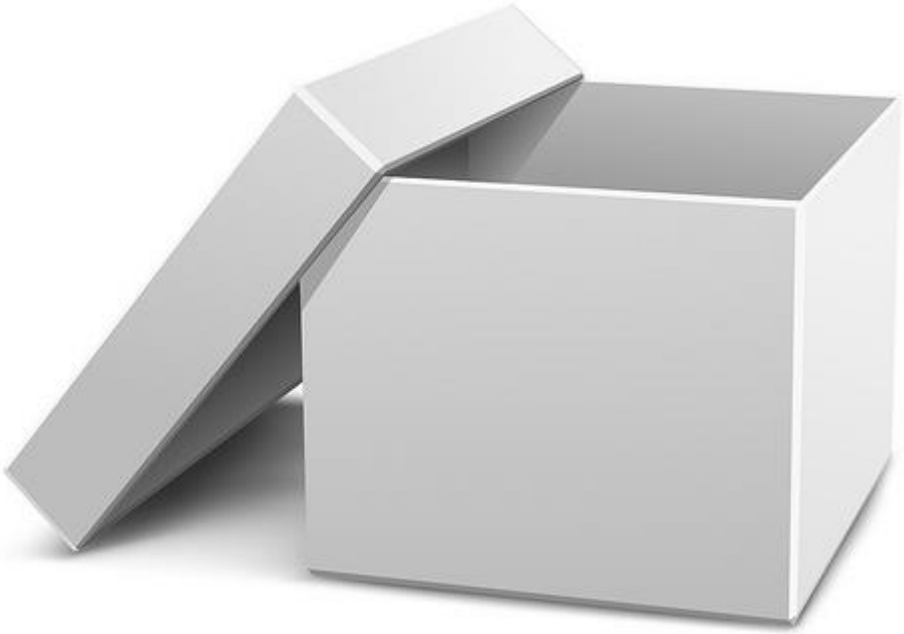
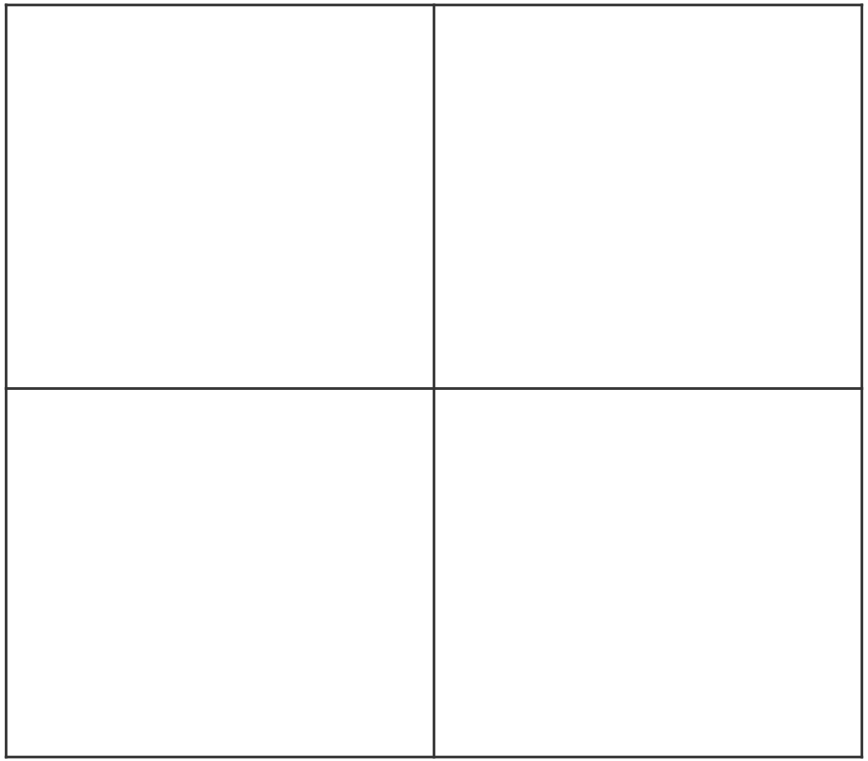
Hypothesis Testing & Statistical Comparisons

Jeff Sauro, PhD

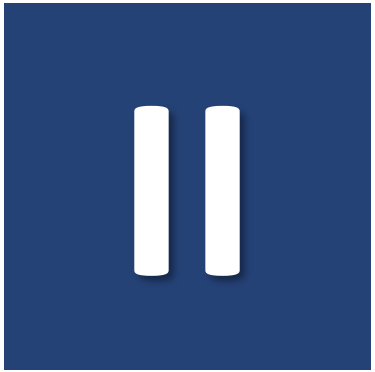




HYPOTHESIS TESTING OVERVIEW

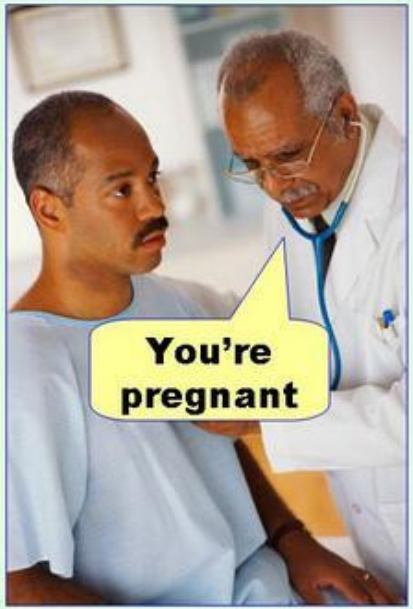
Hypothesis Testing: What can go wrong?

Saying there is a difference when one does not exist

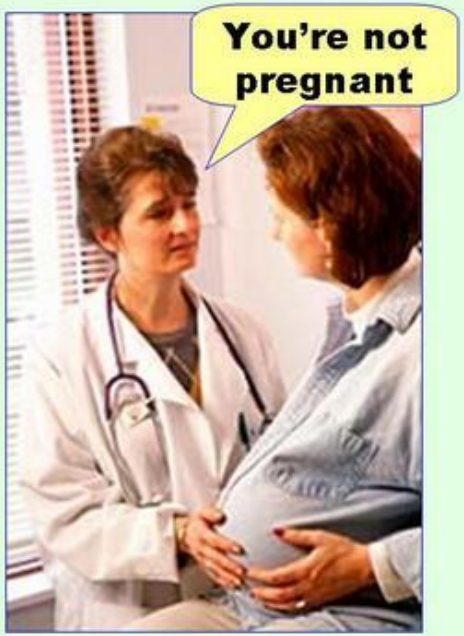


Saying there is **not** a difference when one **does** exist

Type I error
(false positive)



Type II error
(false negative)



Type 1 Error

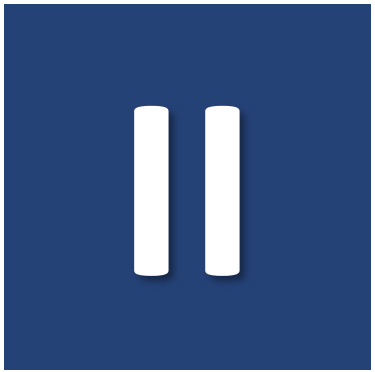
Convicting an innocent person

A false positive

Incorrectly reject the Null

Saying a design IS better, when there really is no difference

Type II Error



Letting the bad guy go free

A false negative

Incorrectly failing to reject the Null

Saying there is no difference between designs, when one really is better

Test your error knowledge

You get a p-value of .01 and conclude the designs are different.

What error can you make?

You get a p-value of .044 and conclude the designs are NOT different.

What error can you make?

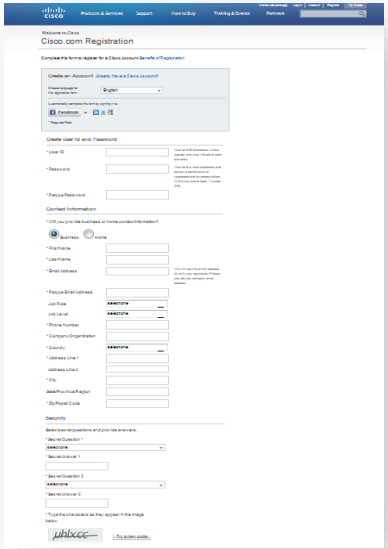
You get a p-value of .049 and conclude the designs are still NOT different.

What error can you make?

COMPARING TWO COMPLETION RATES

Do more users register on the new form?

OLD DESIGN



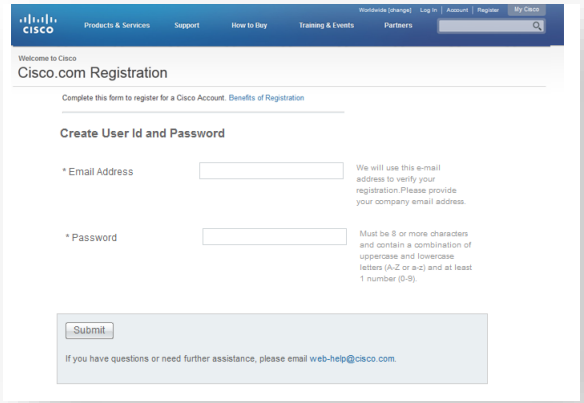
8 Successful

12 Attempted

67%

Completion Rate

NEW DESIGN



11 Successful

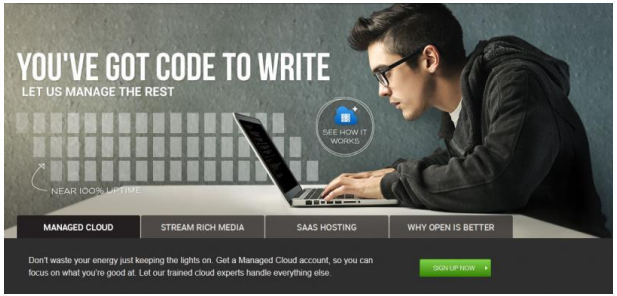
11 Attempted

100%

Completion Rate

Which design generates more purchases?

A/B Test



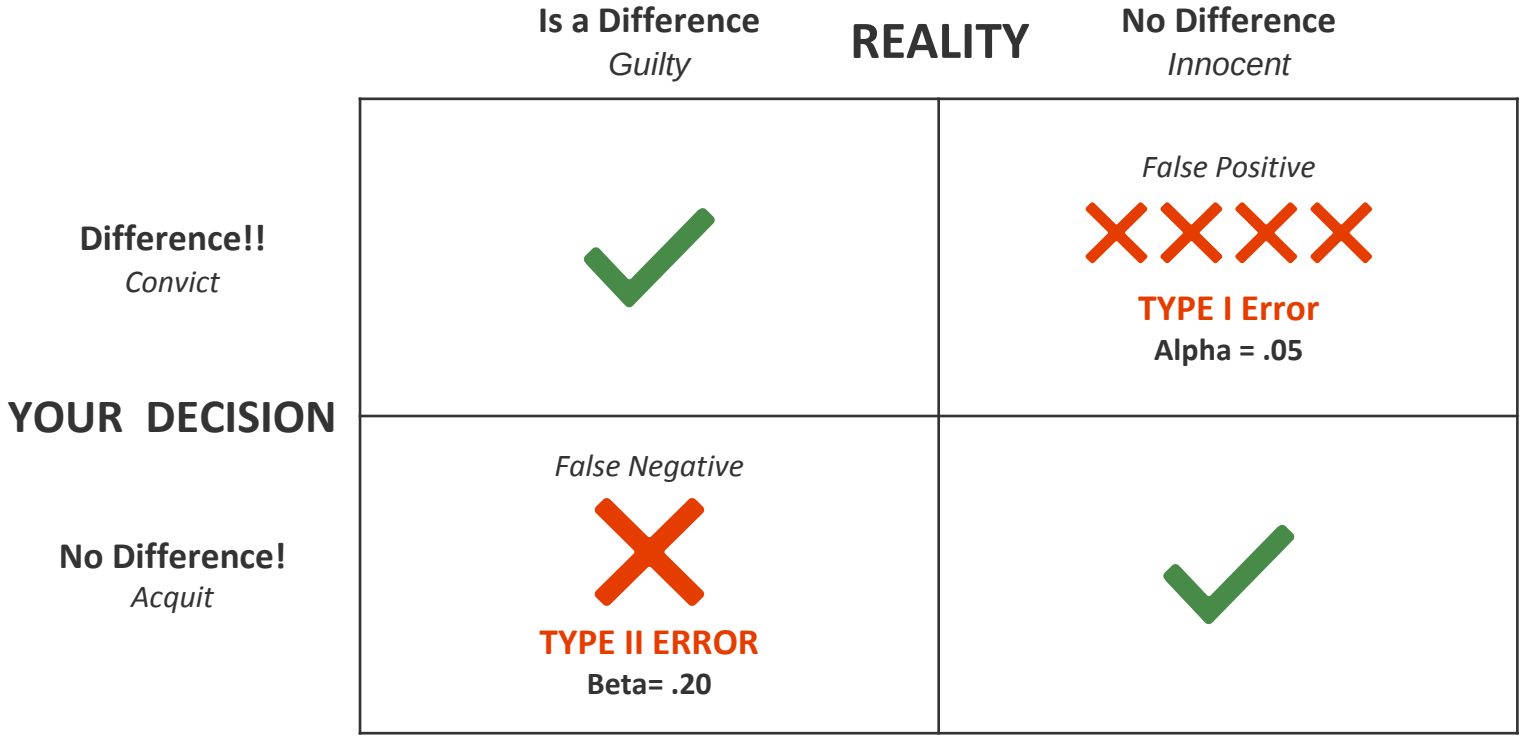
13 Clicks | 3,125,760 Views

.0004%

10 Clicks 1,041,515 Views

.001%

Hypothesis testing errors

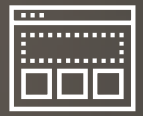


About MeasuringU

MeasuringU is a quantitative research firm based in Denver, Colorado focusing on quantifying the user experience.

MUIQ

Remote UX Testing Platform (Desktop & Mobile)



UX Research

SUPR-Q

Measurement & Statistical Analysis

Eye Tracking & Lab Based Testing

DenverUX BOOTCAMP

UX Boot Camp Aug 16th-18th
denverux.com

MeasuringU.com
@MeasuringU