

Common errors in the interpretation of survey data

1) Quoting percentages only

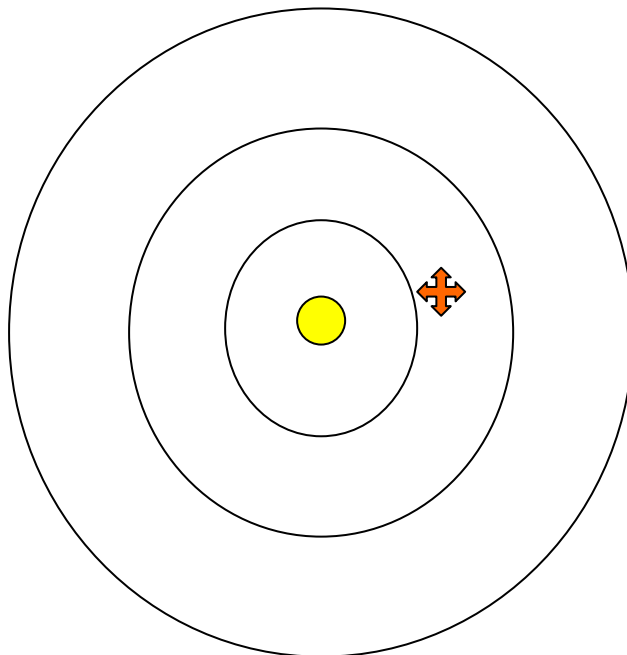
In general, it is irritating, if not unacceptable, for a report to be written quoting only percentages. A person reading a report on findings from survey data must be easily able to determine the base (i.e., the number of cases) on which percentages have been calculated. It can be quite misleading to present percentages and especially changes in percentages when the base for the percentage is very small.

2) Quoting unreliable results – remember the standard error

Statistical inference is the process of ‘guessing’ some attribute of a population from information contained in the sample. The attribute may be, for example, the percentage of Queenslanders who approve of the current premier. This is called a *population parameter* and the only way that we can determine its exact value is by taking a *census* of all Queenslanders. This is usually inconvenient so we aim to guess (or “infer”) this number by taking a *sample* of the population through some sort of survey process and calculating the corresponding *statistic*.

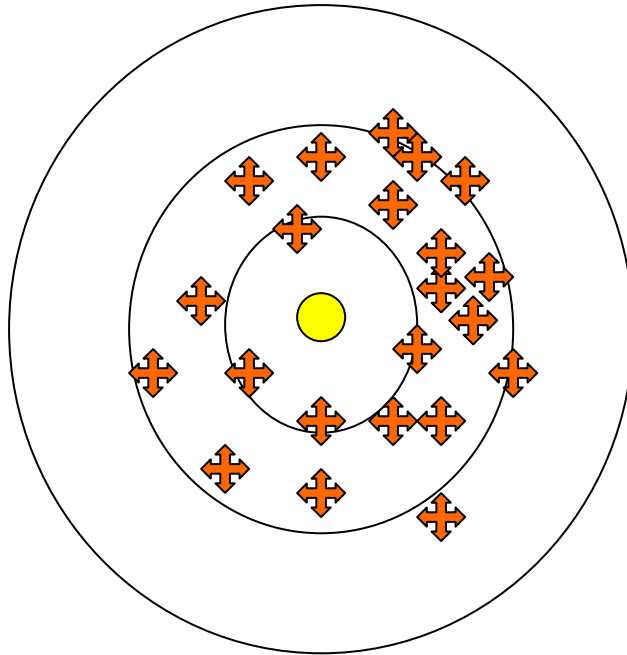
The ‘archery analogy’

The inference process can be likened to an archer aiming at his/her target. The ‘bulls-eye’ on the target represents the estimate (or “parameter”) and our sample statistic is the arrow. Our aim is to get the arrow as close as possible to the ‘bulls-eye’.



The standard error

If many such samples were taken, then some would result in larger estimates and some smaller estimates. The standard error is an average ‘distance’ of each from the real parameter.



Relative standard error

Ideally, we would like all of our errors to be small, thus improving the reliability of our results. However, due to problems of low sample sizes and other issues, small standard errors are not always the case. This raises the question of how big an error is too big?

To answer this question, it is normal practice to compare the standard error with the actual estimate. To make this comparison, we divide the standard error by the estimate obtained, and convert it to a percentage.

For example, in order to estimate the percentage of Queenslanders who prefer the current premier, I take a sample and obtain an estimate of 71% who prefer the current premier, with a standard error of 8%. The relative standard error is therefore:

$$\%RSE = \frac{8}{71} \times 100 = 11\% , \text{ which is quite acceptable.}$$

Generally, if the relative standard error is 25% or less, results have reasonable accuracy. However, as the relative standard error increases above this threshold, more caution needs to be taken when interpreting the results. The Office of Economic and Statistical Research usually highlights unreliable results by the use of asterisks (*). For example, if a result has a relative standard error greater than 25% but less than or equal to 50%, we place one asterisk

next to the value. If the relative standard error is great than 50%, we place two asterisks next to the value, and we advise not using the estimate due to its high unreliability.

3) Quoting results which are not of any practical significance

If you are routinely testing for statistical significance and a significant result is obtained the next step should be to consider whether the result is of any **practical** significance. If it isn't actually very important, then maybe it is not worth commenting on.

4) Using too many significant figures

Don't imply that the data are more accurate than they are.

If the standard error on a population estimate of 55412 is 8,400 then there are two significant figures in the error and the last significant figure is in the "100's" place. Therefore, population values should be rounded to the "100's" place.

5) Incorrectly making a comparison between two survey results

You can only say that one result is lower (or higher) than the other if the two results are statistically different at a specified level. Normally the level used is the 5% level. This means that there is only a 5% chance that you accept that the results are different when in fact they are not.

YOU MUST NOT IGNORE THE UNCERTAINTY INHERENT IN SAMPLE SURVEY RESULTS WHEN YOU INTERPRET THEM.

If two results look different, but they are **NOT** statistically different, you **cannot** say that one result is higher or lower than the other.

The apparent difference may just be a result of the particular sample that was selected from one of the populations and there is in fact no difference in the two populations. On the other hand, there could be a real difference in the populations but the sample selected was too small to be able to detect that the difference was real. You don't know.

Often in sample surveys, a researcher is interested in comparing results for different groups within the population of interest. Often, a group of particular interest to the researcher will be a small percentage of the population and hence the sample will only capture a small number in this group. In these situations, it is very tempting to look at percentages and make comparison statements that are not supported by the data.

For example:

Imagine that in a sample of 1000 people interviewed, 20 reported having been the victim of a particular crime. The researcher was interested in finding out whether people who were victims of this particular crime were as satisfied as the rest of the population with the police service.

The results show that

	Satisfaction with Police		
	% Very Satisfied / Satisfied	% Dissatisfied / Very Dissatisfied	Total
Victim of Crime (n = 20)	50	50	100
Not a Victim of Crime (n = 980)	60	40	100

The table below shows the relevant approximate 95% Confidence Intervals assuming a very large population.

	% Very Satisfied / Satisfied	Approximate 95% Confidence Interval on the % Satisfied / Very Satisfied
Victim of Crime	50	50 ± 22.5
Not a Victim of Crime	60	60 ± 3.1

Because the 95% confidence intervals overlap, you cannot say that people who have been a victim of crime are less satisfied with the police service than people who haven't been a victim of that particular crime, *even though* there seems to be a sizeable difference on first glance.

Extreme care must be used in drawing conclusions about subgroups of a population when the number of units captured by the sample in this sub group is very small.

6) Incorrectly comparing a survey result with an absolute value

A sample survey shows that 52% of the Brisbane population think that Lang Park is the best site for the Sports Stadium.

Can you legitimately say that “More than half of the people in Brisbane support Lang Park as the site of the Sports Stadium”?

Have you asked:

- How many people did they survey?
- What is the size of one standard error?
- What is the 95% confidence interval around the estimate?

Under ‘normal’ circumstances, 95% of our ‘estimates’ will lie within 1.96 (approximately 2) standard errors of the true parameter. Let’s say the 95% confidence interval is 42% - 62%.

The 95% confidence interval contains percentages less than 50%. The true value could be anywhere between 42% and 62%. You should not say “More than half”.

So what could you say?

- About half
- Just quote the value in context, in this instance ‘an estimated 52% of adults in Brisbane regard Lang Park as the best site for the Sports Stadium.’

On the other hand, if the 95% confidence interval was 50.5% – 53.5%, you could say that “More than half” because 50% is lower than the lower bound of the confidence interval.

7) **Incorrectly using the word, “Most”**

Let’s say that sample survey results showed that 45% of people thought that smoking should be totally prohibited, 30% thought that smoking should be prohibited in public places, and the remainder had not thought about the issue.

It would be correct to say that the most frequently occurring response, or the most popular response, was that people thought that smoking should be totally prohibited. (You could of course only say this if 45% was significantly higher than 30%.)

It would not be correct to say that most people thought that smoking should be totally prohibited. Most implies the majority in this context, i.e., at least 50%.

8) **Incorrectly assuming that an association between variables implies some causation**

In experimental research, you manipulate some variable(s) and then measure the effects of this manipulation on other variables. For example, a researcher might artificially increase blood pressure and then record cholesterol level. Only experimental data can conclusively demonstrate causal relations between variables.

The vast majority of survey data come not from experimental research, but from what is called correlational research. In correlational research, variables are not influenced (or manipulated). Variables are measured and relationships or associations between variables (e.g. correlations) are explored. As a result, correlational data cannot conclusively prove causality.

How does this affect how you write up survey data? Say, for example, that we estimate from a sample survey that 28% of males and 65% of females used the internet in the last week. We cannot say that the difference in internet use was *caused* by their sex. We can say internet use was *associated* with sex. The difference may be related to something else entirely – for example, the Soccer World Cup may have been on last week.