If samples are sufficiently large (n>30) and random, resulting distribution of sample means will be approximately normal based on the Central Limit Theorem.

$$S_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

The Standard error of the mean (standard deviation of the sample means) is equal to the population standard deviation divided by the square root of the sample size.

A confidence interval can be created based on the standard error which suggests the likelihood of capturing the population mean.

$$\bar{X} - \sigma_{\bar{X}} \leq \mu \leq \bar{X} + \sigma_{\bar{X}}$$

Confidence Intervals are based on the distribution of the standard normal curve, such that the following percent of the population could be found within the specified *z* scores:

68%	1.00
90%	1.64
95%	1.96
99%	2.58

Restructuring the formula for the standard error provides the first step to securing the required sample size:

$$\sqrt{n} = \frac{\sigma}{S_{\bar{X}}} \quad \text{where} \quad n = \frac{\sigma^2}{S_{\bar{X}}^2}$$

The standard error is a measure of error in the estimate – leading to our potential for controlling the resulting error – How much error are you willing to tolerate. This is also an opportunity to secure a certain level of confidence, in the context of being able to estimate a confidence interval:

$$n = \frac{z^2 \, S^2}{e^2} \quad \text{or for dichotomous variables:} \quad n = \frac{z^2 \, p(1-p)}{e^2}$$

Population size is only a concern when the sample size is large relative to the population size. Adjustments are recommended when the estimated sample size is more than 5% of the total population, since the assumption of independence is no longer tenable.

A finite-population correction can be applied:

$$n' = \frac{nN}{N + n - 1}$$