6
Subtotal	0	8	12

Overall Frequency: 20

Total number of siblings: 20

The data in a list: 0,0,0,0,0,0,1,1,1,1,1,1,1,1,2,2,2,2,2,2

Mean: $\frac{\text{total number of siblings}}{\text{Total frequency}} = 1$

Grouped data

x Weight(g)	Frequency	Mid Point	MP x Freq
40 < x ≤ 50	1	45	45
50 < x ≤ 60	3	65	195
60 < x ≤ 70	5	65	325

Overall Frequency: 9

Overall Total: 565

Mean: 62.8g

The data in a list: 45, 55, 55, 55, 65, 65, 65, 65, 65, 65