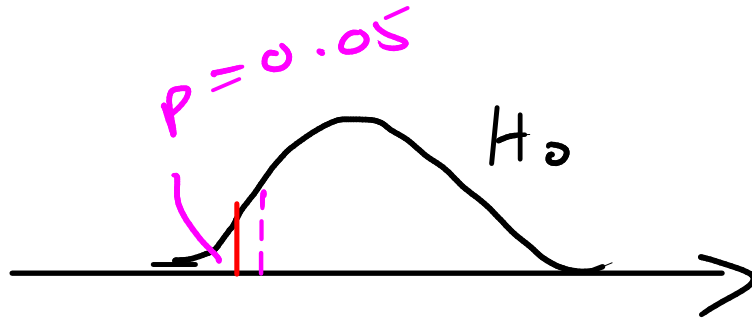


The background of the slide features a stylized globe on the left side, partially obscured by a dense pattern of binary code (0s and 1s) that recedes into the distance, creating a sense of depth and digital connectivity. The colors are primarily light blue and white.

Foundations of Data Science: Hypothesis testing - Issues in hypothesis testing

Type I and Type II errors



Type I error: Rejecting H_0 when it is true

Control by setting α - size of rejection region

Type II error: not rejecting H_0 when it is false.

"Cherry-picking", "data dredging", "p-value hacking"

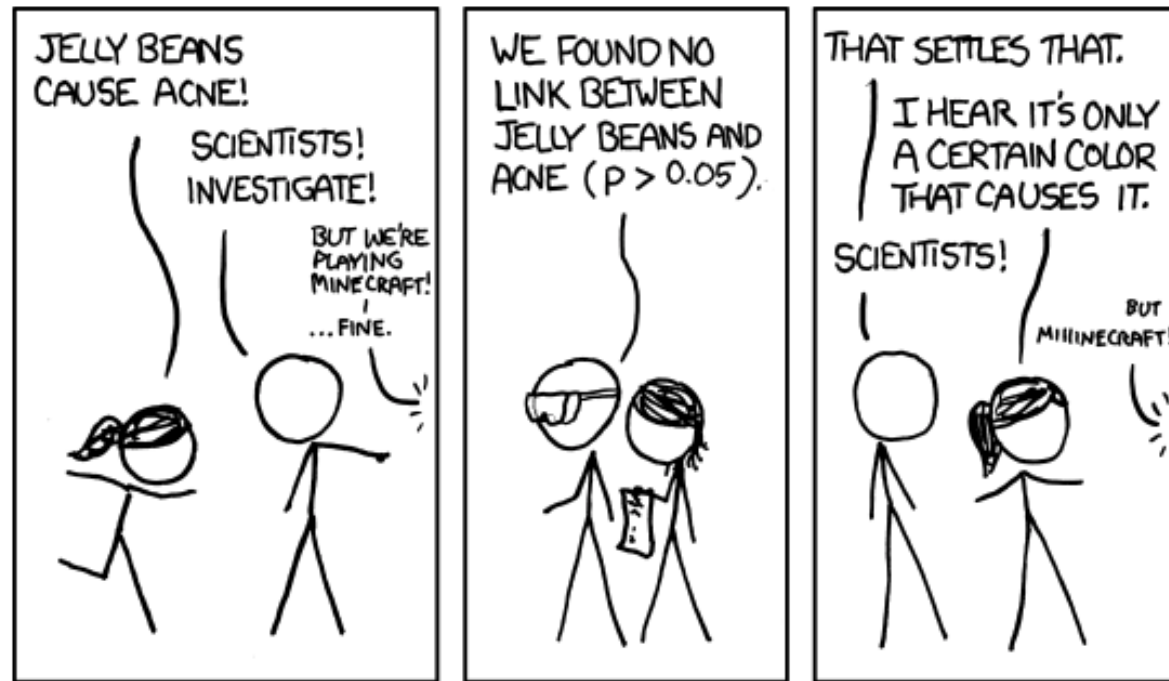
Proper inference requires full reporting and transparency.

P-values and related analyses should not be reported selectively. Conducting multiple analyses of the data and reporting only those with certain *p*-values (typically those passing a significance threshold) renders the reported *p*-values essentially uninterpretable. Cherry-picking promising findings, also known by such terms as data dredging, significance chasing, significance questing, selective inference, and "p-hacking," leads to a spurious excess of statistically significant results in the published literature and should be vigorously avoided... (ASA Statement on Statistical Significance and *P*-values)

Type I error = prob. rejecting H_0 when it is true

Multiple Testing

- 20 tests \Rightarrow 0.05 chance Type I error on each test
- \Rightarrow 0.95 chance of no type I error on e test
- \Rightarrow 0.95^{20} chance no type I errors overall
- \Rightarrow $1 - 0.95^{20} = 0.64$ chance type I error



Read the whole cartoon at:

<https://www>

[.w](https://www)

[explainxkcd.com/wiki/index.php/882:_Significant](https://www.explainxkcd.com/wiki/index.php/882:_Significant)

"A hypothesis is a liability"

