

How Statistics Can Be Misleading



- Recognize misleading statistics in reports, dashboards, and everyday decisions



- Misread data costs analysts, students, and business teams time and money



- Explore key distortion sources: averages, selective reporting, charts, and correlation



- Build habits to question numbers before acting on them



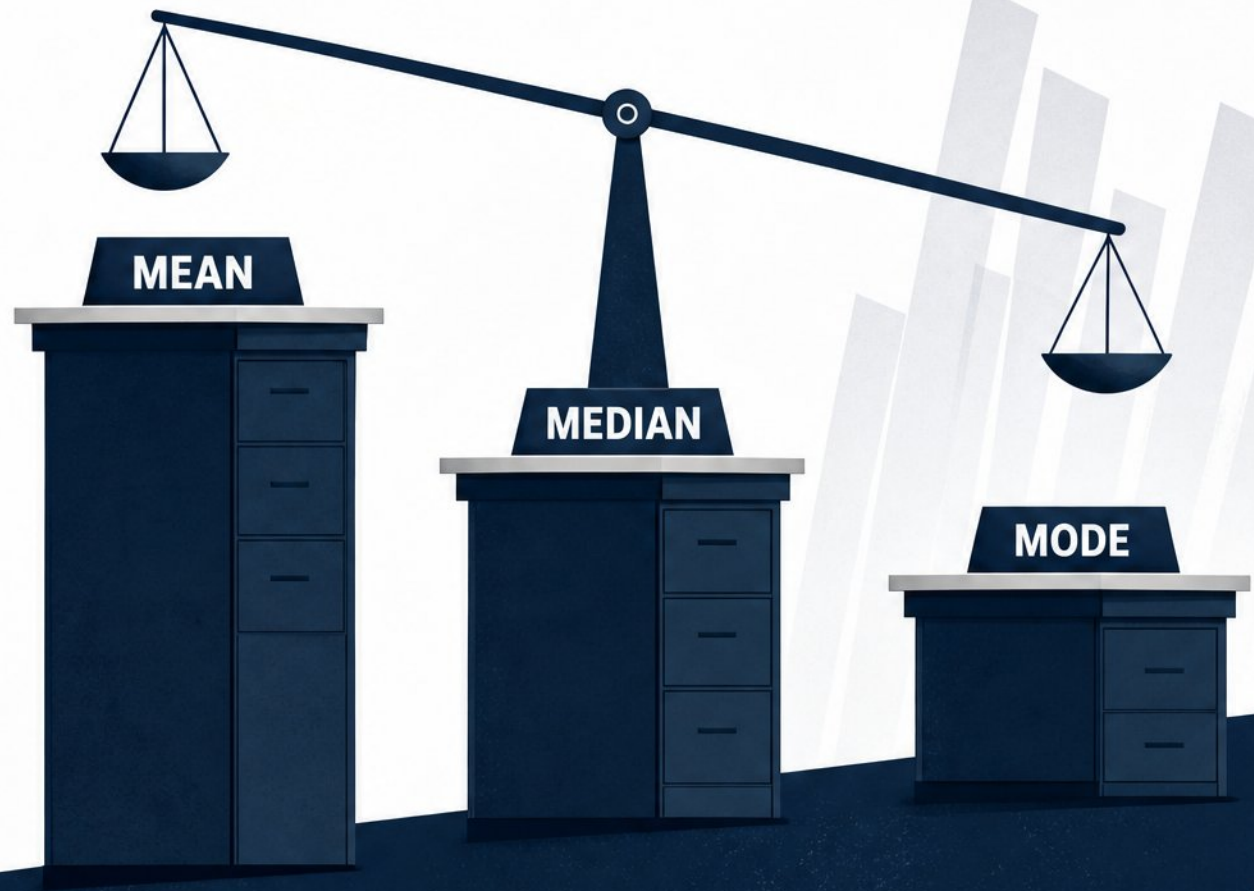
Why Honest Statistics Need Context

- A statistic is never meaningful without context, definition, and comparison
- Accuracy of calculation is not the same as validity of interpretation
- Descriptive summaries are not causal claims
- A single metric can hide distribution, composition, and time-range choices



The Average Trap: When the Mean **Misleads**

- ▶ Outliers and skew pull the mean away from typical values
- ▶ Median stays stable; mean shifts sharply with extremes
- ▶ Salary example: top earners inflate average income
- ▶ Delivery times: few slow orders raise the average
- ▶ Report mean with median and distribution shape



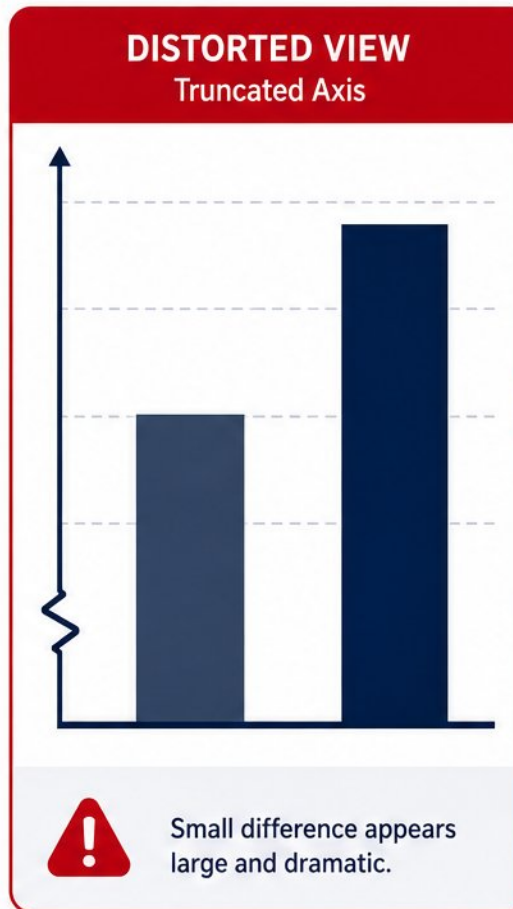
Selective Reporting and Cherry-Picking

-  - Selecting favorable time ranges or subgroups distorts conclusions
-  - P-hacking and multiple comparisons create false confidence
-  - Omitted data and missing denominators hide the full picture
-  - Highlighting positive reviews while ignoring negative feedback misleads
-  - Spot incomplete comparisons before acting on a statistic

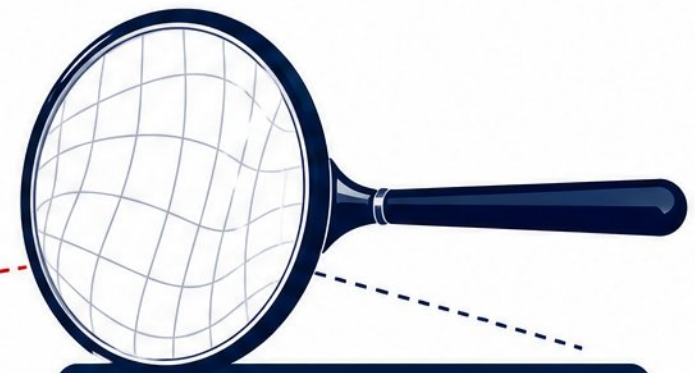
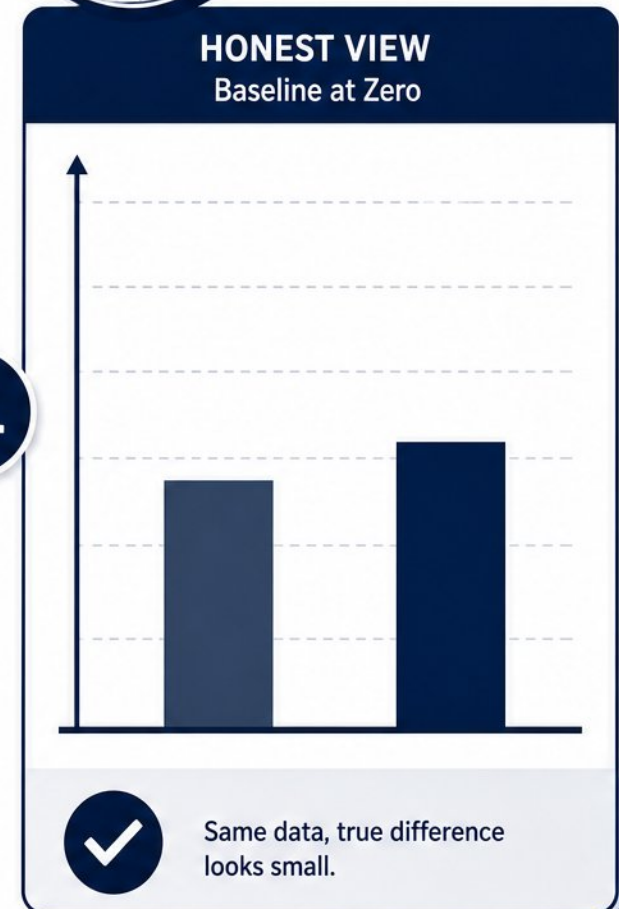


Charts That Distort

- ✗ Truncated axes exaggerate small differences between bar charts
- ✗ 3D effects and area scaling distort true proportions
- ✗ Dual axes can invent patterns that don't exist
- ✗ Bar charts must start at zero; dot plots show small gaps honestly
- ✗ Check axes, baselines, and chart type before trusting a visual

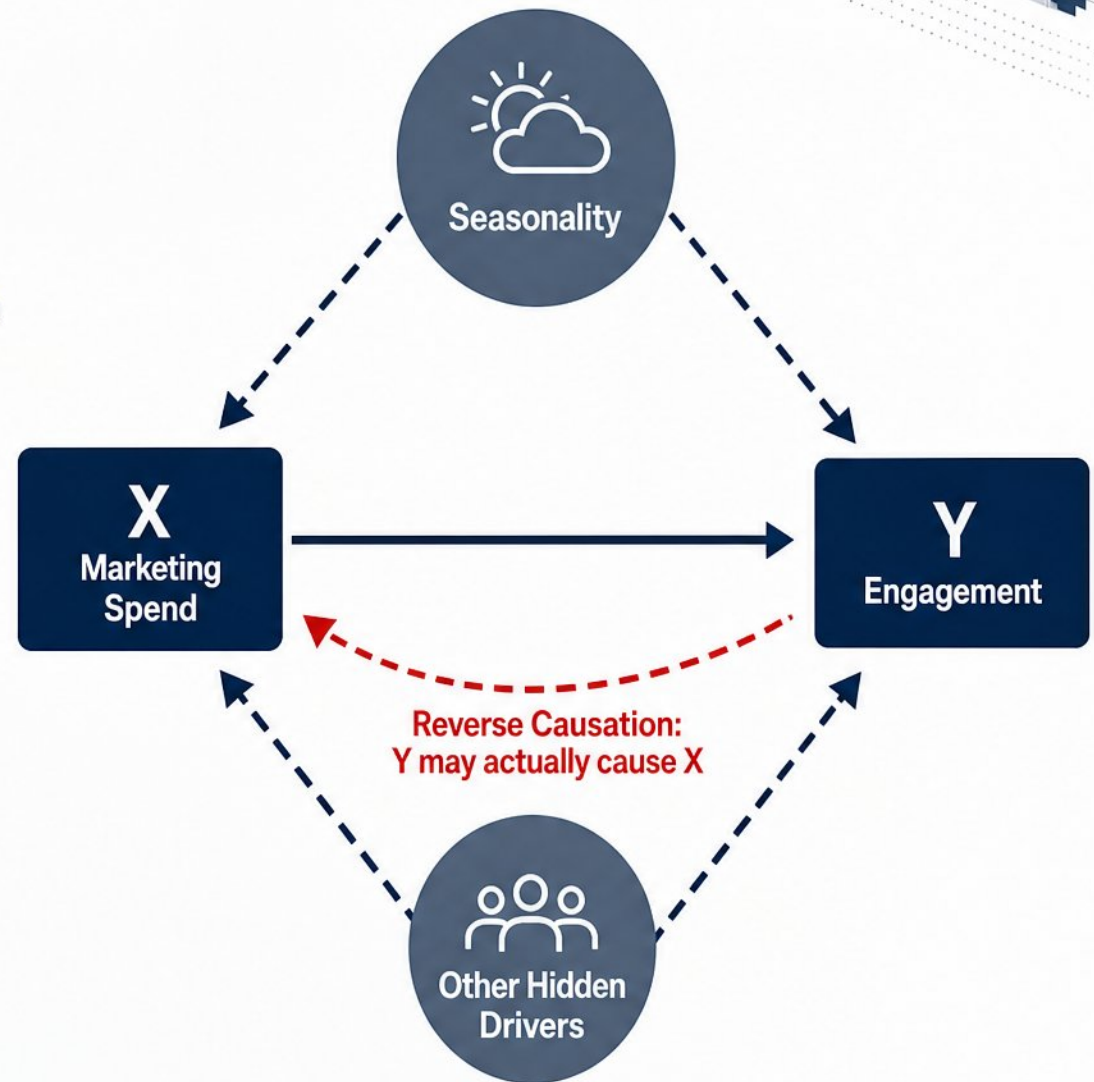


VS.



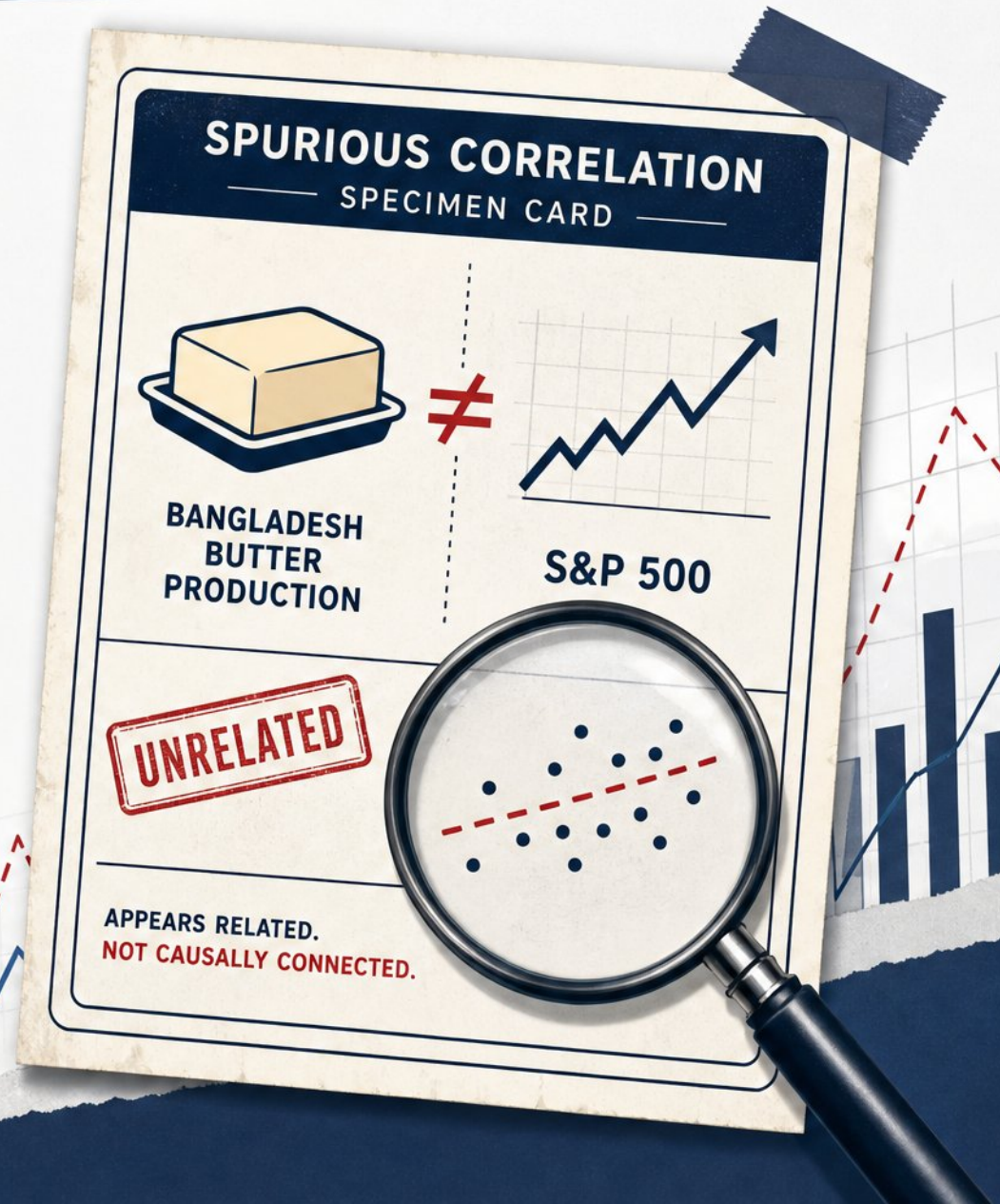
Correlation vs. Causation

- Correlation does not imply causation
- Confounding variables: hidden third factors drive both variables
- Reverse causation: Y may actually cause X
- Marketing spend, engagement, and seasonality often share common drivers
- Always ask: what else could explain this relationship?



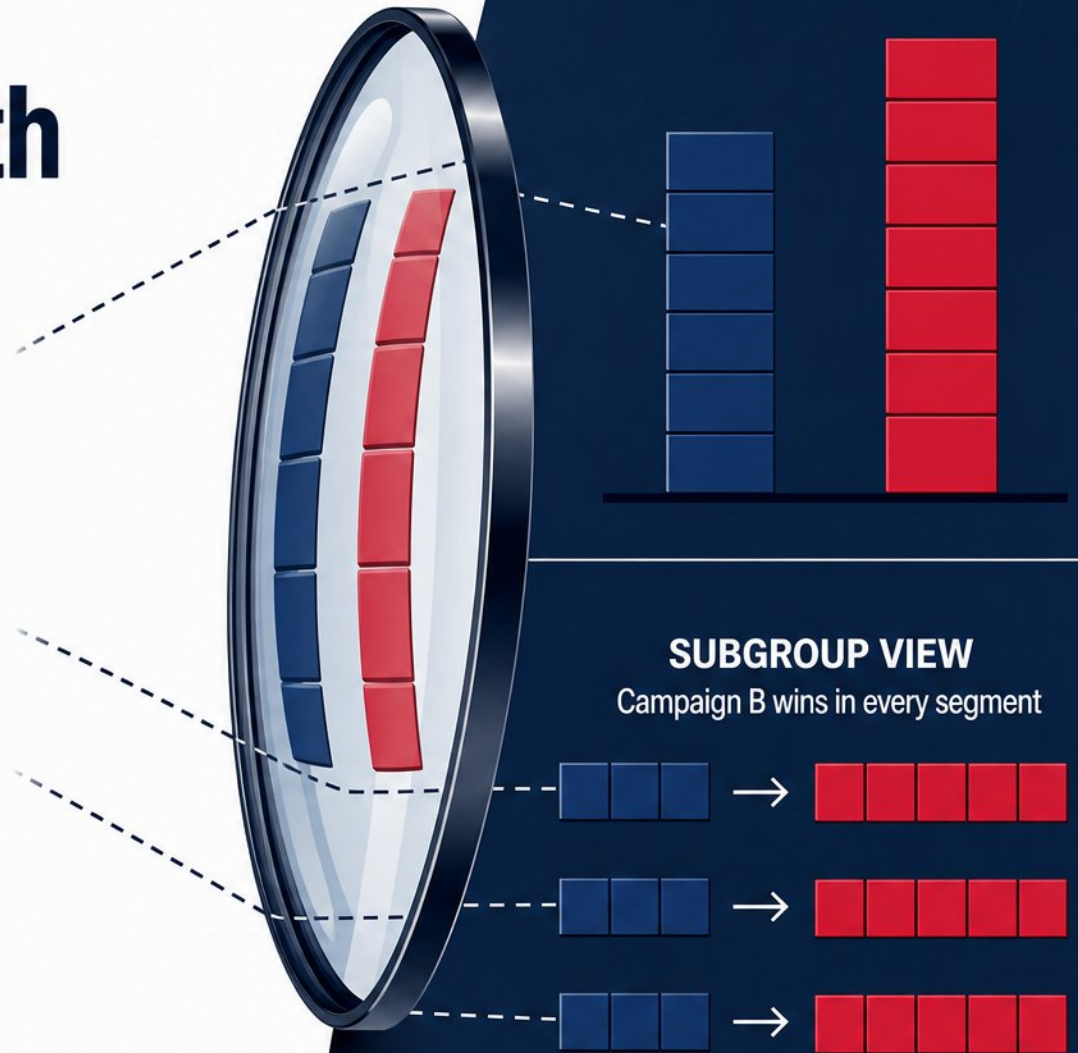
Spurious Correlations in Business Data

- ▶ Data dredging creates coincidental, false relationships in large datasets
- ▶ Bangladesh butter production vs. S&P 500 is a classic fake correlation
- ▶ Ice cream sales and drownings rise together due to hot weather
- ▶ Correlation requires a plausible mechanism before acting
- ▶ Always test out-of-sample and question causal direction



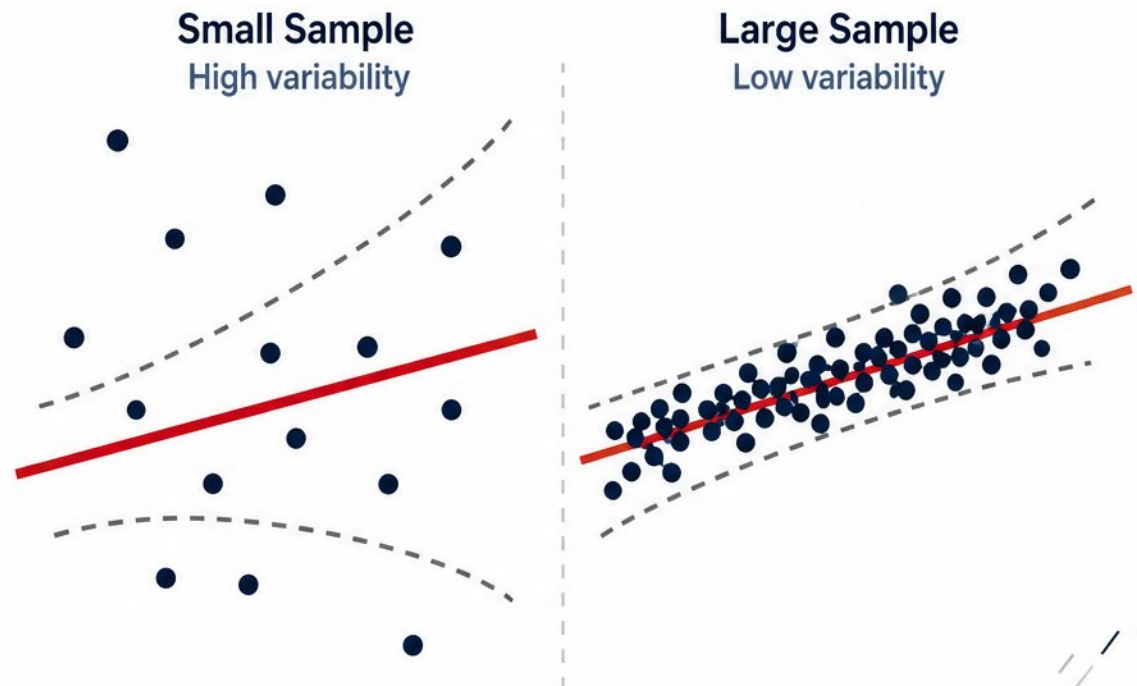
Simpson's Paradox: When Totals Reverse the Truth

- ▶ An overall trend can reverse when data is broken into subgroups
- ▶ Marketing example: Campaign B wins in every segment but loses overall
- ▶ Confounding variable: the mix of high-value and low-value visitors
- ▶ Aggregated dashboards may hide real relationships
- ▶ Always segment data by meaningful dimensions before deciding



Sample Size and Variability

- Small samples produce unstable, imprecise estimates
- Larger samples narrow the margin of error
- Random variation shrinks as sample size grows
- Percentage changes on small bases mislead quickly
- Always ask: How many observations back this number?



Base Rates and Missing Context

- Base-rate fallacy: ignoring how common or rare an event is
- Vivid, specific information overrides general statistical probability
- Relative risk versus absolute risk changes how results feel
- Framing effects: the same number can lead to different decisions



When Good Metrics Go Bad

-  - Conversion rates increase while total volume drops
-  - Aggregates hide product mix and composition shifts
-  - Period-over-period changes distorted by anomalous baseline periods
-  - Fix: pair rates with volumes, averages with medians
-  - Fix: break down aggregates by segment and cohort



How to Spot Data Distortion Quickly

S



– **Source:** Who produced the data, and what are their motives?

C



– **Chart:** Is the chart type appropriate for the data?

A



– **Axes:** Do they start at zero with consistent scales?

M



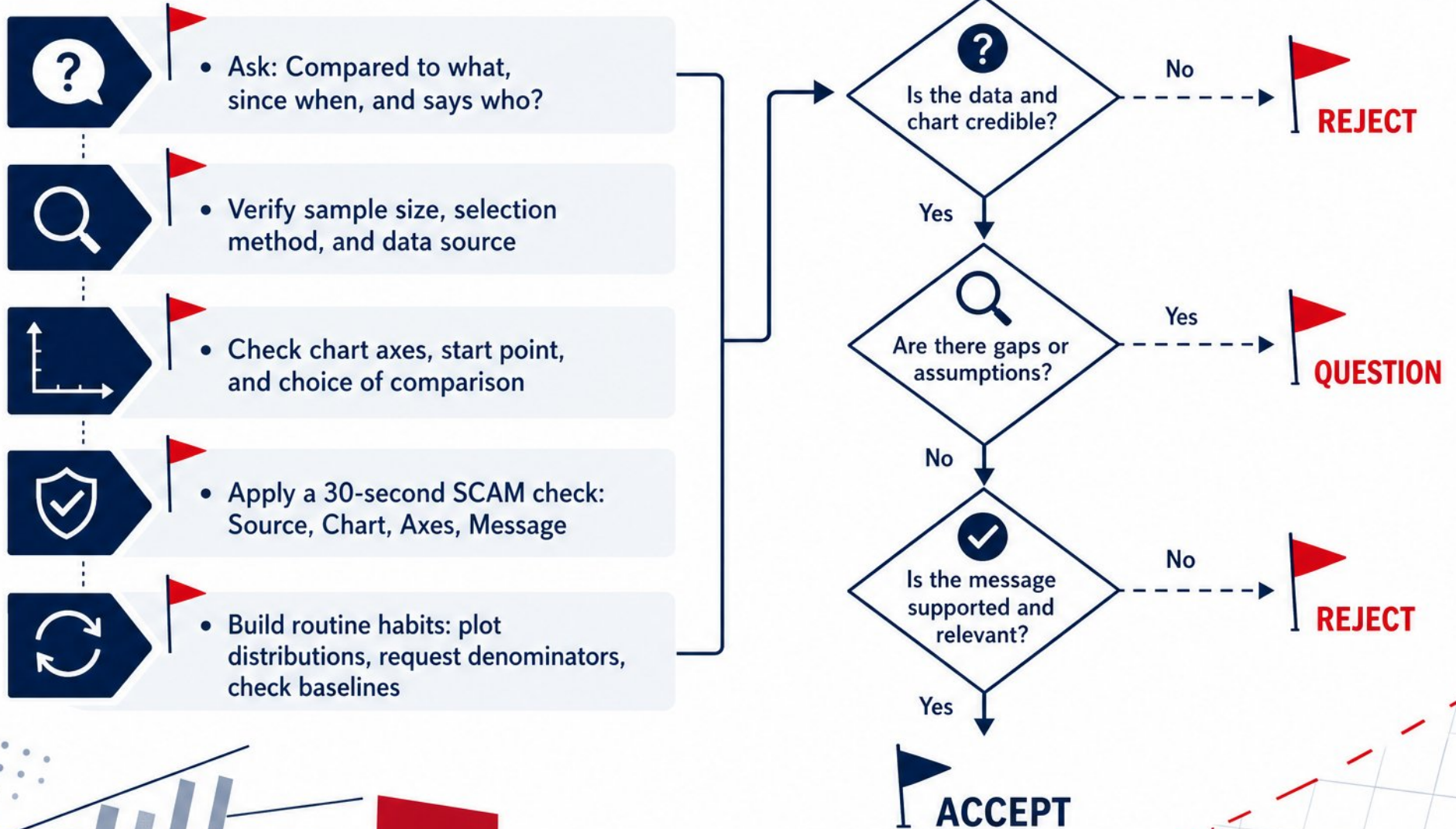
– **Message:** Does the interpretation match the actual data?



– **Red flags:** unsourced numbers, missing denominators, cherry-picked windows

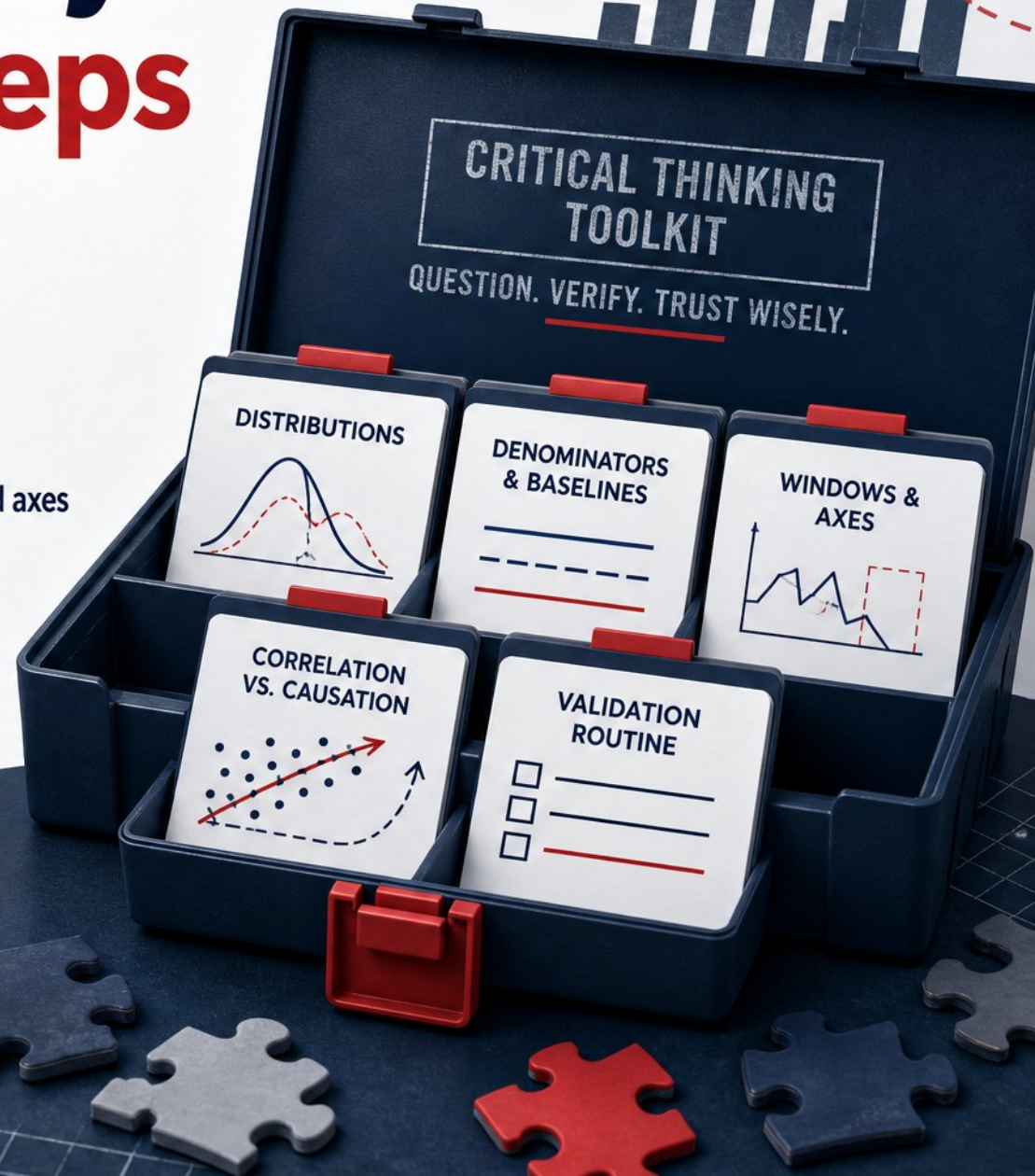


Practical Checks for Analysts and Business Teams



Key Takeaways and **Next Steps**

-  - Plot distributions before trusting averages
-  - Ask for denominators and check baselines
-  - Question cherry-picked windows and truncated axes
-  - Distinguish correlation from causation
-  - Build a personal validation routine for every report



PersonWise.

PersonWise • AI digital-human course platform



Scan the code or click anywhere to watch this course live, with voice Q&A

personwise.ai/t/86ba9ce0/public