

From Pattern to a Contextual Conclusion



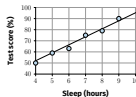
Key idea

State what tends to happen to the **response variable** as the explanatory variable changes. Use “tends to”, because points vary.

Conclusion

Students who sleep longer **tend to** have higher test scores in this sample. The strong positive association supports comparison, not a causal claim.

Conclusion = context + direction + uncertainty



Common mistake

“Sleep causes higher scores” is not justified by a scatter plot alone.

Line of Best Fit: Estimate, Then Interpret



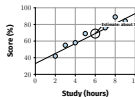
Key idea

A sensible fit follows the centre of a roughly linear cloud, with points balanced above and below. Read across from x , up to the line, then across to y .

Interpolation

At 6 hours, the line predicts about **70%**. This is interpolation because 6 is inside the observed range 2–9 hours. A reasonable graph-reading range is 68–72%.

A prediction is on the line, not necessarily a point



Interpolation Is Safer Than Extrapolation



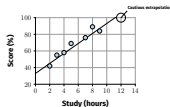
Key idea

Interpolation predicts within the observed x-range. **Extrapolation** extends beyond it, where the pattern may change. The farther outside, the weaker the evidence.

Cautious wording

A prediction at 12 hours extends three hours beyond the data. It may be unrealistic because scores cannot keep rising past 100%.

The fitted line is a model, not a guarantee



Variation, Outliers, and Groups Limit Conclusions



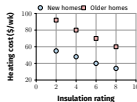
Key idea

Vertical gaps from the line show prediction error. Outliers can pull a fit; groups can make one overall line misleading. Investigate the grouping variable before combining data.

Lurking variable

Older and newer houses form separate groups. House age may influence both insulation rating and heating cost, so insulation alone is not a proven cause.

Separate groups may need separate models



Common mistake

A single overall line can describe neither group well, even when it looks convincing.