

</> FUNCTIONS

IF

Checks if something is true or false, then shows one result if it is and another if it isn't.

`=IF(A1>50, "Pass", "Fail")`

If the score in A1 is above 50, shows "Pass"; otherwise "Fail".

`=IF(C2<>$F$7;"negative";"positive")`

If C2 is different from F7, shows "negative"; if they match, shows "positive".

AVERAGE

Gives you the average of a group of numbers.

`=AVERAGE(B2:B101)`

Averages every value from B2 down to B101.

STDEV.S

Shows how spread out your numbers are. Low means they're close together, high means they're all over the place.

We use STDEV.S (and not STDEV.P) because we always work with a sample, a selected set, not the full population.

`=STDEV.S(B2:B101)`

Spread of the values in that range.

COVARIANCE.S

Checks if two things move together. For example, when it gets hotter, do ice cream sales go up too?

We use COVARIANCE.S (and not COVARIANCE.P) because we always work with a sample, not the full population.

`=COVARIANCE.S(C2:C101;D2:D101)`

Sample covariance between the two ranges.

COUNTIFS

Counts cells that meet **every** condition you give it. A row only counts if it passes all of them.

`=COUNTIFS(A1:A10, "Yes", B1:B10, ">100")`

Counts rows where A is "Yes" and B is over 100.

Only rows 2 and 5 match both conditions, so the count is 2.

	A	B	C
1	Attended?	Sales	Result
2	Yes	150	2
3	No	200	
4	Yes	90	=COUNTIFS(A2:A6, "Yes", B2:B6, ">100")
5	Yes	120	
6	No	50	



Binned data: this counts every value greater than F2 and up to G2 with `=COUNTIFS($C$2:$C$101; ">"&F2; $C$2:$C$101; "<="&G2)`. It's the standard way to count a frequency bin.

SUMPRODUCT

Multiplies matching pairs from two ranges, then adds up the results. Think of it like multiplying quantity by price for each item on a receipt, then totalling the bill.

`=SUMPRODUCT(F2:F7, G2:G7)`

Binned data mean. F holds the relative frequency, G holds the midpoint of each bin.

Each frequency × its midpoint, summed (0.2×5 + 0.5×15 + 0.3×20), gives the weighted mean.

	F	G	H
1	Rel. Freq	Midpoint	Mean
2	0.2	5	14.5
3	0.5	15	
4	0.3	20	
5			=SUMPRODUCT(F2:F4, G2:G4)



Binned data variance: `=SUMPRODUCT(G2:G7;G2:G7;F2:F7)` gives $\Sigma(\text{midpoint}^2 \times \text{frequency})$. Full variance = $(n/(n-1)) \times (\text{that result} - \text{mean}^2)$.

ABSOLUTE REFERENCING (\$)

Absolute Cell Referencing (\$)

A dollar sign (\$) locks part of a cell reference so it stays put when you copy or drag a formula. Without it, Excel quietly shifts every reference by however far you drag.

Why it matters: the formula in B2 multiplies the price by the VAT rate in C1. Drag it down to B3 without a \$, and Excel shifts C1 to C2, which is empty, so the formula breaks. Lock it as `$C$1` and every row keeps pointing at the same VAT cell.

	A	B	C	D
1	Price	Total (wrong)	20%	Total (correct)
2	100	=A2 * C1		=A2 * \$C\$1
3	150	=A3 * C2 		=A3 * \$C\$1



Reading the \$ sign: it locks whatever comes right after it. `$C$1` locks the column *and* the row. `$C1` locks only the column. `C$1` locks only the row.

Example

All four combinations: same formula in B2, dragged one column right (C2) and one row down (B3). Pick a type below, then click a formula cell to see exactly how it's calculated.

A1 — RELATIVE

	A	B	C
1	10	20	30
2	40	=A1 (10)	=B1 (20)
3		=A2 (40)	

\$A\$1 — ABSOLUTE

	A	B	C
1	10	20	30
2	40	=A\$1 (10)	=A\$1 (10)
3		=A\$1 (10)	

\$A1 — COLUMN LOCKED

	A	B	C
1	10	20	30
2	40	=\$A1 (10)	=\$A1 (10)
3		=\$A2 (40)	

A\$1 — ROW LOCKED

	A	B	C
1	10	20	30
2	40	=A\$1 (10)	=B\$1 (20)
3		=A\$1 (10)	



Exam tip: if you're asked to fix a formula like `=COUNTIFS (C2 : C101 ; F2)` so it can be dragged down safely, lock the data range like this: `=COUNTIFS ($C$2 : $C$101 ; F2)`.

DATA ANALYSIS TOOLPAK

Regression

Shows how one thing (X) predicts another (Y). Used for the least-squares line and for making predictions.

SETUP

Input Y Range The outcome you want to predict.

Input X Range The variable you're using to predict it.

READING THE OUTPUT

Multiple R How strongly X and Y are correlated.

R Square % of the change in Y explained by X. Higher = better fit.

Observations Number of data rows used.

Intercept The `a` in `a+bx`. It's the starting value, what Y is when X is 0.

X Variable The `b` in `a+bx`. It's the slope, how much Y changes each time X goes up by 1.