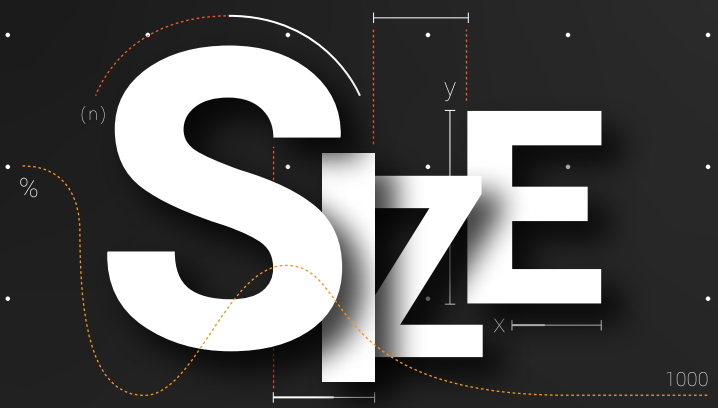
SAMPLE



A Guide to Determining The Right Sample Size for Your Study

AHZUL
BE INSPIRED.

Written by
Rebecca Abravanel, PHD
Head of Quantitative Studio

Clients *often ask us why we recommend the sample sizes we do for quantitative research,*

SO we've developed a short guide to choosing the right sample size based on your audiences, objectives, and needs.

Sample size makes a huge difference in sampling error when sample sizes are small—under 200, for example. Once you have a sample size of 1,000, however, there is relatively little accuracy gained by increasing the sample size. At the 95% level of confidence, a sample size of 1,000 has a sample error of ±3.1%. You can get the sample error down to about 2% by doubling (±2.2%) or tripling (±1.8%) the sample size, but usually the gain in accuracy is not deemed worth the increase in field costs.

AS A RESULT, AT AHZUL WE TYPICALLY RECOMMEND A MINIMUM TOTAL SAMPLE SIZE OF 1,000.

Table 1

Sample Size & Sample Error At the 95% Level of Confidence

Sample Size (n)	Sample Error
100	± 9.8%
200	± 6.9%
500	± 4.4%
1,000	± 3.1%
1,500	± 2.5%
2,000	± 2.2%
3,000	± 1.8%
4,000	± 1.5%
5,000	± 1.4%
10,000	± 1.0%

Once *you have a*

sample size of at least 1,000, the other factors that are just as important, if not more important, to consider are:

1 **What types of respondent groups do you want to be able to analyze separately (if any)? (Once you begin to divide up the sample, subgroup size can get small quickly, even with an n of 1,000.)**

Minimum quotas or sample boosts can be used to increase the number of respondents in key subgroups, but statistical weighting must be used correctly on the back end to ensure that your quotas don't warp your results. For example, in many US government surveys, Asian Americans are oversampled to provide a large enough sample to analyze them separately, but if statistical weighting isn't used on the back end to reduce the impact of the oversample, the results can be skewed (by making the impact of this subgroup greater than it actually is). Another relevant example might be including a minimum quota for your own customers or those of a key competitor. Such groups might be important to be able to analyze separately but estimates of your brand's and/or your key competitor's awareness would be inaccurate if statistical weighting were not used on the back end to adjust for the oversampling/minimum quota.

©2020 AHZUL, ALL RIGHTS RESERVED, PRIVILEGED & CONFIDENTIAL

At Ahzul we use a combination of government data and incidence data from the study to create the statistical weights. This is critical because using government data only will undercut the value of your proprietary study and will make your survey results reflective of the general population even though your target population might be different. For example, say your current customers tend to be relatively young and the research vendor weights the data to the US population age distribution -- your results will not be reflective of your target population because the US population is relatively old.

That is why the statistical weights must account for both qualifying incidence and population distribution. It takes a fair amount of expertise to weight data correctly; **many vendors will assume that population data are sufficient for calculating statistical weights, or may not structure the survey screener appropriately to gather the information needed to calculate weights accurately.**

2 IS YOUR SAMPLE SOURCE LIKELY TO BE REPRESENTATIVE OF YOUR TARGET POPULATION?

Most market research studies are conducted using online panels. With very high online penetration in the US, online sampling is not typically a problem for most product/service categories, but there are exceptions. For example, if you wanted to study welfare recipients, using an online sample source would be skewed, since poorer Americans are systematically less likely to have Internet access. Or, if you are trying to accurately sample some ethnic groups, online sampling should be supplemented with an in-person interview methodology (CAPI intercepts) because the overwhelming majority of online panels are managed only in English, and those who are less comfortable in English are less likely to be panelists.

Summing It Up

- **IF YOU WANT YOUR STUDY TO HAVE LOW SAMPLE ERROR, AIM FOR A STUDY N OF AT LEAST 1,000.**

  If you can afford to do more, then increase the sample, but remember that at least in terms of sample accuracy, there is a diminishing return to increasing sample size (especially after n=1000).

- **WHEN DECIDING ON A SAMPLE SIZE, MAKE SURE TO TAKE IN THE SIZE OF KEY BREAKOUT GROUPS (CUSTOMERS, OR CERTAIN DEMOGRAPHIC GROUPS, ETC.) YOU WANT IN THE ANALYSIS.**

  If you want the data to be broken out by subgroups, you need to make sure you have enough sample within each subgroup to obtain a read, and this will impact your total sample size.

- **MOST IMPORTANTLY, MAKE SURE THAT YOUR STUDY EMPLOYS A METHODOLOGY THAT ALLOWS YOU TO REACH A REPRESENTATIVE SAMPLE OF THE TARGET POPULATION YOU HOPE TO LEARN ABOUT.**

Still hungry **for more?**

Turn the page for an even deeper dive into determining sample size by confidence intervals and sampling error.

©2020 AHZUL. ALL RIGHTS RESERVED, PRIVILEGED & CONFIDENTIAL

An Even Deeper Dive on *Sample Size*

The appropriate size for a sample depends upon 2 factors:

1 **THE LEVEL OF SAMPLING ERROR YOU'RE COMFORTABLE WITH (±1%, ±3%, ±5%, ETC.)**

This is the estimated confidence interval that applies to a specific result, assuming maximum variability. For example, let's say the specific result of concern is the percent aware of your brand. Perhaps you're comfortable with brand awareness being accurate within 3% (either 3% above the true number, or 3% below the true number) but not with a potential error of ±5%.

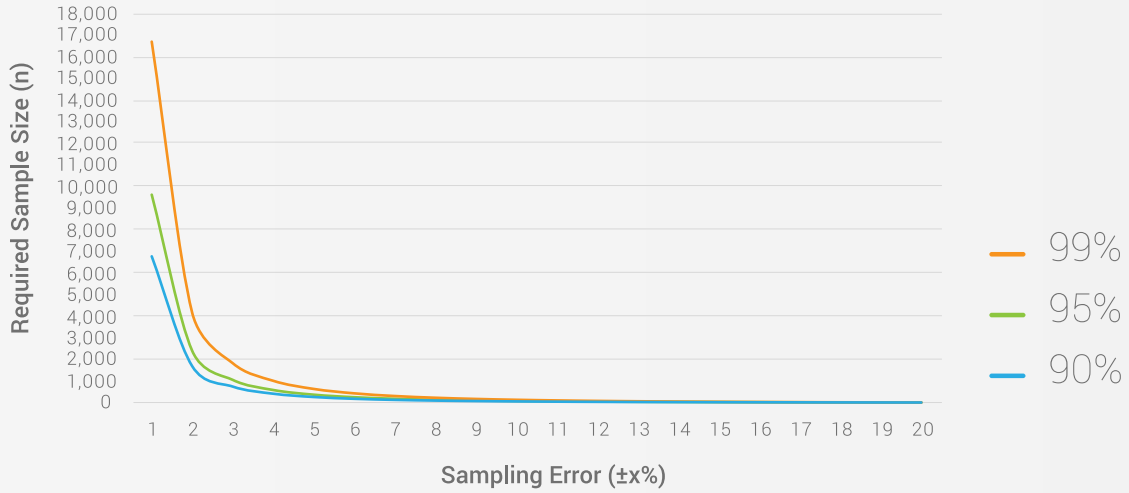
2 **THE LEVEL OF CONFIDENCE YOU'RE COMFORTABLE WITH (USUALLY 90%, 95%, OR 99% ARE USED)**

The level of confidence is an indicator of how frequently your results would fall into the given confidence interval (in #1) were you to repeat the study many times. For example, you might be comfortable with being able to have your estimates of awareness fall into a certain confidence interval 95% of the time (but not 90% of the time). 95% is the level of confidence used in most market research (and many other) studies.

The diminishing return in accuracy gained by increasing sample can be seen in Figure 1 below.

Figure 1

Sampling Error & Required Sample Size at 3 Levels of Confidence



The numbers shown in Figure 1 are provided in Table 2 on the next page, and it's clear that there is relatively little reduction in sampling error above a sample size of 1,000. At the 95% level of confidence, for example, the sampling error with a sample of 1,000 is ±3.1%, at n=1,500 it's ±2.5%, and at n=2,000 it's ±2.2%.

©2020 AHZUL. ALL RIGHTS RESERVED, PRIVILEGED & CONFIDENTIAL

Table 2

Sampling Error & Corresponding Sample Size at 3 Levels of Confidence

Sample Size (n)	Level of Confidence —99%	—95%	—90%
5	57.7	43.8	36.9
10	40.8	31.0	26.1
25	25.8	19.6	16.5
50	18.2	13.9	11.7
75	14.9	11.3	9.5
100	12.9	9.8	8.3
150	10.5	8.0	6.7
200	9.1	6.9	5.8
300	7.4	5.7	4.8
400	6.5	4.9	4.1
500	5.8	4.4	3.7
600	5.3	4.0	3.4
700	4.9	3.7	3.1
800	4.6	3.5	2.9
900	4.3	3.3	2.8
1,000	4.1	3.1	2.6
1,500	3.3	2.5	2.1
2,000	2.9	2.2	1.8
3,000	2.4	1.8	1.5
4,000	2.0	1.5	1.3
5,000	1.8	1.4	1.2
10,000	1.3	1.0	0.8

Table 3

is a **Transposed**

VERSION OF TABLE 1 –SHOWING THE SAMPLE SIZE (N) NEEDED FOR A GIVEN SAMPLING ERROR.

x

y

- For a sampling error of ±3%, for example, a sample of 1,067 is required.
  - In political polls, the sampling error is often ±3%, so most political polls use a sample size of 1,100 (rounding up from 1,067).
- The table allows you to look up the required sample size needed at a given sample error within the desired confidence interval.

Table 3

Required Sample Size for A Given Level of Sampling Error At 3 Levels of Confidence

Sampling Error (±x%)	Level of Confidence —99%	—95%	—90%
1	16,641	9,604	6,806
2	4,160	2,401	1,702
3	1,849	1,067	756
4	1,040	600	425
5	666	384	272
6	462	267	189
7	340	196	139
8	260	150	106
9	205	119	84
10	166	96	68
11	138	79	56
12	116	67	47
13	98	57	40
14	85	49	35
15	74	43	30
16	65	38	27
17	58	33	24
18	51	30	21
19	46	27	19
20	42	24	17

©2020 AHZUL. ALL RIGHTS RESERVED, PRIVILEGED & CONFIDENTIAL

Written by
Rebecca Abravanel, PHD
Head of Quantitative Studio

Contact
Miguel Gomez Winnebrenner
Principal
312.860.3191
miguel@ahzul.com

www.ahzul.com

©2020 AHZUL, ALL RIGHTS RESERVED, PRIVILEGED & CONFIDENTIAL