HOWEVER, the formulae we use for a sample and a population are different!

Most researchers almost never work with population data.

We use samples to infer the population's variance.

	Population	*Estimate Based on a Sample*
Variance	$\sigma^2 = \frac{\Sigma(X-\mu)^2}{N}$ where Σ = to sum X = a score in the distribution μ = the population mean N = the number of cases in the population	$s^2 = \frac{\Sigma(X-\overline{X})^2}{n-1}$ where Σ = to sum X = a score in the distribution $\overline{X}$ = the sample mean n = the number of cases in the sample
Standard Deviation	$\sigma = \sqrt{\frac{\Sigma(X-\mu)^2}{N}}$ where Σ = to sum X = a score in the distribution μ = the population mean N = the number of cases in the Population	$s = \sqrt{\frac{\Sigma(X-\overline{X})^2}{n-1}}$ where Σ = to sum X = a score in the distribution $\overline{X}$ = the sample mean n = the number of cases in the sample

HOWEVER, the formulae we use for a sample and a population are different!

Sample variance tends to be lower than the **population variance** that it purportedly estimates (i.e., undervalues the penalty).

Thus, we divide by ***n-1*** (***n*** is the data in sample) instead of ***n***.

A smaller number in the denominator yields a higher variance.

	Population	*Estimate Based on a Sample*
Variance	$\sigma^2 = \frac{\Sigma(X-\mu)^2}{N}$	$s^2 = \frac{\Sigma(X-\overline{X})^2}{n-1}$
	where Σ = to sum X = a score in the distribution μ = the population mean N = the number of cases in the population	where Σ = to sum X = a score in the distribution $\overline{X}$ = the sample mean n = the number of cases in the sample
Standard Deviation	$\sigma = \sqrt{\frac{\Sigma(X-\mu)^2}{N}}$	$s = \sqrt{\frac{\Sigma(X-\overline{X})^2}{n-1}}$
	where Σ = to sum X = a score in the distribution μ = the population mean N = the number of cases in the Population	where Σ = to sum X = a score in the distribution $\overline{X}$ = the sample mean n = the number of cases in the sample

	Population	Sample
Standard Deviation	$\sigma = \sqrt{\frac{\Sigma(X-\mu)^2}{N}}$	$s = \sqrt{\frac{\Sigma(X-\overline{X})^2}{n-1}}$
	where Σ = to sum X = a score in the distribution μ = the population mean N = the number of cases in the population	where Σ = to sum X = a score in the distribution $\overline{X}$ = the sample mean n = the number of cases in the sample

Thus, we divide by **n-1** (**n** is the data in sample) instead of **n**.

A smaller number in the denominator yields a higher variance.