Survey Analysis for Fake Client

Sean van der Merwe

2022-10-26

Table of Contents

*Activate the navigation pane (View menu) for ease of navigation through the document.*

# Links

[Consultation Unit](https://www.ufs.ac.za/natagri/departments-and-divisions/mathematical-statistics-and-actuarial-science-home/statistical-consultation-unit/statistical-consultation-unit)

[Consultant](https://seanvdm.co.za)

[Frequently asked questions](https://seanvdm.co.za/post/faqconsult/)

# Introduction and disclaimer

Throughout our analysis, we must remember at all times that we are making judgements on the perceptions of people such as those who responded to this survey. [People who choose to respond to surveys may be systematically different from the general population.](https://www.dummies.com/education/math/statistics/statistics-conundrums-dealing-with-survey-nonresponders/) [See the Wikipedia page on Response Bias for more information.](https://en.wikipedia.org/wiki/Response_bias)

Similarly, the survey was only presented to a specific group of people, and we should not try to extend the results beyond that group without seriously considering and accounting for any systematic differences between that group and any broader group.

The data analysed here is inherently random. If the survey were to be repeated then the results will differ. We will try to estimate the extent to which this will be the case, but those estimates are themselves uncertain.

While [the computer software used](https://en.wikipedia.org/wiki/R_%28programming_language%29) is tried and tested, the analysis involves multiple human elements. Both the client and statistician may have introduced human error at various stages of the research process.

**Thus, no guarantee can ultimately be given on the correctness of any findings below.**

Links are given occasionally to sources explaining the statistical concepts addressed. These are colloquial sources to aid understanding, not reference sources. However, most of these sources link to reference sources, should the researcher require academic references for specific topics.

# Statistical topics that might be relevant

## The boringness assumption

* The classical (frequentist) framework of testing has many flaws and alternatives, but remains popular because of its simplicity
* It starts by **assuming that everything is boring**
	+ This is called the null hypothesis
* We assume that there is no difference between cases, treatments, situations, or groups
* We assume that there is no change over time, nor space

## The uncertain experiment

* We do an experiment and **see something that looks interesting**
* No experiment is perfect, there are always errors and uncertainties
* So we ask the question, “How interesting is this thing really?”
	+ Is this just a coincidence?
	+ If we did the experiment again, would the result be as interesting? More interesting?
	+ Or would it be less interesting or even entirely different?

## The p-value

* The [p-value](https://en.wikipedia.org/wiki/P-value) is defined as the probability of seeing something at least as interesting as what we observed, under the assumption that nothing is really interesting in general.
* If the p-value is large then we continue to assume that the null (boring) hypothesis is true. We **DO NOT** accept or conclude that the null hypothesis is true, that would be wrong. We simply say that *we have not found any evidence that contradicts it*.
* If the p-value is small then we reject our null hypothesis and conclude that **there is at least one interesting thing** to see.
* The probability is calculated by saying, “What might have happened had we repeated the experiment under the same settings an infinite number of times?”

## Significance level

* So what is big and what is small for a p-value?
* **Before doing any experiments or collecting any data** we pick a number close to 0 and say that we are happy with that rate of false positives.
* We call this number the significance level $\left(α\right)$.
* The significance level chosen is 0.05. [This means that we only look at results where the p-value is less than](https://en.wikipedia.org/wiki/P-value) 0.05, as other results could easily just be chance variation.
* If you haven’t done a statistical test then don’t use the word *significant*.

## Multiple testing problem

* The more things you try to test the more likely you are to stumble on an ‘interesting’ coincidence and end up talking nonsense
* People who don’t know what they want to know ask a ton of questions in the hope of getting ‘lucky’ and finding out something interesting, but chances are [they are just finding coincidences](https://en.wikipedia.org/wiki/Multiple_comparisons_problem)
* In statistics we often say we are happy with a 5% rate of false positives, but that assumes you are doing 1 test - if you do 40 tests then you expect to get 2 false positives
* It is possible to adjust for this in some ways, but you always lose power.
	+ The most common adjustment is the [Holm-Bonferroni method](https://en.wikipedia.org/wiki/Holm%E2%80%93Bonferroni_method)

## Tabulations

The most basic form of statistics is a simple tabulation of responses into frequencies and relative frequencies.

Should one wish to compare two nominal variables to each other and test for independence then one might do a $χ^{2}$ [(Chi Square) test](https://en.wikipedia.org/wiki/Pearson%27s_chi-squared_test).

## t-Test

* [The t-test](https://en.wikipedia.org/wiki/Student%27s_t-test) is all about averages
* The most popular statistical test for lots of reasons:
	+ People like to think in averages and ask whether things are the same on average
	+ It is easy to do and fairly reliable
	+ It flows from a Central Limit Theorem (and other theorems) that says:
		- If your numbers follow any well-behaved pattern, then the average of a large enough sample will follow a *normal* distribution

## Two-sample t-test

* The standard t-test compares the average of a single set of numbers to a predetermined number
* If you have paired observations (e.g. matched before vs after) then you subtract first to end up with a single set of differences which you can compare to 0 (no difference)
* But what if you have two separate set of numbers (e.g. Free State vs Gauteng)?
	+ You can compare the averages using a two-sample t-test
	+ The two-sample t-test is much less powerful and has a lot more assumptions though

## ANOVA

What if you have more than two sets of numbers?

* [Analysis of Variance](https://en.wikipedia.org/wiki/Analysis_of_variance) (ANOVA) asks whether all the groups have the same average $\left(μ\_{1}=μ\_{2}=μ\_{3}=…\right)$
* Suppose we ask each office of a business to rate their manager on a scale of 0 to 10. Each office operates independently so we can compare the manager ratings using an ANOVA to see if any managers are better or worse liked than the rest.
* ANOVAs assume that the groups have the same variance, so if one manager plays favourites and has more variance in rating then the ANOVA will not work right.
* If the ANOVA gives a small p-value then we reject the null hypothesis that all the averages are equal and conclude that **at least 1 average is different from the rest**.
	+ It does *not* imply that all the averages are different. To determine which averages are different people like to do a *post-hoc* test, such as the [Tukey HSD procedure](https://en.wikipedia.org/wiki/Tukey%27s_range_test)

## Rank-sum and U-test

If your data is badly behaved, skewed, unusual, or just doesn’t really meet the assumptions above then people will recommend using a non-parametric alternative.

**However, non-parametric tests have different null hypotheses, and are not directly interchangeable.**

For comparing a single sample to a fixed value we might use a [Wilcoxon signed-rank test](https://en.wikipedia.org/wiki/Wilcoxon_signed-rank_test). This is *not* testing the average at all, instead it tests the more general question of whether the values tend to be close to or far away from the null value.

For comparing two samples we can use the [Mann-Whitney U test](https://en.wikipedia.org/wiki/Mann%E2%80%93Whitney_U_test). This asks the question of whether one set of numbers is systematically larger or systematically smaller than another.

Because these questions are more general, and make less assumptions, they are also more robust. However, they can have less power to find differences.

More advanced tests also have non-parametric alternatives, e.g. the [Kruskall-Wallis](https://en.wikipedia.org/wiki/Kruskal%E2%80%93Wallis_one-way_analysis_of_variance) approach, which is quite similar to the ANOVA approach, but more robust to violations of the ANOVA assumptions.

## Correlation

Correlation measures the tendency for things to move together.

A positive correlation says that when one thing is above average the other will **tend to** also be above average (below average $\leftrightarrow $ below average). Think of temperature and watermelon consumption say.

**Correlation done not imply causation.**

We calculate the [Pearson correlation coefficients](https://en.wikipedia.org/wiki/Pearson_correlation_coefficient) using the scaled data. Positive correlations indicate responses that tend to vary (move around) in the same direction. Negative correlations are an indication of responses going counter to each other. Correlations that are not statistically different from zero will be marked as such with a cross.

## Congeneric reliability

In this section the goal is to determine to what extent specific sets of question measure the same thing. It is based on [factor analysis](https://en.wikipedia.org/wiki/Factor_analysis) - we try to find an underlying factor explaining the responses to a group of questions as a whole. [The measures calculated here](https://en.wikipedia.org/wiki/Congeneric_reliability) are related to the famous Cronbach’s $α$, but considered more accurate (although alpha is also given).

Reliability measures lower than about 0.7 suggest that multiple concepts are being measured; while measures above 0.95 suggest that you are asking exactly the same thing in different words. [If you are trying to measure a single concept reliably then a value in the target range of 0.7 to 0.95 is desired.](https://en.wikipedia.org/wiki/Cronbach%27s_alpha)

# Univariate analysis

## Frequencies of responses

We create summary tables for each nominal question and then for each scale question as if it was purely categorical.

### Gender

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Female | 76 | 0.38 |
| Male | 120 | 0.60 |
| Other | 4 | 0.02 |

### Age

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| 18-25 | 62 | 0.31 |
| 26-35 | 86 | 0.43 |
| 36+ | 52 | 0.26 |

### Likert1

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 0 | 0.00 |
| Disagree | 6 | 0.03 |
| Neutral | 44 | 0.22 |
| Agree | 75 | 0.38 |
| Strongly Agree | 60 | 0.30 |
| Not Answered | 15 | 0.07 |

### Likert2

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 91 | 0.46 |
| Disagree | 77 | 0.38 |
| Neutral | 20 | 0.10 |
| Agree | 2 | 0.01 |
| Strongly Agree | 0 | 0.00 |
| Not Answered | 10 | 0.05 |

### Likert3

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 91 | 0.46 |
| Disagree | 76 | 0.38 |
| Neutral | 19 | 0.10 |
| Agree | 4 | 0.02 |
| Strongly Agree | 0 | 0.00 |
| Not Answered | 10 | 0.05 |

### Likert4

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 126 | 0.63 |
| Disagree | 50 | 0.25 |
| Neutral | 10 | 0.05 |
| Agree | 1 | 0.00 |
| Strongly Agree | 0 | 0.00 |
| Not Answered | 13 | 0.06 |

### Likert5

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 0 | 0.00 |
| Disagree | 0 | 0.00 |
| Neutral | 6 | 0.03 |
| Agree | 46 | 0.23 |
| Strongly Agree | 143 | 0.72 |
| Not Answered | 5 | 0.03 |

### Likert6

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 0 | 0.00 |
| Disagree | 0 | 0.00 |
| Neutral | 1 | 0.00 |
| Agree | 7 | 0.04 |
| Strongly Agree | 179 | 0.90 |
| Not Answered | 13 | 0.06 |

### Likert7

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 1 | 0.00 |
| Disagree | 19 | 0.10 |
| Neutral | 65 | 0.32 |
| Agree | 70 | 0.35 |
| Strongly Agree | 38 | 0.19 |
| Not Answered | 7 | 0.04 |

### Likert8

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 57 | 0.28 |
| Disagree | 69 | 0.34 |
| Neutral | 46 | 0.23 |
| Agree | 15 | 0.07 |
| Strongly Agree | 0 | 0.00 |
| Not Answered | 13 | 0.06 |

### Likert9

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 46 | 0.23 |
| Disagree | 75 | 0.38 |
| Neutral | 50 | 0.25 |
| Agree | 14 | 0.07 |
| Strongly Agree | 2 | 0.01 |
| Not Answered | 13 | 0.06 |

### Likert10

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 107 | 0.54 |
| Disagree | 73 | 0.36 |
| Neutral | 14 | 0.07 |
| Agree | 1 | 0.00 |
| Strongly Agree | 0 | 0.00 |
| Not Answered | 5 | 0.03 |

### Likert11

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 42 | 0.21 |
| Disagree | 86 | 0.43 |
| Neutral | 44 | 0.22 |
| Agree | 13 | 0.06 |
| Strongly Agree | 2 | 0.01 |
| Not Answered | 13 | 0.06 |

### Likert12

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 20 | 0.10 |
| Disagree | 69 | 0.34 |
| Neutral | 57 | 0.28 |
| Agree | 33 | 0.16 |
| Strongly Agree | 11 | 0.06 |
| Not Answered | 10 | 0.05 |

### Likert13

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 57 | 0.28 |
| Disagree | 74 | 0.37 |
| Neutral | 44 | 0.22 |
| Agree | 17 | 0.09 |
| Strongly Agree | 1 | 0.00 |
| Not Answered | 7 | 0.04 |

### Likert14

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 68 | 0.34 |
| Disagree | 93 | 0.47 |
| Neutral | 24 | 0.12 |
| Agree | 2 | 0.01 |
| Strongly Agree | 0 | 0.00 |
| Not Answered | 13 | 0.06 |

### Likert15

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 2 | 0.01 |
| Disagree | 9 | 0.04 |
| Neutral | 36 | 0.18 |
| Agree | 77 | 0.38 |
| Strongly Agree | 69 | 0.34 |
| Not Answered | 7 | 0.04 |

### Likert16

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 3 | 0.01 |
| Disagree | 13 | 0.06 |
| Neutral | 56 | 0.28 |
| Agree | 81 | 0.41 |
| Strongly Agree | 40 | 0.20 |
| Not Answered | 7 | 0.04 |

### Likert17

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 2 | 0.01 |
| Disagree | 14 | 0.07 |
| Neutral | 56 | 0.28 |
| Agree | 80 | 0.40 |
| Strongly Agree | 37 | 0.18 |
| Not Answered | 11 | 0.06 |

### Likert18

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Strongly Disagree | 0 | 0.00 |
| Disagree | 0 | 0.00 |
| Neutral | 1 | 0.00 |
| Agree | 21 | 0.10 |
| Strongly Agree | 170 | 0.85 |
| Not Answered | 8 | 0.04 |

## Scale questions summary

We summarise the scale responses in stages.

### Quantiles

First we start by using quantiles (based on sorting in ascending order). Specifically, we note the count of responses, the minimum, the first quartile, the median, the third quartile, and the maximum.

*This is the least interesting and least relevant of all the results.* It is generally just included in an appendix for interest.

| Ques | Count | Min | Q1 | Median | Q3 | Max | Description |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Q3 | 185 | 2 | 3 | 4 | 5 | 5 | Likert1 |
| Q4 | 190 | 1 | 1 | 2 | 2 | 4 | Likert2 |
| Q5 | 190 | 1 | 1 | 2 | 2 | 4 | Likert3 |
| Q6 | 187 | 1 | 1 | 1 | 2 | 4 | Likert4 |
| Q7 | 195 | 3 | 4 | 5 | 5 | 5 | Likert5 |
| Q8 | 187 | 3 | 5 | 5 | 5 | 5 | Likert6 |
| Q9 | 193 | 1 | 3 | 4 | 4 | 5 | Likert7 |
| Q10 | 187 | 1 | 1 | 2 | 3 | 4 | Likert8 |
| Q11 | 187 | 1 | 2 | 2 | 3 | 5 | Likert9 |
| Q12 | 195 | 1 | 1 | 1 | 2 | 4 | Likert10 |
| Q13 | 187 | 1 | 2 | 2 | 3 | 5 | Likert11 |
| Q14 | 190 | 1 | 2 | 3 | 3 | 5 | Likert12 |
| Q15 | 193 | 1 | 1 | 2 | 3 | 5 | Likert13 |
| Q16 | 187 | 1 | 1 | 2 | 2 | 4 | Likert14 |
| Q17 | 193 | 1 | 4 | 4 | 5 | 5 | Likert15 |
| Q18 | 193 | 1 | 3 | 4 | 4 | 5 | Likert16 |
| Q19 | 189 | 1 | 3 | 4 | 4 | 5 | Likert17 |
| Q20 | 192 | 3 | 5 | 5 | 5 | 5 | Likert18 |

### Numeric summary

Then we summarise the numeric scales by their mean, standard deviation, standard error of the mean, confidence interval for the mean, expected value under the null hypothesis that the responses are purely noise, and p-value of testing that null hypothesis.

*This is only a check that people were not neutral in their responses to the questions.* If we put the responses on a scale of 1 to 5 then 3 is a neutral response and we would like to see that people did not just answer 3 on average.

| Ques | Mean | Std.Dev | SE | Lower | Upper | NullExp | pvalue | p\_adj | Description |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Q3 | 4.022 | 0.834 | 0.061 | 3.901 | 4.143 | 3.0 | 0 | 0 | Likert1 |
| Q4 | 1.647 | 0.710 | 0.052 | 1.546 | 1.749 | 2.5 | 0 | 0 | Likert2 |
| Q5 | 1.663 | 0.743 | 0.054 | 1.557 | 1.770 | 2.5 | 0 | 0 | Likert3 |
| Q6 | 1.390 | 0.616 | 0.045 | 1.302 | 1.479 | 2.5 | 0 | 0 | Likert4 |
| Q7 | 4.703 | 0.521 | 0.037 | 4.629 | 4.776 | 3.0 | 0 | 0 | Likert5 |
| Q8 | 4.952 | 0.238 | 0.017 | 4.917 | 4.986 | 3.0 | 0 | 0 | Likert6 |
| Q9 | 3.648 | 0.924 | 0.067 | 3.516 | 3.779 | 3.0 | 0 | 0 | Likert7 |
| Q10 | 2.102 | 0.931 | 0.068 | 1.967 | 2.236 | 2.5 | 0 | 0 | Likert8 |
| Q11 | 2.203 | 0.934 | 0.068 | 2.068 | 2.338 | 3.0 | 0 | 0 | Likert9 |
| Q12 | 1.533 | 0.652 | 0.047 | 1.441 | 1.625 | 2.5 | 0 | 0 | Likert10 |
| Q13 | 2.182 | 0.897 | 0.066 | 2.052 | 2.311 | 3.0 | 0 | 0 | Likert11 |
| Q14 | 2.716 | 1.056 | 0.077 | 2.565 | 2.867 | 3.0 | 0 | 0 | Likert12 |
| Q15 | 2.124 | 0.955 | 0.069 | 1.989 | 2.260 | 3.0 | 0 | 0 | Likert13 |
| Q16 | 1.786 | 0.701 | 0.051 | 1.685 | 1.887 | 2.5 | 0 | 0 | Likert14 |
| Q17 | 4.047 | 0.909 | 0.065 | 3.918 | 4.176 | 3.0 | 0 | 0 | Likert15 |
| Q18 | 3.736 | 0.917 | 0.066 | 3.606 | 3.866 | 3.0 | 0 | 0 | Likert16 |
| Q19 | 3.720 | 0.900 | 0.065 | 3.590 | 3.849 | 3.0 | 0 | 0 | Likert17 |
| Q20 | 4.880 | 0.341 | 0.025 | 4.832 | 4.929 | 3.0 | 0 | 0 | Likert18 |

### Box plots

We try to plot each question group using [box plots](https://en.wikipedia.org/wiki/Box_plot).

*This is a neat visual summary of the responses, showing the variation clearly.*

### Question group 1

### Question group 2

### Question group 3

## Tick Multiple Options Sections

Here we analyse those questions where people could tick multiple options. These must be analysed separately as they must be handled in groups. Note that percentages here should add up to more than 100%.

*This type of question does not lend itself to any statistical analysis and its use is generally discouraged.*

### Programs

| Response | Frequency | Relative.Frequency |
| --- | --- | --- |
| Excel | 79 | 0.16 |
| Other | 77 | 0.16 |
| Outlook | 81 | 0.17 |
| Powerpoint | 83 | 0.17 |
| Teams | 78 | 0.16 |
| Word | 82 | 0.17 |

## Text questions

The text questions are summarised with a crude word cloud. This is not a substitute for a formal thematic analysis, and is mostly for interest.

*Word clouds can be enlightening if differentiated based on a binary explanatory variable.*

### Comments

# Cross tabulations

We test whether the responses to each question are dependent on the response to select key questions.

## Nominal versus Nominal

Here we do tests of independence on the frequency tables, similar to the classical Chi-square test of independence, but using simulated p-values to account for the imbalance in groups.

### Analysis of nominal variables by Gender

#### Age

|  | Female | Male | Other |
| --- | --- | --- | --- |
| 18-25 | 24 | 37 | 1 |
| 26-35 | 33 | 53 | 0 |
| 36+ | 19 | 30 | 3 |

#### Summary

| Ques | raw\_p\_value | adj\_p\_value | Description |
| --- | --- | --- | --- |
| Q2 | 0.236 | 0.236 | Age |

## Scale versus Nominal

Here we would use a Kruskal-Wallis, ANOVA, or similar analysis.

### Analysis of scale variables by Gender

#### Summary

| Ques | raw\_p\_value | adj\_p\_value | Description |
| --- | --- | --- | --- |
| Q3 | 0.987 | 1 | Likert1 |
| Q4 | 0.465 | 1 | Likert2 |
| Q5 | 0.845 | 1 | Likert3 |
| Q6 | 0.811 | 1 | Likert4 |
| Q7 | 0.984 | 1 | Likert5 |
| Q8 | 0.107 | 1 | Likert6 |
| Q9 | 0.412 | 1 | Likert7 |
| Q10 | 0.174 | 1 | Likert8 |
| Q11 | 0.508 | 1 | Likert9 |
| Q12 | 0.673 | 1 | Likert10 |
| Q13 | 0.479 | 1 | Likert11 |
| Q14 | 0.102 | 1 | Likert12 |
| Q15 | 0.624 | 1 | Likert13 |
| Q16 | 0.177 | 1 | Likert14 |
| Q17 | 0.259 | 1 | Likert15 |
| Q18 | 0.329 | 1 | Likert16 |
| Q19 | 0.507 | 1 | Likert17 |
| Q20 | 0.662 | 1 | Likert18 |

### Analysis of scale variables by Age

#### Summary

| Ques | raw\_p\_value | adj\_p\_value | Description |
| --- | --- | --- | --- |
| Q3 | 0.528 | 1.000 | Likert1 |
| Q4 | 0.017 | 0.311 | Likert2 |
| Q5 | 0.324 | 1.000 | Likert3 |
| Q6 | 0.804 | 1.000 | Likert4 |
| Q7 | 0.338 | 1.000 | Likert5 |
| Q8 | 0.902 | 1.000 | Likert6 |
| Q9 | 0.695 | 1.000 | Likert7 |
| Q10 | 0.699 | 1.000 | Likert8 |
| Q11 | 0.039 | 0.665 | Likert9 |
| Q12 | 0.414 | 1.000 | Likert10 |
| Q13 | 0.454 | 1.000 | Likert11 |
| Q14 | 0.556 | 1.000 | Likert12 |
| Q15 | 0.841 | 1.000 | Likert13 |
| Q16 | 0.274 | 1.000 | Likert14 |
| Q17 | 0.691 | 1.000 | Likert15 |
| Q18 | 0.204 | 1.000 | Likert16 |
| Q19 | 0.254 | 1.000 | Likert17 |
| Q20 | 0.700 | 1.000 | Likert18 |

## Correlations

### 1

### 2

### 3

## Congeneric reliability

### 1

| raw\_alpha | std.alpha | G6(smc) | average\_r | S/N | ase | mean | sd | median\_r |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0.283 | 0.277 | 0.26 | 0.06 | 0.383 | 0.076 | 3.391 | 0.408 | 0.055 |

### 2

| raw\_alpha | std.alpha | G6(smc) | average\_r | S/N | ase | mean | sd | median\_r |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0.283 | 0.277 | 0.26 | 0.06 | 0.383 | 0.076 | 3.391 | 0.408 | 0.055 |

### 3

| raw\_alpha | std.alpha | G6(smc) | average\_r | S/N | ase | mean | sd | median\_r |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0.283 | 0.277 | 0.26 | 0.06 | 0.383 | 0.076 | 3.391 | 0.408 | 0.055 |

Reliability calculations assume that all questions are asked in the same direction. Should this not be the case then the researcher should indicate on the metadata spreadsheet (Info) which questions need to be flipped (using a column of indicators) before doing this analysis.

# Index construction

Sometimes it is meaningful to collapse sections into a reduced number of latent dimensions. When a number of responses is collapsed to a single dimension this is often called an index.

*In this survey the groups have low reliability. Thus, we will not attempt to collapse into reduced dimensions, nor analyse such dimensions or underlying factors at this time.*

There are a few methods by which this might be done in general. One approach is to use a factor analysis, another is to use [principal component analysis](https://en.wikipedia.org/wiki/Principal_component_analysis). PCA seeks to find a linear combination of responses that captures as much as possible of the variation in as few dimensions as possible.

Often it makes sense to use an average, or weighted average with weights chosen manually. This is particularly useful when there are weights suggested by theory or external data.

*Sometimes you will see the use of simple sums to construct indices. This should be avoided as even a single missing value, or a single group with a different number of questions, will render all results not interpretable.*

# Referencing the software

In case it is useful, I give the referencing guides from the software itself.

To cite R in publications use:

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

A BibTeX entry for LaTeX users is

@Manual{, title = {R: A Language and Environment for Statistical Computing}, author = {{R Core Team}}, organization = {R Foundation for Statistical Computing}, address = {Vienna, Austria}, year = {2022}, url = {<https://www.R-project.org/>}, }

We have invested a lot of time and effort in creating R, please cite it when using it for data analysis. See also ‘citation(“pkgname”)’ for citing R packages.