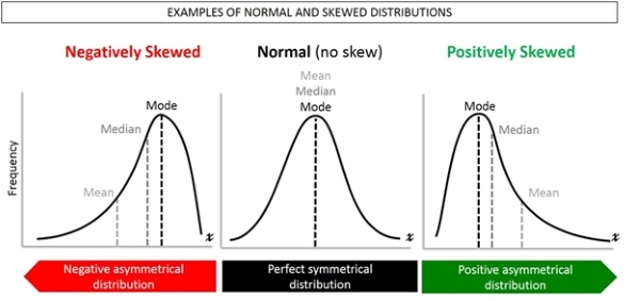
**Basic Concepts of Statistics**

**Central tendency**

* central value for a probability distribution
* the tendency for the values of a random variable to cluster round its mean, mode, or median
* median is used primarily for skewed distributions

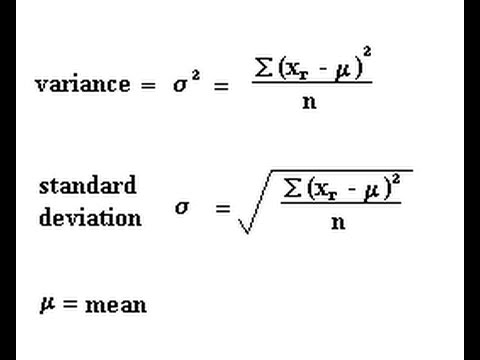


**Standard deviation**

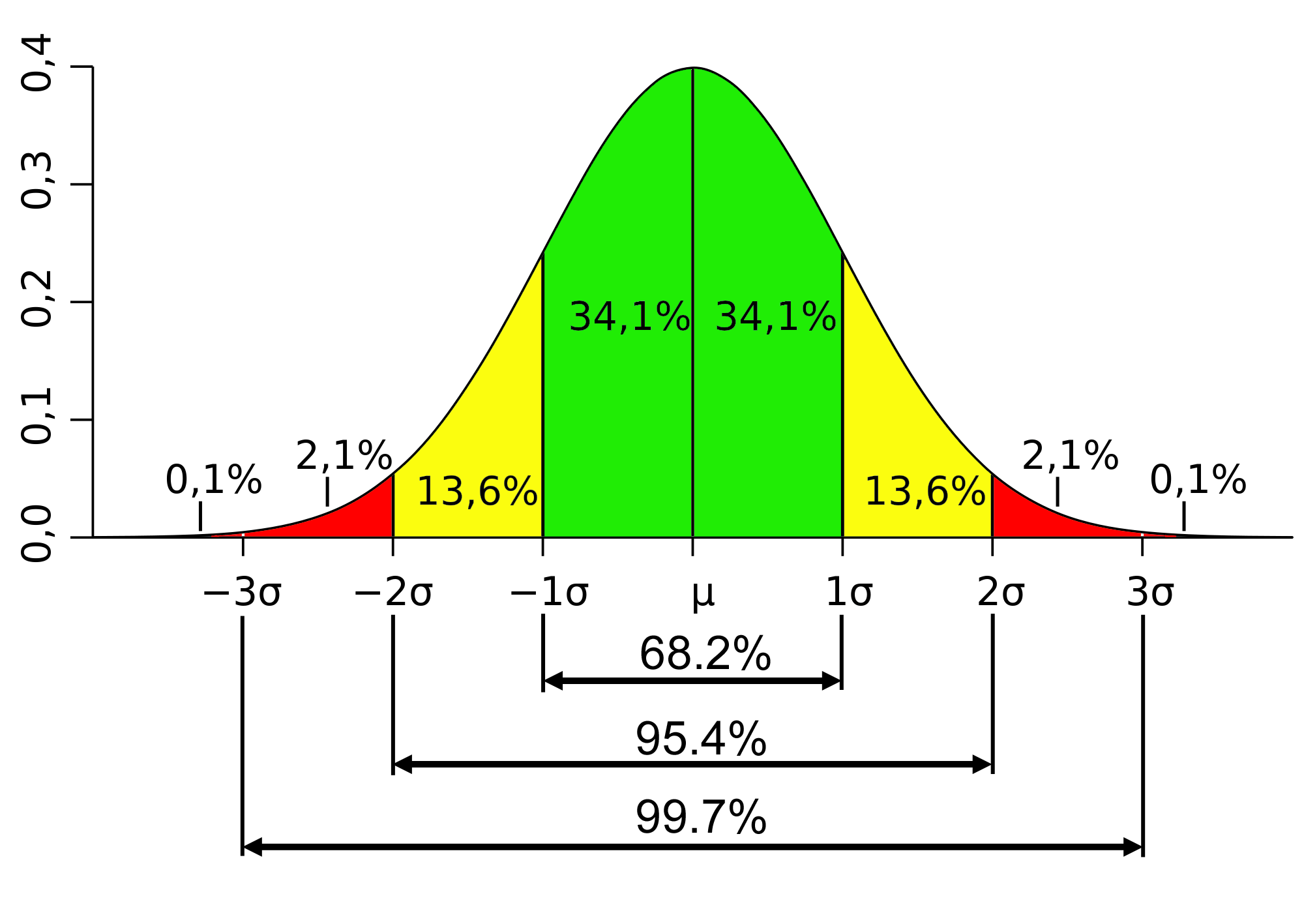
* amount of variation or dispersion of a set of data values
* the square root of its variance

**Variance**

* the expectation of the squared deviation of a random variable from its mean.
* Informally, it measures how far a set of (random) numbers are spread out from their average value



**Normal distribution / Gaussian distribution**



* bell curve
* mean = median = mode

**Covariance**

* the tendency in the linear relationship between the variables
* when the variables tend to show similar behaviour, the covariance is

positive

* when the variables tend to show opposite behaviour, the covariance is

negative

**Correlation**

* The normalized version of the covariance, the correlation coefficient, however, shows by its magnitude the strength of the linear relation.
* It has a value between +1 and −1, where 1 is total positive linear correlation, 0 is no linear correlation, and −1 is total negative linear correlation.

**Pearson's correlation coefficient** is the covariance of the two variables divided by the product of their standard deviations

