A FIELD GUIDE
INCOGNATI · SPECIMEN SERIES No. 2

Numbers that *want to convince you* but *need* to be *checked*

Eight ways data gets misread, misrepresented, or misinterpreted before it ever reaches you — real, documented cases, with the tell that gives each one away.

BEFORE YOU BEGIN

How This Guide Works

Every specimen here is a real, documented case — from published research, government data, or extensively verified history, not invented for the lesson. Sources are cited at the bottom of each plate so you can check the checking.

Where Field Guide No. 1 covered how claims get framed and spread, this edition covers what happens once you're already looking at a number, a chart, or a study — the point where accurate data quietly becomes a misleading conclusion.

OBSERVED SPECIMEN
The real data, study, or reported figure — and how it was actually used.

FIELD MARK
The one question that exposes the pattern, every time it resurfaces.

SPECIMENS IN THIS EDITION

STATISTICAL SLEIGHT
SPECIMEN · NO. 01

PLATE I

Correlation Presented as Causation

Causa Ficta

First observed 2014 **Status** Widely cited teaching example

OBSERVED SPECIMEN — "SPURIOUS CORRELATIONS" PROJECT, TYLER VIGEN

Data analyst Tyler Vigen's **Spurious Correlations** *project charts the number of Nicolas Cage films released each year against U.S. drownings in swimming pools — two entirely unrelated trends that move together with a correlation coefficient of roughly 0.66, purely by chance.*

THE HOOK

Two lines that rise and fall together read as connected, even when nothing actually links them.

THE MECHANISM

Compare enough variables and some will correlate by pure coincidence — the correlation itself proves nothing about causation.

Field mark: ask whether a plausible mechanism connects the two trends, or whether a third shared factor — or nothing at all — could explain the pattern.

CASE VERIFIED — Tyler Vigen, Spurious Correlations project, tylervigen.com, since 2014.

STATISTICAL SLEIGHT
SPECIMEN · NO. 02

PLATE II

Misleading Average

Media Fallax

First observed Recurring pattern **Status** 2022 Federal Reserve data

OBSERVED SPECIMEN — FEDERAL RESERVE SURVEY OF CONSUMER FINANCES, 2022

The Fed's 2022 survey found the **average** *U.S. household net worth was $1.06 million — while the* **median** *household, the one actually in the middle, held just $192,700. A small number of ultra-wealthy households pull the average far above what a typical family owns.*

THE HOOK

"Average" sounds like it describes a typical case, but a mean can be dragged far from typical by a handful of extreme values.

THE MECHANISM

The median — the true midpoint — is often the more honest number, especially when wealth or income is highly concentrated.

Field mark: when you see "average," ask whether the median tells a different story — a big gap between the two is itself informative.

CASE VERIFIED — Federal Reserve, "Changes in U.S. Family Finances from 2019 to 2022," Oct. 2023.

STATISTICAL SLEIGHT
SPECIMEN · NO. 03

PLATE III

Percentage vs. Percentage-Point

Punctum Percentuale

First observed 2004 case **Status** Common, ongoing reporting error

OBSERVED SPECIMEN — ABC NEWS "WHO'S COUNTING," 2004

Mathematician John Allen Paulos flagged coverage describing a Social Security proposal as diverting **"2% of taxes"** *into private accounts. The actual change was from 6.2% to 4.2% of taxable income — a 2 percentage-point cut, but a 32% relative cut in the money involved, far larger than "2%" suggested.*

THE HOOK

"2%" and "2 percentage points" sound almost interchangeable but describe very different sizes of change.

THE MECHANISM

Reporting the smaller-sounding figure, instead of the percentage-point or relative change, can make a substantial shift look minor.

Field mark: when a rate moves from one percentage to another, check whether "percent" or "percentage points" is meant — the relative change can be many times larger.

CASE VERIFIED — John Allen Paulos, ABC News, 2004; explained further by The Journalist's Resource, Oct. 2022.

STATISTICAL SLEIGHT
SPECIMEN · NO. 04

PLATE IV

Unrepresentative Sample

Specimen Distortum

First observed 1936 **Status** Standard textbook case since

OBSERVED SPECIMEN — LITERARY DIGEST PRESIDENTIAL POLL, 1936

The Literary Digest mailed **10 million** *straw-poll ballots and predicted a landslide for Alf Landon over Franklin Roosevelt. Roosevelt won in one of the largest landslides in U.S. history — the sample, drawn from telephone directories and car registrations, had skewed toward wealthier, more Republican-leaning households during the Depression.*

THE HOOK

A sample of 10 million people looks enormous and authoritative, regardless of who those 10 million actually are.

THE MECHANISM

A sample can be huge and still unrepresentative if the method used to gather it systematically excludes or oversamples particular groups.

Field mark: ask how a sample was gathered, not just how big it was — a smaller, well-drawn random sample beats a massive skewed one.

CASE VERIFIED — widely documented in polling history; the Literary Digest folded within two years of the failed prediction.

INTERPRETATION TRAP
SPECIMEN · NO. 05

PLATE V

Survivorship Bias

Superstes Fallacia

First observed 1943 **Status** Cited across fields today

OBSERVED SPECIMEN — STATISTICAL RESEARCH GROUP MEMORANDUM, WWII

Reviewing damage on Allied bombers that **returned** *from missions, statistician Abraham Wald was asked where to add armor based on the parts showing the most bullet holes. He concluded the opposite: those areas needed the least reinforcement, since planes hit there still made it home — the truly vulnerable spots were the ones with no holes on returning planes, because aircraft hit there never returned at all.*

THE HOOK

The data in front of you — returning planes — feels complete, but it's only the surviving half of the picture.

THE MECHANISM

Whatever didn't survive to be observed is invisible in the dataset, and its absence can point to exactly the wrong conclusion.

Field mark: ask what didn't make it into the dataset at all — the missing cases, not the visible ones, are often where the real answer is.

CASE VERIFIED — Abraham Wald, Statistical Research Group memorandum, 1943; documented by W. Allen Wallis, analyzed in the AMS Feature Column, 2016.

INTERPRETATION TRAP
SPECIMEN · NO. 06

PLATE VI

Base Rate Neglect

Frequentia Basica

First observed 1995 study **Status** Persistent in medical decision-making

OBSERVED SPECIMEN — GIGERENZER & HOFFRAGE, PSYCHOLOGICAL REVIEW, 1995

A widely cited study found most physicians badly overestimated the odds that a **positive mammogram** *meant a patient actually had breast cancer — because they focused on the test's accuracy while ignoring how rare the condition is in the screened population to begin with.*

THE HOOK

A test described as "90% accurate" sounds like a 90% chance the result is correct, regardless of context.

THE MECHANISM

When the condition tested for is rare, even an accurate test produces far more false positives than true positives in absolute terms.

Field mark: for any test result, ask how common the underlying condition actually is in the group tested — the base rate changes the math entirely.

CASE VERIFIED — Gigerenzer & Hoffrage, "How to Improve Bayesian Reasoning Without Instruction," Psychological Review, 1995.

INTERPRETATION TRAP
SPECIMEN · NO. 07

PLATE VII

Extrapolation Beyond the Data

Extrapolatio Temeraria

First observed 1968 **Status** Predictions did not materialize as stated

OBSERVED SPECIMEN — "THE POPULATION BOMB," PAUL EHRLICH, 1968

Ehrlich extrapolated then-current population growth rates forward and predicted **hundreds of millions** *would starve in the 1970s and '80s. The trend line had held for years — but the prediction didn't account for the Green Revolution's agricultural yield gains, already underway, which would outpace population growth for decades.*

THE HOOK

A trend that has held steady for years feels safe to project forward indefinitely.

THE MECHANISM

Trends change when underlying conditions change — a line fit to past data assumes nothing else will shift, which is rarely true over decades.

Field mark: ask how far past the actual measured range a projection is being pushed, and what would have to stay constant for it to hold.

CASE VERIFIED — Paul R. Ehrlich, The Population Bomb, 1968; widely analyzed in retrospective coverage of 20th-century population forecasting.

INTERPRETATION TRAP
SPECIMEN · NO. 08

PLATE VIII

False Precision

Praecisio Ficta

First observed 1965 **Status** Still marketed as a research-based target today

OBSERVED SPECIMEN — ORIGIN OF THE "10,000 STEPS" HEALTH TARGET

The now-famous **"10,000 steps a day"** *target didn't come from a clinical study — it originated in 1965 when a Japanese company marketed a pedometer called "manpo-kei" ("10,000-steps meter"), chosen partly because the character for 10,000 resembled a person walking. Six decades later it's still cited as a precise, research-derived medical threshold.*

THE HOOK

A specific, round number sounds like it was derived from careful measurement, even when its origin was something else entirely.

THE MECHANISM

Once a number gets repeated enough times attached to health or science, its actual origin stops getting checked.

Field mark: ask where a suspiciously specific or round number actually came from — marketing copy and clinical research produce very different kinds of numbers.

CASE VERIFIED — pedometer origin widely corroborated; reviewed by I-Min Lee et al., JAMA Internal Medicine, 2019.

Closing Plate · Field Method

Reading the Numbers: A Data Gut Check

five questions to ask before a chart or statistic changes your mind

1. **Find the baseline**
   Does an axis start at zero, and does the range shown cover the full picture or a cherry-picked slice?

2. **Mean or median?**
   Ask which measure is being used — a big gap between the two is itself the story.

3. **Check the sample**
   How many were measured, and how were they chosen? Big isn't the same as representative.

4. **Look for the missing cases**
   What didn't make it into this dataset at all? The absence can matter more than what's shown.

5. **Trace the number to its source**
   Was it measured, modeled, or just repeated until it sounded official?

INCOGNATI

The Atlas Goes Deeper

This guide covers eight specimens; Field Guide No. 1 covers twelve more on how false claims get framed and spread. The Incognati Atlas catalogs thousands more — biases, fallacies, distortions, and manipulation patterns, each one documented the same way: real case, real source, real tell.
