Is there a relationship?

Correlation in your internal assessment: what to do, what to write, how to evaluate

Use this when you have measured **two variables at the same sites** (e.g. distance downstream and channel width). Your calculator does the arithmetic. The marks come from **choosing the right test, stating it properly, and interpreting and evaluating the result**.

Step 1 Plan before you look at the data

- **Variables.** x = the one you think does the influencing (e.g. distance from source). y = the one that responds (e.g. channel width).
- **Hypotheses.** H_0: there is no correlation between x and y.
  H_1: y increases as x increases *(one-tailed)*, **or** x and y are correlated *(two-tailed)*.
- **One-tailed** only if geographical theory predicts the direction (e.g. the Bradshaw model). Decide now; never switch after seeing the result.
- **Significance level** $\alpha = 0.05$. **Count your pairs,** n**.** Fewer than 5: no test possible. 5–7: even a clear pattern may not be significant.

Step 2 Draw the scatter graph and look at it

Put x on the horizontal axis and y on the vertical axis, with a title and units.

What you see	What to do
Points close to a **straight line**	Pearson may be possible (Step 4)
A **curve that keeps going one way**	Spearman
One or two points far away from the rest	Spearman; explain the outlier, never delete it
Points **bunched at one end** or unevenly spread	Spearman
A **shapeless cloud**	Test anyway (expect no significance); **no** trend line
A **U-shape or hump** (up then down)	**No** correlation test; describe and explain the shape

Step 3 Check the shape: is it a straight line? Are there outliers?

Is it linear? The ruler test. Lay a ruler along the general trend.

- **Linear:** points fall either side of the ruler *all the way along*, roughly equally.
- **Not linear:** points sit on one side at both ends and the other side in the middle (a curve), or the pattern bends or levels off.
- *Optional GDC check:* after linear regression, plot the residuals (RESID) against x. Random scatter = linear; a curve or U-shape = not linear.

Are there outliers? There are two kinds.

(a) Extreme in one variable. Use 1-Var Stats on your GDC to get Q_1 and Q_3. A value is an outlier if

$$\text{value} < Q_1 - 1.5 \times \text{IQR} \quad \text{or} \quad \text{value} > Q_3 + 1.5 \times \text{IQR}, \qquad \text{IQR} = Q_3 - Q_1.$$

Check each variable separately.

(b) Off the pattern. A site whose x and y are each ordinary, but whose *combination* does not fit the trend. The IQR rule will not find it: you must spot it on the scatter graph.

What to do with an outlier: first check it is not a recording error (e.g. a flowmeter count of 2664 among values near 200, or a value in the wrong units). Report any error you find; do not silently fix it. If the value is genuine: **keep it**, label it on the graph, use Spearman, and **explain it** geographically. You may also give the coefficient with and without it to show how much it matters, but it stays in your data table.

Step 4 Choose: Pearson or Spearman?

Pearson (r): only if *all* are true:

- both variables are measurements on a scale (m, m s^{-1}, ℃)
- the pattern is roughly a straight line
- no extreme outliers
- roughly 10 or more points

Spearman's rank (r_s): everything else.

- Asks only: does y tend to rise (or fall) as x rises?
- Copes with curves, outliers, uneven spread.
- **If in doubt, and always for small n: Spearman.**

Geography: Spearman is the usual technique.

Say why you chose it in one sentence, referring to your scatter graph and your outlier check.

Step 5 Calculate and decide using the p-value

On your GDC run the **linear regression t-test**. It gives both the coefficient r and the p-value. For Spearman, enter the *ranks* of each variable, not the raw values; the r it reports is then r_s.

TI-84 Plus CE	stat → TESTS → LinRegTTest. Xlist L1, Ylist L2, choose β & ρ: $\neq 0$ (two-tailed), > 0 or < 0 (one-tailed). Read r and p.
TI-Nspire CX	menu → Statistics → Stat Tests → Linear Reg t Test; choose the alternative hypothesis. Read r and PVal.
Casio fx-CG50	Statistics → TEST → REG → LinearReg t-test; choose β & ρ $\neq$, $<$ or $>$. Read r and p.

The alternative you choose on the calculator must match H_1: $\neq 0$ for two-tailed, > 0 for a predicted positive correlation, < 0 for a predicted negative one.

Decide by comparing p with $\alpha = 0.05$:

- $p < 0.05$**:** reject H_0: there is evidence of a correlation (in the predicted direction, if one-tailed).
- $p \geq 0.05$**:** do not reject H_0: no evidence of a correlation. This does *not* prove there is no relationship.

What p means: if there were really no correlation, p is the chance of getting a coefficient at least this far from zero. It is **not** the probability that H_0 is true, and a small p does **not** mean a strong relationship.

Precision: give r or r_s to 3 decimal places and p to 3 decimal places (or $p < 0.001$). Name the calculator test in your method.

For Spearman on ranks the calculator p-value is a close approximation, which is acceptable for IA purposes. You may check it with an online calculator such as *Social Science Statistics* (listed in the ESS guide, Tool 2, p. 27), but cite it:

> Social Science Statistics. (n.d.). *Spearman's rho (ρ) rank correlation calculator*. Retrieved [date], from `https://www.socscistatistics.com/tests/spearman/`
> Pearson version: `https://www.socscistatistics.com/tests/pearson/`

Step 6 Interpret the result: say what it means for the river

1. **State the result formally:** test, n, one/two-tailed, coefficient, p-value, α, decision on H_0.
2. **Direction:** positive (y rises with x) or negative (y falls as x rises)?
3. **Strength** (as a guide): about 0.7 or more strong, 0.4–0.7 moderate, below 0.4 weak. Significance and strength are *different*: with few sites a strong coefficient can have $p > 0.05$; with thousands of data points a tiny one can have $p < 0.05$.
4. **Link to theory:** does it match what the model (e.g. Bradshaw) predicts? Why, or why not?
5. **Anomalies:** which sites sit away from the pattern, and what local factor might explain them? Explain; never delete.
6. **Cause:** correlation is not causation. What actually drives the change (e.g. discharge, tributaries, geology)?

Step 7 Evaluate: how much can you trust the result?

- **Sample size:** how many sites? Would one site changing rank change the decision?
- **Independence:** are some sites very close together, so they are really measuring the same stretch of river?
- **Measurement reliability:** did other groups get different values at the same sites? Units, instruments, recording errors?
- **Other variables:** what else changes downstream that could produce the same pattern?
- **Improvement:** a specific change (more or better-spaced sites, repeat readings) *and* how it would make the conclusion more reliable.

Mistakes that lose marks

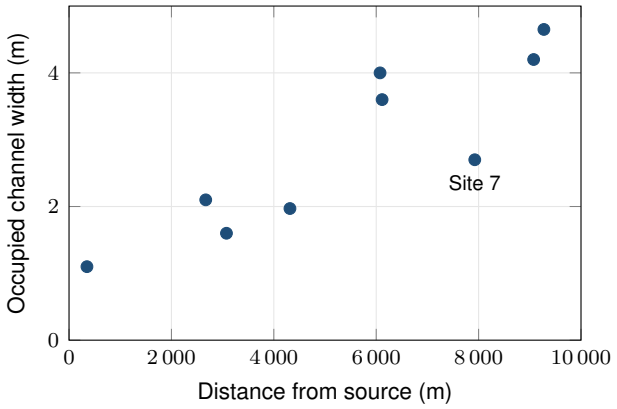
- A **trend line through a shapeless cloud**, or a trend line that is never discussed.
- Treating **Excel's R^2 as a test**. It describes the fit; it does not test significance.
- Claiming significance **without** stating the test, n, p and α.
- Saying the test **proves** a cause, or that not rejecting H_0 **proves** there is no relationship.
- Running a test that **does not answer your question**.

Worked example: does channel width increase downstream?

Data: June 2026 fieldwork, **Group 1** sheet, column C (distance from source, m) and column G (occupied channel width, m), 9 sites, values as recorded.

Plan. H_0: no correlation between distance and width. H_1: width increases downstream (**one-tailed**, as the Bradshaw model predicts). $\alpha = 0.05$, $n = 9$.
Checks. *Linear?* Points lie either side of a ruler along the whole length: roughly linear. *Outliers?* IQR fences: width -1.69 to 7.57 m; distance -5569 to $16\,941$ m. No value is extreme in either variable. Site 7 is type (b): off the pattern, not extreme.
Choice. Spearman, because n is small and Site 7 sits off the pattern.
Result. GDC LinRegTTest on the ranks, $\rho > 0$: $r_s = 0.883$, $p = 0.001$ (3 d.p.). $p < 0.05$ and r_s is positive, so H_0 is rejected.

Fig. 1: Channel width downstream (Group 1)



Interpretation

There is a **strong, positive** and **significant** correlation ($r_s = 0.883$, $n = 9$, one-tailed, $p = 0.001 < 0.05$) between distance from source and channel width. Width rises from about 1 m near the source to about 4.5 m at Sites 8–9, as the Bradshaw model predicts: tributaries join and discharge increases, so the channel widens. Distance does not cause the widening itself; it stands in for discharge.

Anomaly: Site 7 (2.7 m) is narrower than Sites 5, 6 and 8. A local factor such as bedrock, reinforced banks or a measurement issue should be investigated and explained; the site stays in the analysis.

Evaluation

- **Independence:** Sites 5 and 6 are only 40 m apart and Sites 8 and 9 only 200 m apart, so there are effectively fewer than nine independent sites, which makes the result look stronger than it is.
- **Reliability:** other groups recorded quite different widths at the same sites (e.g. Site 7: 2.7 m in Group 1, 7.4 m in Group 2), so the width measurements themselves are uncertain.
- **Robustness:** $p = 0.001$ is far below 0.05, so the decision would survive one or two sites changing rank. But a small p does not remove the independence and reliability problems above.
- **Improvement:** evenly spaced sites along the whole river, with repeat width readings at each, would give independent data and show whether Site 7 is a real feature or a recording error.