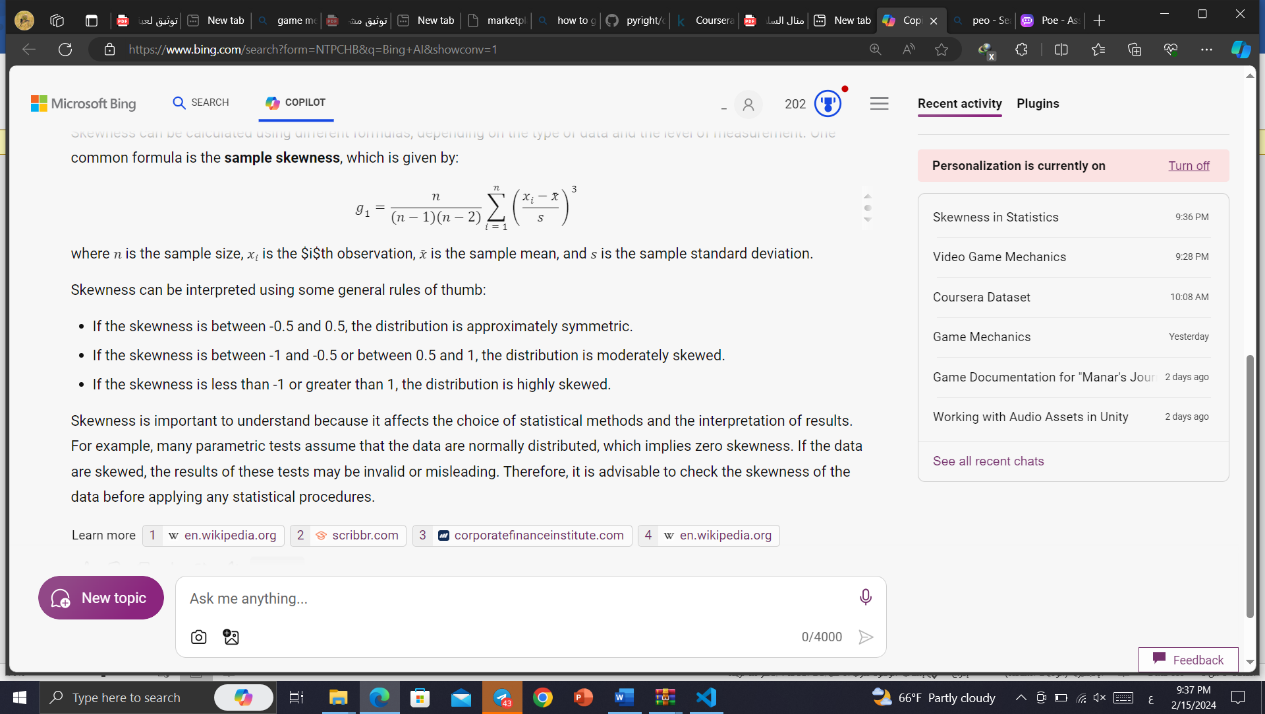
**Skewness** is a measure of how asymmetric a distribution is. It indicates how much the distribution deviates from a normal distribution, which is symmetric and has no skew. Skewness can be positive, negative, or zero.

* **Positive skew** means that the right tail of the distribution is longer than the left tail, and the mean is greater than the median. For example, income distribution is usually positively skewed, as there are a few people with very high incomes that pull the mean up.
* **Negative skew** means that the left tail of the distribution is longer than the right tail, and the mean is less than the median. For example, exam scores are often negatively skewed, as there are a few students with very low scores that pull the mean down.
* **Zero skew** means that the distribution is symmetric, and the mean and median are equal. For example, a normal distribution has zero skew, as both tails are equally balanced.

Skewness can be calculated using different formulas, depending on the type of data and the level of measurement. One common formula is the **sample skewness**, which is given by:

where n is the sample size, xi​ is the $i$th observation, xˉ is the sample mean, and s is the sample standard deviation.

Skewness can be interpreted using some general rules of thumb:

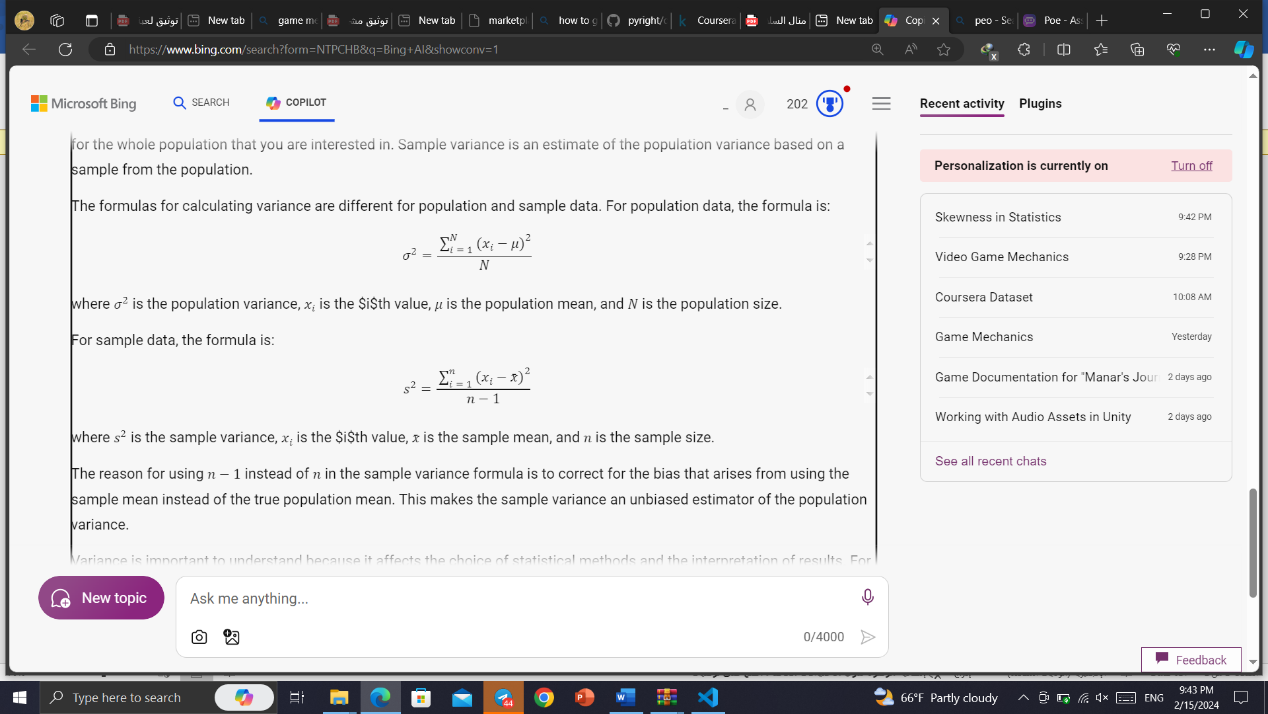
* If the skewness is between -0.5 and 0.5, the distribution is approximately symmetric.
* If the skewness is between -1 and -0.5 or between 0.5 and 1, the distribution is moderately skewed.
* If the skewness is less than -1 or greater than 1, the distribution is highly skewed.

Skewness is important to understand because it affects the choice of statistical methods and the interpretation of results. For example, many parametric tests assume that the data are normally distributed, which implies zero skewness. If the data are skewed, the results of these tests may be invalid or misleading. Therefore, it is advisable to check the skewness of the data before applying any statistical procedures.

**Variance** is a measure of how much the data values differ from the mean. It is calculated by taking the average of squared differences between each value and the mean. Variance tells you the degree of spread in your data set. The more spread the data, the larger the variance is in relation to the mean.

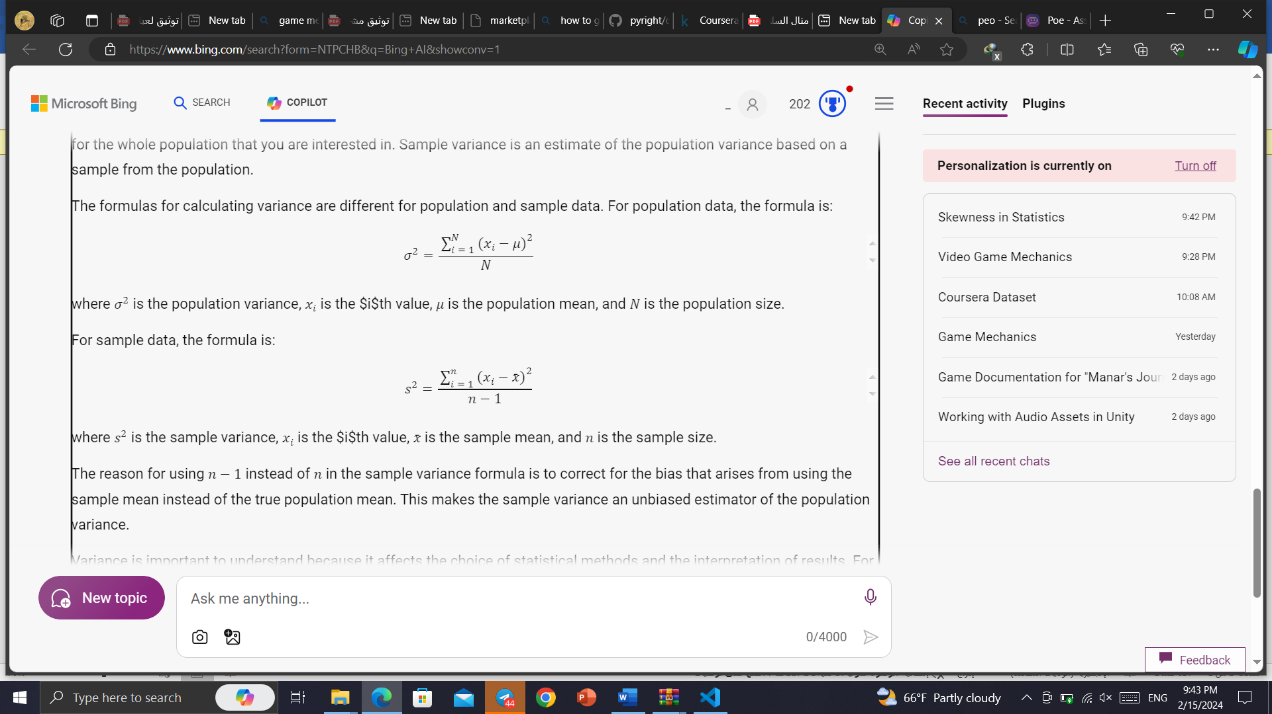
There are two types of variance: population variance and sample variance. Population variance is the exact value of variance for the whole population that you are interested in. Sample variance is an estimate of the population variance based on a sample from the population.

The formulas for calculating variance are different for population and sample data. For population data, the formula is:



where σ2 is the population variance, xi​ is the $i$th value, μ is the population mean, and N is the population size.

For sample data, the formula is:

​

where s2 is the sample variance, xi​ is the $i$th value, xˉ is the sample mean, and n is the sample size.

The reason for using n−1 instead of n in the sample variance formula is to correct for the bias that arises from using the sample mean instead of the true population mean. This makes the sample variance an unbiased estimator of the population variance.

Variance is important to understand because it affects the choice of statistical methods and the interpretation of results. For example, many parametric tests assume that the data are normally distributed, which implies zero skewness and a certain value of variance. If the data are not normally distributed or have a different variance, the results of these tests may be invalid or misleading. Therefore, it is advisable to check the variance of the data before applying any statistical procedures.

"**Deviation**" is a term that can have different meanings depending on the context in which it is used. Here are a few common interpretations of the term:

1. Statistical Deviation: In statistics, deviation refers to the amount by which a value or observation differs from the average or expected value. It is commonly measured using standard deviation, which quantifies the dispersion or spread of data points around the mean.
2. Deviation in Mathematics: In mathematics, deviation often refers to the absolute difference between a value and a reference point. For example, the deviation of a number from the mean of a set of numbers can be calculated by taking the absolute difference between the number and the mean.
3. Deviation in Project Management: In project management, deviation refers to a departure from the planned or expected course of action. It can occur when there are variations in project scope, timeline, budget, or quality. Project managers often monitor and manage deviations to ensure that projects stay on track.
4. Deviation in Psychology: In psychology, deviation can refer to behaviors or characteristics that deviate from what is considered normal or typical. For example, abnormal behaviors or psychological disorders are often described as deviations from the norm.
5. Deviation in Ethics: In ethical discussions, deviation can refer to actions or behaviors that deviate from accepted moral or ethical standards. It can involve violating established rules, norms, or principles.
6. Deviation in Engineering: In engineering, deviation can refer to the difference between the desired or intended specifications of a product or system and its actual performance. It is often used to assess the quality or accuracy of a product or process