ADVANCED TECHNIQUES IN EXCEL

edunet
foundation

STATISTICAL FUNCTIONS

AVERAGE

- The AVERAGE function in Excel calculates the average (arithmetic mean) of a group of numbers. The AVERAGE function ignores logical values, empty cells and cells that contain text.

- To calculate the average of a group of numbers, use the AVERAGE function.

AVERAGE

Instead of using the AVERAGE function, use SUM and COUNT.

1. For example, the AVERAGE function below calculates the average of the numbers in cells A1 through A3.

2. The formula below produces the exact same result.

AVERAGE

3. The following AVERAGE function calculates the average of the numbers in cells A1 through A3 and the number 8.

A4		⋮	✕	✓	f_x	=AVERAGE(A1:A3,8)				
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	
1	8									
2	4									
3	9									
4	7.25									
5										

4. The AVERAGE function ignores logical values (TRUE or FALSE), empty cells and cells that contain text.

A7		⋮	✕	✓	f_x	=AVERAGE(A1:A6)				
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	
1	8									
2	4									
3	9									
4	TRUE									
5										
6	USA									
7	7									
8										

edunet
foundation

AVERAGEA

The AVERAGEA function also returns the average (arithmetic mean) of a group of numbers. However, the logical value FALSE and cells that contain text evaluate to 0 and the logical value TRUE evaluates to 1. The AVERAGEA function also ignores empty cells.

1. For example, take a look at the AVERAGEA function below.

2. You can use the normal AVERAGE function to check this result.

AVERAGEIF

- The AVERAGEIF function in Excel calculates the average of cells that meet one criteria. AVERAGEIFS calculates the average of cells that meet multiple criteria.

- To average cells based on one criteria, use the AVERAGEIF function. For example, to calculate the average excluding zeros.

AVERAGEIF

1. For example, the AVERAGEIF function below (two arguments) calculates the average of all values in the range A1:A7 that are greater than 0.

2. The AVERAGEIF function below (three arguments, last argument is the range to average) calculates the average of all values in the range B1:B7 if the corresponding cells in the range A1:A7 contain exactly Apple.

A9			✕	✓	fx	=AVERAGEIF(A1:A7,">0")

	A	B	C	D	E	F	G	H	I
1	0								
2	10								
3	0								
4	0								
5	20								
6	0								
7	0								
8									
9	15								
10									

B9			✕	✓	fx	=AVERAGEIF(A1:A7,"Apple",B1:B7)

	A	B	C	D	E	F	G	H	I
1	Banana	70							
2	Strawberry	1							
3	Apple	4							
4	Pear	60							
5	Kiwi	20							
6	Raspberry	5							
7	Apple	8							
8									
9		6							
10									

AVERAGEIF

3. The AVERAGEIF function below calculates the average of all values in the range B1:B7 if the corresponding cells in the range A1:A7 do not contain exactly Banana.

4. The AVERAGEIF function below calculates the average of all values in the range B1:B7 if the corresponding cells in the range A1:A7 contain a series of zero or more characters + berry. An asterisk (*) matches a series of zero or more characters.

| B9 | | fx | =AVERAGEIF(A1:A7,"<>Banana",B1:B7) |

	A	B	C	D	E	F	G	H	I
1	Banana	70							
2	Strawberry	1							
3	Apple	4							
4	Pear	60							
5	Kiwi	20							
6	Raspberry	5							
7	Apple	8							
8									
9		16.33333							
10									

| B9 | | fx | =AVERAGEIF(A1:A7,"*berry",B1:B7) |

	A	B	C	D	E	F	G	H	I
1	Banana	70							
2	Strawberry	1							
3	Apple	4							
4	Pear	60							
5	Kiwi	20							
6	Raspberry	5							
7	Apple	8							
8									
9		3							
10									

AVERAGEIF

5. The AVERAGEIF function below calculates the average of all values in the range B1:B7 if the corresponding cells in the range A1:A7 contain exactly 4 characters. A question mark (?) matches exactly one character.

B9			×	✓	fx	=AVERAGEIF(A1:A7,"????",B1:B7)				
	A	B	C	D	E	F	G	H	I
1	Banana	70							
2	Strawberry	1							
3	Apple	4							
4	Pear	60							
5	Kiwi	20							
6	Raspberry	5							
7	Apple	8							
8									
9		40							
10									

The AVERAGEIFS function (with the letter S at the end) in Excel calculates the average of cells that meet multiple criteria.

6. The AVERAGEIFS function below calculates the average of all values in the range A1:A7 that are greater than or equal to 500 and less than or equal to 1000.

A9			×	✓	fx	=AVERAGEIFS(A1:A7,A1:A7,">=500",A1:A7,"<=1000")				
	A	B	C	D	E	F	G	H	I
1	58								
2	1000								
3	4								
4	1200								
5	12								
6	600								
7	9								
8									
9	800								
10									

Note: first argument is the range to average, followed by two or more range/criteria pairs.

AVERAGEIF

- 7. The AVERAGEIFS function below calculates the average of all values in the range C1:C7 if the corresponding cells in the range A1:A7 contain exactly Apple and the corresponding cells in the range B1:B7 contain exactly Red.

- Note: again, the first argument is the range to average, followed by two or more range/criteria pairs.

| C9 | | fx | =AVERAGEIFS(C1:C7,A1:A7,"Apple",B1:B7,"Red") |

	A	B	C	D	E	F	G	H	I
1	Apple	Green	58						
2	Banana	Yellow	1000						
3	Banana	Yellow	4						
4	Apple	Red	1200						
5	Apple	Green	12						
6	Apple	Red	600						
7	Banana	Yellow	9						
8									
9			900						
10									

MEDIAN

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
A3							f_x	=MEDIAN(A1:O1)							
1	0	7	8	6	5	9	8	7	4	8	0	3	5	6	8	
2																
3	6															
4																

- To find the median (or middle number), use the MEDIAN function.
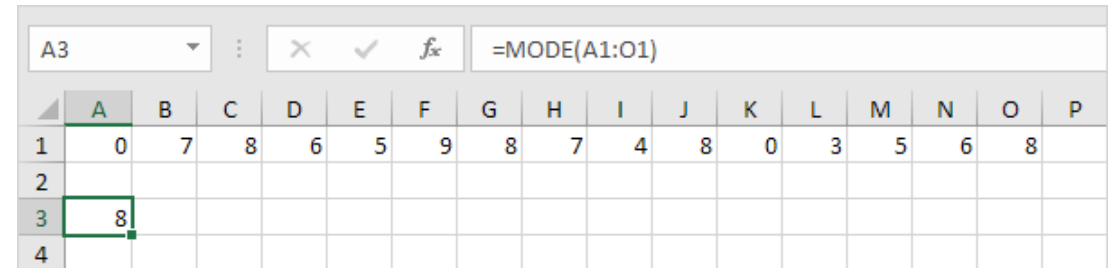
Check:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	0	0	3	4	5	5	6	6	7	7	8	8	8	8	9	

MODE

- Use the MODE function in Excel to find the most frequently occurring number in a list of numbers. Use MODE.MULT to find multiple modes.

- To find the most frequently occurring number, use the MODE function.

- Note: visit our page about the MODE function to learn more about this Excel function.

A3					f_x	=MODE(A1:O1)									
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	0	7	8	6	5	9	8	7	4	8	0	3	5	6	8	
2																
3	8															
4																

edunet
foundation

MODE

1. The MODE function below returns the most frequently occurring number (8).

2. The new MODE.SNGL function (SNGL stands for single) produces the exact same result.

MODE

3. Change the value in cell A2 to 5. In this example, there are multiple modes (5 and 8). MODE and MODE.SNGL always return a single mode.

4. If you have Excel 365 or Excel 2021, simply use the MODE.MULT function to find multiple modes.

Note: the MODE.MULT function, entered into cell A8, fills multiple cells. Wow! This behavior in Excel 365/2021 is called spilling.

MODE

5. Select multiple cells.
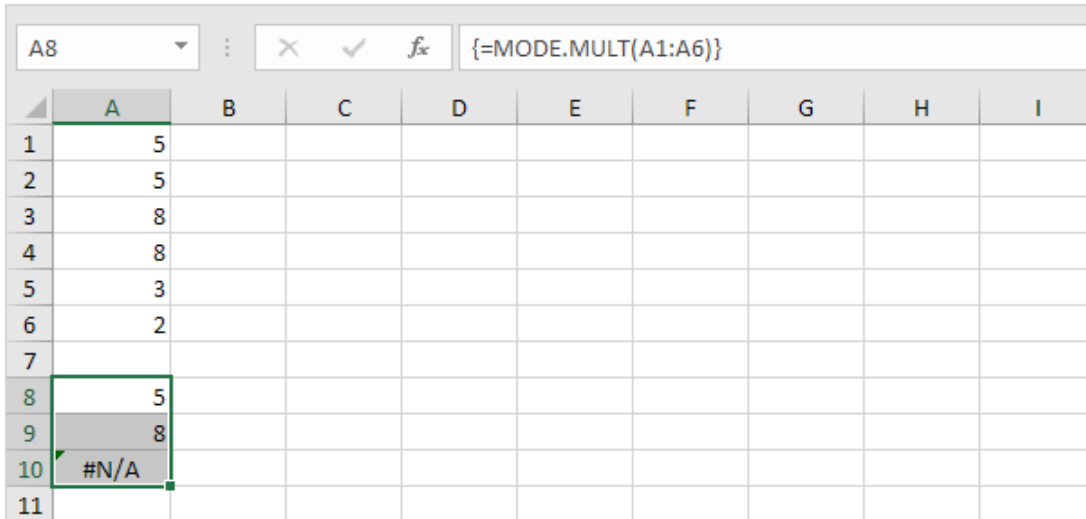
6. Enter the MODE.MULT function.
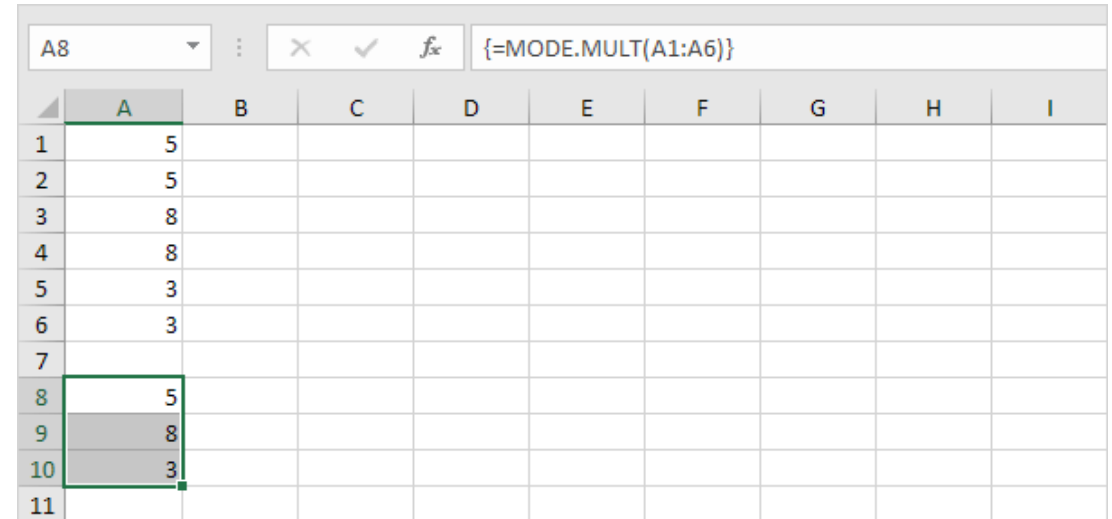
MODE

7. Finish by pressing CTRL + SHIFT + ENTER.

8. Change the value in cell A6 to 3.

A8			×	✓	*fx*	{=MODE.MULT(A1:A6)}

	A	B	C	D	E	F	G	H	I
1	5								
2	5								
3	8								
4	8								
5	3								
6	2								
7									
8	5								
9	8								
10	#N/A								
11									

A8			×	✓	*fx*	{=MODE.MULT(A1:A6)}

	A	B	C	D	E	F	G	H	I
1	5								
2	5								
3	8								
4	8								
5	3								
6	3								
7									
8	5								
9	8								
10	3								
11									

Note: the formula bar indicates that this is an array formula by enclosing it in curly braces {}. To delete this array formula, select the range A8:A10 and press Delete. In this example, there are only two modes (5 and 8) so we could have selected two cells at step 5.

MODE

9. The MODE function only works with numbers. Use INDEX, MODE and MATCH to find the most frequently occurring word in Excel.

STANDARD DEVIATION

- Standard deviation is a number that tells you how far numbers are from their mean.

- To calculate the standard deviation, use the STEDV function.

- Note: standard deviation is a number that tells you how far numbers are from their mean. Learn more about this topic on our page about standard deviation.

A3				f_x	=STDEV(A1:O1)										
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	0	7	8	6	5	9	8	7	4	8	0	3	5	6	8	
2																
3	2.82															
4																

STANDARD DEVIATION

For example, the numbers below have a mean (average) of 10.

Explanation: the numbers are all the same which means there's no variation. As a result, the numbers have a standard deviation of zero. The STDEV function is an old function. Microsoft Excel recommends using the new STEDV.S function which produces the exact same result.

STANDARD DEVIATION

2. The numbers below also have a mean (average) of 10.

Explanation: the numbers are close to the mean. As a result, the numbers have a low standard deviation.

STANDARD DEVIATION

3. The numbers below also have a mean (average) of 10.

Explanation: the numbers are spread out. As a result, the numbers have a high standard deviation.

	A	B	C	D	E	F	G	H	I
1	0								
2	20								
3	1								
4	27								
5	2								
6	12.58968								
7									

A6 — =STDEV(A1:A5)

STDEV.P

- The STDEV.P function (the P stands for Population) in Excel calculates the standard deviation based on the entire population. For example, you're teaching a group of 5 students. You have the test scores of <u>all</u> students. The entire population consists of 5 data points. The STDEV.P function uses the following formula:

$$\sigma = \sqrt{\frac{\Sigma(x_i - \mu)^2}{N}}$$

STDEV.P

In this example, $x_1 = 5$, $x_2 = 1$, $x_3 = 4$, $x_4 = 6$, $x_5 = 9$, $\mu = 5$ (mean), $N = 5$ (number of data points).

1. Calculate the mean (μ).

A6			$\times$	$\checkmark$	f_x	=AVERAGE($A$1:$A$5)			
	A	B	C	D	E	F	G	H	I
1	5								
2	1								
3	4								
4	6								
5	9								
6	5								
7									

2. For each number, calculate the distance to the mean.

B1			$\times$	$\checkmark$	f_x	=A1-AVERAGE($A$1:$A$5)			
	A	B	C	D	E	F	G	H	I
1	5	0							
2	1	-4							
3	4	-1							
4	6	1							
5	9	4							
6									

edunet
foundation

STDEV.P

3. For each number, square this distance.

	A	B	C	D	E	F	G	H	I
	B1			× ✓ f_x	=(A1-AVERAGE($A$1:$A$5))^2				
1	5	0							
2	1	16							
3	4	1							
4	6	1							
5	9	16							
6									

4. Sum (∑) these values.

	A	B	C	D	E	F	G	H	I
	B6			× ✓ f_x	=SUM(B1:B5)				
1	5	0							
2	1	16							
3	4	1							
4	6	1							
5	9	16							
6		34							
7									

STDEV.P

5. Divide by the number of data points (N = 5).

B6					f_x	=SUM(B1:B5)/COUNT(A1:A5)			

	A	B	C	D	E	F	G	H	I
1	5	0							
2	1	16							
3	4	1							
4	6	1							
5	9	16							
6		6.8							
7									

6. Take the square root.

B6					f_x	=SQRT(SUM(B1:B5)/COUNT(A1:A5))			

	A	B	C	D	E	F	G	H	I
1	5	0							
2	1	16							
3	4	1							
4	6	1							
5	9	16							
6		2.607681							
7									

STDEV.P

7. Fortunately, the STDEV.P function in Excel can execute all these steps for you.

STDEV.S

The STDEV.S function (the S stands for Sample) in Excel estimates the standard deviation based on a sample. For example, you're teaching a large group of students. You <u>only</u> have the test scores of 5 students. The sample size equals 5. The STDEV.S function uses the following formula:

$$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n - 1}}$$

STDEV.S

In this example, $x_1=5$, $x_2=1$, $x_3=4$, $x_4=6$, $x_5=9$ (same numbers as above), $\bar{x}=5$ (sample mean), $n=5$ (sample size).

1. Repeat steps 1-5 above but at step 5 divide by n-1 instead of N.

2. Take the square root.

B6			f_x	=SUM(B1:B5)/(COUNT(A1:A5)-1)

	A	B	C	D	E	F	G	H	I
1	5	0							
2	1	16							
3	4	1							
4	6	1							
5	9	16							
6		8.5							
7									

B6			f_x	=SQRT(SUM(B1:B5)/(COUNT(A1:A5)-1))

	A	B	C	D	E	F	G	H	I
1	5	0							
2	1	16							
3	4	1							
4	6	1							
5	9	16							
6		2.915476							
7									

STDEV.S

3. Fortunately, the STDEV.S function in Excel can execute all these steps for you.

Note: why do we divide by n - 1 instead of by n when we estimate the standard deviation based on a sample? Bessel's correction states that dividing by n-1 instead of by n gives a better estimation of the standard deviation.

	A	B	C	D	E	F	G	H	I
A6	=STDEV.S(A1:A5)								
1	5								
2	1								
3	4								
4	6								
5	9								
6	2.915476								
7									

VARIANCE

Variance is the square of the standard deviation. It's that simple. Sometimes it's easier to use the variance when solving statistical problems.

1. The VAR.P function below calculates the variance based on the entire population.

Note: you already knew this answer (see step 5 under STDEV.P). Take the square root of this result to find the standard deviation based on the entire population.

A6		⋮	✕	✓	*fx*	=VAR.P(A1:A5)			
	A	B	C	D	E	F	G	H	I
1	5								
2	1								
3	4								
4	6								
5	9								
6	6.8								
7									

VARIANCE

2. The VAR.S function below estimates the variance based on a sample.

Note: you already knew this answer (see step 1 under STDEV.S). Take the square root of this result to find the standard deviation based on a sample.

	A	B	C	D	E	F	G	H	I
1	5								
2	1								
3	4								
4	6								
5	9								
6	8.5								
7									

A6   =VAR.S(A1:A5)

VARIANCE

3. VAR and VAR.S produce the exact same result.

Note: Microsoft Excel recommends using the new VAR.S function.

MIN

- To find the minimum value, use the MIN function.

A3			×	✓	fx		=MIN(A1:O1)									
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**	**P**
1	0	7	8	6	5	9	8	7	4	8	0	3	5	6	8	
2																
3	0															
4																

MAX

- To find the maximum value, use the MAX function.

A3				fx	=MAX(A1:O1)										
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	0	7	8	6	5	9	8	7	4	8	0	3	5	6	8	
2																
3	9															
4																

edunet
foundation

MAXIFS AND MINIFS

1. For example, the MAXIFS function below finds the highest female score.

Note: the first argument (D2:D12 in this example) is always the range in which the maximum or minimum will be determined. This MAXIFS function has 1 range/criteria pair (B2:B12/Female).

	F2				fx	=MAXIFS(D2:D12,B2:B12,"Female")			
	A	B	C	D	E	F	G	H	I
1	Name	Gender	Country	Score					
2	Richard	Male	United States	74		96			
3	Jennifer	Female	United Kingdom	92					
4	James	Male	United States	65					
5	Lisa	Female	Canada	82					
6	Sharon	Female	Australia	50					
7	Elizabeth	Female	Canada	91					
8	Carol	Female	United States	96					
9	Mark	Male	United States	58					
10	John	Male	Canada	67					
11	Susan	Female	United Kingdom	54					
12	David	Male	United States	83					
13									

MAXIFS AND MINIFS

2. The MINIFS function below finds the lowest female score.

MAXIFS AND MINIFS

3. For example, the MAXIFS function below finds the highest female score in Canada.

Note: this MAXIFS function has 2 range/criteria pairs (B2:B12/Female and C2:C12/Canada). MAXIFS and MINIFS functions can handle up to 126 range/criteria pairs.

	A	B	C	D	E	F	G	H	I
						=MAXIFS(D2:D12,B2:B12,"Female",C2:C12,"Canada")			
1	Name	Gender	Country	Score					
2	Richard	Male	United States	74		91			
3	Jennifer	Female	United Kingdom	92					
4	James	Male	United States	65					
5	Lisa	Female	Canada	82					
6	Sharon	Female	Australia	50					
7	Elizabeth	Female	Canada	91					
8	Carol	Female	United States	96					
9	Mark	Male	United States	58					
10	John	Male	Canada	67					
11	Susan	Female	United Kingdom	54					
12	David	Male	United States	83					
13									

MAXIFS AND MINIFS

4. The MAXIFS function below finds the highest score below 60.

Note: this MAXIFS function only uses the range D2:D12.

F2			f_x	=MAXIFS(D2:D12,D2:D12,"<60")					
	A	B	C	D	E	F	G	H	I
1	Name	Gender	Country	Score					
2	Richard	Male	United States	74		58			
3	Jennifer	Female	United Kingdom	92					
4	James	Male	United States	65					
5	Lisa	Female	Canada	82					
6	Sharon	Female	Australia	50					
7	Elizabeth	Female	Canada	91					
8	Carol	Female	United States	96					
9	Mark	Male	United States	58					
10	John	Male	Canada	67					
11	Susan	Female	United Kingdom	54					
12	David	Male	United States	83					
13									

edunet
foundation

LARGE

- To find the third largest number, use the following LARGE function.

Check:

SMALL

- To find the second smallest number, use the following SMALL function.

| A3 | | | fx | =SMALL(A1:O1,2) |

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	0	7	8	6	5	9	8	7	4	8	0	3	5	6	8	
2																
3	0															
4																
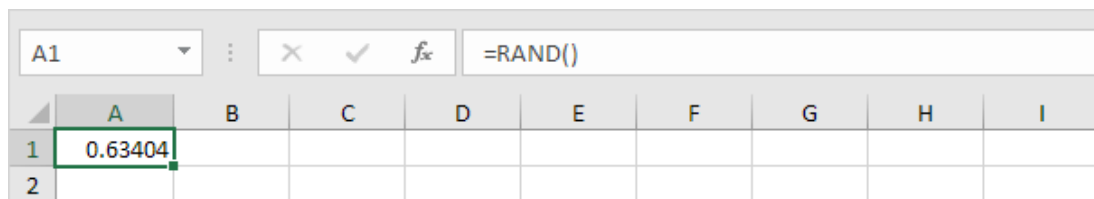
Check:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	0	0	3	4	5	5	6	6	7	7	8	8	8	8	9	

RANDOM NUMBERS - RAND

The RAND function generates a random decimal number between 0 and 1.
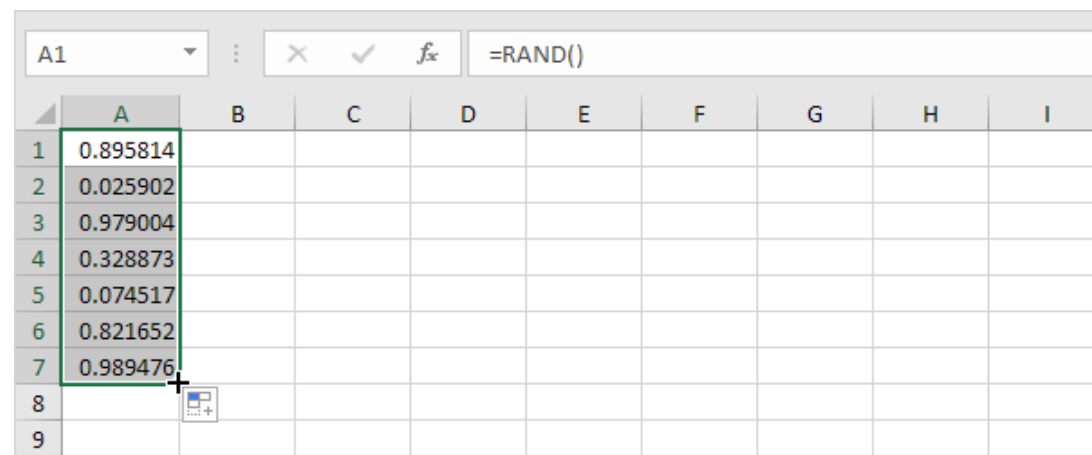
1. Select cell A1.
2. Type =RAND() and press Enter. The RAND function takes no arguments.

A1	▼	⋮	✕	✓	*fx*	=RAND()

	A	B	C	D	E	F	G	H	I
1	0.63404								
2									

3. To generate a list of random numbers, select cell A1, click on the lower right corner of cell A1 and drag it down.
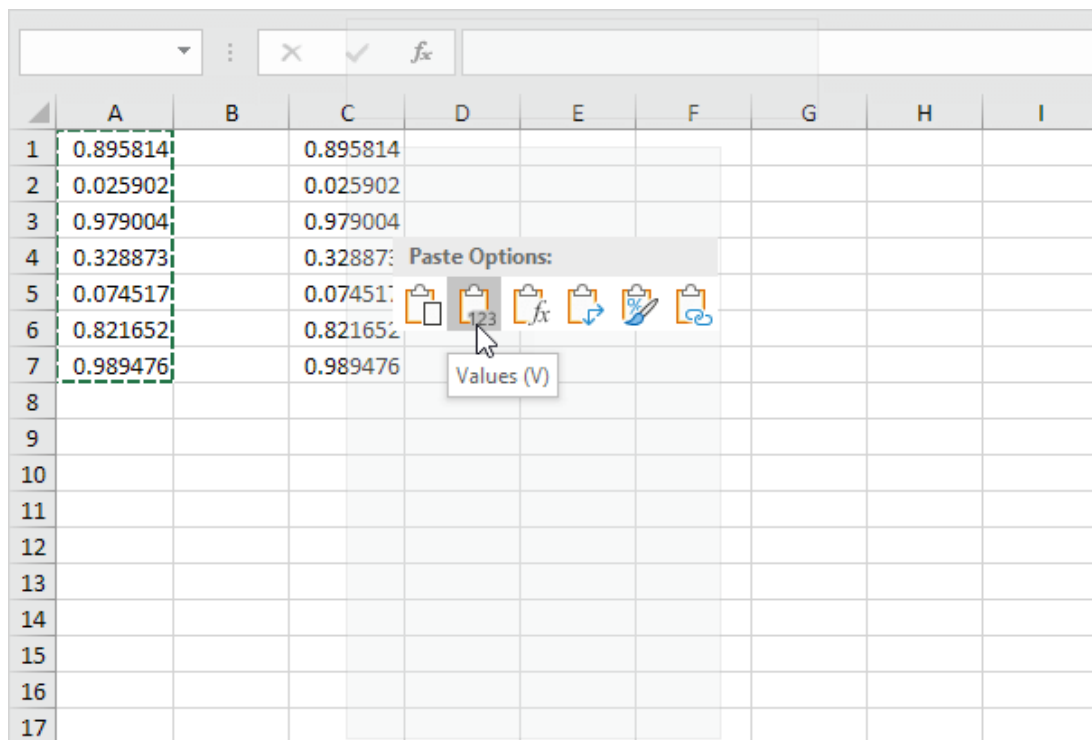
A1	▼	⋮	✕	✓	*fx*	=RAND()

	A	B	C	D	E	F	G	H	I
1	0.895814								
2	0.025902								
3	0.979004								
4	0.328873								
5	0.074517								
6	0.821652								
7	0.989476								
8									
9									

Note that cell A1 has changed. That is because random numbers change every time a cell on the sheet is calculated.
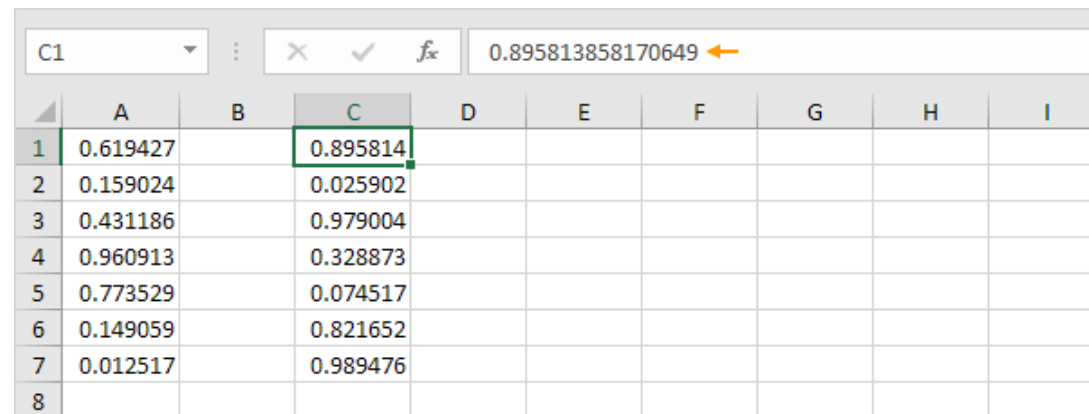
RANDOM NUMBERS - RAND

4. If you want random numbers that remain constant once generated, simply copy the random numbers and paste them as values.

5. Select cell C1 and look at the formula bar. This cell holds a value now and not the RAND function.

RANDOM NUMBERS - RANDBETWEEN

The RANDBETWEEN function generates a random whole number between two boundaries.

1. Select cell A1.

2. Type =RANDBETWEEN(50,75) and press Enter.

A1		⋮	✕ ✓	*fx*	=RANDBETWEEN(50,75)			
	A	B	C	D	E	F	G	H	I
1	59								
2									

3. If you want to generate random decimal numbers between 50 and 75, modify the RAND function as follows:

A1		⋮	✕ ✓	*fx*	=50+25*RAND()			
	A	B	C	D	E	F	G	H	I
1	61.3715								
2									

RANK

The RANK function in Excel returns the rank of a number in a list of numbers. Use RANK.AVG to return the average rank if more than one number has the same rank.

1. If the third argument is omitted (or 0), Excel ranks the largest number first, second largest number second, etc.

2. If the third argument is 1, Excel ranks the smallest number first, second smallest number second, etc.

B1			fx	=RANK(A1,$A$1:$A$9)				
	A	B	C	D	E	F	G	H	I
1	1455	6						
2	1534	5						
3	1108	9						
4	2109	3						
5	2109	3						
6	2917	1						
7	1427	7						
8	1220	8						
9	2298	2						
10								

B1			fx	=RANK(A1,$A$1:$A$9,1)				
	A	B	C	D	E	F	G	H	I
1	1455	4						
2	1534	5						
3	1108	1						
4	2109	6						
5	2109	6						
6	2917	9						
7	1427	3						
8	1220	2						
9	2298	8						
10								

Note: when we drag this formula down, the absolute reference ($A$1:$A$9) stays the same, while the relative reference (A1) changes to A2, A3, A4, etc.

Note: the RANK function is an old function. Microsoft Excel recommends using the new RANK.EQ function in Excel 2010 or later. There's no difference at all between these two functions.

RANK

3. The RANK.AVG function in Excel 2010 or later returns the average rank if more than one number has the same rank.

| B1 | | | | f_x | =RANK.AVG(A1,$A$1:$A$9,1) |

	A	B	C	D	E	F	G	H	I
1	1455	4							
2	1534	5							
3	1108	1							
4	2109	6.5 ←							
5	2109	6.5 ←							
6	2917	9							
7	1427	3							
8	1220	2							
9	2298	8							
10									

4. To create a clean RANK function, use a named range instead of an absolute range reference.

| B1 | | | | f_x | =RANK.AVG(A1,Data,1) |

	A	B	C	D	E	F	G	H	I
1	1455	4							
2	1534	5							
3	1108	1							
4	2109	6.5							
5	2109	6.5							
6	2917	9							
7	1427	3							
8	1220	2							
9	2298	8							
10									

Explanation: the named range Data refers to the range A1:A9.

edunet foundation

RANK

5. We want to show the top 3 results. Enter the value 3 into cell D2.

D2				fx	3				
	A	B	C	D	E	F	G	H	I	J
1	Student	Score		Top		Student	Score			
2	Linda	86		3						
3	Joseph	64								
4	Susan	80								
5	David	61								
6	Jessica	95								
7	Betty	79								
8	Thomas	91								
9										

6. To find the third largest score, use the LARGE function below. The named range Scores refers to the range B2:B8.

F2				fx	=LARGE(Scores,D2)				
	A	B	C	D	E	F	G	H	I	J
1	Student	Score		Top		Student	Score			
2	Linda	86		3		86				
3	Joseph	64								
4	Susan	80								
5	David	61								
6	Jessica	95								
7	Betty	79								
8	Thomas	91								
9										

RANK

7. The FILTER function below shows all students with a score greater than or equal to the third largest score.

	A	B	C	D	E	F	G	H	I	J
						=FILTER(A2:B8,Scores>=LARGE(Scores,D2))				
1	Student	Score		Top		Student	Score			
2	Linda	86		3		Linda	86			
3	Joseph	64				Jessica	95			
4	Susan	80				Thomas	91			
5	David	61								
6	Jessica	95								
7	Betty	79								
8	Thomas	91								
9										

Note: the FILTER function, entered into cell F2, fills multiple cells. Wow!

8. Finally, add the SORT function to sort by the second column (second argument), in descending order (third argument).

	A	B	C	D	E	F	G	H	I	J
						=SORT(FILTER(A2:B8,Scores>=LARGE(Scores,D2)),2,-1)				
1	Student	Score		Top		Student	Score			
2	Linda	86		3		Jessica	95			
3	Joseph	64				Thomas	91			
4	Susan	80				Linda	86			
5	David	61								
6	Jessica	95								
7	Betty	79								
8	Thomas	91								
9										

Note: use 1 to sort in ascending order, use -1 to sort in descending order.

edunet
foundation

RANK

- 9. This RANK formula works for every N. For example, enter the value 5 into cell D2 to show the top 5 results.

- Note: this dynamic array formula, entered into cell F2, fills multiple cells. This behavior in Excel 365/2021 is called spilling.

F2			×	✓	fx	=SORT(FILTER(A2:B8,Scores>=LARGE(Scores,D2)),2,-1)				
	A	B	C	D	E	F	G	H	I	J
1	Student	Score		Top		Student	Score			
2	Linda	86		5		Jessica	95			
3	Joseph	64				Thomas	91			
4	Susan	80				Linda	86			
5	David	61				Susan	80			
6	Jessica	95				Betty	79			
7	Betty	79								
8	Thomas	91								
9										

PERCENTILES AND QUARTILES

- 1. Use the PERCENTILE function shown below to calculate the 30th percentile. Excel returns the value 12.7. This means that 30% (6 out of 20) of the scores are lower or equal to 12.7

- Note: The second argument of the PERCENTILE function must be a decimal number between 0 and 1. Excel uses a slightly different algorithm to calculate percentiles and quartiles than you find in most statistics books. If you're interested, download the Excel file.

PERCENTILES AND QUARTILES

- 2. Use the PERCENTILE function shown below to calculate the 90th percentile. Excel returns the value 61.7. This means that 90% (18 out of 20) of the scores are lower or equal to 61.7

PERCENTILES AND QUARTILES

- 3. Use the QUARTILE function shown below to calculate the 1st quartile. Excel returns the value 11.25. This means that 25% (5 out of 20) of the scores are lower or equal to 11.25

PERCENTILES AND QUARTILES

Note: The second argument of the QUARTILE function must be a number between 0 and 4. PERCENTILE(A1:A20,0.25) gives the exact same result. It's up to you which function you would like to use for quartiles. Below you can find an overview.

Minimum Value	PERCENTILE(A1:A20,0)	QUARTILE(A1:A20,0)	MIN(A1:A20)
1st Quarter	PERCENTILE(A1:A20,0.25)	QUARTILE(A1:A20,1)	
Median	PERCENTILE(A1:A20,0.50)	QUARTILE(A1:A20,2)	MEDIAN(A1:A20)
3rd Quarter	PERCENTILE(A1:A20,0.75)	QUARTILE(A1:A20,3)	
Maximum Value	PERCENTILE(A1:A20,1)	QUARTILE(A1:A20,4)	MAX(A1:A20)

FREQUENCY

- The FREQUENCY function in Excel calculates how often values occur within the ranges you specify in a bin table. You can also use the COUNTIFS function to create a frequency distribution.

1. First, enter the bin numbers (upper levels) in the range C4:C8.

	A	B	C	D	E	F	G	H
1	Number of students							
2	22							
3	29							
4	40		20					
5	30		25					
6	48		30					
7	24		35					
8	21		40					
9	19							
10	24							
11	22							
12	25							
13	52							
14	35							
15	40							
16	31							
17	37							
18	21							
19	23							
20								

FREQUENCY

2. Select the range D4:D9 (extra cell), enter the FREQUENCY function shown below (without the curly braces) and finish by pressing CTRL + SHIFT + ENTER.

Note: the formula bar indicates that this is an array formula by enclosing it in curly braces {}. To delete this array formula, select the range D4:D9 and press Delete.

| D4 | | fx | {=FREQUENCY(A2:A19,C4:C8)} |

	A	B	C	D	E	F	G	H
1	Number of students							
2	22							
3	29							
4	40		20	1				
5	30		25	8				
6	48		30	2				
7	24		35	2				
8	21		40	3				
9	19			2				
10	24							
11	22							
12	25							
13	52							
14	35							
15	40							
16	31							
17	37							
18	21							
19	23							
20								

FREQUENCY

3. If you have Excel 365 or Excel 2021, simply select cell D4, enter the FREQUENCY function and press Enter. Bye bye curly braces.

Note: the FREQUENCY function, entered into cell D4, fills multiple cells. Wow! This behavior in Excel 365/2021 is called spilling.

	A	B	C	D	E	F	G	H
1	Number of students							
2	22							
3	29							
4	40		20	1				
5	30		25	8				
6	48		30	2				
7	24		35	2				
8	21		40	3				
9	19			2				
10	24							
11	22							
12	25							
13	52							
14	35							
15	40							
16	31							
17	37							
18	21							
19	23							
20								

D4 =FREQUENCY(A2:A19,C4:C8)

FREQUENCY

4. Hide the column with the bin numbers (upper levels) and insert a column with proper bin labels.

Explanation: 1 value is less than or equal to 20, 8 values are greater than or equal to 21 and less than or equal to 25, etc. At step 2, we selected the range D4:D9 (instead of the range D4:D8). As a result, the FREQUENCY function also counts the number of values that are greater than 40.

	A	B	C	E	F	G	H	I
1	Number of students							
2	22							
3	29		Bin	Count				
4	40		0-20	1				
5	30		21-25	8				
6	48		26-30	2				
7	24		31-35	2				
8	21		36-40	3				
9	19		40+	2				
10	24							
11	22							
12	25							
13	52							
14	35							
15	40							
16	31							
17	37							
18	21							
19	23							
20								

FREQUENCY

5. You can also use the Analysis Toolpak to create a histogram.

Bin	Frequency
0-20	1
21-25	8
26-30	2
31-35	2
36-40	3
40+	2
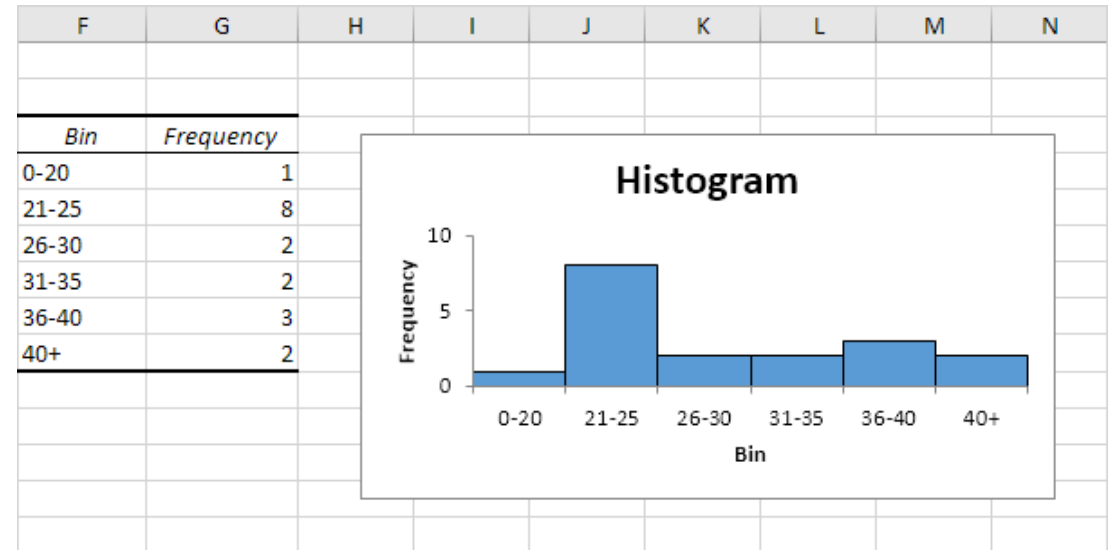
FREQUENCY

6. Change the bin numbers. Select the range D4:D9 (no extra cell), enter the FREQUENCY function shown below (without the curly braces) and finish by pressing CTRL + SHIFT + ENTER.

Explanation: the last bin number is greater than or equal to the maximum value (52). As a result, we don't need an extra cell to count the number of values that are greater than 60.

D4			×	✓	f_x	{=FREQUENCY(A2:A19,C4:C9)}

	A	B	C	D	E	F	G	H
1	Number of students							
2	22							
3	29							
4	40		10	0				
5	30		20	1				
6	48		30	10				
7	24		40	5				
8	21		50	1				
9	19		60	1				
10	24							
11	22							
12	25							
13	52	←						
14	35							
15	40							
16	31							
17	37							
18	21							
19	23							
20								

FREQUENCY

7. You can also use the COUNTIFS function to create a frequency distribution.

Explanation: the COUNTIFS function in Excel counts cells based on two or more criteria. The COUNTIFS function shown above has 2 range/criteria pairs. The & operator joins ">=" with the value in cell C4 and "<=" with the value in cell D4. You can easily copy this formula to the other cells.

| E4 | f_x | =COUNTIFS($A$2:$A$19,">="&C4,$A$2:$A$19,"<="&D4) |

	A	B	C	D	E	F	G	H
1	Number of students							
2	22							
3	29		Min	Max	Count			
4	40		0	10	0			
5	30		11	20	1			
6	48		21	30	10			
7	24		31	40	5			
8	21		41	50	1			
9	19		51	60	1			
10	24							
11	22							
12	25							
13	52							
14	35							
15	40							
16	31							
17	37							
18	21							
19	23							
20								

ARRAY FORMULAS

DYNAMIC ARRAYS

1. Fix the reference to cell E2 by placing a $ symbol in front of the column letter and row number.

2. To quickly copy the formula in cell C2 to the other cells, select cell C2, click on the lower right corner of cell C2 and drag it down to cell C7.

DYNAMIC ARRAYS

Check:

3. If you have Excel 365 or Excel 2021, simply enter the dynamic array formula shown below into cell C2.

Explanation: the absolute reference ($E$2) stays the same, while the relative reference (B2) changes to B3, B4, B5, B6 and B7.

DYNAMIC ARRAYS

4. Press Enter.

Explanation: this dynamic array (DA) formula, entered into cell C2, fills multiple cells. This behavior in Excel 365/2021 is called spilling. If you select a cell in the range C2:C7, a blue border appears.

5. If you select a cell in the range C3:C7, the formula in the formula bar will be greyed out.

DYNAMIC ARRAYS FUNCTIONS - SORT

1. The SORT function below sorts by the second column, in ascending order.

F2			f_x	=SORT(A2:D15,2)				
	A	B	C	D	E	F	G	H	I
1	Last Name	Sales	Country	Quarter		Last Name	Sales	Country	Quarter
2	Smith	$16,753.00	UK	Qtr 3		Jones	$1,390.00	USA	Qtr 3
3	Johnson	$14,808.00	USA	Qtr 4		Brown	$3,255.00	USA	Qtr 2
4	Williams	$10,644.00	UK	Qtr 2		Brown	$4,865.00	USA	Qtr 4
5	Jones	$1,390.00	USA	Qtr 3		Jones	$7,433.00	UK	Qtr 1
6	Brown	$4,865.00	USA	Qtr 4		Jones	$9,213.00	USA	Qtr 4
7	Williams	$12,438.00	UK	Qtr 1		Johnson	$9,339.00	UK	Qtr 2
8	Johnson	$9,339.00	UK	Qtr 2		Smith	$9,698.00	USA	Qtr 1
9	Smith	$18,919.00	USA	Qtr 3		Williams	$10,644.00	UK	Qtr 2
10	Jones	$9,213.00	USA	Qtr 4		Williams	$12,438.00	UK	Qtr 1
11	Jones	$7,433.00	UK	Qtr 1		Johnson	$14,808.00	USA	Qtr 4
12	Brown	$3,255.00	USA	Qtr 2		Williams	$14,867.00	USA	Qtr 3
13	Williams	$14,867.00	USA	Qtr 3		Smith	$16,753.00	UK	Qtr 3
14	Williams	$19,302.00	UK	Qtr 4		Smith	$18,919.00	USA	Qtr 3
15	Smith	$9,698.00	USA	Qtr 1		Williams	$19,302.00	UK	Qtr 4
16									

DYNAMIC ARRAYS FUNCTIONS - FILTER

2. The FILTER function below extracts all USA records.

DYNAMIC ARRAYS FUNCTIONS - RANDARRAY

3. The RANDARRAY function below generates random decimal numbers between 0 and 1. The array below consists of 5 rows and 2 columns.

DYNAMIC ARRAYS FUNCTIONS - SEQUENCE

4. The SEQUENCE function below generates a two-dimensional array. Rows = 7, Columns = 4, Start = 0, Step = 5.

DYNAMIC ARRAYS FUNCTIONS - UNIQUE

5. The UNIQUE function below (with no extra arguments) extracts unique values.

DYNAMIC ARRAYS FUNCTIONS - XLOOKUP

6. The XLOOKUP function below looks up the ID and returns the first name, last name and salary.

DYNAMIC ARRAYS FUNCTIONS - SORTBY

7. The SORTBY function sorts a range based on the values in a corresponding range.

Note: use 1 to sort in ascending order, use -1 to sort in descending order.

ARRAY MANIPULATION - VSTACK

- The VSTACK function in Excel 365 vertically stacks multiple arrays into a single array.

- Note: the VSTACK function, entered into cell B5, fills multiple cells. This behavior is called spilling.

ARRAY MANIPULATION - HSTACK

■ The HSTACK function in Excel 365 horizontally stacks multiple arrays into a single array.

ARRAY MANIPULATION - TOROW

- The TOROW function in Excel 365 converts a 2D array into a single row.

ARRAY MANIPULATION - TOCOL

■ The TOCOL function in Excel 365 converts a 2D array into a single column.

ARRAY MANIPULATION - WRAPROWS

- The WRAPROWS function converts a 1D array into a 2D array. WRAPROWS starts a new row after a specified number of elements (second argument).

ARRAY MANIPULATION - WRAPCOLS

- The WRAPCOLS function converts a 1D array into a 2D array. WRAPCOLS starts a new column after a specified number of elements (second argument).

THANK YOU