Understanding quantitative data

What is quantitative data?
Quantitative data is any information that can be reduced to a set of numbers. Any information from which you can create averages, differences, or totals is quantitative data. Many forms of qualitative data can be turned into quantitative data by assigning numbers to categories (for example, letter grades, clusters in the literacy and numeracy continua, or EAL/D phases).

Types of common quantitative data in schools might include:
- Student assessment scores and other student outcomes (for example School Excellence Framework (SEF) data reports).
- Aggregates of responses from surveys (for example Tell Them From Me).
- Financial or Human Resources (HR) information.

What are the benefits of quantitative analysis?
Quantitative analysis is focused on combining information across a group or multiple groups of people, and using this to draw out wider trends. High quality quantitative analysis can be generalised to tell you about more than the specific individuals you are collecting information on (for example, election polls of just over a thousand people can give an estimate of how the whole country will vote).

Quantitative data can also allow you to be more precise with particular statements. As well as indicating whether student learning is different between groups or is changing over time (which qualitative analysis can also tell you), quantitative analysis can tell you by how much student learning differs and by how much it has changed.

What are the limitations of quantitative analysis?
Because quantitative analysis is focused on drawing out trends across a number of people, it is less effective at providing a lot of detail about an individual, or explaining why these trends have come about. Often, analysis using quantitative data answers broad questions, but may lead you to ask more specific questions. Answering these more specific questions may require follow-up qualitative investigation. For example:

<table>
<thead>
<tr>
<th>Broad question (before)</th>
<th>Specific question (after)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why is growth for students in my class this year lower than last year?</td>
<td>Why don’t these five students appear to fully understand these particular topics?</td>
</tr>
<tr>
<td>Why is our attendance rate low?</td>
<td>Why are our Year 9 girls less engaged?</td>
</tr>
<tr>
<td>How can we improve our teaching practices?</td>
<td>What is happening in these particularly effective classrooms, and how do we translate them to other classrooms?</td>
</tr>
</tbody>
</table>

Analysing quantitative data

Analysing quantitative data can take many forms, from looking at a graph to complex statistical research. This page sets out a few key principles that are relevant to all forms of quantitative analysis.

Variability is normal, and not all changes in data are meaningful
The single most important thing about statistics is that they must be treated with caution and thoughtfulness.

Most of the measures we are using in an educational context (like test scores) are imperfect indicators of something we are interested in but can’t directly observe – how much a student knows or has learned or a student’s engagement or wellbeing.
Our measures are imperfect because they can be affected by many unpredictable factors or events. For example:

- A student has recently experienced a traumatic event in their home life
- It was 40 degrees on the day of the HSC exam
- A teacher has had to take three months emergency leave, resulting in a series of different casual teachers taking their classes
- The school has been connected to a high profile incident.

Some of these are likely to affect outcomes of individual students, while others might affect classes or the entire school. If we are interested in things like how effective the school is in imparting skills and knowledge, then none of these causes are especially meaningful to us. We should pay attention to what the data is telling us, but be prepared to take it with a grain of salt.

The more students making up a measure, the more we can trust our data

Sometimes changes in individual students can have large effects on aggregate scores for the whole school. For example, take a very small school that only has six students in each year – two students around the minimum standard, two students in the middle of the NAPLAN band range and two near the top. Suppose that two things happen:

1. In 2013, a student who typically gets lower scores is away sick on NAPLAN day.
2. In 2014, a student who typically gets higher scores is away on holiday on NAPLAN day.

Because there are only six students, the absence of one student is going to make a big difference to the average score in those years, particularly because the absent student had very different scores from the average. It is important to note that the scores for the participating students might still tell you some useful things about those individuals and their learning. The issue we have is generalising from these results to how the school is doing as a whole.

Now consider the same situation, but there are 50 students in the year in the same school. In this case, the absence of any one student doesn’t really affect the school’s average. The average score will still go up and down from year to year, but the change won’t be as large. In this instance, the unpredictable events happening to individual students don’t really affect the measure we are looking at, and so we are better able to assess how all of the students as a group are doing. When we have larger numbers of students, the differences between individual students become less important (the ‘outliers’ tend to cancel each other out), and it’s easier to tell how the whole group is doing. This is why the data reports you might come across tend to emphasise the number of students making up a measure (or the measure’s sample size).

The more consistent the trend over time, the more we should trust our data

In schools with small sample sizes, it’s dangerous to place too much emphasis on any particular year. However, we might still be able to learn something by looking at the trend over time. If the result in any particular year is being caused by unpredictable events, these will tend to go away the next year. If we see a consistent pattern across a longer period of time, it is more likely to be indicative of something real that is going on.

The smaller the confidence interval, the more we should trust our data

Measures like test scores are imprecise estimates of the learning we are trying to measure. Statisticians have a way of quantifying exactly how imprecise any particular measure is: the confidence interval. Confidence intervals are usually represented visually by a set of error bars above and below a particular value (see Figure 1).

![Figure 1. Example of a confidence interval](image-url)

If the same students were to sit a test like NAPLAN or the HSC on many different days, they might get a slightly different result each time (due to unpredictable factors like how they are feeling on any particular day). This might result in different average scores for a school on each of 100 different days. The confidence interval is a range of scores that we are 95 per cent confident contains the true value for the school. The true value will lie outside the range only 5 per cent of the time – on 5 of the 100 days.

Note that My School uses 90 per cent confidence intervals, rather than 95 per cent. The interpretation is exactly the same, it just means that the true value for the school will lie outside of the confidence intervals on My School about 10 per cent of the time, rather than 5 per cent of the time.

When the confidence interval is very narrow, it means that we are very sure about the estimated number, and it is unlikely to be affected much by unpredictable events. This usually happens when there are a lot of students and the students are very similar to one another. When the confidence interval is very wide, it means that we aren’t at all sure about our estimate – the true value could be very far away from what this data is telling us. This is most often the case in very small schools.
Confidence intervals are useful because they tell us whether any observed differences in values (between results from one year and the next, or results from one school to a group of similar schools) are likely to be due to normal variation, or whether something more meaningful is going on. As rules of thumb:

1. If a confidence interval for a school does not include a certain value, then the school’s results are statistically significantly different from that value.
2. If the confidence intervals for two different groups do not overlap, then those groups are statistically significantly different from each other (see Figure 2).

Figure 2. Example of confidence intervals showing statistically significantly different scores

These statements mean that we are quite sure the two values or groups are actually different from one another in a tangible way, and the observed differences are not simply due to scores varying in regular fashion from year to year.

Make as many comparisons as possible (as long as they are fair comparisons).

It can be difficult to make value judgements about quantitative student performance data without additional pieces of information. For example, if students scored an average of 80 on a particular maths assessment, how do you know if this is a good or a bad thing? One easy way to analyse your students’ data is to compare it to some sort of benchmark. In doing this, successes and opportunities for future growth can become clearer. Depending on your data, you might be able to make three types of common comparisons:

- **Comparison to self over time:** If a school or a class is seeing results from the same type of assessment increase year-on-year, then it could be a good indication that things are improving. However, this comparison method has some limitations. The first is that at a student level, simple year-on-year growth may not be sufficient to demonstrate desirable progress. If a student learns a small fraction of the curriculum over one year, they will have made progress from the start of the year, but would likely not be meeting expectations that we would have of their learning. The second limitation is that, at a class or school level, this comparison does not tell us how much room there is for improvement.

- **Comparison to normative or ‘expected’ standards:** The learning of a student at a point in time or over a period of time can be assessed by comparing it to defined external benchmarks. Often this is based on population norms, or the average student (such as expected growth measures reported in SMART). Other benchmarks can include an expected level of proficiency (such as NAPLAN’s National Minimum Standards or the Literacy and Numeracy Continua) or system-level targets (such as the Premier’s priority of students in the top two NAPLAN bands). These comparisons can be limited when referring to students who are very different from the average student, such as gifted and talented students or students with some disabilities.

- **Comparison to other students or schools:** One way to gauge student performance and growth over time is by comparing it to a different set of students operating in a different class or school over the same period of time. When making this comparison you should take care that the comparison is a fair one. For example, if your school has students from very high or very low-SES backgrounds, then a comparison to the state average is not likely to be informative.
The more similar the comparison group is to your students, the more useful the comparison will be. Because these comparisons tend to be relative, you should be careful when making value judgements (it must always be the case that some schools and some students are below average). However, these comparisons can often enable investigation to see how much scope there is for improvement, and what practices might be effective in getting you there.

Each of these comparisons have strengths and limitations. However, they are not mutually exclusive – you can strengthen your ability to extract meaning from your data by using several of these comparisons at the same time. The best types of data will allow you to track your school’s student performance over time, relative to the performance of similar students, as well as an unchanging external benchmark or target.

**Always take into account context, and ask ‘what else might be going on?’**

A large body of research has established that particular characteristics of students and schools are associated with different outcomes:

- High-SES students tend to have higher scores than low-SES students.
- Students who are not Aboriginal tend to have higher scores than Aboriginal students.
- Girls tend to have higher scores than boys (except in maths or science).
- Students in selective schools and opportunity classes tend to have higher scores than students in non-selective schools or classes.
- Students who have had higher scores in the past will tend to have higher scores in the future.

This means that it is very difficult for some schools to assess their effectiveness. If a school has lower scores than other schools, does this mean that it has a lot of room for improvement, or simply that the students in the school face more educational disadvantage than in the typical school? There are two main approaches to getting at this issue.

The first approach is to compare the results of the school to a group of statistically similar schools. This comparison is included in My School, SMART, and the SEF reports. In all three examples, schools are ranked according to their ICSEA value, and the school is compared to the average of the schools with similar ICSEA rankings. For example, in SEF reports and SMART, your school is compared to the 20 schools with ICSEAs just below your school. Note that this group is just statistically similar – it may be very different in all sorts of ways that aren’t captured by the ICSEA value (for example, in terms of its size, location, or cultural background of students). Your SEF data reports contain a Contextual Information report, which compares exactly how similar your school is to its similar school group using a range of indicators.

The second approach is a new type of measure called a value-added measure. This is a statistical method that attempts to adjust for external factors outside of the control of schools, in order to isolate the individual impact on learning an individual school is having. For example, suppose that girls tended to outscore boys by 20 points, but otherwise all other differences in results were due to different approaches schools have taken. In this scenario, it would be unfair to directly compare scores for a boys-only school to a girls-only school – the girls school scores might look better simply because it enrols different students, not because it is doing anything differently. We could make the comparison more fair by adding 20 points on to the boys school, or taking 20 points away from the girls school.

Value-added scores are a more sophisticated form of this simple adjustment, which adjust for many different factors outside of the control of schools, all at once. They also adjust for the previous scores of each student, and so become a measure of growth between two points (for example, between Year 3 and Year 5). A value-added score represents the growth between the two relevant years that we would expect if the average NSW government school student was in your school. The actual growth of the average student is shown as a horizontal line on the graphs in your report (for VA3-5, the average growth is about 85 NAPLAN points). If your school is above this line, then your school is estimated to be adding more to scores than the average school. (See Figure 3).

The current value-added models take the following factors into account:

- Student SES (parental education and occupation)
- Student Aboriginality
- Concentration of SES at the school (school FOEI)
- Opportunity Class enrolment [VA5-7 only]
- Student gender [VA9-12 only]
- Fully academic selective schools [VA7-9 and VA9-12 only]
- Co-educational, girls only, or boys only schools [VA7-9 and VA9-12 only].

For more information about value-added measures, how they are calculated for NSW government schools, and how to interpret them, see the CESE value-added Learning Curve.