



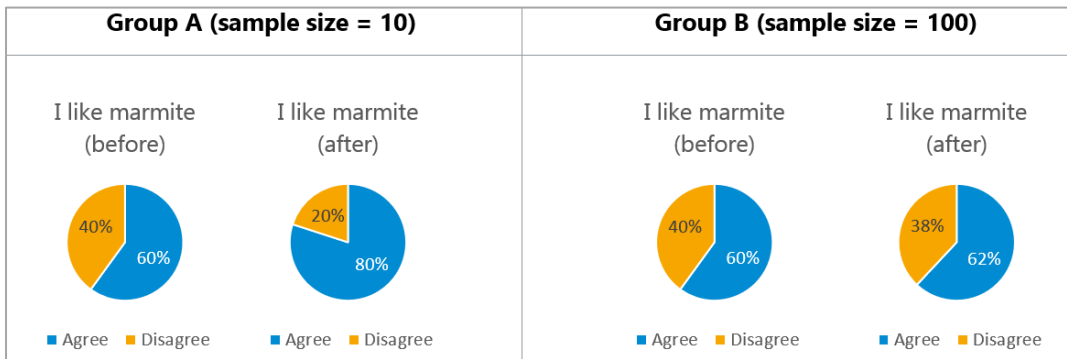
# Collecting, analysing and reporting data best practice

## Avoid average numbers that don't make sense in the context

- One or two musicians were present, not 1.5 musicians
- People attended an average of 3 or 4 times, not 3.6 times

## If your sample size is less than 50, use numbers instead of percentages

- Percentages on small sample sizes can be misleading
- Look what happens to the 'after' pie chart if 2 people change their mind in a small or big group:



## Be careful how you report percentages

- Percentage increase: "There was a 200 percent increase in people feeling optimistic"  
Calculate: the increase as a fraction of the original value:  $20/10=2 \quad 2 \times 100=200\%$
- Percentage points: "The proportion of people feeling optimistic increased by 50 percentage points"  
Calculate: difference between original percentage and increased percentage:  $75\%-25\% = 50\%$

What is being measured? The change in level of your variable or the people agreeing with your variable?

- Level of variable: "on a scale of 1 – 5, the average level of reported optimism increased from 2.2 to 4.1"
- Proportion of people: "the proportion of people feeling optimistic increased from 25% to 75%"

## Use the correct base figure

- Use a full dataset. E.g. only those with start and end data.
- In multi-select, use the number of people who have answered a question, not the number of answers.  
E.g. 20 of the 60 participants played piano; not 20 of the 100 answers.

## Investigate unused data

- Are there similarities in the drop-out group? Look at your user or engagement data to look for patterns. If they have a high/low level of initial optimism are they likely to drop out. If they are on a short sentence/ their age?
- Why do people skip a question? Do they understand it, are your response options correct, is your routing correct (online), is it visible, is it relevant?

## Be mindful of factors that can skew your data

- Response rate: do you have sample size big enough to generalise across the larger population (min.30%)
- Is it representative: are you asking a cross section? Those likely to have positive and negative views.
- External factors: what are you not in control of that may affect your service. E.g. universal credit/ food banks.
- Timing: survey given at awards night may generate an inflated response, survey about cycling administered in Dec likely to give different results than Jul.
- Halo effect: are users telling you what they think you want to hear. Minimise through anonymity, time elapsed.