AP Statistics: Learn Like A Genius
Correlation vs. Causation — Oxford Method Video Guide
Unit 2, Lesson 5

Core Principle

AI should challenge you, not do the work for you!
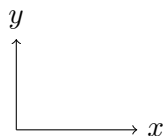This guide uses retrieval practice, Socratic questioning, and progressive complexity.
Remember: The struggle to understand is where learning happens.

Name: _____   **Date:** _____   **Pre-Confidence:** ___/10

Pre-Video: Knowledge Activation

Socratic: Before watching, activate what you know:

a. Without looking it up, what do you think correlation means? _____

b. Draw what a "strong positive relationship" might look like:

c. Complete: "Just because two things happen together doesn't mean..." _____

Video 1: Calculate the Correlation r [Total: 6 min]

Part A: Learning Objectives [0:00–0:26]

1. The three key questions we'll answer (fill in as you hear them):

   - How do we _____ the correlation r?
   - How do we _____ it?
   - What _____ can we learn from r alone?

Part B: Context and Hypothesis [0:27–1:21]

2. The hypothesis chain school leaders believe:

$$\text{Poverty} \rightarrow \underline{\qquad\qquad} \rightarrow \underline{\qquad\qquad}$$

3. **PAUSE VIDEO Make a prediction:** If we fix attendance, will test scores improve?

   - My prediction: _____ because _____
   - *Confidence (1–5):* ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5

Part C: Understanding the Data [1:22–2:00]

4. Sample size: $n =$ _____

5. Describe the relationship using three adjectives: _____, _____, _____

Part D: The Correlation Coefficient [2:01–3:32]

6. The correlation r gives the _____ and quantifies the _____ of a <u>linear</u> relationship.

7. Range of r: from _____ to _____ (inclusive)

8. **PAUSE VIDEO [2:34–2:58] Analyze the six scatterplots on screen:**

    a. Rank plots by $|r|$ from smallest to largest: _____

    b. Write "+" or "–" under each to mark the sign of r

    c. Circle any plot where a single outlier could be inflating/deflating r. Why? _____

    d. *Confidence (1–5):* ⚪ 1 ⚪ 2 ⚪ 3 ⚪ 4 ⚪ 5

9. **Socratic:** Instead of memorizing the formula, think conceptually:

    - Why might r be unitless? _____

    - What would r equal if all points fell perfectly on a line? _____

Part E: Interpreting Our Data [3:33–4:17]

10. Our correlation: $r =$ _____

11. In plain English: "Students who _____ tend to _____"

12. **Critical Question:** Does $r = 0.95$ prove attendance CAUSES better scores? _____

Fix the Statistical Error

A student says: "Because $r = 0.95$, the slope of the regression line must be 0.95 and raising attendance **causes** higher scores."
Fix both errors:
Error 1 (about slope): _____
Error 2 (about causation): _____
Confidence (1–5): ⚪ 1 ⚪ 2 ⚪ 3 ⚪ 4 ⚪ 5

Part F: The Surprising Result [4:18–5:09]

13. Three initiatives schools tried:

    - _____
    - _____
    - _____

14. Results:

- Attendance: _____
- Test scores: _____

15. **Socratic:** Why is this surprising given $r = 0.95$? (2-3 sentences)

_____

_____

Part G: Key Takeaways [5:28–5:55]

16. Check all TRUE statements about r:

☐ The sign tells direction of a linear relationship

☐ The magnitude tells strength of a linear relationship

☐ r alone reveals curved patterns

☐ r alone identifies outliers

☐ You need the scatterplot to assess form and unusual features

☐ r changes if we convert x from feet to meters

17. **Multi-Level Explanation:** Explain correlation r at three levels:

- To a 10-year-old: _____
- To your AP Stats classmate: _____
- To a statistician (use "linear," "strength," "direction"): _____

Video 2: Correlation Causation [Total: 7 min]

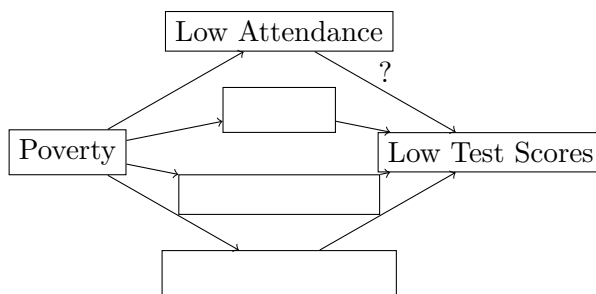## Part A: The Achievement Gap [0:00–1:05]

1. What is the "achievement gap"? _____

2. Important distinction: Achievement on exams _____

Part B: The Attendance Theory [1:06–2:44]

3. Two reasons higher-income students might have better attendance:

- _____
- _____

4. **PAUSE VIDEO** Before continuing, predict: Will attendance initiatives work? Why/why not?

_____

Part C: Alternative Causal Model [3:18–4:26]

5. **Draw the complete causal web.** Fill in the boxes and add arrows:

6. In one sentence, why might raising attendance alone fail? _____

7. *Confidence (1–5):* ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5

Part D: The Fundamental Principle [4:27–5:08]

8. Write the principle three ways:

- Mathematical: Correlation _____ Causation
- In your words: _____
- Real example: _____

Part E: Spurious Correlations [5:30–6:10]

9. Maine divorce rate correlates with per capita consumption of _____

10. This is obviously _____ (coincidental/causal/meaningful)

Two Truths and a Lie

(1) A large $|r|$ can occur even when there is no causal relationship.
(2) If $r = 0$, then X and Y are independent.
(3) r is unitless and unchanged by rescaling units of X or Y.

Circle the lie and correct it: _____

Part F: Call to Action [6:11–7:05]

11. What two things should you do about educational inequity?

- _____
- _____

12. Causal inference will be covered in Unit _____

Synthesis: Creating Deep Understanding

Create Your Understanding

A. Design Your Own Example
Create a scenario with strong correlation but no causation:

- Variable 1: _____

- Variable 2: _____

- Hidden cause: _____

- Sketch it:

B. Action Plan
You're a school administrator. Based on these videos:

1. What would you target instead of just attendance? _____

2. How would you test if it's actually causal? _____

Spaced Retrieval — Ask yourself tomorrow (no video

A. Define r and list two things it does *not* tell you that a scatterplot does.

B. Give a real-world example of strong correlation that is plausibly non-causal. Explain the lurking variable.

C. Sketch a scatterplot with $r \approx 0$ but a clear curved relationship. Why does r miss it?

Self-Assessment

Check Your Understanding:

☐ I can interpret r values

☐ I understand r's limitations

☐ I can identify lurking variables

☐ I know correlation causation

☐ I can spot spurious correlations

☐ I can think critically about data

Reflection:

- Post-video confidence: ___/10

- One thing that surprised me:
  _____

- Question I still have:
  _____

Remember: You do the thinking, AI does the challenging
Be critical • Be cautious • Be compassionate • Avoid BS