

Name _____

Date _____

Unit 6 · Collecting data — revision sheet

Everything you need from every section. Read it before the end-of-unit test.

6.1 Conducting an investigation

- A **statistical question** is one whose answer varies, so you need data to answer it.
- Write the question and your **prediction** first — then you know what data to collect.
- **Primary** data you collect yourself; **secondary** data somebody else collected.
- **Categorical** data is words. **Discrete** data you count. **Continuous** data you measure.
- A **tally chart** counts reliably in fives; the totals become the **frequency** column.
- Group continuous data into **class intervals** that do not overlap and leave no gaps: $150 \leq h < 160$, then $160 \leq h < 170$.
- A fair question is neutral, and its answer boxes cover every case exactly once.
- A prediction that the data does not support is still a real result. Report it.

Key words: data · statistical question · prediction · primary data · secondary data · categorical data · discrete data · continuous data · tally chart · frequency · class interval

x What is my shoe size?
 ✓ What is the most common shoe size in Year 7?
 ✓ Do learners who walk to school arrive earlier?
 Question: which sport is most popular in Year 7?
 Prediction: football, because most clubs are football clubs
 Data needed: each learner's favourite sport (categorical)
 favourite colour → categorical

6.2 Taking a sample

- The **population** is everyone you want to know about; the **sample** is who you actually ask.
- A **census** asks everyone — most reliable, rarely practical.
- A larger **sample size** generally gives a more reliable result, because one unusual answer matters less.
- A larger sample **does not** fix bias. Bias is about who is chosen, not how many.
- **Random**: everyone has an equal chance. **Systematic**: a fixed rule such as every 10th.
- **Convenience**: whoever is nearest — usually biased.
- Ask where you are sampling from: a bus stop samples bus users, a football club samples football fans.
- To estimate a whole population, find the proportion in the sample and apply it: 24 out of 60 is 0.4, so $0.4 \times 900 = 360$.
- Always state your sample size when you report a result — it tells the reader how much to trust it.

Key words: population · sample · census · sample size · random sample · systematic sample · convenience sample · bias · representative

Population: all 800 learners in the school

Sample: 40 of them

Census: all 800 — accurate, but a whole day's work

10 asked → 7 yes → 70%

one more yes → 80% a huge jump

500 asked → 260 yes → 52%

one more yes → 52.2% barely moves