Data Dredging

Data dredging (also known as p-hacking, fishing expedition, or post-hoc data mining) refers to the misuse of statistical analysis by performing multiple comparisons or subgroup analyses until a statistically significant result is found — often by chance alone — without a pre-specified hypothesis.

 In scientific critique, data dredging implies: Searching for significance rather than testing a meaningful, pre-defined question.

Multiple subgroup analyses without correction for multiple testing.

Highlighting spurious associations as if they were genuine discoveries.

Retrospective storytelling: interpreting random patterns as clinically or biologically relevant.

 Why it’s problematic: Increases the false positive rate.

Produces results that cannot be replicated.

Undermines the credibility of the research.

Often used to inflate the perceived value of weak or null findings.

 Example (in a critical review): “The authors engage in data dredging, slicing the dataset until statistical significance emerges — not to test a hypothesis, but to justify one retroactively.”