

Thinly Sliced: A Researcher's Guide to Answering Non-Researchers' Sample Size Questions

By: Chandler Mercer & Edward Staples

Prepared for the Consumer Research Conference

About Us



Chandler Mercer | Head of Content Insights

Chandler is the Head of Content Insights at Spotify, a multi-disciplinary team of researchers that produce insights to help the company create and acquire the world's best audio content.

He is the co-founder of an internal quantitative consumer research guild and is a board member of the research practitioners group, ensuring quality research practices across the company.

Nearly a decade in the research industry has led Chandler to working across multiple industries and fortune 500 companies, both internally and externally.

Chandler also holds a master's degree in market research and has a passion for progressing the discipline internally, while helping other researchers produce methodologically rigorous insights.



Edward Staples | Sr. Director, Business Development

Edward has 15+ years of client- and supplier-side experience in marketing and market research, including 5 years with Kantar. His clients include Market Research firms and Fortune 500 companies across a spectrum of industries.

At Prodege, Edward is responsible for supporting new and existing client relationships through a consultative approach, and with a best-in-class support team.

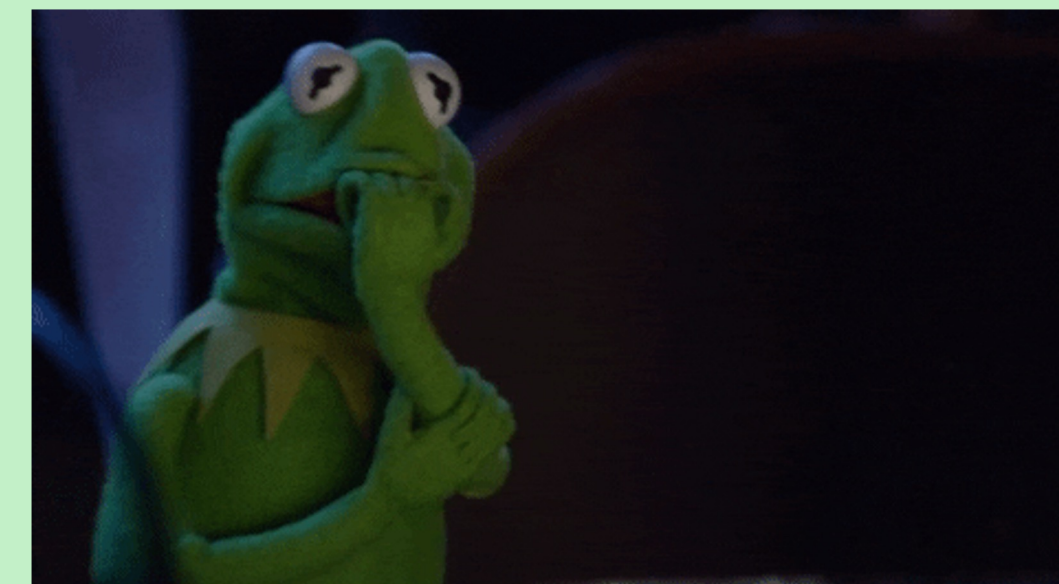
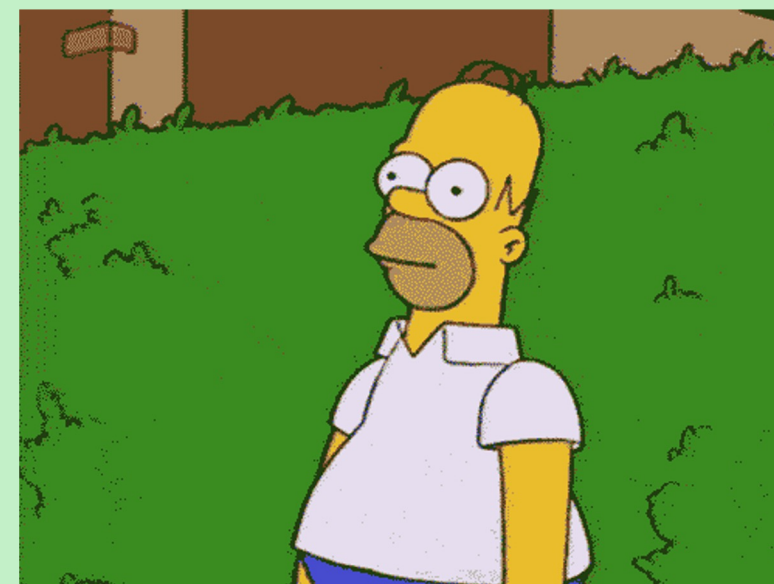
Prodege is a full-service market research firm whose proprietary panel provides high quality results and advanced targeting solutions to ensure the right respondents for your research. We offer self-service options in addition to full service solutions that include research design consultation and insights curation, all available in an easy-to-use data visualization platform. With no subscriptions or license fees, we welcome projects large and small.

**In the true spirit of Halloween,
today we are going to cover a
very scary topic...**

Raise your hand if...

you've ever laid awake at night debating how to answer the age old stakeholder question:

“How many responses do we need for the data to be statistically significant?”



But fear not!

**With a little foresight &
communication, we can
rest easy at night!**



So today we want to provide a playbook for answering this question.

We'll cover...

1. Getting to the bottom of their question
2. The two most common underlying questions
3. How to answer both of them
4. Real world Spotify examples

**The most important step
happens to be the easiest!**

**Just start by simply asking them to
clarify what they really want to know.**



1. Is the sample size large enough for the results to be statistically *defensible*?



MORE COMMON

2. Is the sample size large enough to *detect a significant/meaningful result* of a statistical test?



MORE COMPLEX

**Each question demands
a much different response.**

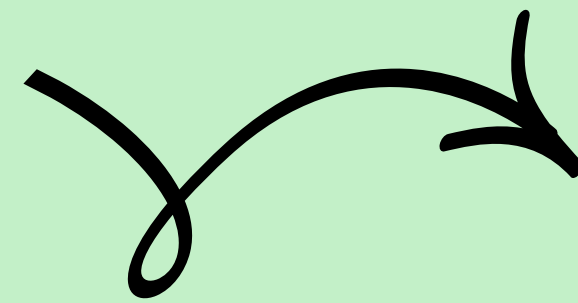
Let's dive into each!

Starting with the first question...

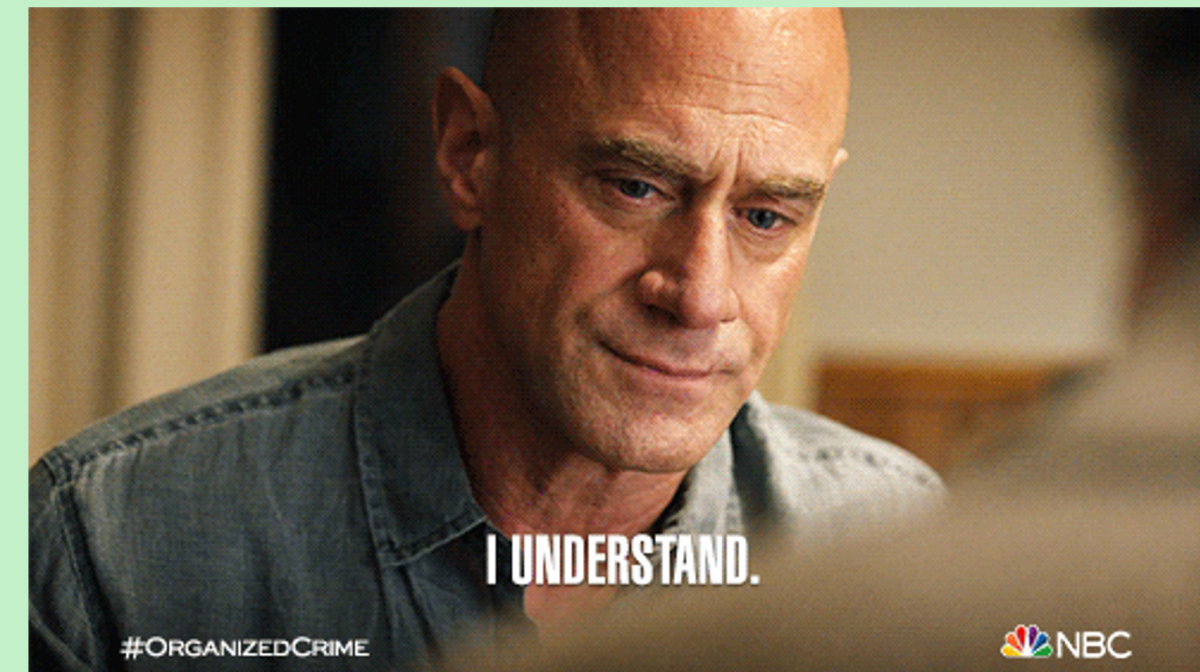
“is the sample size large enough for the results to be statistically defensible?”

What they're really asking.

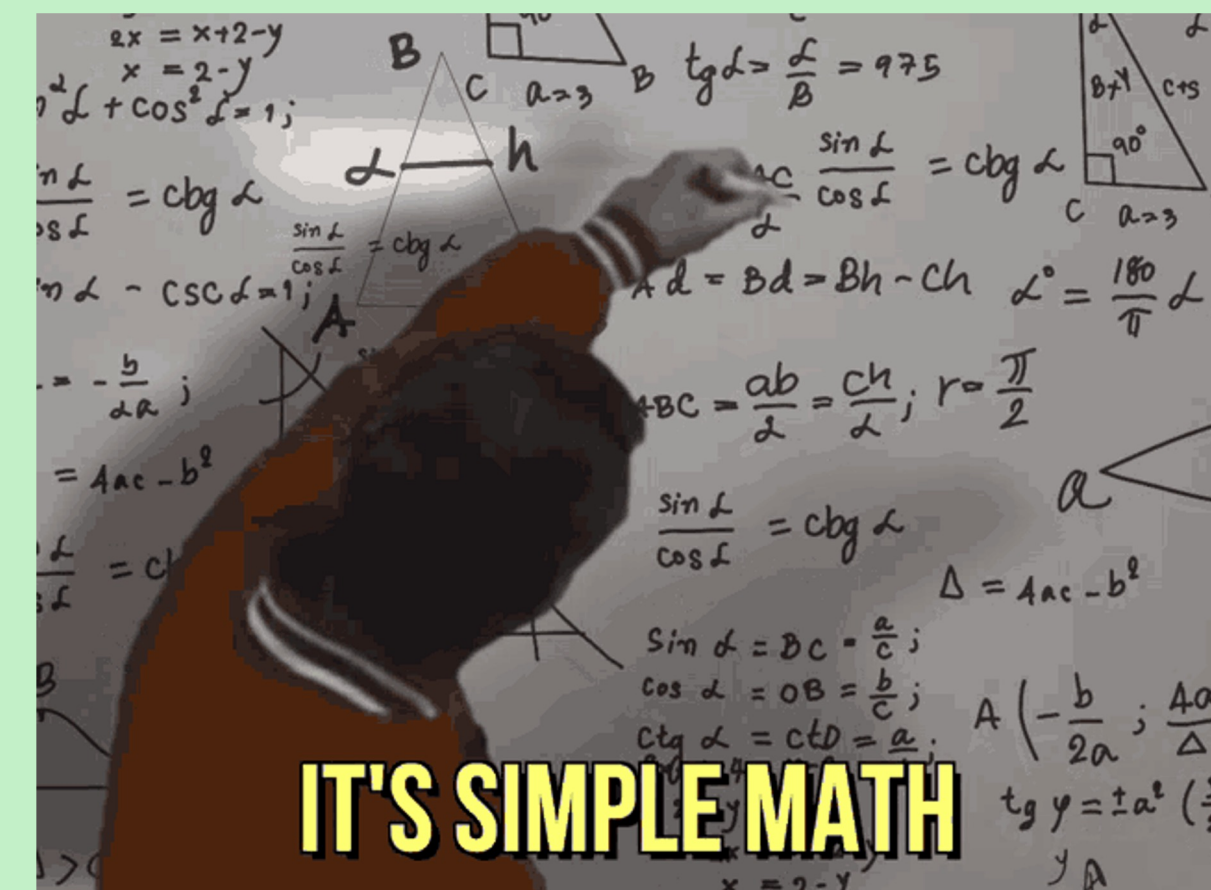
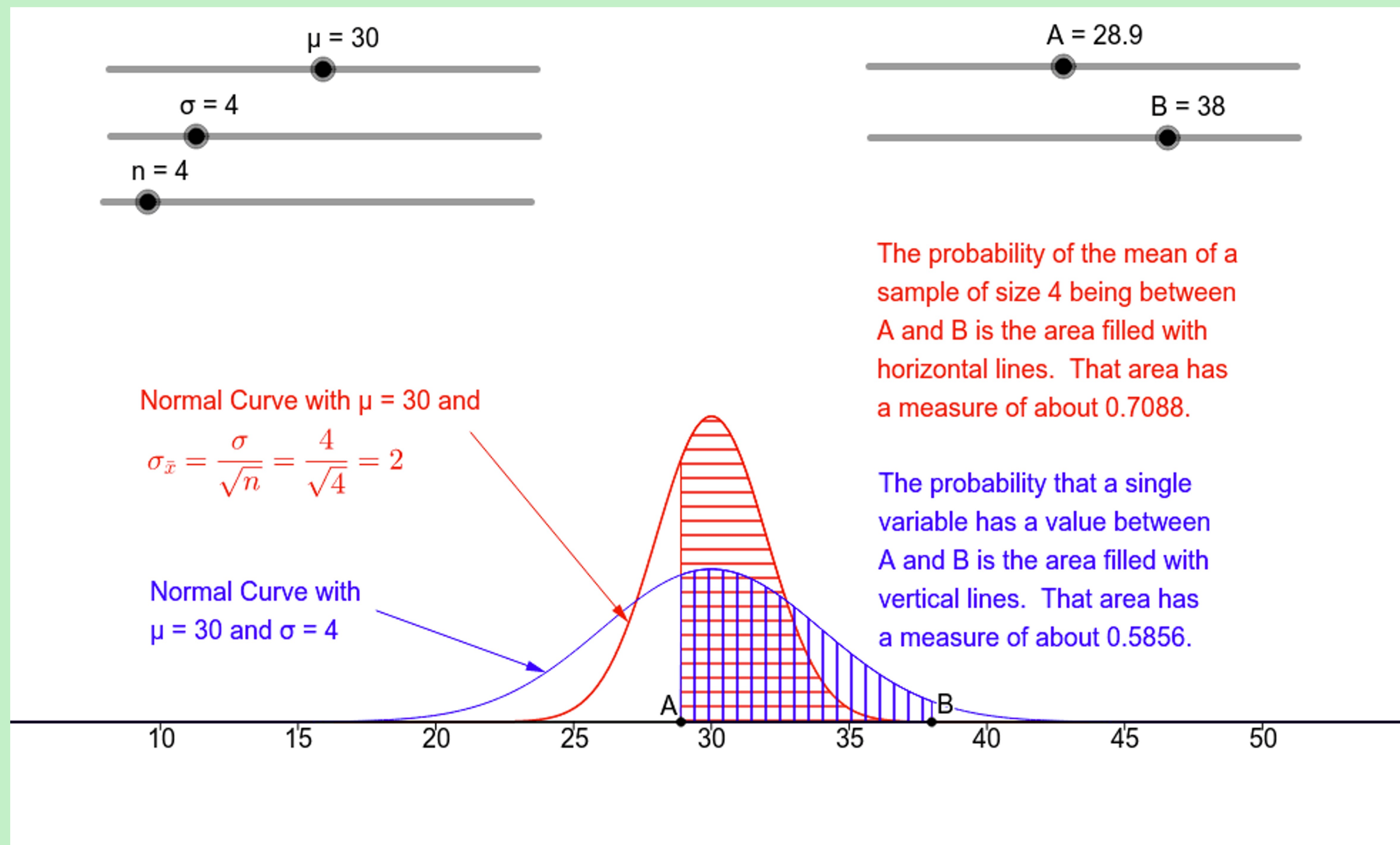
“Will the results be significant?”



“Can I trust the results?”

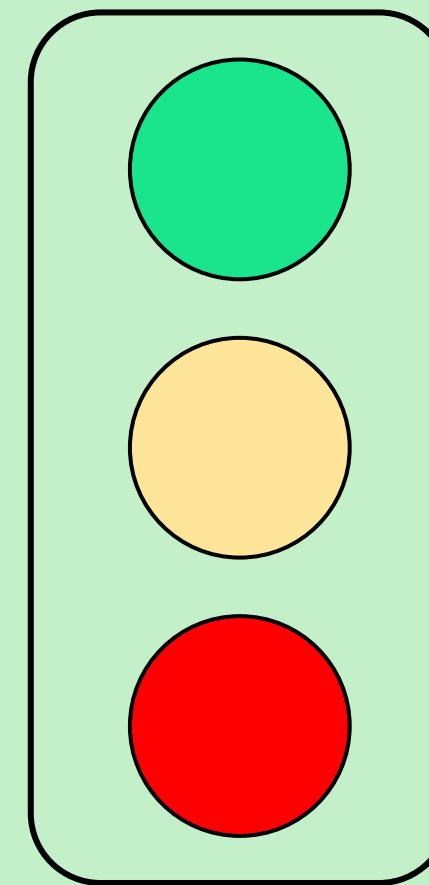
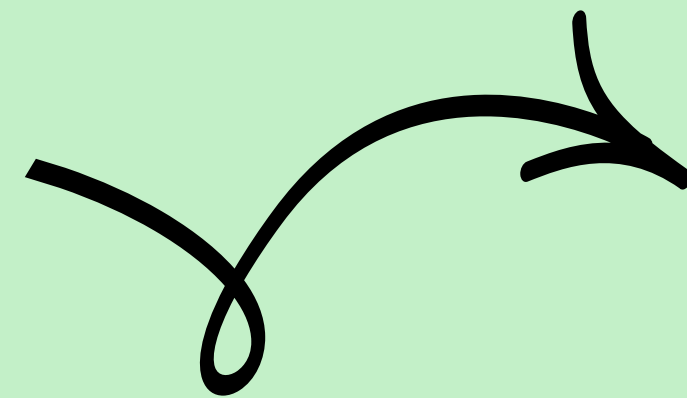


So they want confidence, give them confidence!



In other words...

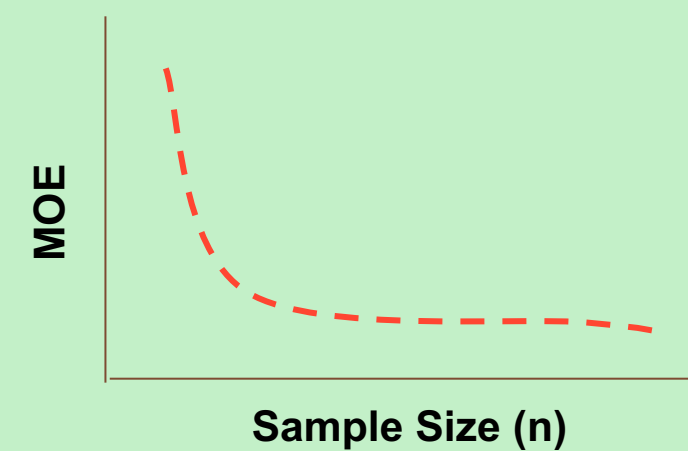
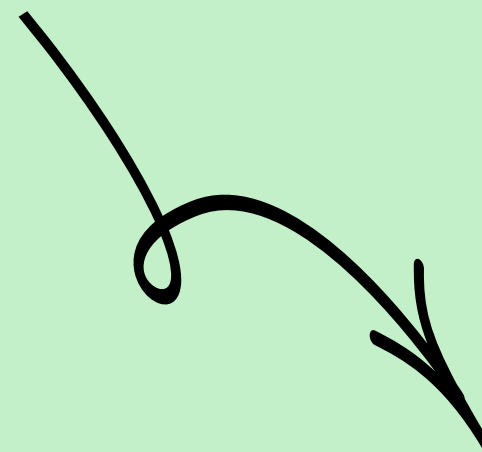
The Central Limit Theorem (CLT)
provides us with
rules of thumb 👍



$n > 100$
(*confidence**)

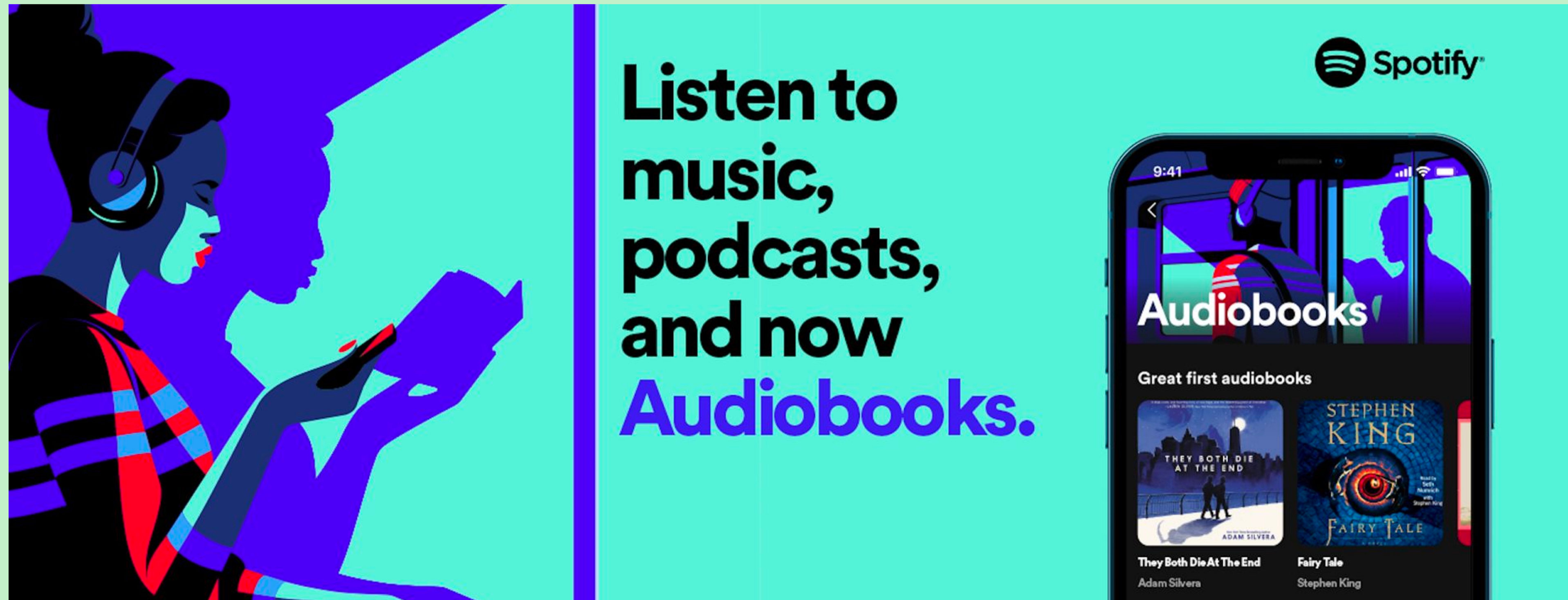
$n = 30 - 99$
(*strong pulse*)

$n < 30$
(*directional*)



**Well conducted studies will
see diminishing returns
with larger samples**

A Spotify example.



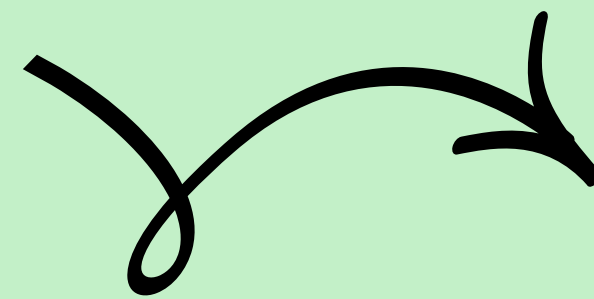
Tracking study with [prodege](#)

Now, that tricky second question...

“is the sample size large enough to detect a significant/meaningful result of a statistical test?”

What they're really asking.

“Will the results be significant?”



“Will we be able to detect a difference if it exists?”

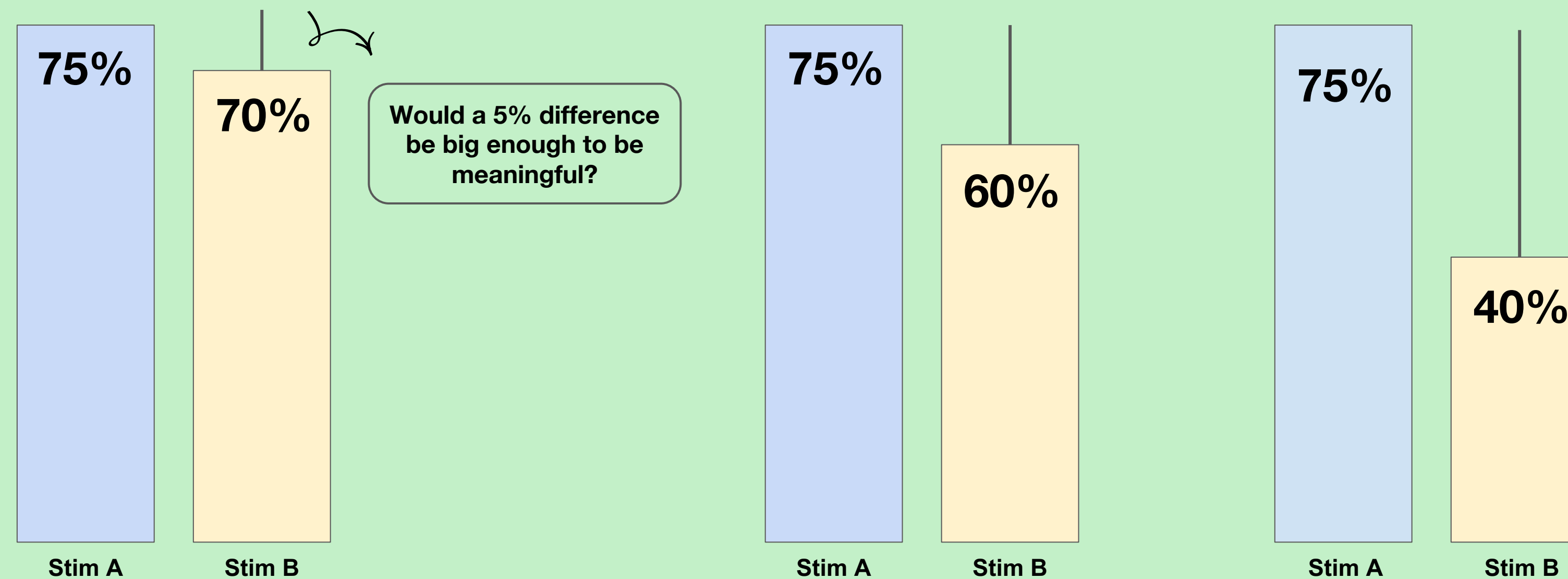


**With enough sample size,
tiny differences can be
detected as significant.**

But is that really helpful?

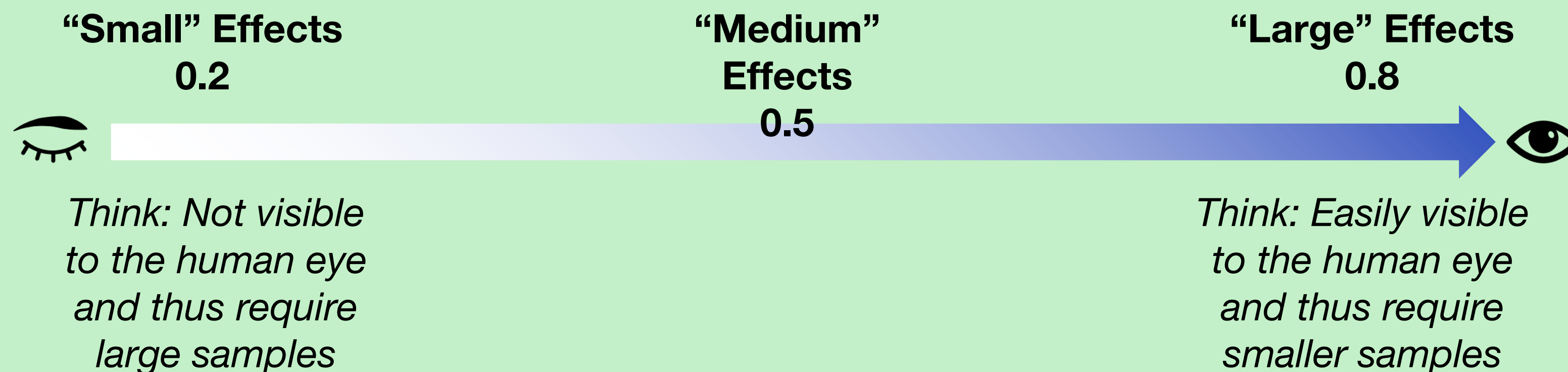
Focus on what is meaningful.

Have your stakeholder imagine all of the below results were significant. Which one would your stakeholder need to measure to make a decision?



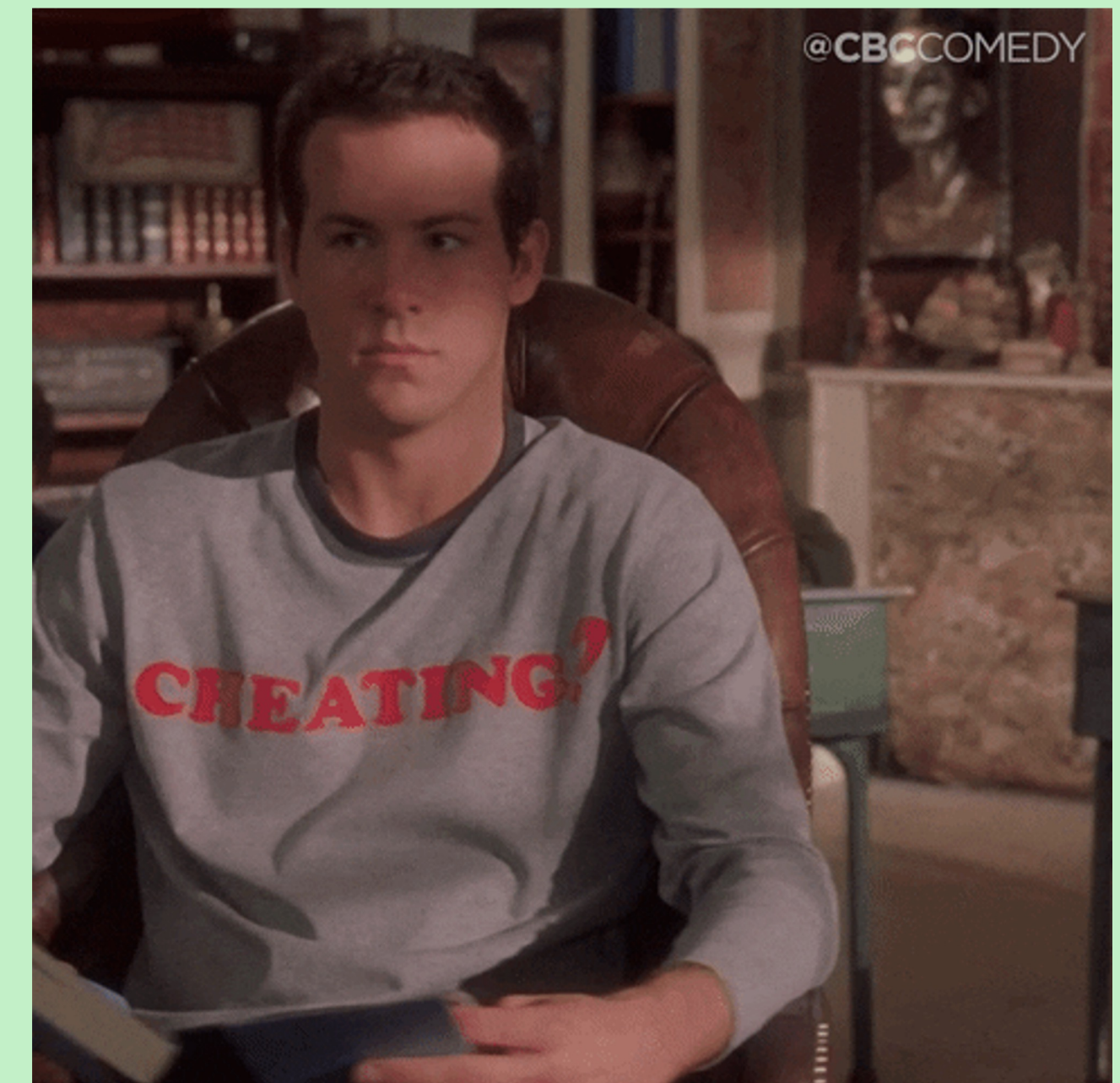
These guidelines help us make an informed decision about effect size.

Effect size (aka Cohens D) is used to measure how big or small the difference is between observations. This is represented by a decimal between 0 & 1 and roughly interpreted by the following:



**Time for the
cheat code...**

Enter Power Analysis!



Power Analysis determines what sample size will ensure a high probability of detecting a meaningful difference between two or more groups if it exists.

To use the cheat code, you must first know what buttons to press. 🎮

 **Power**

Likelihood a significance test will detect an effect when there actually is one.

Use 80-90%!

 **Effect Size**

Now you know what this one is all about!

Unless otherwise directed, use .5 to start.

α Significance

You know the drill!

Use 95% significance level!
(aka .05)

Now Google, plug, and chug!

Website A!

Results

The total number of subjects required: 128 (64 in each group)

Test family t-test

Sample groups Independent groups

Number of tails Two

 Effect size 0.5

α Significance level (α) 0.05

 Power 0.8

Submit

Website B!

A-priori Sample Size Calculator for Student t-Tests

This calculator will tell you the minimum required total sample size and per-group sample size for a one-tailed or two-tailed t-test study, given the probability level, the anticipated effect size, and the desired statistical power level.

Please enter the necessary parameter values, and then click 'Calculate'.

Anticipated effect size (Cohen's d): 0.5

Desired statistical power level: 0.8

Probability level: 0.05

Calculate!

Minimum total sample size (one-tailed hypothesis): 102

Minimum sample size per group (one-tailed hypothesis): 51

Minimum total sample size (two-tailed hypothesis): 128

Minimum sample size per group (two-tailed hypothesis): 64

Voila!

Return to your stakeholder and inform that you've done the math and know exactly what sample size you need for the test.



A Spotify example.



Ad-hoc study with [prodege](#)

Questions?

Why Prodege?

Cutting-Edge
Consumer
Intelligence
Platform

3,500+
Satisfied
Global Clients

Leading brands & agencies value Prodege's flexible "research for all" approach and impressive consumer reach, which makes it easy to uncover the answers to their business questions, big or small.

Our Unparalleled Client Support Services By the Numbers:

- 3,050+ surveys programmed in 2021
- 100+ global Project Managers with diversified experience; including firms such as Hall & Partners, Millward Brown and Nielsen
- 50+ onshore and offshore programmers providing 24/7 support and quick turnaround
- 20+ data operations experts to code brand & full open-ends, create complex tables & banners and provide weighting support

Awards & Affiliations





Thank You!

edward.s@prodege.com