

TIP SHEET: DETERMINING YOUR SAMPLE SIZE

GOALS

This will help you to:

- ✓ *Identify your population.*
- ✓ *Determine your desired boundaries for the analyses.*
- ✓ *Estimate the number of people needed for your assessment.*

You know what question you want to study. You know how you want to conduct the study. Now the only thing standing between you and starting your research is figuring out who and how many people need to actually be involved in your research. In order to do that, there are **three (3) things** you should keep in mind:

- 1) What population am I studying?
- 2) What level of precision can I accept for my analyses/results?
- 3) How do I know how many people I need?

Identify Your Population

Let's begin at the beginning. By now, you probably know a little bit about the people that you would like to assess, especially as you know what it is that you want to examine and how you are going to examine it. The most important things to know about your population at this point can be summed up as:

- ❖ **Who exactly are they?**
- ❖ **How many of them are available?**

Who are they? This seems easy enough, but there are a few things that need to be considered. For example, if you want to conduct an assessment of a course, will your population be everyone who enrolls in that course during the Fall or over the entire academic year? Are you interested in assessing certain students (only males, students at program benchmarks, potential term graduates, etc.)? A few factors that should be considered in determining your correct population may include:

- ❖ How long do I plan to administer the assessment?
- ❖ Who am I interested in studying?
- ❖ Are there particular participant characteristics I am interested in?

Once you can answer these questions, it's time to move on to the next question.

How many of them are available? To answer this, there are resources available. One such resource is the IR data packages (<https://www.mdc.edu/ir/>) which provide useful information by school and particular areas of interest (e.g. Honors College students, Aviation, ENC1101 students, etc.).



Institutional Research

TIP SHEET: DETERMINING YOUR SAMPLE SIZE

Determine Desired Boundaries for Analysis

Now that you know who your population is that you are targeting and how many individuals there are, you then need to figure out three “boundaries” related to the analysis:

- ❖ **Margin of Error**
- ❖ **Confidence Level**
- ❖ **Variability**

All three of these will affect the sample size that is needed; typically, ***the higher the boundary is set, the larger the sample size*** that will be needed.

Margin of Error: This refers to how closely the results obtained from the sample will predict the true values in the population. Typically, $\pm 5\%$ margin of error is used. This means that if our results show that 60% of students use cell phones to check email, a 5% margin of error means that the true value lies between 55% and 65% of students from the population.

Confidence Level: This refers to how close the results from your sample approximate the normal distribution found in the population. Most people use a 95% confidence interval (CI) meaning that if the population was sampled 100 times, 95% of these samples would have a value within the margin of error selected (see above).

Variability: The degree to which the concepts or attributes being assessed are found within the population is defined as the variability. For example, if you wanted to measure student use of cell phones to access email, (using the above exemplar) we would define the variability as 60% (students using cell phone)-40% (students not using cell phones). If the variability is extremely high (e.g. 90%-10%), a larger sample size may be needed to get an accurate result.

Select Sample

Once you have set the boundaries for the analysis and you know the population that you will be assessing, the last step is to combine these things and select your sample. There are numerous ways to do so. You can conduct an online search for a sample size calculator or use this sample size calculator (<http://www.raosoft.com/samplesize.html>) which incorporates many of the aspects discussed.

Alternatively, you can look for tables that provide an estimate based on the boundaries selected. Most of these tables estimate a base sample size; when using a table for base sample size estimation, you will need to consider the response/retention rate of your sample to calculate a final sample size (see attached table).

TIP SHEET: DETERMINING YOUR SAMPLE SIZE

Base Sample Size Table*

		Confidence=	
		95.0%	
Population Size	Degree of Accuracy/ Margin of Error	Population Size	Degree of Accuracy/ Margin of Error
		5.0%	
10	10	7,500	365
20	19	8,500	368
30	28	10,000	370
50	44	11,000	371
75	63	12,000	372
100	80	13,000	373
150	108	14,000	374
200	132	15,000	375
250	152	16,000	375
300	169	17,000	376
400	196	18,000	376
500	217	19,000	377
600	234	20,000	377
700	248	25,000	378
800	260	30,000	379
900	269	35,000	380
1,000	278	40,000	381
1,200	291	45,000	381
1,500	306	50,000	381
2,000	322	75,000	382
2,500	333	100,000	383
3,500	346	125,000	383
4,000	351	150,000	384
5,000	357	175,000	384
6,000	361	200,000	384

*Modified from Krejcie, R.V. and Morgan, D.W. (1970). *Determining sample size for research activities*. Educational and Psychological Measurement, 30, 607-610.

- Note:**
- 1) This table assumes a confidence level of 95% and a variability of 50% (providing the most conservative estimates for sample size).
 - 2) For populations of 100 or less, the whole population should be used as normality assumptions are weaker.
 - 3) To calculate the needed sample size, use the base estimate from the table and multiple by the level of response, i.e. an 80% response rate for a population of 500 students would require a sample of $217 / 0.8 \approx 271$.



Institutional Research

TIP SHEET: DETERMINING YOUR SAMPLE SIZE

Finally, the following equation can be used to compute the final sample size needed:

$$n = \frac{\left(\frac{P [1 - P]}{A^2} + \frac{P [1 - P]}{N} \right)}{R}$$

Where:

- n = sample size required
- N = Number of people in population
- P = Estimated variability in population (use as a decimal and use smaller variability, i.e. from example 0.4 for non-cell phone email user students)
- A = Margin of error (use as decimal – 0.05 for 5%, 0.03 for 3%, etc.)
- Z = Confidence level (z-scores; use **1.96** for 95% confidence)
- R = Estimated retention/response rate (use as a decimal)