

Presenting survey results – Report writing

Introduction

Report writing is one of the most important components in the survey research cycle. Survey findings need to be presented in a way that is readable and technically acceptable. Good writing is essential to ensure the importance and usefulness of the findings are not lost.

This paper has been designed to provide guidance on how to write a report so that the information is easy to understand, and so that the findings are not unintentionally misrepresented.

The overall aim of the written report is to communicate the survey findings. A report provides a formal record of the survey research, and can provide a foundation for future research efforts.

The written report is often the only aspect of a survey research project that gets exposure. Consequently, the overall quality of a project is often judged on how well the report is written and presented.

A good report should:

- meet the brief and address the objectives
- present clear findings with strong conclusions (and recommendations if appropriate)
- be accurate and free of errors
- use tables, graphs and text appropriately
- be written in a way that will convey the results without overwhelming or boring the reader.

This paper provides hints and tips on writing for research reports. Tables and graphs are also important components of a research report. Information on presenting data in tables and graphs can be found in the following companion papers produced by the Queensland Government Statistician's Office (QGSO):

- *Presenting statistical information — Graphs*
- *Presenting statistical information — Tables*

Please note: the data contained in this paper are fictitious and are provided as examples only.

Key terms

The following statistical terms have been used in this paper and a brief definition is given for each. For a fuller definition and further information about these and other such terms, please see *Assessing survey data* available on the QGSO website.

Term	Brief definition
Confidence interval	provides a range of values around the estimate, within which the true value can be expected to fall. The smaller the confidence interval is for a particular estimate, the more precise the estimate is.
Relative standard error (RSE)	a measure of accuracy of a survey estimate — calculated by dividing the standard error by the estimate obtained — often expressed as a percentage.
Standard error	a statistical term for measuring the accuracy with which a statistic taken from a sample represents an entire population. The smaller the standard error, the more representative the sample will be.
Statistical significance	a statement about the likelihood of findings being due to chance.

Report structure

Most survey research reports follow a fairly standard structure with the following inclusions:

Front page

This is the window into the report. Choose a short, relevant and interesting title. Include the date of publication/release, author and contact details such as address and phone number. Copyright and/or licence information should also be included, if report is published.

(see <http://www.ausgoal.gov.au/copyright>)

Table of contents

This allows the reader to gain an overview of what is contained in the document, and quickly locate sections they may wish to read.

Executive summary

The executive summary should summarise all the main points discussed in the report, including the methods, key results (preferably grouped by objectives or themes) and conclusions/recommendations. The executive summary should ideally be limited to one or two pages. It should highlight the key messages that the decision makers will need to know, but not necessarily every result that is of statistical significance. The summary should be the most polished part of the document, as it will almost certainly be the part that is most read.

Background and objectives

This section provides the reader with context to understand the background to the survey, and the reasons for which the research has been conducted. It would typically also include information on how the research results are intended to be used. The amount of detail provided in this section will depend on the nature and complexity of the survey.

Survey method

This part discusses who has been included in the survey and why, the source material or list from which respondents from the target population were drawn (the list may include individuals, households or institutions), how many people were surveyed, how they were contacted, and the method of data collection. Individual sub-sections may be required to adequately cover all the details, depending on the complexity of the survey. Highly technical information regarding survey method and operations can be included in an appendix.

Survey results

This is the main body of the report, containing details of the survey results. Consider structuring this section according to the objectives of the project or themes, rather than by question order. Refer back to the research objectives and ensure they are all covered.

The information contained in this section shows the key results from the survey. It does not need to include every possible breakdown of the data. Rather, it should highlight results that are of both statistical and practical significance.

All sources of data in the findings should be referenced so it is clear where they have come from, and so their credibility can be assessed. Any dubious content should be put into an appendix of supporting material, rather than included in this section.

Appendices

The survey questionnaire should always be included in the appendices.

The appendices may also contain detailed technical information about survey methodology, collection processes, data processing and analysis techniques.

Appendices commonly contain a glossary of technical terms and a list of figures and graphs. Alternatively, a glossary and list of figures and graphs can be provided at the front of the document, to introduce and explain technical terms the reader will encounter.

Style

Agencies and organisations will have their own preferences in terms of style. Regardless of specific style preferences, the report should be consistent in application of the following components:

- font
- spacing
- abbreviations
- dot points
- formatting of headings and subheadings
- page numbering
- quotes and references.

Writing for reports

The purpose of report writing is to communicate the findings of the research. The report should tell the whole story — what the objectives of the research were, how the data were collected, what the data say and what the implications of the findings are.

Every individual has their own style of writing. While there is no single 'right' style for report writing, there are some basic principles that should be followed.

Plain English

Report writing often involves communicating abstract and potentially difficult concepts (such as those involved in statistical modelling and analysis). Careful use of language is especially important in this context. Poor quality writing makes it difficult for the reader to work out the intended meaning, and can lead to the results being misinterpreted. Use of plain English helps to ensure the report is easily understood by the target audience. Plain English emphasizes clarity, brevity, and the avoidance of jargon and technical language. While the use of technical language is appropriate for some audiences, it should never be assumed that all readers will understand all terms. If technical terms are included, they should each be defined.

Consider the audience

The information needs to be appropriately conveyed to the key decision makers. There is a high probability that most people who will read the report are not trained statisticians. Assume that readers will not have a strong grasp of statistical concepts. Provide definitions for terms such as statistical significance, confidence interval etc. in layman's terms.

Use clear simple language and short sentences

Write in a way that will be easily understood by the target audience. Use clear, unambiguous language that is appropriate to their reading skills and knowledge.

Try to include just one complete, clearly presented point per sentence. If connecting more than one point in a sentence, make sure they are directly and clearly linked.

Readers should feel as though they're getting good value in return for the number of words they're required to read. Don't use three words when two will do — use precise word choices, and use an online thesaurus to help find the most appropriate word if necessary. Also work to eliminate unnecessary words. For example, much of the time the word 'that' is totally unnecessary. When in doubt, leave it out.

Don't include clichés, colloquialisms, vogue words/expressions or needless jargon.

Use dot points where appropriate. They can be an effective way of summarising a large amount of information.

Correct grammar and punctuation

It's very important to ensure that grammar, spelling and punctuation are correct. Careless errors can be distracting and damage credibility.

Plain English examples

Original:

More Mazdas than Holdens had automatic steering. Mazdas were also more likely than Holdens to have airbags. Additionally, significantly higher proportions of Mazdas than Holdens had leather seats.

Plain English:

Mazdas were more likely than Holdens to have:

- *automatic steering*
- *airbags*
- *leather seats.*

Original:

When the data are demarcated by age, it can be seen that the likelihood of purchasing lollipops is similar for all three age groups. There are no notable findings here.

Plain English:

The likelihood of purchasing lollipops is similar for all three age groups.

Original:

When looking at expenditure and amount of time spent shopping, females tended to have a higher mean and a higher median than males.

Plain English:

Females tended to spend more and take more time when shopping than males.

Significant figures

Figures should be rounded as appropriate, so that the level of accuracy of the data is correctly represented.

The significant figures of a number are those digits that carry meaning contributing to its precision. For example, if the standard error on a population estimate of 55,412 is 8,400 then there are two significant figures in the error, and the last significant figure is in the 100s place. Therefore, population values should be rounded to the 100s place. In this instance, the figure quoted should be 55,400, rather than 55,412.

Using percentages

Data users must be able to easily determine the base (i.e. the number of cases) on which percentages have been calculated. It is misleading to present percentages, and especially changes in percentages, when the base for the percentage is very small. For example:

Imagine there are 10 respondents to an “agree/disagree” question in a survey. Nine respondents agree, and one disagrees. If percentages were to be quoted on such a small sample, they would show that 90% of respondents agree and 10% disagree. If just one respondent answered differently (ie if eight respondents agreed and two disagreed), then percentages quoted would show that 80% of respondents agree and 20% disagree. This looks like a large shift (10%), but in fact represents only one person changing their mind.

Imagine there were 1,000 respondents to the same question, Nine hundred respondents agree, and 100 disagree. Percentages quoted would again show that 90% of respondents agree and 10% disagree. In order to obtain a shift of 10% in your quoted results, you would need 100 people to provide a different response.

Generally, quoting percentages on less than 30 cases is considered to be misleading and is strongly discouraged. It is preferable to quote percentages when you have at least 100 cases.

Reporting on the average (mean)

When providing a measure of an average (mean) value, appropriate context should be given. The reason behind the mean value is often more important than the number itself. For example, if a student takes four tests perfectly, and fails to take a fifth test (thus earning a zero), the mean reflects this. If the tests are all worth 100 points, the mean score is 80%. However this does not really reflect the average performance level for this student, because of the outlier of zero. When reporting on this information, it should be stated that the mean score of 80% is affected by the student not taking one of the tests, and that the student's mean score based on the tests actually taken was 100%.

Comparing survey results and subgroups using the confidence interval

It is correct to say that one result is lower (or higher) than the other only if the two results are statistically significantly different.

If two results look different, but they are not statistically different, it is incorrect to state that one result is higher or lower than the other.

It can be tempting to look at percentages and make comparison statements that are not supported by the data.

For example:

Imagine that in a sample of 1,000 people interviewed, 20 reported having been the victim of a crime. The researcher was interested in finding out whether people who were victims of crime were as satisfied as the rest of the population with the police service.

The results show:

	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

It could seem appropriate to state that people who were victims of a crime were less likely to be satisfied with the police service (50%) than people who were not victims of a crime (60%). However, this does not take into account the confidence interval around the estimate. The 95% confidence interval shows the range within which we can be 95% sure that the true value lies.

The table below shows the 95% confidence intervals for the two sample groups.

	% 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

Even though there seems to be a sizeable difference (10%) on first glance, it can be seen that the 95% confidence intervals overlap. It would therefore be incorrect to 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 crime.

Note: the confidence interval will be bigger for small sample sizes than for larger sample sizes. Extreme care must be taken in drawing conclusions about subgroups of a population when the number of units captured by the sample in the sub group is very small.

Using the phrase 'more than half'

A sample survey shows that 52% of the Brisbane population think that Lang Park is the best site for the sports stadium. It could seem appropriate to report that 'More than half of the Brisbane population thinks that Lang Park is the best site for the sports stadium'.

Before using this phrase, consider the 95% confidence interval for the statistic. If the 95% confidence interval for the survey result of 52% is $\pm 8\%$, then the true value could be anywhere between 46% and 60%. The response of 52% could be reported as:

- 'About half', or
- '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 $\pm 1.5\%$, then the true value lies somewhere between 50.5% – 53.5%. In that instance, it would be appropriate to say 'More than half' because 50% is lower than the lower bound of the confidence interval.

Using the word 'most'

Results from a sample survey show that 45% of people think that smoking should be totally prohibited, 30% think 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 think that smoking should be totally prohibited (assuming that 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, i.e. more than 50%.

Statistical significance vs practical significance

Statistical significance refers to whether any differences observed between groups being studied are 'real' or whether they are simply due to chance.

If the analysis reveals a statistically significant result, the next step should be to consider whether the result is of any practical significance. It is usually only worth commenting on statistically significant results that are of practical importance in the context of the research question.

Remember the relative standard error (RSE)

Survey results are estimates of population values and always contain some error because they are based on samples, and not the entire population. The RSE is one tool for assessing the precision of survey estimates.

Reference should be made to the RSE of an estimate before reporting on the data. Generally, if the RSE is 25% or less, results have reasonable accuracy. As the RSE increases above this threshold more caution should be used when interpreting and reporting on the results. If the RSE is greater than 50%, the result has very low reliability and should not be used in the report.

Association vs causation

In experimental research, a variable is manipulated and the effect(s) of this manipulation on other variables is measured. For example, a researcher might artificially increase blood pressure and then record cholesterol level. In most cases, only experimental data can conclusively demonstrate causal relationships between variables. The vast majority of survey data is correlational research, not experimental 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.

For example, estimates from a sample survey might show that that 28% of males and 65% of females used the internet in the last week.

It would be correct to report that internet use was *associated* with sex. It would not be correct to say that the difference in internet use was *caused* by their sex. The difference may be related to something else entirely — for example, the Football World Cup may have been televised that week.

There are some cases in which a correlation between variables is considered sufficient evidence to prove some form of relationship. For example, young drivers are correlated with more accidents, so insurance companies tend to charge them more for car insurance. It would not be ethical to design an experiment to test this, so correlational data are used as a proxy for experimental data in this instance.

Referencing the data

All sources of data in the findings should be referenced so that:

- it is clear where they have come from
- their credibility can be assessed.

Checking the report

All reports should be thoroughly checked before being presented or published. The report should be proofread by someone other than the author. Proofreading the report should include checking the following items:

- Has an appropriate report structure been used?
- Does the title page show who prepared the report, when it was prepared and contact details?
- Is an up-to-date table of contents included?
- Is the presentation of the report consistent in every respect, including use of fonts, margins, heading styles?
- Are page numbers included?
- Are all figures and graphs labelled and numbered?
- Are tables and graphs understandable and appropriately and consistently labelled?
- Do figures in the tables and graphs match figures quoted in the text?
- Is the report free of typographical and interpretation errors?
- Does the report cover all the research objectives?
- Is the language clear and concise, and free of grammatical errors?

Client input

In instances where research has been commissioned by a third party (the 'client'), the client should be given the opportunity to request changes to the report. The first version of the report issued to the client should therefore be clearly labelled as a 'draft for comment'.

The client may make suggestions that can enhance the usefulness and quality of the report. These suggestions should be incorporated into the report if at all possible. Sometimes, however, a client may request changes that could affect the integrity of the report. For example, they may request a change of wording which would make a particular finding appear more important than it actually is. This can happen if the client has a particular agenda in mind, or lacks understanding of how the results should be interpreted. If the requested change affects the integrity of the reporting, then the change should not be incorporated. In such instances, the report writer should provide the client with appropriate information to inform and educate them as to why the requested changes are inappropriate.

The role of a good researcher is to report the findings as they actually are, not as the client wishes them to be.

Summary

The overall quality of a research project is usually judged on how well the report is written and presented. It is therefore essential to ensure the report is useful, accurate and error free.

A good report:

- addresses the objectives
- uses an appropriate and logical structure
- is consistent in style and application of formatting
- presents findings in a clear and succinct manner, using plain English
- provides relevant context around statistical findings
- focuses on findings that are of statistical and practical significance
- is free of errors
- meets the brief
- uses tables, graphs and text appropriately
- incorporates client feedback where appropriate.

While technical knowledge helps, report writing is an art which is generally best learnt through practice and experience.

Further information and advice

Queensland Government Statistician's Office has extensive experience in the design and conduct of statistical surveys and is able to provide specialist consultation and advice on request. For more information please see the QGSO website, www.qgso.qld.gov.au, or email govstat@treasury.qld.gov.au.

