

## Glossary of Terms:

**Mean:** A commonly used measure of the center of a batch of numbers, which is also called the average. It is the sum of all observations divided by the number of (non-missing) observations.

**SE of mean:** An estimate of the dispersion in the distribution of samples means that you would obtain if you took repeated samples from the same population. The standard error of the mean is calculated by dividing the standard deviation by the square root of n.

**Standard deviation:** The most common measure of dispersion, or how spread out the data are, about the mean.

**Variance:** A measure of dispersion, or how spread out the data are, about the mean.

**Coefficient of variation:** A measure of relative variability, usually calculated only when all values are positive, as in weight, sales, or exam data. The coefficient of variation expresses the standard deviation of the data as a percentage of the mean.

**First quartile:** Every group of data has three quartiles. If you sort the data from smallest to largest, the first 25% of the data is less than or equal to the first quartile. The second quartile is the median; 50% of the data is less or equal to the median, and 50% is greater than or equal to the median. The first 75% of the data is less than or equal to the third quartile, and 25% of the data is greater than or equal to the third quartile.

**Median:** The middle of the data: half the observations are less than or equal to it. Suppose the column contains n values. If n is odd, the median is the value in the middle. If n is even, the median is the average of the two middle values. The median is equivalent to the second quartile.

**Third quartile:** Every group of data has three quartiles. If you sort the data from smallest to largest, the first 25% of the data is less than or equal to the first quartile. The second quartile is the median; 50% of the data is less or equal to the median, and 50% is greater than or equal to the median. The first 75% of the data is less than or equal to the third quartile, and 25% of the data is greater than or equal to the third quartile.

**Interquartile range:** Every group of data has three quartiles. If you sort the data from smallest to largest, the first 25% of the data is less than or equal to the first quartile. The second quartile is the median; 50% of the data is less or equal to the median, and 50% is greater than or equal to the median. The first 75% of the data is less than or equal to the third quartile, and 25% of the data is greater than or equal to the third quartile.

**Trimmed mean:** Calculates a 5% trimmed mean. Minitab removes the smallest 5% and the largest 5% of the values (rounded to the nearest integer), and then averages the remaining values.

**Sum:** Calculates the sum.

### Note

Missing values are omitted from the calculation of the function Sum.

**Minimum:** Calculates the smallest number.

### Note

Missing values are omitted from the calculation of the function Minimum.

**Maximum:** Calculates the largest number.

### Note

Missing values are omitted from the calculation of the function Maximum.

**Range:** Calculates the difference between the largest and smallest data values.

**.N nonmissing:** Choose to display the number of nonmissing column entries.

**N missing:** Choose to display the number of missing column entries.

**N total:** Choose to display the total (nonmissing and missing) number of column entries.

**Cumulative N:** Choose to display the cumulative number of entries.

**Percent:** Choose to display the percent of observations that a group constitutes. The percent will be 100 unless you use a By variable.

**Cumulative percent:** A cumulative percent is the sum of all the percentage values up to that category.

**Sum of squares:** Choose to display the sum of the squared data values. This is the uncorrected sums of squares, without first subtracting the mean. Squares each value in the column, and computes the sum of those squared values. That is, if the column contains  $x_1, x_2, \dots, x_n$ , then sum of squares calculates  $(x_1 + x_2 + \dots + x_n)$ .

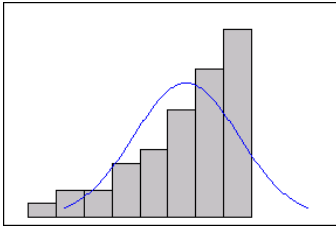
### Note

Missing values are omitted from the calculation of the function Sum of squares, uncorrected.

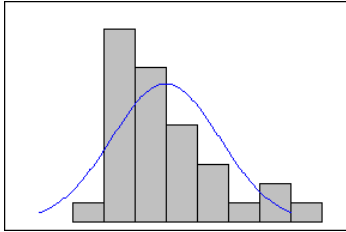
**Skewness:** A lack of symmetry. A distribution is skewed if one tail extends farther than the other. When you calculate skewness Minitab provides a skewness value:

- A value close to 0 indicates symmetric data.
- Negative values indicate negative/left skew.
- Positive values indicate positive/right skew.

Example of a negative or left-skewed distribution (skewness = -1.01)



Example of a positive or right-skewed distribution (skewness = 1.08)

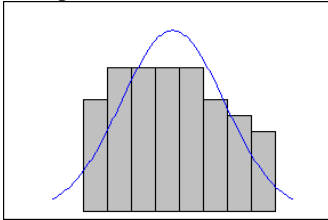


**Kurtosis:** How sharply peaked a distribution is.

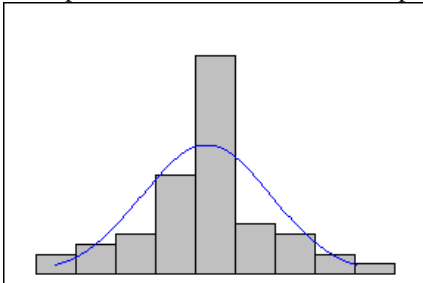
When you calculate kurtosis Minitab provides a kurtosis value:

- Values close to 0 indicate normally peaked data.
- Negative values indicate a distribution that is flatter than normal.
- Positive values indicate a distribution with a sharper than normal peak.

Example of a flatter than normal distribution (kurtosis = - 1.03)



Example of a distribution with a sharper than normal peak (kurtosis = 0.76)



**MSSD:** MSSD computes half the Mean of the Squared Successive Differences of a batch of numbers. For example, suppose a column contains 1, 2, 4, and 10. The successive differences are  $2 - 1 = 1$ ,  $4 - 2 = 2$ , and  $10 - 4 = 6$ . Then:  
 $MSSD = \text{mean}(1, 2, \text{ and } 6) / 2 = 6.83333$

Source: © 1972 - 2004 Minitab Inc.