3.1 Collecting Data

- The first stage is always to decide on a *hypothesis*. Everything from then on is geared to testing the hypothesis. We probably won't "prove" or "disprove" it completely, but we'll get strong/weak evidence in favour of or against it. A hypothesis is a *statement*, not a question or a "title".
- Making a good plan before you start saves lots of time and wasted effort. Key questions to ask pupils are "How exactly are you going to do that?"; "What problems might you run into?"; "What might go wrong with that?"; "Why might that not be easy?"; "How long do you expect that part to take?", etc.
- An interesting hypothesis can make all the difference to pupils' motivation. Topical or local ideas work best so long as they're not so personal that they cause embarrassment. Ethics is obviously an issue, and relates to how the data is collected and to how it's used; anonymity, confidentiality, etc.
- Questionnaires can be fun to write, but if every pupil wants every other pupil to fill in a questionnaire then the photocopying costs will be astronomical! One workable strategy is to invent a questionnaire on the board as a class exercise containing all the questions people want answers to for testing their hypotheses. Everyone copies it down as we go (because the data will make no sense unless we remember the questions). Before the next lesson the teacher prepares a results table (A3, perhaps) with space for every pupil's answers, one pupil per line, and during the next lesson this is passed around the classroom while the pupils are doing some other task. Every pupil fills in his/her answers and then this sheet is photocopied (perhaps reduced to A4) and given to pupils, maybe one between two, for processing the data.

3.1.1 **NEED** "Skateboarding Questionnaire" sheet.

What makes for a good/bad questionnaire?

Choose a hypothesis and think of good and bad questions for a questionnaire.

Bad questions are
- *unclear;*
- *irrelevant;*
- *biased;*
- *embarrassing; or*
- *give results which are hard to interpret / process.*

Tips for good questions include these:
- *use tick boxes;*
- *specify ranges of values (get the "filler-in" to do the grouping of data so that we don't have to);*
- *remember to include "other", where applicable;*
- *if amounts are requested, state the units (e.g., years, cm, £, etc.).*

A common error is to include the same value at the start of one tick-box range as is at the end of the previous one; e.g., "0-2", "2-4", etc. (Here it would not be clear which box to tick for "2".)

You can collect badly produced questionnaires written by "important" people/organisations, such as the deputy head, exam boards, or government departments!

(You might need to conceal their origins to protect the guilty!)

3.1.2 Pupils write a questionnaire to test their own hypotheses.
Popular topics include sport, food (especially sweets), smoking, television, music, shops, leisure facilities, often comparing boys and girls.

This works best if the whole class discuss what makes a good hypothesis and what makes a good questionnaire before pupils work independently on this.

You may need to be cautious about reinforcing stereotypes: a hypothesis may be a "prejudice".

You will need to be alert to hypotheses that can be tested by asking one simple question; e.g., "Most pupils think that school uniform should be abolished" could lead to a rather short "questionnaire", so the pupil could develop an additional hypothesis or make this one a bit more nuanced.

3.1.3 **NEED** photocopies of a page of text from two different books (aimed at different age pupils):

You can begin by discussing what might be a suitable hypothesis.

novels may be best.

This book is aimed at year 6 (say) and this one at year 8 (say). How would you expect them to be different?
Which of these ideas could we test using statistics? Can you suggest a suitable hypothesis?

The book aimed at older people might be expected to have a more complex storyline, more mature themes, longer words, harder words, smaller print, more pages, be more expensive, etc.

e.g., "Books aimed at older pupils will have, on average, longer words."

Write the ideas on the board. It may be best to agree on one hypothesis to pursue, or different pupils or groups of pupils could work on different hypotheses.

Exactly what would we have to do to test this hypothesis?

Pupils need to work out a clear step-by-step plan before they start, dealing with questions such as "how many words will you include in your sample?", "how will you choose them?", "what will you do about hyphenated words?", etc.

100 words from each text may be a sensible sample size.

3.1.4 Methods of Travel to School.
This well-worn idea can still be of interest provided that there is a hypothesis and some clear purpose to the survey.
e.g., "Is it true that people nowadays are walking less and driving more?";
"A politician claims that nationally most children still travel the main part of their journey to school on foot. Is that true for this class?"

Why is walking less popular than it once was? safety concerns, too busy, live further from school than people used to, etc.

What are the advantages of walking?
healthy (but is it if the air is full of vehicle fumes?), sociable, no traffic/parking problems, cheap, independent, etc.

What should we do about people who use a variety of means to get to school or come different ways on different days?

There are several possible approaches; e.g., "What was the main method you used to get to school this morning?" (Remember to include "other" in case someone rode in by horse or landed by helicopter!)

A more complicated type of hypothesis would be "Most people in this class took less than 20 minutes to get to school this morning."

You could group the data (e.g., 1-10, 11-20, 21-30 min, etc.) and draw a bar chart. It may be better to ask people to tick a range for how long it took rather than simply to offer either "$<$ 20" or "$\geq$ 20", because you could then use the same data to test a similar hypothesis ($<$ 10 mins, say).

3.1.5 **NEED** stopwatches (you could perhaps borrow some from the Science department).
Do something and time how long it takes;
e.g.,

- saying the alphabet backwards;
- estimating 1 minute: with eyes closed (so that clocks/other pupils aren't an influence) pupils face a clock with a seconds hand (perhaps held by the teacher). They open their eyes when they think 1 minute is up and write down the time showing on the clock;
- doing a tangram (or similar) puzzle;
- putting a list of names into alphabetical order.

These are relatively quick and easy ways of getting some continuous data to use for testing a hypothesis.

This list could easily be extended.

3.1.6 Subtleties of Sampling.
Imagine I stand on a motorway bridge and have some way of measuring the speeds of the vehicles on the motorway (say I borrow a radar gun from the police). Imagine I count the number of vehicles that go under the bridge in one direction (assume the motorway isn't too busy) and I keep a record of the number of *speeding* vehicles.
Say 20% of the vehicles are breaking the speed limit.

There are lots of factors to consider;
e.g., is this a typical road?, did I sample enough vehicles?, was it during "rush hour"?, etc. But also, the sampling method *is actually biased, because if you stand in a stationary place you are more likely to be passed by a speeding vehicle than a slower one, because they are travelling faster. You inevitably over-sample the faster vehicles by the method of sampling. (Imagine that throughout the motorway*

Is that a valid test?

there were lots of cars all going at 5 mph – it's quite unlikely one will pass under your bridge during the time you're doing your counting.)

3.1.7 **NEED** photocopy of A and B lines (see sheet) – or you can use two readily available lengths; e.g., length of classroom and width of exercise book.
Estimate the lengths of the two lines to the nearest mm. Write down in rough – you can't change your mind later!
Which length do you think we will estimate more accurately? (That's our hypothesis; e.g., "On average people are better at estimating short lengths than long lengths".)

Ask pupils beforehand to put away rulers etc., and to close exercise books that contain squared paper. Emphasise that "estimate" just means "best guess" in this context.

People usually think that they can estimate the length of the shorter line better because they can more easily imagine dividing it into centimetres if it's only a few cm long.

Make a table on the board, beginning

length (to nearest mm)	frequency
0-9	

and going as far as necessary. (Different ranges may be needed, depending on the sizes of the lengths estimated.) Hands up for each group.
Check that total frequency is the same as the number of people in the room

Draw grouped bar charts and/or frequency polygons and compare the distributions.
Is there any evidence to support our hypothesis?
Can you estimate the true lengths of the lines from our data?

Actual lengths: A is 53 mm; B is 128 mm.

People probably find it easiest to estimate "medium-sized" lengths that are within our common experience.

3.1.8 **NEED** rulers or metre sticks.
One route to continuous data is to measure "reaction times". It's best to find the mean of three attempts, discounting freak results ("I wasn't ready!").
One pupil holds the zero end of a ruler between two open fingers of another pupil and without warning drops the ruler. The other pupil catches the ruler as quickly as possible and the measured length is an indication of reaction time.

Hold the ruler over a bag or coat so that it isn't damaged by hitting the floor.
30 cm rulers may not be long enough!

Remember that long length indicates long reaction time (slow reactions), but the two are not proportional. If the length dropped is l cm, then the reaction time t (in seconds) is $t = \sqrt{\frac{l}{490}}$.

Typical values are around 0.1 to 0.3 seconds.

(Multiply by 1000 to convert to milliseconds.)

For what jobs might you need fast reaction times?
What things would affect someone's reaction time, do you think?

aeroplane pilot, driving instructor, train driver, operator of machinery, zoo-keeper

tiredness, alcohol/drugs, stress, age

3.1.9 **NEED** tape measures or long rulers.
Is there any correlation between hand-span and arm length (wrist to elbow)?

(Of course, within any school year the range of different ages is likely to be almost 12 months and possibly more.)

These statistics are less problematic than height and weight, which are too sensitive for many pupils.

Generally there is reasonably strong correlation within a class of similar-age pupils.

Skateboarding Questionnaire

Here are two versions of a questionnaire about skateboarding.
The aim is to find out about pupils' attitudes to skateboarding at school.

What is different about the questionnaires?
Why do you think these pupils wrote the questionnaires the way they did?

Andy's questionnaire

1. Do you think that keeping healthy is important? Yes/No
2. Do you think that skateboarding is a good form of exercise? Yes/No
3. Do you think that the school should encourage pupils to do things that promote their health? Yes/No
4. Do you agree that there should be a skateboarding area at school? Yes/No

Billie's questionnaire

1. Do you think that keeping healthy is important? Yes/No
2. Do you think that skateboarding can be dangerous? Yes/No
3. Do you think that the school should discourage pupils from injuring themselves on pointless and dangerous activities? Yes/No
4. Do you agree that skateboarding should be banned at school? Yes/No

What makes these questionnaires biased?

Can you write a better version.

Can you make a version that other people agree is unbiased?

A ________________

B ______________________________________

A ________________

B ______________________________________

A ________________

B ______________________________________

A ________________

B ______________________________________

A ________________

B ______________________________________

A ________________

B ______________________________________

A ________________

B ______________________________________

A ________________

B ______________________________________

Sampling

- **Population**
  This refers to the whole set of data values (they don't have to be people to be a population).
  It has a mean μ (mu) and variance σ^2 (sigma squared), but the population is usually so big that it wouldn't be practicable to work these out exactly using every piece of data.
  So instead you take a **sample** and find the mean $\bar{x}$ and variance $\text{var}(x)$ of your sample and use these as estimates of μ and σ^2.

- **Sample**
  To give accurate answers, the sample you take must be
  ✓ **large enough**, so that your results aren't distorted by a few unusual values.
  The **sampling fraction** is a way of saying how big your sample is relative to the population:

  $$\textbf{sampling fraction} = \frac{\textbf{sample size}}{\textbf{population size}}$$

  ✓ **representative**, so that what's true of the sample is true of the population generally.
  This means having a fair way of choosing which values go into the sample.

- **Simple random sampling**
  Here, every possible sample of that size is equally likely to be chosen, and therefore every member of the population is just as likely to be picked.
  (Calculator random numbers are often useful.)

- **Stratified sampling**
  Use this when you expect there to be different sub-groups of the population (called **strata**) with *different* properties, so you want to make sure that you have data from each stratum (in the right proportions).
  e.g., voting patterns may be different among low, medium and high income groups, so you could sample from each of these strata in the same proportions that they occur in the population as a whole.

- **Cluster sampling**
  This is useful in the opposite situation, when you have different sub-groups of the population that you expect to have *similar* properties. So you just pick one or more of these and choose your sample from among them.
  e.g., milk yields from a particular breed of cows are likely to be similar from similar farms in the same area, so you randomly choose several farms and sample from just them.

- **Quota sampling**
  This involves deciding in advance how many from each strata you want to include in your sample.
  e.g., you ask ten year 10 pupils and ten year 11 pupils what is their favourite subject at school; this method is often used in market research (clipboard people).

- **Systematic sampling**
  This involves choosing every 3rd (or whatever) item from a list. So long as the list is arranged in a random order this will give a random sample.
  e.g., picking every 5th person from a random page of the telephone directory would be a way of systematically sampling people whose names are in the phone book.