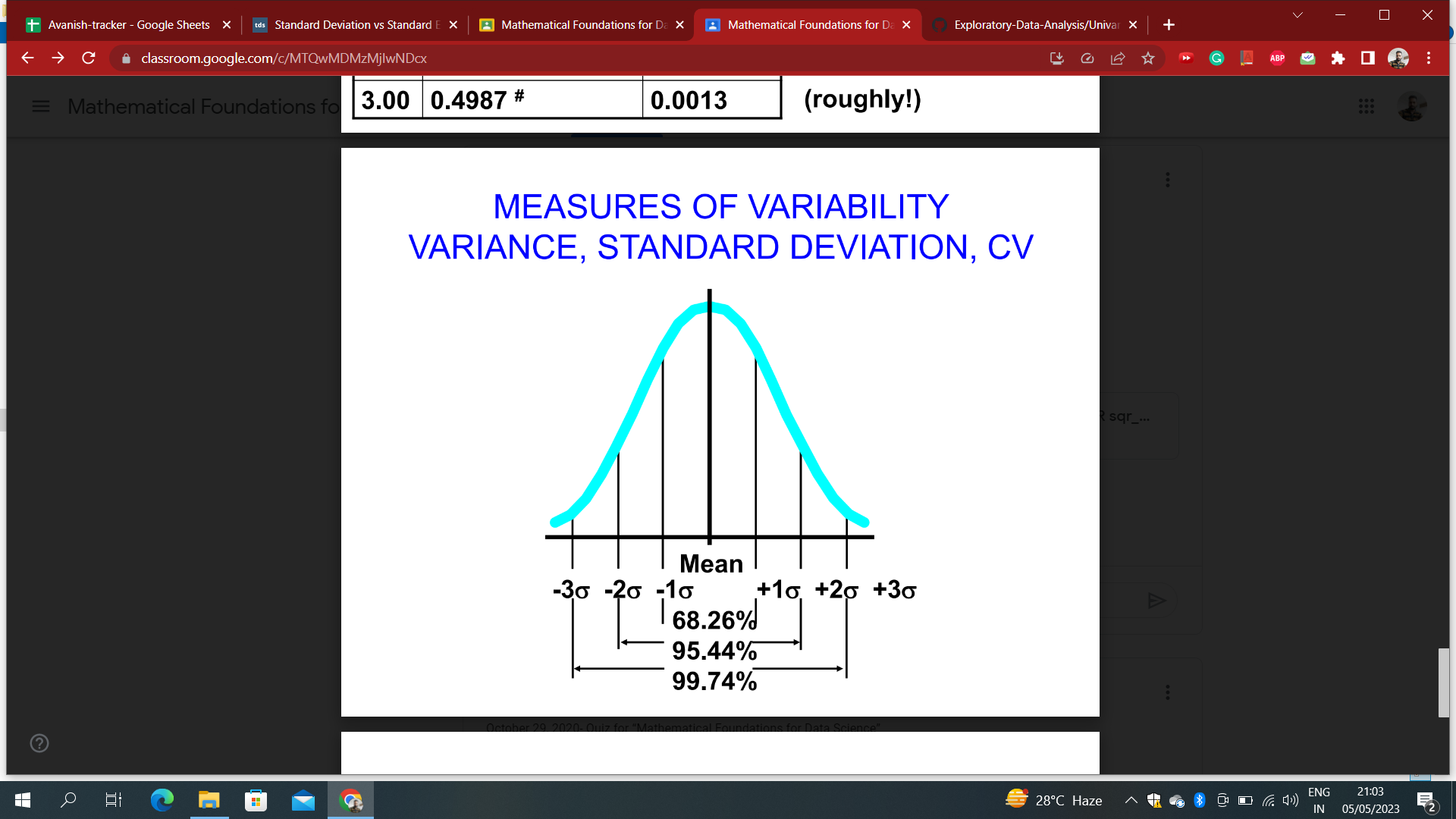
**Standard Deviation and Central Limit Theorem**

The **standard deviation** measures the **variability**(aka, the **spread**) of data points around the **mean**in a given dataset. In other words, it tells us, on average, how far each data point is away from the mean.

**Interpretation:**



Example: suppose that the final marks has a bell-shaped distribution, with a mean of 75 and a standard deviation of 7. Then,

• approximately 68% marks fall between (75-7) = 68 and (75+7) = 82.

• approximately 95% marks fall between (75-2\*7) = 61 and (75+2\*7) = 89, and

• virtually all the measurements fall between (75-3\*7) = 54 and (75+3\*7) = 96

**Calculation:**

 Compute the deviation (distance from the mean or difference between the score and the mean) for each score.

 Square each deviation.

 Compute the mean of the squared deviations.

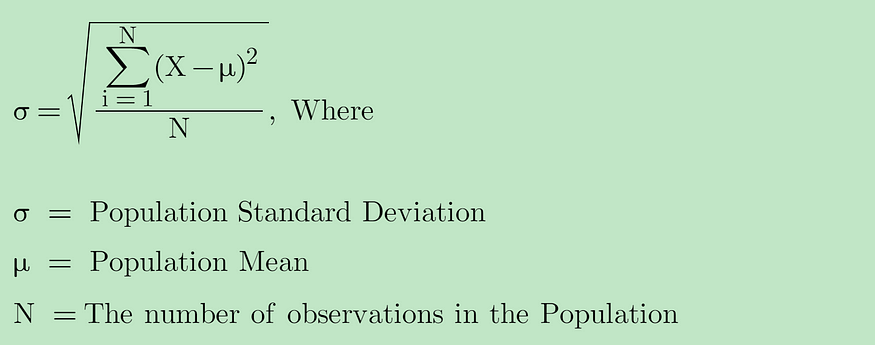
 Take the square root.

 Finally, take the square root of the variance to obtain the standard deviation.

## **Population Standard Deviation:**

In the real world, we’re interested in estimating a certain characteristic in a **population**. Standard deviation is anexample of these characteristics.

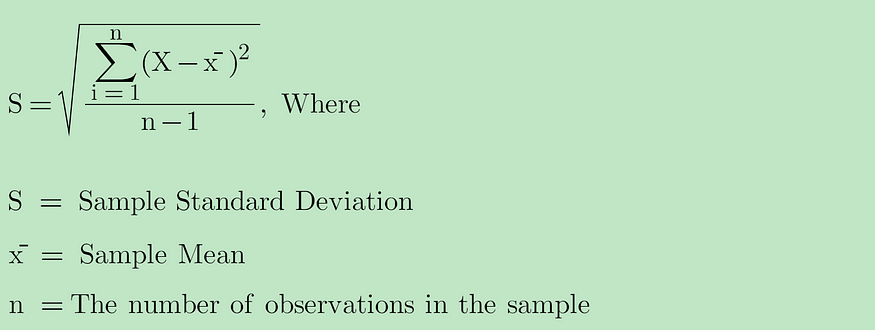
When you have **ALL** the data points from a population, you can compute the **TRUE**value of the population standard deviation using the following formula



## **Sample Standard Deviation:**

Oftentimes, it is difficult to collect all the data points from the population due to time, financial, or technical limitations. For example, if we would like to compute the **TRUE** standard deviation of household income in Los Angeles, we would need to get income from all the households in Los Angeles, which is almost impossible to do.

Instead, we can collect random samples from the population and make inferences about the population standard deviation using **Sample Standard Deviation.**The formula for sample standard deviation is



## **Why use n-1 for sample standard deviation?**

You will notice that we are using the sample mean (x̄) instead of the population mean (μ) for the sample standard deviation because we don’t know anything about the population mean. x̄ is a reasonable estimate for μ. Since the sample mean is based on the data, it will get drawn toward the center of mass for the data. Therefore, any value X in the sample dataset would be closer to x̄ than to μ. The numerator in the sample standard deviation would get artificially smaller than it is supposed to be. As a result, the sample standard deviation would be **underestimated**. To correct this **bias**in the sample standard deviation, we would use **“n-1” instead of “n” (**aka, **Bessel’s correction**) for sample standard deviation**.** Using n-1 would make the sample standard deviation larger than otherwise using n. Therefore, we have a less biased estimate of the population standard deviation, giving us a conservative estimate of variability.

**Note**: The **variance, σ2** is equal to the average squared deviation from the mean, μ.

**Central Limit Theorem:**

The Central Limit Theorem states that the [**sampling distribution**](https://www.statisticshowto.com/probability-and-statistics/sampling-in-statistics/sampling-distribution/)**of the sample means** approaches a normal distribution as the sample size gets larger — no matter what the shape of the *population* distribution. This fact holds especially true for sample sizes over 30.

All this is saying is that as you take more *samples*, especially large ones, your graph of the *sample means* will look more like a normal distribution.

As a rough estimate, the Central Limit Theorem predicts a roughly normal distribution under any of the following conditions:

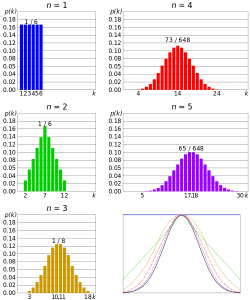
• The population distribution is normal; or

• The sampling distribution is symmetric and the sample size is ≤15; or

• The sampling distribution is moderately skewed and the sample size is 16 ≤ n ≤ 30; or

• The sample size is greater than 30, without outliers.

Here’s what the Central Limit Theorem is saying, graphically. The picture below shows one of the simplest types of test: rolling a fair die. The **more times you roll the die**, the more likely the shape of the distribution of the means tends to look like a**normal distribution graph**.

[](https://www.statisticshowto.com/probability-and-statistics/normal-distributions/central-limit-theorem-definition-examples/)

**The Central Limit Theorem and Means**

An essential component of the Central Limit Theorem is that the **average of your sample means will be the population mean**. In other words, add up the means from all of your samples, find the average and that average will be your actual population mean. Similarly, if you find the average of all of the standard deviations in your sample, you’ll find the actual standard deviation for your population. It’s a pretty useful phenomenon that can help accurately predict characteristics of a population.