

Significant Difference User's Guide

Understanding how to determine whether the figures on the SAARF AMPS survey have increased or decreased significantly from survey to survey is an extremely important task as all sample survey results are, unavoidably, subject to a margin of error. Margin of error refers to the level of correspondence between the sample and the universe it represents. For probability samples such as the one used for SAARF AMPS, the margin of error can be calculated. Two factors determine the size of the error – sample size and penetration percentage. When determining whether there is a significant difference between survey results the first step is to calculate the margin of error.

The equation for calculating margin of error is as follows:

$$s = \sqrt{\frac{p \times (100-p)}{n}} \times 1.96$$

s = standard error (margin of error)

p = penetration percentage

n = sample size

1.96 = constant value to translate the margin of error to the 95% confidence level (most market research is based on 95%)

Once the margin of error has been calculated, the interval around the estimated value (penetration percentage) can be calculated in which the true value falls within a confidence level of 95%. The final step is to determine whether the intervals/ranges overlap at all. If there is no overlap between the calculated ranges then they are said to be statistically significantly different.

Example:

The AMPS data table below shows the incidence of people aged 16+ years who own/use/maintain a motor vehicle for AMPS 2004/2005 (Jan-Jun '04 and Jan-Jun '05), AMPS 2005/2006 (Jan-Jun '05 and Jan-Jun '06), and AMPS 2006/2007 (Jan-Jun '06 and Jan-Jun '07) for the total adult population.

		AMPS 2004/2005	AMPS 2005/2006	AMPS 2006/2007
		Total	Total	Total
Total	Audience(000)	30,656	30,903	31,109
	Resps	24,407	24,813	24,812
	%Col	100	100	100
	%Row	100	100	100
Own Use Maintain - Personal Vehicle	Audience(000)	4,796	4,954	5,212
	Resps	6,841	7,092	7,151
	%Col	15.6	16	16.8
	%Row	100	100	100

The figure framed in red is the sample size or *n*.

The figure framed in blue is the penetration percentage or *p*.

If you are not working with the total adult population but with a specific demographic for example Gauteng province, you would use the sample size and penetration percentage for the specific demographic you are working with (see table below).

		AMPS 2004/2005 - Main Total	AMPS 2005/2006 - Main Branded Total	AMPS 2006/2007 - Main Branded Total	AMPS 2004/2005 - Main Gauteng	AMPS 2005/2006 - Main Branded Gauteng	AMPS 2006/2007 - Main Branded Gauteng
Total	Audience(000)	30,656	30,903	31,109	6,043	6,360	6,402
	Resps	24,407	24,813	24,812	6,004	6,349	6,336
	%Col	100	100	100	100	100	100
	%Row	100	100	100	19.7	20.6	20.6
Own Use Maintain - Personal Vehicle	Audience(000)	4,796	4,954	5,212	1,672	1,765	1,864
	Resps	6,841	7,092	7,151	2,075	2,270	2,303
	%Col	15.6	16	16.8	27.7	27.8	29.1
	%Row	100	100	100	34.9	35.6	35.8

AMPS 2004/2005 (Total Adults)

15.6% is the penetration (p) for this release.

24,407 is the sample size (n) for this release.

So then, using the formula to calculate the margin of error:

$$s = \sqrt{\frac{p \times (100-p)}{n}} \times 1.96$$

$$s = \sqrt{\frac{15.6 \times (100-15.6)}{24,407}} \times 1.96$$

$$s = \sqrt{\frac{15.6 \times (84.4)}{24,407}} \times 1.96$$

$$s = \sqrt{\frac{1316.64}{24,407}} \times 1.96$$

$$s = \sqrt{0.05394517966} \times 1.96$$

$$s = 0.23226101622 \times 1.96$$

$$s = 0.45523159179$$

$$s = \underline{0.46} \text{ (rounded)}$$

In order to calculate the 95% confidence range, take the penetration (15.6%) and subtract and add the standard error (margin of error) (0.46%).

$$15.6\% - 0.46\% = 15.14\%$$

$$15.6\% + 0.46\% = 16.06\%$$

What this means is that at a 95% confidence level, the incidence of people aged 16+ years that own/use/maintain a motor vehicle for AMPS 2004/2005 is within the range of 15.14% to 16.06%.

AMPS 2005/2006 (Total Adults)

16% is the penetration (p) for this release.

24,813 is the sample size (n) for this release.

So then, using the formula to calculate the margin of error:

$$s = \sqrt{\frac{p \times (100-p)}{n}} \times 1.96$$

$$s = \sqrt{\frac{16 \times (100-16)}{24,813}} \times 1.96$$

$$s = \sqrt{\frac{16 \times (84)}{24,813}} \times 1.96$$

$$s = \sqrt{\frac{1344}{24,813}} \times 1.96$$

$$s = \sqrt{0.05416515536} \times 1.96$$

$$s = 0.23273408723 \times 1.96$$

$$s = 0.45615881097$$

$$s = \underline{0.46 \text{ (rounded)}}$$

In order to calculate the 95% confidence range, take the penetration (16%) and subtract and add the standard error (margin of error) (0.46%).

$$16\% - 0.46\% = 15.54\%$$

$$16\% + 0.46\% = 16.46\%$$

What this means is that at a 95% confidence level, the incidence of people aged 16+ years that own/use/maintain a motor vehicle for AMPS 2005/2006 is within the range of 15.54% to 16.46%.

AMPS 2006/2007 (Total Adults)

16.8% is the penetration (p) for this release.

24,812 is the sample size (n) for this release.

So then, using the formula to calculate the margin of error:

$$s = \sqrt{\frac{p \times (100-p)}{n}} \times 1.96$$

$$s = \sqrt{\frac{16.8 \times (100-16.8)}{24,812}} \times 1.96$$

$$s = \sqrt{\frac{16.8 \times (83.2)}{24,812}} \times 1.96$$

$$s = \sqrt{\frac{1397.76}{24,812}} \times 1.96$$

$$s = \sqrt{0.05633403192} \times 1.96$$

$$s = 0.23734791324 \times 1.96$$

$$s = 0.46520190995$$

$$s = \underline{0.47} \text{ (rounded)}$$

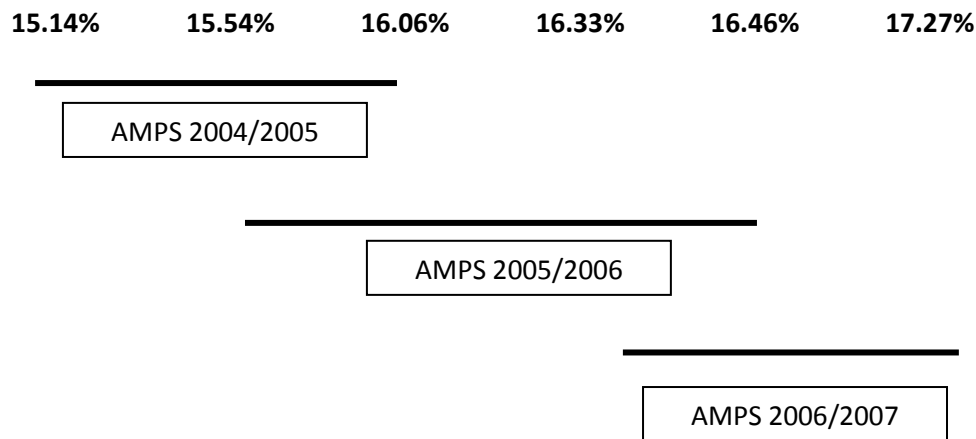
In order to calculate the 95% confidence range, take the penetration (16.8%) and subtract and add the standard error (margin of error) (0.47%).

$$16.8\% - 0.47\% = 16.33\%$$

$$16.8\% + 0.47\% = 17.27\%$$

What this means is that at a 95% confidence level, the incidence of people aged 16+ years who own/use/maintain a motor vehicle for AMPS 2006/2007 is within the range of 16.33% to 17.27%.

Represented diagrammatically, the above calculated ranges are as follows:



In the above diagram, the ranges for AMPS 2004/2005 and AMPS 2005/2006 overlap. This means that there was not a statistically significant difference for the incidence of people aged 16+ years who own/use/maintain a motor vehicle between AMPS 2004/2005 and AMPS 2005/2006.

Similarly, the ranges for AMPS 2005/2006 and AMPS 2006/2007 overlap. This means that there was not a statistically significant difference for the incidence of people aged 16+ years who own/use/maintain a motor vehicle between AMPS 2005/2006 and AMPS 2006/2007.

The ranges for AMPS 2004/2005 and AMPS 2006/2007, however, do not overlap. This means that the incidence of people aged 16+ years that own/use/maintain a motor vehicle between AMPS for AMPS 2006/2007 has significantly increased from AMPS 2004/2005.

For the purposes of this example, the range has been calculated in percentages and not in thousands. The range can be calculated in thousands by simply multiplying the calculated range percentages by the universe figure. For the purposes of comparison however, it is recommended to rather use the range percentage. The reason for this is that for AMPS the universe size is updated each year in accordance with the changes in the South African population. If you look at the above example for instance, the universe (audience or population) for AMPS 2004/2005 is 30,656,000; the universe for AMPS 2005/2006 is 30,903,000; and the universe for AMPS 2006/2007 is 31,109,000. This means that if the thousands are used the baseline will vary from year to year and as such the thousands are not directly comparable. However, because the percentages consistently have a baseline of 100 this difficulty can be avoided.