MODULE 3

Statistical Functions

TOPICS

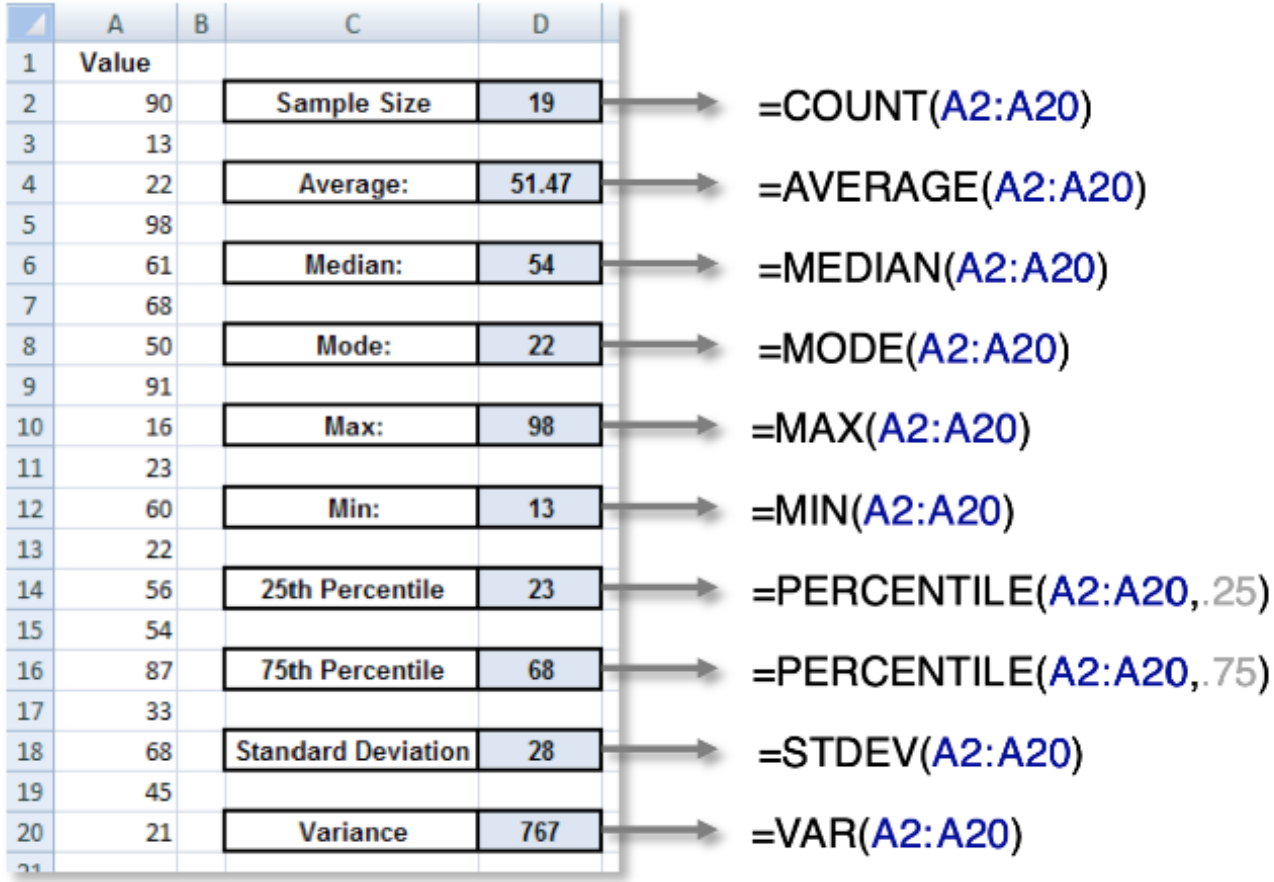
- Basic Stats Functions
- SMALL/LARGE, RANK, PERCENTRANK
- RAND(), RANDBETWEEN
- SUMPRODUCTS
- COUNTIF/SUMIF/AVERAGEIF
- EXERCISES

Commonly Used Statistical Functions

Function	Output
ABS	The absolute value of a number
AVERAGE	The average or arithmetic mean for a group of numbers
COUNT	The number of cell locations in a range that contain a numeric character
COUNTA	The number of cell locations in a range that contain a text or numeric character
MAX	The highest numeric value in a group of numbers
MEDIAN	The middle number in a group of numbers (half the numbers in the group are higher than the median and half the numbers in the group are lower than the median)
MIN	The lowest numeric value in a group of numbers
MODE	The number that appears most frequently in a group of numbers
PRODUCT	The result of multiplying all the values in a range of cell locations
SQRT	The positive square root of a number
STDEV.S	The standard deviation for a group of numbers based on a sample
SUM	The total of all numeric values in a group

Basic Stats Functions

The Count, Average, Median, Mode, Max/Min, Percentile and Standard Deviation/Variance functions are used to perform basic calculations on a data array



	A	B	C	D	
1	Value				
2	90		Sample Size	19	=COUNT(A2:A20)
3	13				
4	22		Average:	51.47	=AVERAGE(A2:A20)
5	98				
6	61		Median:	54	=MEDIAN(A2:A20)
7	68				
8	50		Mode:	22	=MODE(A2:A20)
9	91				
10	16		Max:	98	=MAX(A2:A20)
11	23				
12	60		Min:	13	=MIN(A2:A20)
13	22				
14	56		25th Percentile	23	=PERCENTILE(A2:A20,.25)
15	54				
16	87		75th Percentile	68	=PERCENTILE(A2:A20,.75)
17	33				
18	68		Standard Deviation	28	=STDEV(A2:A20)
19	45				
20	21		Variance	767	=VAR(A2:A20)

RANK / SMALL / LARGE

	A
1	Value
2	90
3	13
4	22
5	98
6	61
7	68
8	50

RANK(A2,A2:A8) = 2
RANK(A3,A2:A8) = 7 (lowest)
RANK(A4,A2:A8) = 6
RANK(A5,A2:A8) = 1 (highest)
RANK(A6,A2:A8) = 4
RANK(A7,A2:A8) = 3
RANK(A8,A2:A8) = 5

The RANK function returns the rank of a particular number among a list of values

The SMALL/LARGE functions return the nth smallest/largest values within an array

	A
1	Value
2	90
3	13
4	22
5	98
6	61
7	68
8	50

LARGE(A2:A8,2) = 90
(the 2nd largest number in the array is 90)

SMALL(A2:A8,3) = 50
(the 3rd smallest number in the array is 50)

PERCENT RANK

	A	B
1	Value	Percent Rank
2	2,717	18%
3	3,485	24%
4	5,202	76%
5	3,612	29%
6	4,432	59%
7	2,699	12%
8	4,585	65%
9	6,003	94%
10	4,820	71%
11	2,550	6%
12	5,795	88%
13	4,240	41%
14	6,827	100%
15	4,359	53%
16	2,320	0%
17	5,775	82%
18	4,241	47%
19	3,966	35%

PERCENTRANK returns the rank of a value as a percentage of a given array or dataset

=PERCENTRANK(array, x)

What range of data are you looking at?

Which value within the range are you looking at?

PERCENTRANK(A14,$A$2:$A$19) = 100% (highest)

PERCENTRANK(A14,$A$2:$A$19) = 0% (lowest)

RAND / RAND BETWEEN

RAND() and RANDBETWEEN act like random number generators in Excel:

	A	B	C	D	E
1	0.5173	0.4091	0.7560	0.9012	0.2167
2	0.0906	0.2317	0.0906	0.5856	0.8646
3	0.1544	0.8240	0.4279	0.8782	0.7795
4	0.0097	0.0872	0.7740	0.9137	0.7815
5	0.2089	0.7028	0.0449	0.8173	0.9983
6	0.0761	0.4388	0.4056	0.5639	0.0668

The RAND() function returns a random value between 0 and 1 (to 15 digits)

The RANDBETWEEN function returns an integer between two values that you specify

	A	B	C	D	E
1	83	23	64	62	92
2	59	45	40	50	91
3	24	37	70	30	32
4	54	85	69	55	3
5	73	12	36	53	2
6	29	72	68	59	99

=RANDBETWEEN(0,100)

SUM PRODUCT

The SUMPRODUCT formula multiplies corresponding cells from multiple arrays and returns the sum of the products (Note: all arrays must have the same dimensions)

=SUMPRODUCT(array1, array2 … array_N)

Example: Total Revenue

	A	B	C	D
1	Product	Quantity	Price	Revenue
2	Apple	2	$0.50	$1.00
3	Banana	4	$1.00	$4.00
4	Orange	3	$0.80	$2.40
5	Total			$7.40

Without using SUMPRODUCT, you could multiply quantity*price in each row and sum the products

	A	B	C	D
1	Product	Quantity	Price	Revenue
2	Apple	2	$0.50	
3	Banana	4	$1.00	
4	Orange	3	$0.80	
5	Total			$7.40

SUMPRODUCT(B2:B4,C2:C4) = $7.40

SUM PRODUCT

SUMPRODUCT is often used with filters to calculate products only for rows that meet certain criteria:

	A	B	C	D
1	Store	Product	Quantity	Price
2	Stop & Shop	Apple	2	$0.50
3	Shaws	Banana	4	$1.00
4	Market Basket	Banana	3	$1.00
5	Trader Joe's	Pineapple	8	$2.50
6	Stop & Shop	Orange	2	$0.80
7	Shaws	Apple	1	$0.50
8	Market Basket	Apple	5	$0.50
9	Trader Joe's	Banana	6	$1.00
10	Market Basket	Pineapple	3	$2.50
11	Trader Joe's	Orange	8	$0.80
12	Stop & Shop	Pineapple	3	$2.50
13	Shaws	Pineapple	5	$2.50
14	Stop & Shop	Banana	2	$1.00
15	Shaws	Orange	6	$0.80
16	Market Basket	Orange	7	$0.80
17	Trader Joe's	Apple	3	$0.50

Quantity of goods sold at Shaws:

SUMPRODUCT((A2:A17="Shaws")*C2:C17) = 16

Total revenue from Shaws:

SUMPRODUCT((A2:A17="Shaws")*C2:C17*D2:D17) = $21.80

Revenue from apples sold at Shaws:

SUMPRODUCT((A2:A17="Shaws")*(B2:B17="Apple")*C2:C17*D2:D17) = $0.50

PRO TIP:

When you add filters to a SUMPRODUCT, you need to change the commas to multiplication signs

SUM PRODUCT

Great, but how does it really work?

SUMPRODUCT((A2:A17="Shaws")*(B2:B17="Apple")*C2:C17*D2:D17) = $0.50

	A	B	C	D
1	Store	Product	Quantity	Price
2	Stop & Shop	Apple	2	$0.50
3	Shaws	Banana	4	$1.00
4	Market Basket	Banana	3	$1.00
5	Trader Joe's	Pineapple	8	$2.50
6	Stop & Shop	Orange	2	$0.80
7	Shaws	Apple	1	$0.50
8	Market Basket	Apple	5	$0.50
9	Trader Joe's	Banana	6	$1.00
10	Market Basket	Pineapple	3	$2.50
11	Trader Joe's	Orange	8	$0.80
12	Stop & Shop	Pineapple	3	$2.50
13	Shaws	Pineapple	5	$2.50
14	Stop & Shop	Banana	2	$1.00
15	Shaws	Orange	6	$0.80
16	Market Basket	Orange	7	$0.80
17	Trader Joe's	Apple	3	$0.50

What YOU see

What EXCEL sees

When you apply a condition or filter to a column, Excel translates those cells as **0's *(if false)* and 1's *(if true)***

If you multiply all four columns, ONLY ROWS THAT SATISFY ALL CONDITIONS WILL PRODUCE A NON-ZERO SUM

	A	B	C	D
1	Store	Product	Quantity	Price
2	0	1	2	$0.50
3	1	0	4	$1.00
4	0	0	3	$1.00
5	0	0	8	$2.50
6	0	0	2	$0.80
7	1	1	1	$0.50
8	0	1	5	$0.50
9	0	0	6	$1.00
10	0	0	3	$2.50
11	0	0	8	$0.80
12	0	0	3	$2.50
13	1	0	5	$2.50
14	0	0	2	$1.00
15	1	0	6	$0.80
16	0	0	7	$0.80
17	0	1	3	$0.50

COUNTIF / SUMIF / AVERAGEIF

The COUNTIF, SUMIF, and AVERAGEIF formulas calculate a sum, count, or average based on specific criteria

	A	B
1	Name	Age
2	George	90
3	Maria	13
4	Ryan	22
5	Tim	98
6	George	61
7	Tim	68
8	Tim	50
9	Maria	91
10	George	16
11	Maria	23
12	Tim	60
13	Ryan	22
14	Maria	56
15	George	54
16	George	87
17	Ryan	33
18	Ryan	68
19	Ryan	45
20	George	21

=COUNTIF(range, criteria)

=SUMIF(range, criteria, sum_range)

=AVERAGEIF(range, criteria, average_range)

Which cells need to match your criteria?

Under what condition do I want to sum, count, or average?

Where are the values that I want to sum or average?

COUNTIF(B2:B20,22) = 2

SUMIF(A2:A20,"Ryan",B2:B20) = 190

SUMIF(A2:A20,"<>Tim",B2:B20) = 702

AVERAGEIF(A2:A20,"Maria",B2:B20) = 45.75

COUNTIFS / SUMIFS / AVERAGEIFS

COUNTIFS, SUMIFS, and **AVERAGEIFS** are used when you want to evaluate a count, sum, or average based on multiple conditions or criteria

=COUNTIFS(criteria_range1, criteria1, criteria_range2 , criteria2…)

=SUMIFS(sum_range, criteria_range1, criteria1, criteria_range2 , criteria2…)

=AVERAGEIFS(average_range, criteria_range1, criteria1, criteria_range2 , criteria2…)

	A	B	C	D
1	Month	Tactic	Campaign	Clicks
2	Jan	Search	Google	166
3	Jan	Search	MSN	263
4	Jan	Display	Contextual	289
5	Jan	Display	Retargeting	137
6	Feb	Search	Google	124
7	Feb	Search	MSN	311
8	Feb	Display	Contextual	350
9	Feb	Display	Retargeting	384
10	Mar	Search	Google	168
11	Mar	Search	MSN	358
12	Mar	Display	Contextual	347
13	Mar	Display	Retargeting	390

COUNTIFS(B2:B13,“Search”, D2:D13,“>200”) = 3

SUMIFS(D2:D13, A2:A13,“Feb”,B2:B13,“Display”) = 734

AVERAGEIFS(D2:D13, A2:A13,“Jan”,C2:C13,“MSN”) = 263

PRO TIP:

If you use < or >, you need to add quotation ***marks as you would with text (i.e. “>200”)***