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

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.

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

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

$$\bar{X} - z_{\frac{\alpha}{2}} SE(\bar{X}) \leq \mu \leq \bar{X} + z_{\frac{\alpha}{2}} SE(\bar{X})$$

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:

Confidence	→	***z***
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. Here we use *S* as the sample estimate of the population standard deviation:

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

The standard error is a measure of sampling error in the estimate – useful in controlling the resulting error. How much error are you willing to tolerate? We can also specify a certain level of confidence (based on the *z*-score associated with levels of confidence as above), to create a confidence interval for our sample mean (using *S*):

$$n = \frac{S^2}{S_{\bar{X}}^2} = \frac{S^2}{SE(\bar{X})^2} = \frac{S^2}{e^2}$$

$$n = \frac{z^2 S^2}{e^2}$$

or for dichotomous variables where we estimate proportions (using *p*):

$$n = \frac{z^2 \; p(1-p)}{e^2}$$

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:

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