School of Agriculture, Business, and Technology

Common Excel Functions for Summarizing Data

Developed By

Dr. Walid H. Shayya

A. CONCATENATE: Joins two or more text strings (or numbers) into one text string.

Syntax: CONCATENATE(text1,text2, ...)
Text1, text2, ... are 1 up to 255 items that may be text, numbers, cell references, or a combination of those.

Example 1: CONCATENATE(120, " up to ", 125) equals *120 up to 125*
Example 2: If cells A1:B1 contain 5 and 60, then
CONCATENATE(A1,B1) equals *560*
CONCATENATE(A1," to ",B1) equals *5 to 60*

B. COUNT: Counts the number of cells that contain numbers and also numbers within the list of arguments.

Syntax: COUNT(number1,number2, ...)
Number1, number2, ... are 1 to 30 arguments that can contain or refer to a variety of different types of data, but only numbers are counted.

Example 1: COUNT(9, 5) equals *2*
Example 2: If cells A2:E2 contain 5, 25, 30, 40, and 60, then
COUNT(A2:C2) equals *3*
COUNT(B2:E2, 3) equals *5*

C. SUM: Adds all the numbers in a range of cells.

Syntax: SUM(number1,number2, ...)
Number1, number2, ... are 1 to 30 arguments for which you want the total value or sum. Arguments that are error values or text that cannot be translated into numbers cause errors.

Example 1: SUM(9, 5) equals *14*
Example 2: If cells A2:E2 contain 5, 25, 30, 40, and 60, then
SUM(A2:C2) equals *60*
SUM(B2:E2, 3) equals *158*

D.   AVERAGE: Returns the average (arithmetic mean) of the arguments.

Syntax: AVERAGE(number1,number2, ...)
Number1, number2, ... are 1 to 30 numeric arguments for which you want the average. If an array or reference argument contains text, logical values, or empty cells, those values are ignored; however, cells with the value zero are included. Therefore, keep in mind the difference between empty cells and those containing the value zero when averaging cells (empty cells are not counted, but zero values are).

Example 1:   AVERAGE(10, 7, 9, 27,2) equals *11*
Example 2:   If A1:A5 and contains the numbers 10, 7, 9, 27, and 2, then:
AVERAGE(A1:A5) equals *11*
AVERAGE(A1:A5, 5) equals *10*
AVERAGE(A1:A5) equals SUM(A1:A5)/COUNT(A1:A5) equals *11*

E.   MIN: Returns the smallest number in a set of values.

Syntax: MIN(number1,number2, ...)
Number1, number2, ... are 1 to 30 numbers for which you want to find the minimum value. You can specify arguments that are numbers, empty cells, logical values, or text representations of numbers. Arguments that are error values or text that cannot be translated into numbers cause errors. Also, MIN returns 0 if the arguments contain no numbers.

Example 1:   MIN(10, 7, 9, 27,2) equals *2*
Example 2:   If A1:A5 contains the numbers 10, 7, 9, 27, and 2, then:
MIN(A1:A5) equals *2*
MIN(A1:A5, 0) equals *0*

F.   MAX: Returns the largest number in a set of values.

Syntax: MAX(number1,number2, ...)
Number1, number2, ... are 1 to 30 numbers for which you want to find the maximum value. You can specify arguments that are numbers, empty cells, logical values, or text representations of numbers. Arguments that are error values or text that cannot be translated into numbers cause errors. Also, MAX returns 0 if the arguments contain no numbers.

Example 1:   MAX(10, 7, 9, 27,2) equals *27*
Example 2:   If A1:A5 contains the numbers 10, 7, 9, 27, and 2, then:
MAX(A1:A5) equals *27*
MAX(A1:A5, 29) equals *29*

G.   MEDIAN: Returns the median of the given numbers. The median is the number in the middle of a set of numbers; that is, half the numbers have values that are greater than the median, and half have values that are less.

Syntax: MEDIAN(number1,number2, ...)
Number1, number2, ... are 1 to 30 numbers for which you want the median.   The arguments should be either numbers or names, arrays, or references that contain

numbers. Microsoft Excel examines all the numbers in each reference or array argument. If an array or reference argument contains text, logical values, or empty cells, those values are ignored; however, cells with the value zero are included. Also, the MEDIAN function calculates the average of the two numbers in the middle if there is an even number of numbers in the set (see the following second example).

Example 1:   MEDIAN(1, 2, 3, 4, 5) equals *3*
Example 2:   If A1:A6 contains the numbers 3, 1, 2, 6, 5, and 4, then:
               MEDIAN(A1:A6) equals *3.5*, i.e., the average of 3 and 4

H.   MODE: Returns the most frequently occurring (or repetitive) value in an array or range of data (the same as the **MODE.SNGL** function in MS Excel 2010 or later). Like the MEDIAN function, the MODE is a location measure.

Syntax: MODE(number1,number2, ...)
               Number1, number2, ... are 1 to 30 arguments for which you want to calculate the mode. The arguments should be numbers, names, arrays, or references that contain numbers. If an array or reference argument contains text, logical values, or empty cells, those values are ignored; however, cells with the value zero are included. If the data set contains no duplicate data points, MODE returns the "#N/A" error value. If the data set contains equal duplicate of data points, MODE returns the one it encounters first in the data set.

Example 1:   MODE({5.6, 4, 4, 3, 2, 4}) equals *4*
Example 2:   MODE({5.6, 6, 3, 3, 2, 6}) equals *6*
Example 3:   If A1:A6 contains the numbers 4, 3, 4, 2, 5.6, and 4, then:
               MODE(A1:A6) equals *4*

I.   MODE.MULT (only starting with version 10 of MS Excel): Returns the multiple modes of the most frequently occurring (or repetitive) values in an array or range of data. Because the function returns an array, it should be entered as an array formula.

Syntax: MODE.MULT(number1,number2, ...)
               Number1, number2, ... are 1 to 30 arguments for which you want to calculate the multiple modes. The arguments should be numbers, names, arrays, or references that contain numbers. If an array or reference argument contains text, logical values, or empty cells, those values are ignored; however, cells with the value zero are included. If the data set contains no duplicate data points, function returns the "#N/A" error value. If the data set contains equal duplicate of data points, the function those in the highlighted array of cells.

Example 1:   MODE.MULT({5.6, 4, 4, 3, 3, 2}) equals *{4,3}*
Example 2:   MODE.MULT({5.6, 6, 3, 3, 2, 2, 6}) equals *{6,3,2}*
Example 3:   If A1:A8 contains the numbers 4, 3, 4, 3, 2, 5.6, 3, and 4, then:
               MODE.MULT(A1:A8) equals *{4,3}*

J.   AVEDEV: Estimates the average (or mean) of the absolute deviations of data points from their mean. The mean deviation is a measure of how widely values are dispersed from the average value (the mean).

Syntax: AVEDEV(number1,number2,...)

Number1,number2, ... are 1 to 30 number arguments corresponding to a sample of a population. You can also use a single array or a reference to an array instead of arguments separated by commas. STDEV uses the following formula:

$$\frac{1}{n}\sum |x - \bar{x}|$$

Example 1: AVEDEV(1345, 1301, 1368, 1322, 1310, 1370, 1318, 1350, 1303, 1299) equals *23.72*

Example 2: If A1:A10 contains the numbers 1345, 1301, 1368, 1322, 1310, 1370, 1318, 1350, 1303, and 1299, then:

AVEDEV(A1:A10) equals *23.72*

K. STDEV: Estimates the standard deviation based on a sample (the same as the **STDEV.S** function in MS Excel 2010 or later). The standard deviation is a measure of how widely values are dispersed from the average value (the mean).

Syntax: STDEV(number1,number2,...)

Number1,number2, ... are 1 to 30 number arguments corresponding to a sample of a population. You can also use a single array or a reference to an array instead of arguments separated by commas. STDEV assumes that its arguments are a sample of the population. If your data represents the entire population, then compute the standard deviation using STDEVP. The standard deviation is calculated using the "nonbiased" or "n-1" method. STDEV uses the following formula:

$$\sqrt{\frac{n\sum x^2 - \left(\sum x\right)^2}{n(n-1)}}$$

Example 1: STDEV(1345, 1301, 1368, 1322, 1310, 1370, 1318, 1350, 1303, 1299) equals *27.46*

Example 2: If A1:A10 contains the numbers 1345, 1301, 1368, 1322, 1310, 1370, 1318, 1350, 1303, and 1299, then:

STDEV(A1:A10) equals *27.46*

L. STDEVP: Calculates the standard deviation based on the entire population given as arguments (the same as the **STDEV.P** function in MS Excel 2010 later). The standard deviation is a measure of how widely values are dispersed from the average value (the mean).

Syntax: STDEVP(number1,number2,...)

Number1,number2, ... are 1 to 30 number arguments corresponding to a sample of a population. You can also use a single array or a reference to an array instead of arguments separated by commas. STDEVP assumes that its arguments are the entire population. If your data represents the a sample of the population, then compute the standard deviation using STDEV. The standard deviation is calculated using the "biased" or "n" method. STDEVP uses the following formula:

$$\sqrt{\frac{n\sum x^2 - \left(\sum x\right)^2}{n^2}}$$

Example 1: STDEVP(1345, 1301, 1368, 1322, 1310, 1370, 1318, 1350, 1303, 1299) equals *26.05*

Example 2: If A1:A10 contains the numbers 1345, 1301, 1368, 1322, 1310, 1370, 1318, 1350, 1303, and 1299, then:
STDEVP(A2:E3) equals *26.05*

M. VAR: Estimates the variance based on a sample (the same as the **VAR.S** function in MS Excel 2010 or later). The variance is another measure of how widely values are dispersed from the average value (the mean).

Syntax: VAR(number1,number2,...)
Number1,number2, ... are 1 to 30 number arguments corresponding to a sample of a population. You can also use a single array or a reference to an array instead of arguments separated by commas. VAR assumes that its arguments are a sample of the population. If your data represents the entire population, then compute the variance using VARP. The variance is calculated using the "nonbiased" or "n-1" method. VAR uses the following formula:

$$\frac{\sum(x-\bar{x})^2}{(n-1)}$$

Example 1: VAR(1345, 1301, 1368, 1322, 1310, 1370, 1318, 1350, 1303, 1299) equals *754.27*

Example 2: If A1:A10 contains the numbers 1345, 1301, 1368, 1322, 1310, 1370, 1318, 1350, 1303, and 1299, then:
VAR(A1:A10) equals *754.27*

N. VARP: Calculates the variance based on the entire population given as arguments (the same as the **VAR.P** function in MS Excel 2010 or later). The variance is another measure of how widely values are dispersed from the average value (the mean).

Syntax: VARP(number1,number2,...)
Number1,number2, ... are 1 to 30 number arguments corresponding to a sample of a population. You can also use a single array or a reference to an array instead of arguments separated by commas. VARP assumes that its arguments are the entire population. If your data represents the a sample of the population, then compute the variance using VAR. The variance is calculated using the "biased" or "n" method. VARP uses the following formula:

$$\frac{\sum(x-\bar{x})^2}{n}$$

Example: VARP(1345, 1301, 1368, 1322, 1310, 1370, 1318, 1350, 1303, 1299) equals *678.84*

O. QUARTILE: Returns the quartile of a data set (the same as the **QUARTILE.INC** function in MS Excel 2010 or later). The 1st quartile is the median of the lower half of the data while the 3rd quartile is the median of the upper half of the data. The QUARTILE (and QUARTILE.INC) function includes the median (the central value in the ordered list) in both halves if there are an odd number of data points in the original ordered data set. However, the function will split the data set exactly in half if there are an even number of data points in the original ordered data set but produce a value for the 1st quartile that is slightly higher than the median of the 1st half of the ordered data set and a value for the 3rd quartile that is slightly lower than the median of the 2nd half of the ordered data set.

Syntax: QUARTILE(array,quart)
array is an array of (or reference to) a set of values for which you want the quartile value.
quart indicates which value to return (0 - minimum value, 1 - 1st quartile, 2 - 2nd quartile or median, 3 - 3rd quartile, and 4 - maximum value).

Example: If A1:A8 contains the numbers 7, 12, 4, 1, 8, 9, 10, and 2, then:
QUARTILE(A1:A8,1) equals *3.5*
QUARTILE(A1:A8,2) equals *7.5*
QUARTILE(A1:A8,3) equals *9.25*

QUARTILE.EXC is another function available in MS Excel 2010 and later where the median (the central value in the ordered list) is excluded from both halves of the data set if there are an odd number of data points in the original ordered data set. However, the function will split the data set exactly in half (as will the QUARTILE and QUATILE.INC functions) if there are an even number of data points in the original ordered data set but produce a value for the 1st quartile that is slightly higher than the median of the 1st half of the ordered data set and a value for the 3rd quartile that is slightly lower than the median of the 2nd half of the ordered data set.

Example: If A1:A8 contains the numbers 7, 12, 4, 1, 8, 9, 10, and 2, then:
QUARTILE.EXC(A1:A8,1) equals *3.25*
QUARTILE.EXC(A1:A8,2) equals *7.5*
QUARTILE.EXC(A1:A8,3) equals *9.75*

P. FREQUENCY: Calculates how often values occur within a range of values, and returns a vertical array of numbers. Because FREQUENCY returns an array, it must be entered as an array formula.

Syntax: FREQUENCY(data_array,bins_array)
Data_array is an array of (or reference to) a set of values for which you want to count frequencies. If data_array contains no values, FREQUENCY returns an array of zeros.
Bins_array is an array of or reference to intervals into which you want to group the values in data_array. If bins_array contains no values, FREQUENCY returns the number of elements in data_array.

Remarks: FREQUENCY is entered as an array formula after you select a range of adjacent cells into which you want the returned distribution to appear. In order to enter this formula correctly, you need to push the "Ctrl" and "Shift" simultaneously and hold while clicking on the "OK" option of the equation window (or clicking the "Enter" key). The number of elements in the returned array is one more than the

---

number of elements in bins_array. The extra element in the returned array will hold the count of any values above the highest interval. For example, when counting three ranges of values (intervals) that are entered into three cells, be sure to enter FREQUENCY into four cells for the results. The extra cell returns the number of values in data_array that are greater than the third interval value. *FREQUENCY ignores blank cells and text.*

Example: Suppose a worksheet lists scores for an exam. The scores are 85, 85, 78, 79, 83, 80, 95, 88, and 97 and are entered into A1:A9, respectively. The data_array would contain a column of these test scores. The bins_array would be another column of intervals by which the test scores are grouped. In this example, bins_array will be saved in cells C4:C6 and would contain the values 70, 80, and 90. When the FREQUENCY function is entered as an array, the software will return a count of each of the scores ranges (i.e., those corresponding to the grade ranges of 70 or less, more than 70 but less than or equal to 80, more than 80 but less than or equal to 90, and more than 90). If the following formula is entered as an array formula after selecting four vertical cells adjacent to the data_array, then:

FREQUENCY(A1:A9,C4:C6) equals *{0;3;4;2}*

The fourth number (i.e., 2) is the count of the two values (95 and 97) that are greater than the highest interval (i.e., more than 90).