



## Distorting the Data

*Persuasion & Influence · Tier 3 · Component 1 — instructor overview*

**Grades:** 11–12    **Time:** ~50 min (flexible)    **Format:** whole-class reading, then individual or pairs

At this level, outright fabrication is rare; the sophisticated distortions use *real* data, selected and arranged to mislead. This component gives students a short list of moves — cherry-picking, survivorship, false precision, missing denominators, correlation-as-cause — and the single question that exposes each.

### WHAT STUDENTS WILL BE ABLE TO DO

- Explain how an accurate statistic can still support a false conclusion.
- Name five distortions (selection, survivorship, false precision, missing denominator, correlation-as-cause) and the question that exposes each.
- Demand the denominator and the comparison a statistic conceals.
- Distinguish a correlation from a defended causal claim.

### WHY IT MATTERS

Students at this age will vote, sign contracts, and read research. The manipulations they'll meet are rarely lies — they're true numbers stripped of the context that would change their meaning. A graduate who reflexively asks "compared to what, out of how many, and who was left out?" can dismantle most statistical spin on contact.

The running example — **Apex Institute**, a for-profit college advertising "94% of our graduates are employed" — is chosen because a single sentence hides three separate distortions at once.

### PACING

| SEGMENT                 | TIME      | WHAT HAPPENS                                            |
|-------------------------|-----------|---------------------------------------------------------|
| Reading                 | 12 min    | Read the Student pages; work the Apex example together. |
| Name the distortion     | 12–15 min | Activity Part A in pairs.                               |
| Restore the denominator | 12 min    | Activity Parts B–C.                                     |
| Debrief                 | 10 min    | Discussion prompts (Instructor p. 4).                   |

### MATERIALS

The 4 Student pages per student or pair. Optional: a real statistic from an ad or headline to dissect together.



## Distorting the Data

*The mechanism in depth — part 1 of 2*

Each move below uses real data; what is manipulated is the selection, the frame, or the comparison, not the number. That is precisely what makes statistical distortion so effective at this level — it survives a naive fact-check, because every figure quoted is accurate. The defense is a short list of questions, one per move. The running example is **Apex Institute** and its billboard: "94% of our graduates are employed."

### SELECTION AND SURVIVORSHIP: WHO GOT COUNTED

---

Cherry-picking and selection bias report the subset that flatters the claim and quietly drop the rest (selection bias). Survivorship bias is the sharpest version: study only the winners who remain, and the failures — being absent — never register (survivorship bias). "94% employed" almost certainly counts only graduates who *responded* to a survey, and only those who *graduated* at all. Students who dropped out, or who ignored the survey because they were unemployed, are invisible. The honest question is not "is 94% true?" but "94% of *whom* — and who was left out of the count?"

### FALSE PRECISION AND MISSING DENOMINATORS

---

An oddly exact figure borrows the authority of rigor it may not possess (overconfidence in precision); a raw count with no base rate hides whether a number is large or trivial (base-rate neglect). Apex's "94%" also hides a denominator problem of a different kind: "employed" is undefined. Employed at what — a job requiring the degree, or any job at all, including the one the student already had before enrolling? A precise-looking percentage rests on a vague, unstated definition. Ask: *out of how many, and compared to what?*



## Distorting the Data

*The mechanism in depth — part 2 of 2*

### CORRELATION DRESSED AS CAUSE

---

A genuine association is presented as proof that one thing produced the other, with no mechanism offered and no confounders ruled out (correlation-causation confusion). Apex's implied claim — *enroll here and you'll be employed* — treats a correlation (graduates tend to be employed) as a causal promise (Apex causes employment). But the people motivated and able to finish a program were likely more employable to begin with; the school may be selecting for the outcome it takes credit for. Ask: *what else could explain this, and is a causal story actually defended?*

### READING APEX WHOLE

---

Notice that no single number in "94% of our graduates are employed" is necessarily false. The sentence is a small machine that runs three distortions at once — a selected, survivor-only sample; a precise percentage over an undefined term; and a correlation sold as a causal promise. This is the texture of sophisticated misinformation: not a lie, but an arrangement of true parts engineered to produce a false whole.

### ONE LINE TO CARRY

---

**Real data can still lie.** A number's accuracy says nothing about a claim's honesty — the honesty lives in the sample it was drawn from, the denominator it's measured against, and the comparison it invites. Teach the questions, not the outrage.



## Distorting the Data

*Discussion guide — modular; assemble to fit your period*

These blocks are timed so you can build a lesson from them: a single period, or a full unit day. Run the activity with an opener before and a debrief after — pick what fits.

### TIME BUDGET — MIX AND MATCH

| BLOCK                  | TIME        | USE                |
|------------------------|-------------|--------------------|
| <b>Opener</b> (before) | 5–8 min     | Pick one, below.   |
| <b>Activity</b>        | 20–25 min   | The student pages. |
| <b>Debrief</b> (after) | ~5 min each | Pick 1–4 prompts.  |
| <b>Close</b>           | 3 min       | The takeaway.      |

**Single period:** opener + activity + one debrief + close. **Full lesson:** opener + activity + two or three debriefs + an extension.

### BEFORE — THE OPENER (PICK ONE · 5–8 MIN)

FORMAT: think-pair-share or a quick hands-up. GOAL: surface what students already notice and pose the question. END RESULT: every student has committed to a prediction or named a real example you can return to in the debrief.

- **Predict.** Put up “94% of our graduates are employed.” Ask: convincing? Take a quick read; revisit after the five moves.
- **From their world.** Ask for one statistic from the news this week; keep two on the board to interrogate in the debrief.

### THE ACTIVITY (BUDGET 20–25 MIN)

CONCRETE GOAL — STUDENTS CAN: expose the distortion in a real-sounding statistic and state the exact question that reveals it.

### AFTER — THE DEBRIEF (PICK 1–4 · ~5 MIN EACH)

FORMAT: whole-class; call on the examples students generated in the opener and activity. END RESULT: students can state the takeaway in their own words and back it with one concrete example from their own life or the activity.

- How can a statistic be completely accurate and still mislead?
- Why is “out of how many?” such a powerful question?
- What’s the difference between two things correlating and one causing the other?
- Take one stat from the board: which distortion, and what question exposes it?

### CLOSE (3 MIN)

Have a few students state the takeaway in their own words: **Compared to what, out of how many, and who got left out?**



## Distorting the Data

*Leading the discussion — pacing, redirects, and warm-ups*

The previous page is the plan; this page is for running it. Budget the period from the total below, and keep the redirect moves handy — most discussions falter in one of these four ways.

### DOES IT FIT THE PERIOD?

ONE PERIOD (~45–50 MIN): opener 6 + activity 20–25 + one debrief 5 + close 3 = **34–39 min.** BLOCK (~90 MIN): add two more debriefs and a differentiation extension from the facilitation page.

### WHEN DISCUSSION STALLS OR DERAILS

| IF...                       | MOVE                                                                                                                                        |
|-----------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| Silence, or “I don’t know”  | Shrink the question: “Just name the first thing you noticed.” Give 30 seconds of silent think-time, then call on a pair, not an individual. |
| One or two voices dominate  | “Let’s hear from someone who hasn’t gone yet.” Run it as think-pair-share first, so every student has an answer ready to offer.             |
| It turns personal or heated | Move the trial from the person to the message: “What in the text makes you say that?” Keep the claim on trial, never the classmate.         |
| Answers stay on the surface | Push for evidence: “Where exactly — quote the line.” Then “What would change your mind?” to surface the reasoning underneath.               |

### WARM-UP BANK (SWAP IN FOR ANY OPENER)

Interchangeable with the opener on the previous page. Vary them across a unit so the hook stays fresh.

- **Two-minute hunt.** Find one real example of today’s move on your phone right now; keep the best one for the debrief.
- **Steelman it.** What’s the most honest reason someone would use this technique? Persuasion isn’t always dishonest — name the legitimate version first.
- **Rate and defend.** Put one message on the board; everyone rates it 1–5 for how persuasive, then defends the spread of answers.
- **Odd one out.** Three short examples, one isn’t the move — which, and how can you tell?

## Distorting the Data

Facilitation, anticipated moves & answer key

### ANTICIPATED STUDENT RESPONSES & MISCONCEPTIONS

- **"If the number is true, the claim is true."** The central misconception. Have them locate the true number inside a misleading claim and articulate why accuracy isn't honesty.
- **"They just made the statistic up."** Usually not — the distortion is in selection and framing. Fabrication is easier to catch; this is the harder, more common case.
- **Correlation reflex.** Some will label everything "correlation not causation" without a mechanism. Push them to propose a plausible confounder, not just recite the phrase.
- **Denominator blindness.** Students accept raw counts. Drill the habit: every count needs "out of how many."

### DIFFERENTIATION & EXTENSIONS

**Support:** give the five moves and their questions as a reference card. **Stretch:** students find a real advertised statistic and write a short "what this number hides" analysis; or locate the methodology behind a cited study. **Cross-curricular:** sampling, base rates, and confounders in statistics; source evaluation in research writing.

### ANSWER KEY — ACTIVITY

| ITEM                                              | SAMPLE RESPONSE                                                                  |
|---------------------------------------------------|----------------------------------------------------------------------------------|
| A1 "our grads earn \$95,000"                      | <i>Survivorship — only responding, employed grads. Out of how many enrolled?</i> |
| A2 "reports up 200% this year"                    | <i>Missing base rate — from 1 to 3? 200% of what?</i>                            |
| A3 "73.4% more effective"                         | <i>False precision — exactness borrowing rigor it may not have.</i>              |
| A4 "laptop owners get better grades — so buy one" | <i>Correlation as cause; wealthier students may have both.</i>                   |
| B ("50 side effects")                             | <i>Need the denominator — 50 out of how many doses?</i>                          |



---

## Distorting the Data

*How true numbers tell false stories*

A billboard for Apex Institute reads: "**94% of our graduates are employed.**" It's a great number. It might also be completely true — and still a near-perfect lie. Because packed inside that one accurate-sounding sentence are three separate distortions, none of which requires a single false figure. The most convincing misinformation you'll meet as an adult won't be made up. It will be real data, selected and arranged to mislead.

Five moves do most of the work, and each one dies to a single question. Learn the questions and you can take apart most of the statistics thrown at you.

### WHO GOT COUNTED?

---

**Cherry-picking** shows the data that fits and drops the data that doesn't. **Survivorship bias** is the sharpest version: it studies only the winners who are still around. "Every successful founder dropped out of college" ignores the thousands who dropped out and failed — they're not in the room to be counted. Apex's "94% employed" almost certainly means 94% of the graduates who *answered a survey* — not the students who dropped out, and not the graduates who were unemployed and quietly ignored the email. ASK: WHO WAS LEFT OUT OF THE COUNT?

### OUT OF HOW MANY? COMPARED TO WHAT?

---

A super-precise number ("73.4% more effective") *looks* like science but can be a guess in a lab coat. A big raw count ("2,000 cases!") is meaningless without the total — 2,000 out of 10,000 is a crisis; 2,000 out of 10 million is a rounding error. And Apex hides one more thing: what does "employed" even mean? Any job at all, including the one you had before enrolling? A precise percentage sitting on a vague word is not the rigor it pretends to be. ASK: OUT OF HOW MANY, AND MEASURED HOW?



---

## Distorting the Data

*Correlation, cause, and reading a number whole*

### DID A ACTUALLY CAUSE B?

---

Two things moving together is not proof that one caused the other. Ice-cream sales and drownings rise together — not because ice cream drowns people, but because it's summer, and heat drives both. Apex's whole pitch runs on this: it wants "our graduates tend to be employed" to feel like "enrolling at Apex will get you employed." But the people with the drive and resources to finish a program were probably more employable to begin with. Apex may simply be taking credit for selecting people who were going to succeed anyway. ASK: WHAT ELSE COULD EXPLAIN THIS — AND DID ANYONE ACTUALLY MAKE THE CAUSAL CASE?

### READING THE WHOLE MACHINE

---

Here's the key move at this level: go back and notice that no number in "94% of our graduates are employed" has to be false. The sentence is a little machine running three distortions at once — a survivor-only sample, a precise percentage over an undefined word, and a correlation sold as a promise. That's what sophisticated misinformation actually looks like. Not a lie you can catch by fact-checking a figure, but an *arrangement* of true parts built to produce a false whole.

### THE ONE LINE

---

**Real data can still lie.** A number being accurate tells you nothing about whether the claim is honest. The honesty lives somewhere else — in the sample it came from, the denominator it's measured against, and the comparison it's quietly inviting you to make. When a statistic is handed to you as proof, don't check whether the number is true. Check what it's hiding.





## Distorting the Data

Name: \_\_\_\_\_ Date: \_\_\_\_\_

**The questions:** Who got left out? · Out of how many? · Compared to what? · What else could explain it?

### PART A — NAME THE DISTORTION

Name the move (selection / survivorship / false precision / missing denominator / correlation-as-cause) and the question that exposes it.

| STATISTIC                                                     | MOVE · EXPOSING QUESTION |
|---------------------------------------------------------------|--------------------------|
| 1. "Our graduates earn \$95,000 on average."                  |                          |
| 2. "Reports of the problem jumped 200% this year!"            |                          |
| 3. "This method is 73.4% more effective."                     |                          |
| 4. "Students who own laptops get better grades — so buy one." |                          |

### PART B — RESTORE THE DENOMINATOR

"Our vaccine caused 50 side effects!" What single number would tell you whether 50 is alarming or trivial? Write the fuller, honest version.

|             |
|-------------|
| <hr/> <hr/> |
|-------------|

**PART C — TAKE APART APEX**

"94% of our graduates are employed." Name all three distortions hidden in this one sentence, and for each, the question you'd ask.

| DISTORTION | THE QUESTION |
|------------|--------------|
| 1.         |              |
| 2.         |              |
| 3.         |              |

**PART D — CORRELATION COURTROOM**

A study finds towns with more bookstores have lower crime. A politician says: "Build bookstores to cut crime." Give one other explanation, and what evidence would be needed before believing the causal claim.

|             |
|-------------|
| <hr/> <hr/> |
|-------------|

**PART E — IN THE WILD**

Find one real statistic in the news. Name any distortion it might contain and the question you'd ask to check it.

|             |
|-------------|
| <hr/> <hr/> |
|-------------|