



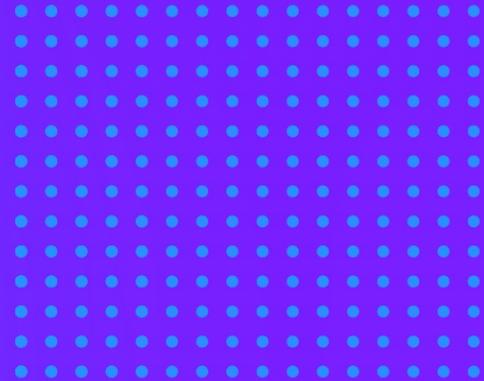
To Stat Sig and Beyond: Expand your experiment horizons today with bandit tests

Rebecca R. Carter

Principal Statistician
Product Manager - Statistics

Cam Hayes

Lead Strategy Manager
Experimentation Services



Who are we?

Rebecca

[linkedin.com/in/rebeccarcarter/](https://www.linkedin.com/in/rebeccarcarter/)



Friendly neighborhood statistician
Product roadmap: statistics
Powerful experiment designs

Cam

[linkedin.com/in/camhayes3/](https://www.linkedin.com/in/camhayes3/)



Lead Strategy Consultant
Experimentation Strategy Expert
MAB Fan

OPTICON

A blue-tinted photograph of a busy trade show floor. The word "OPTICON" is overlaid in large, white, bold, sans-serif capital letters across the center. The background shows a large exhibition space with many people networking, some sitting at small round tables. There are large, curved architectural elements in the background, possibly part of a stage or display. The ceiling has a grid of lights and a rig of stage lights. The overall atmosphere is professional and social.

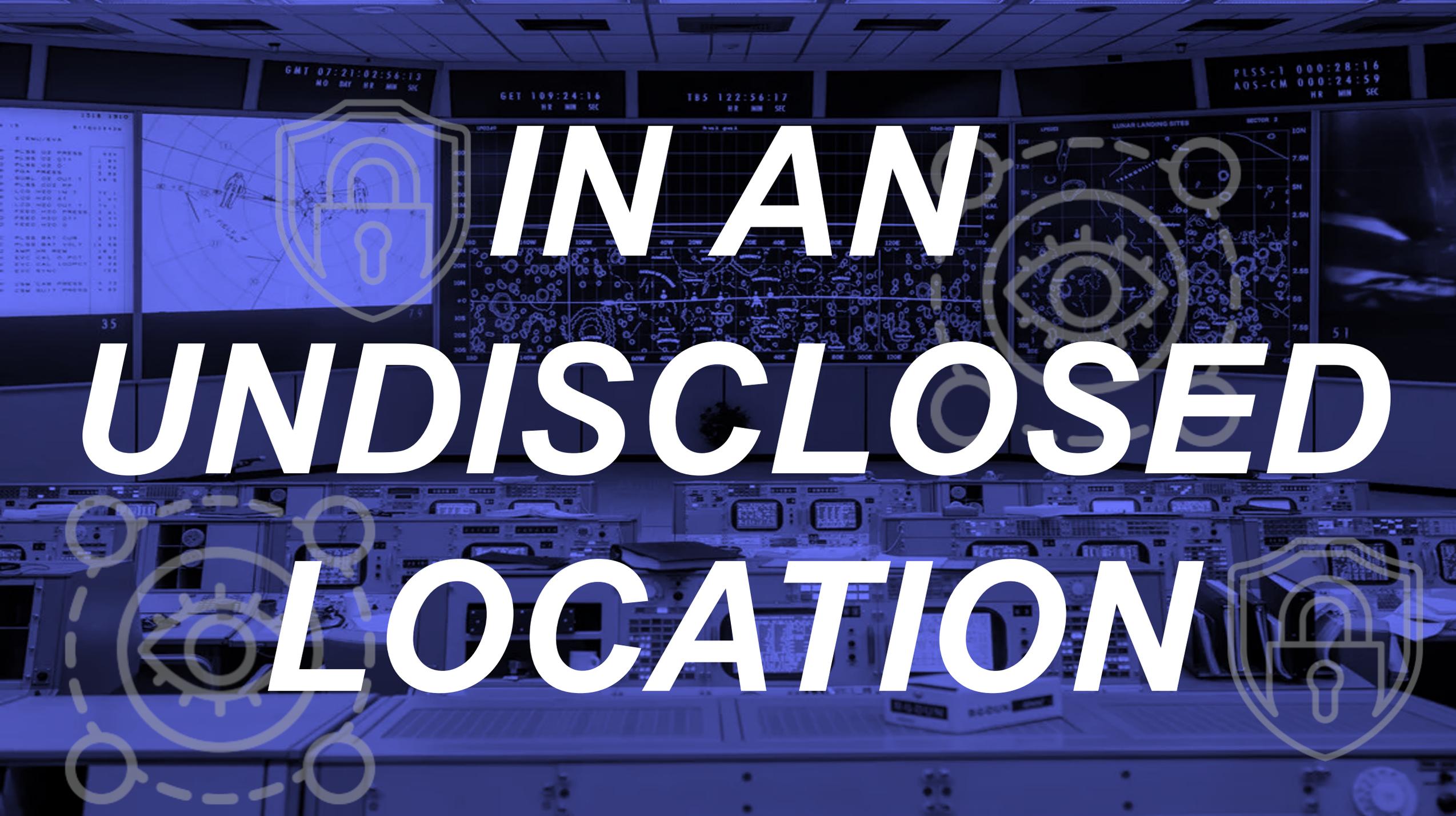


SANDIEGO



MISSION

CONTROL



GMT 07:21:02:56:13
MO DAY HR MIN SEC

GET 109:24:16
HR MIN SEC

TBS 122:56:17
HR MIN SEC

PLSS-1 000:28:16
AOS-CM 000:24:59
HR MIN SEC

IN AN

UNDISCLOSED

LOCATION

WELCOME

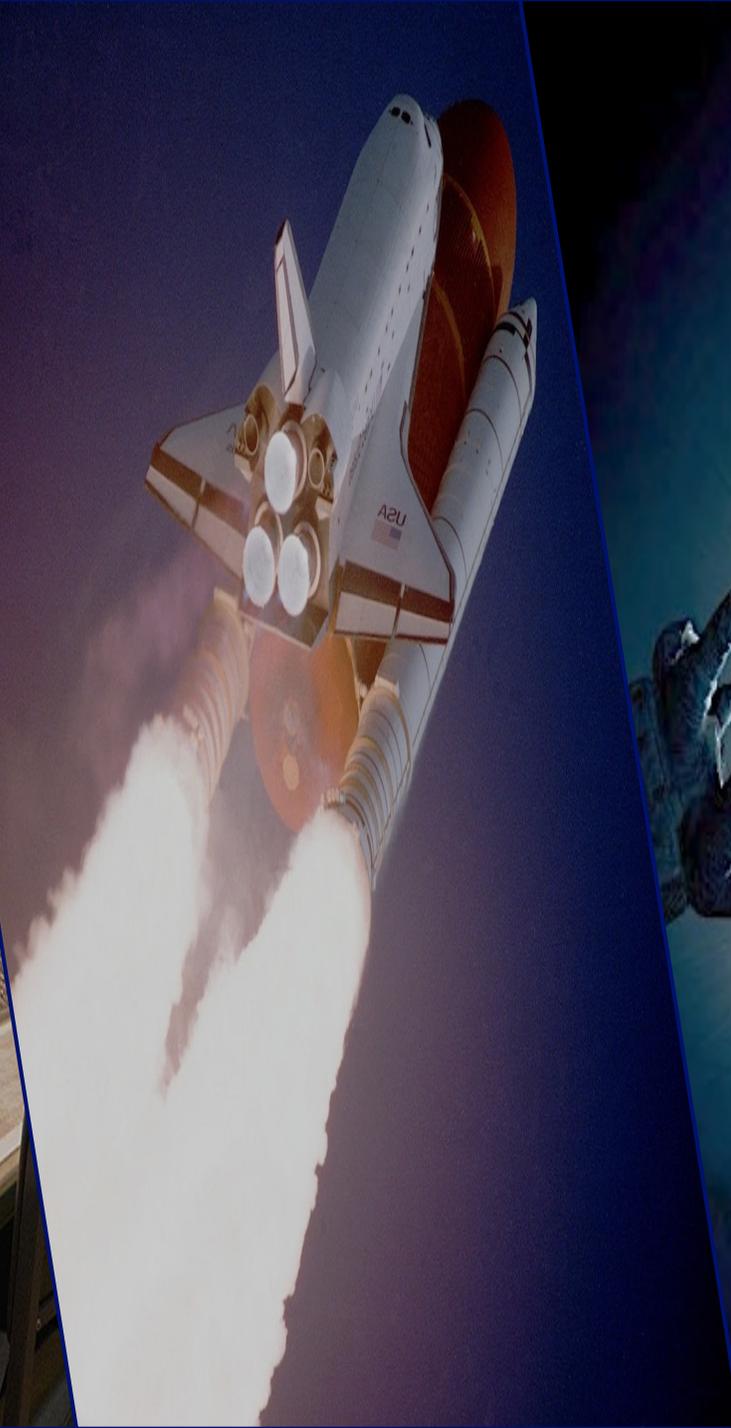
INTRERPID

HEROS





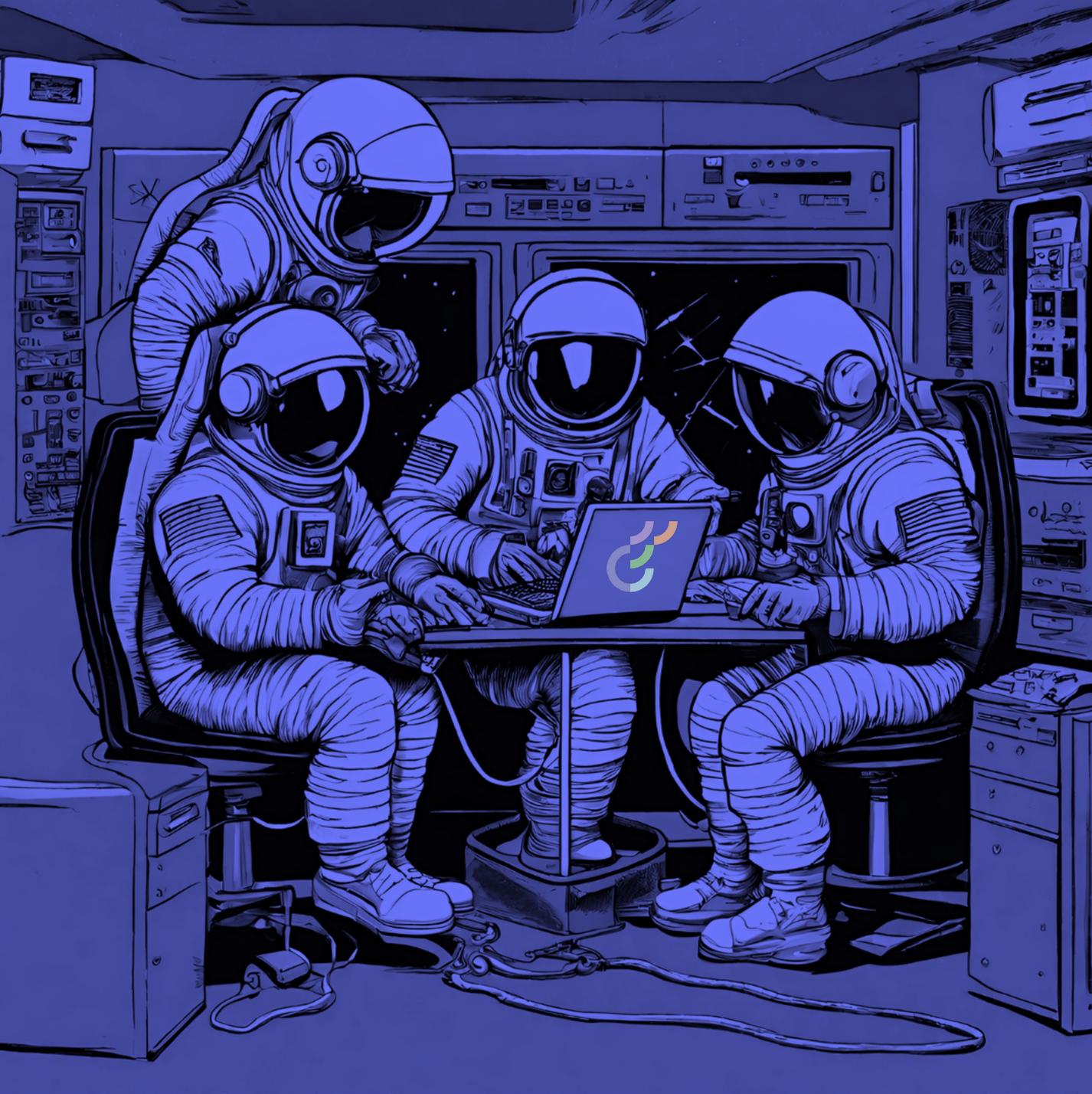
THIS IS YOU



01

LOOKING TO NEW HORIZONS

Leaving our comfort
zone of A/B testing



A/B tests: our home base

Nearly 8 out of 10
experiments run
on Optimizely are
with Stats Engine



We love A/B tests

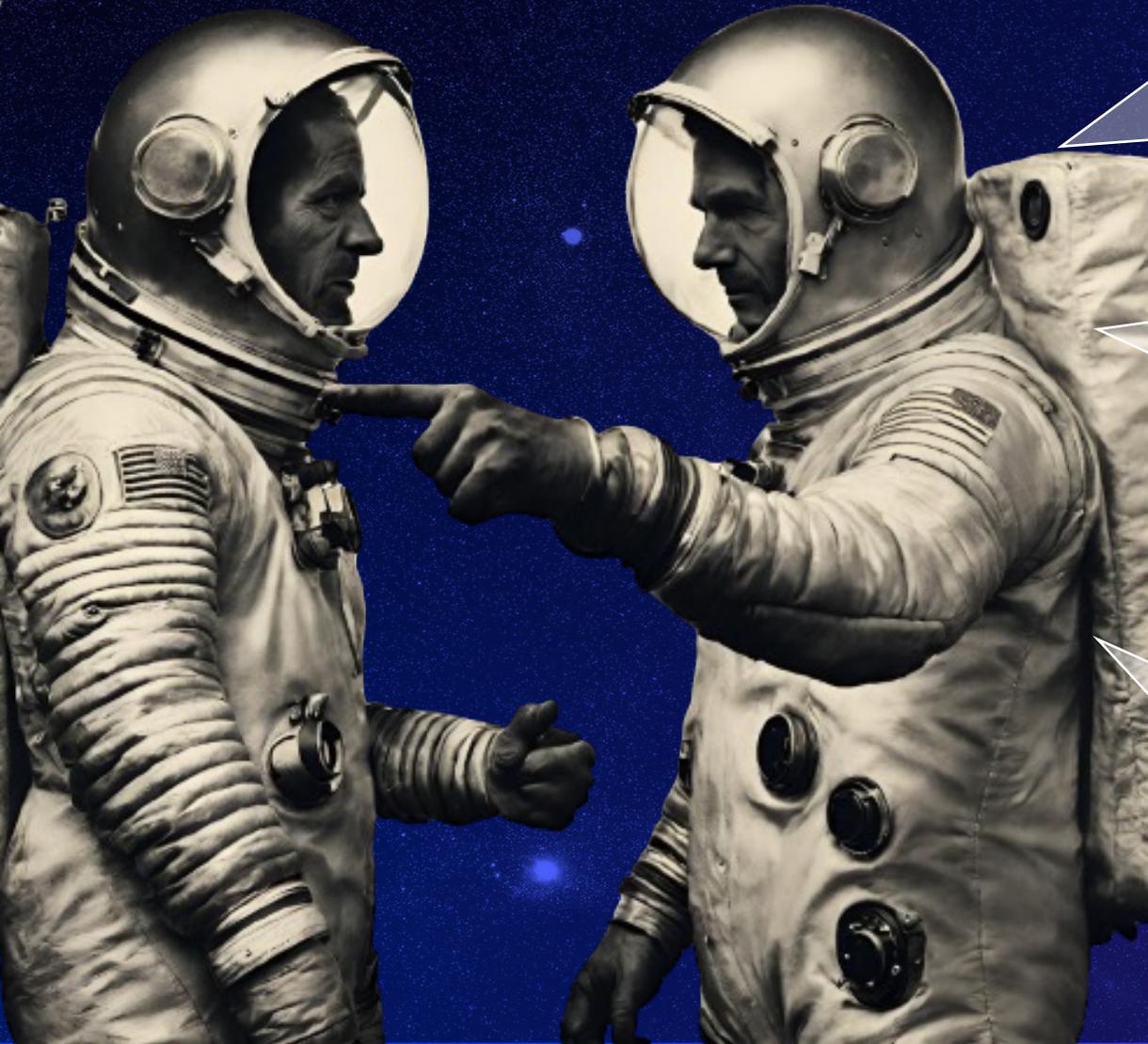
But we have **dreams**
of bigger, bolder,
faster experiments

But we hear about the many

uncertainties

that keep us **grounded** from launching a bandit test

BANDIT FUDS



**ITS TOO
RISKY**

**WE CAN'T RUN A TEST
DURING OUR BIGGEST
PROMO OF THE YEAR**

**IF THERE'S NO STAT
SIG IN THIS THING
WHAT IS THE POINT?!**



Experiment Results

[Michiel]
Experimentation vs A/...

Date Range: Oct 13, 2021 - Oct 25, 2021 | Segment: All Visitors | Baseline: A/B Testing and Feature Manag... | Print View | Export CSV | Share | Manage Metrics | Edit Experiment

A/B Test • Archived

Results last updated: October 25, 2021 at 3:55 PM

Last Published
Oct 25, 2021

Days Running
12

Audiences
Everyone

Page
URL Targeting for [Michiel]
Experimentation vs A/B testing in menu

Visitors
20,680

Description
From PM: People experimental testing to web and people to

Summary

Variations	Visitors	Visit Page: Visited ...	Web clicks	Full stack clicks
A/B Testing Feature Management	10,307 49.84%	--	0.99%	0.25%
Experimentation	10,373 50.16%	--	-55.19% 0.44%	-23.57% 0.19%

Statistical Significance

Baseline
>99%
Loser

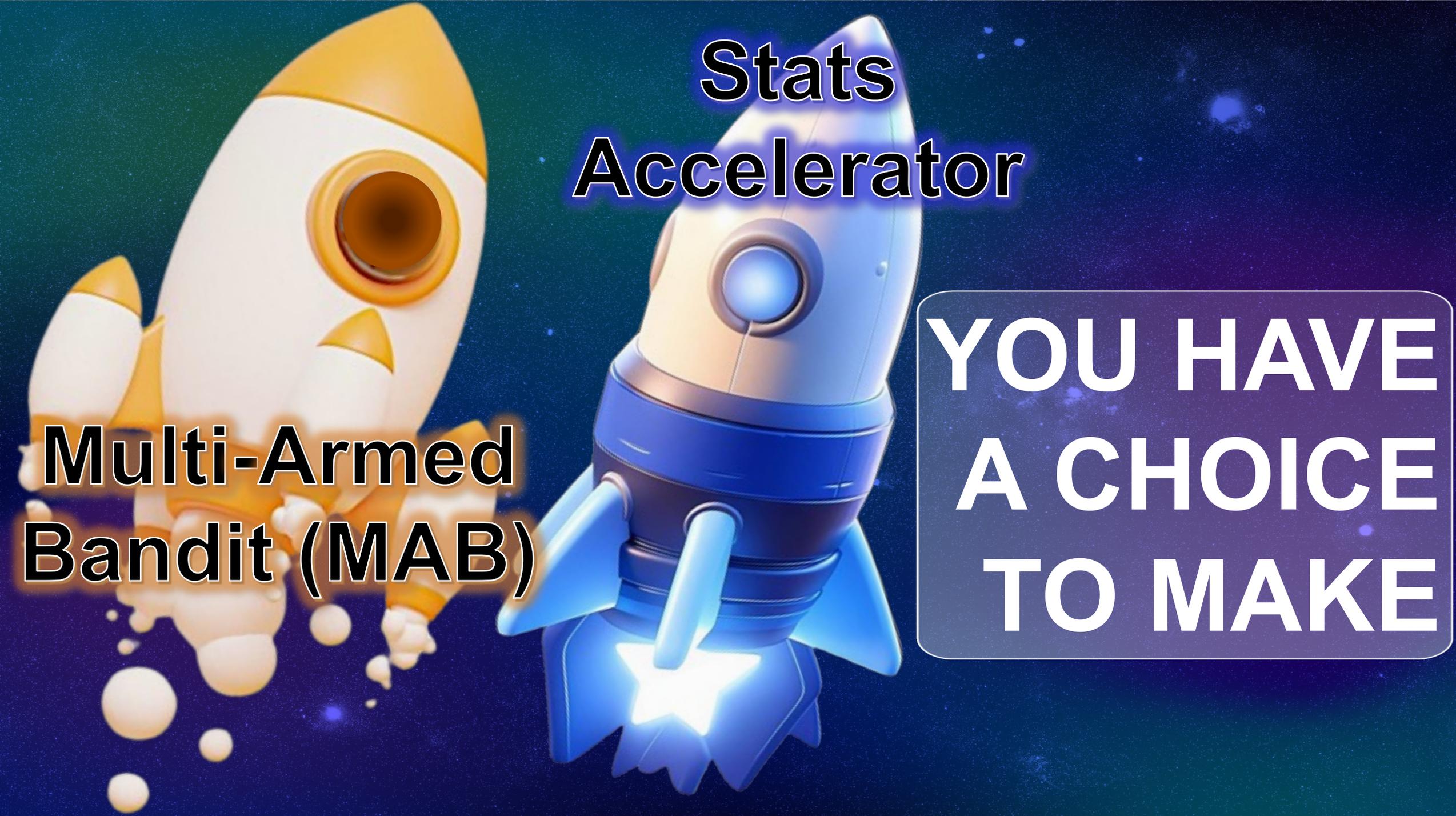
Improvement Over Time

Reset Results

02

LIFTOFF

**Launching with the right
tool for the right reason**



**Stats
Accelerator**

**Multi-Armed
Bandit (MAB)**

**YOU HAVE
A CHOICE
TO MAKE**



MULTI-ARMED BANDIT (MAB)

Adjusting its approach based on cumulative evidence

It's building a traffic distribution strategy

MAB's goal: Drive MORE conversions



START YOUR
EXPLORATIONS INTO
BANDITS WITH THE **MAB**

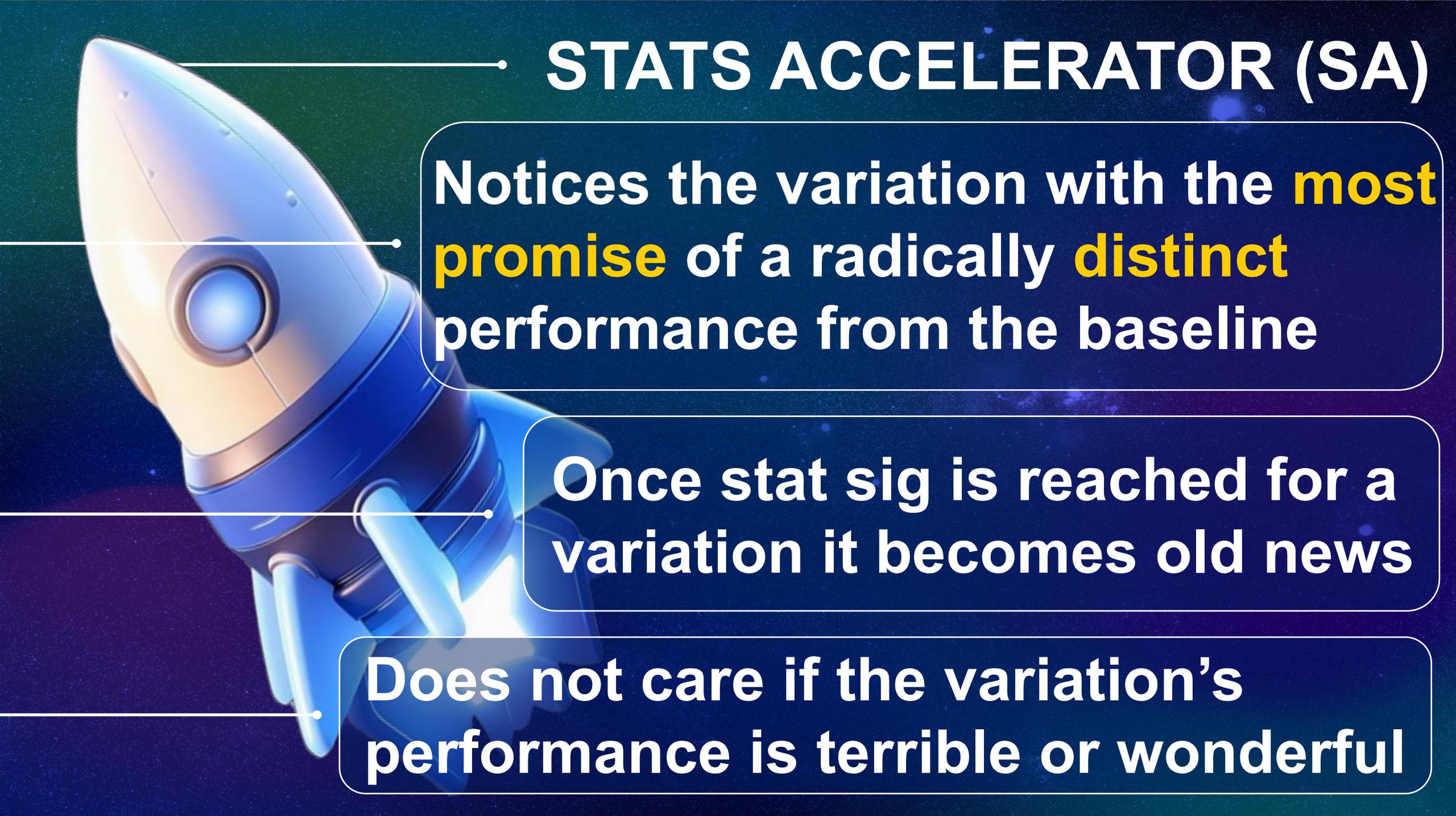


What can MAB do for you?

Maximizes short term gains

**Perfect for capturing value from
MASSIVE bursts of traffic**

Great for LP copy & form fills



STATS ACCELERATOR (SA)

Notices the variation with the **most promise** of a radically **distinct** performance from the baseline

Once stat sig is reached for a variation it becomes old news

Does not care if the variation's performance is terrible or wonderful



What can Stats Accelerator do for you?

Great for *bias to ship* situations

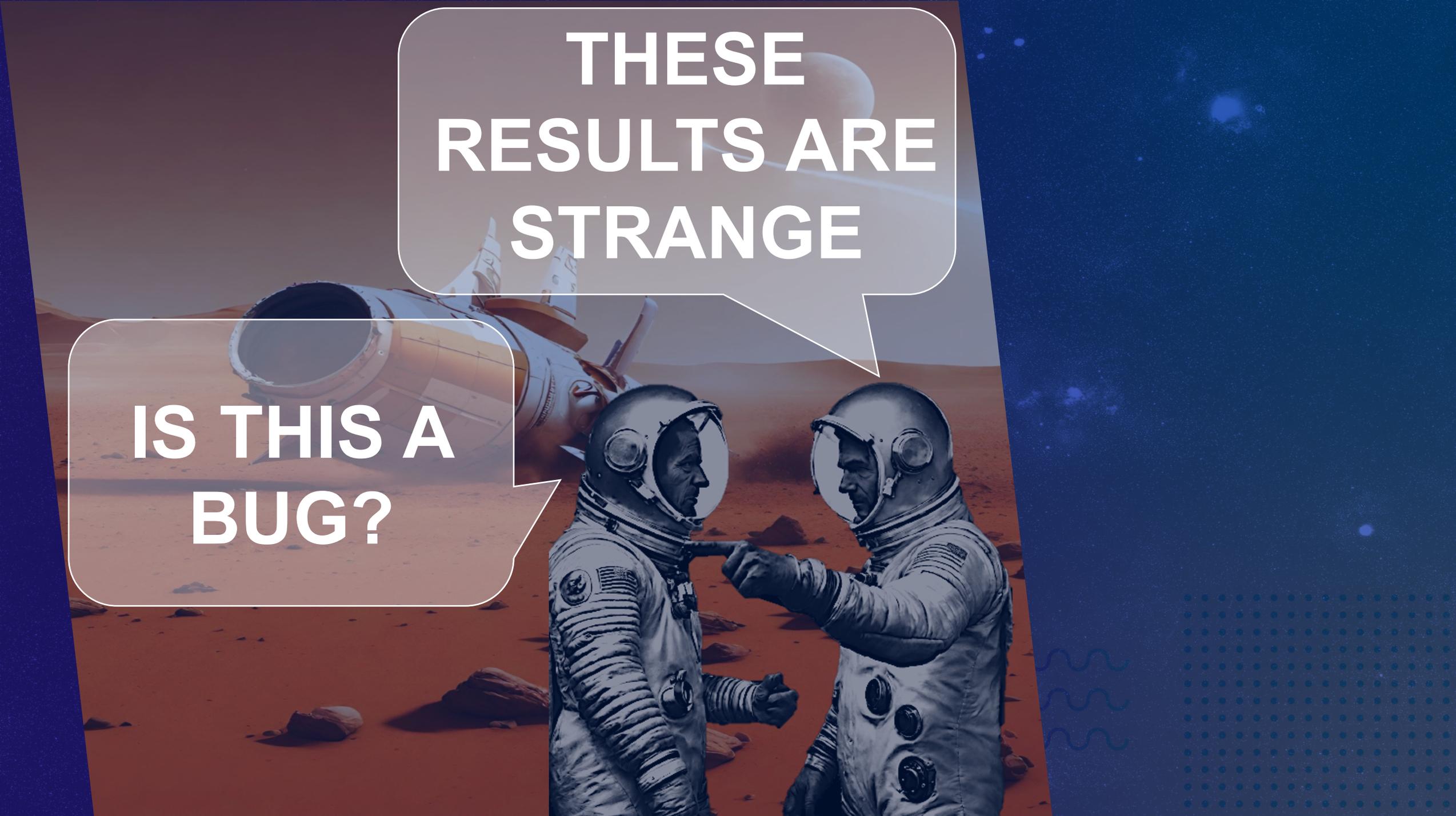
Excellent for choosing between multiple high performing variations

Useful when your decision will impact multiple areas of your site

03

BURNED BY BANDITS

Houston, we have a problem...

The image shows two astronauts in white space suits standing on a reddish-brown, rocky surface, likely Mars. In the background, a large white and orange spacecraft is partially visible. The scene is overlaid with a semi-transparent blue filter. A large white speech bubble is positioned at the top center, containing the text 'THESE RESULTS ARE STRANGE'. A smaller white speech bubble is on the left side, containing the text 'IS THIS A BUG?'. The right side of the image features a dark blue background with a pattern of small white dots and wavy lines.

**THESE
RESULTS ARE
STRANGE**

**IS THIS A
BUG?**

We need to talk about Stats Accelerator

Stats Accelerator identifies
a **sequence** of the most
interesting variations

If one variation has less traffic than
another despite *both* hitting stat sig...

Don't freak out!
That's totally normal here



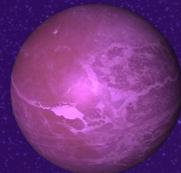
How Stats Accelerator navigates



Baseline is
first comparison



This variation's performance estimates are the **second-most** optimistic compared to the baseline



This variation's performance estimates are the **most** optimistic compared to the baseline

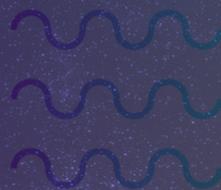


Here the performance estimates are kinda **similar** to the baseline

STATS ACCELERATOR



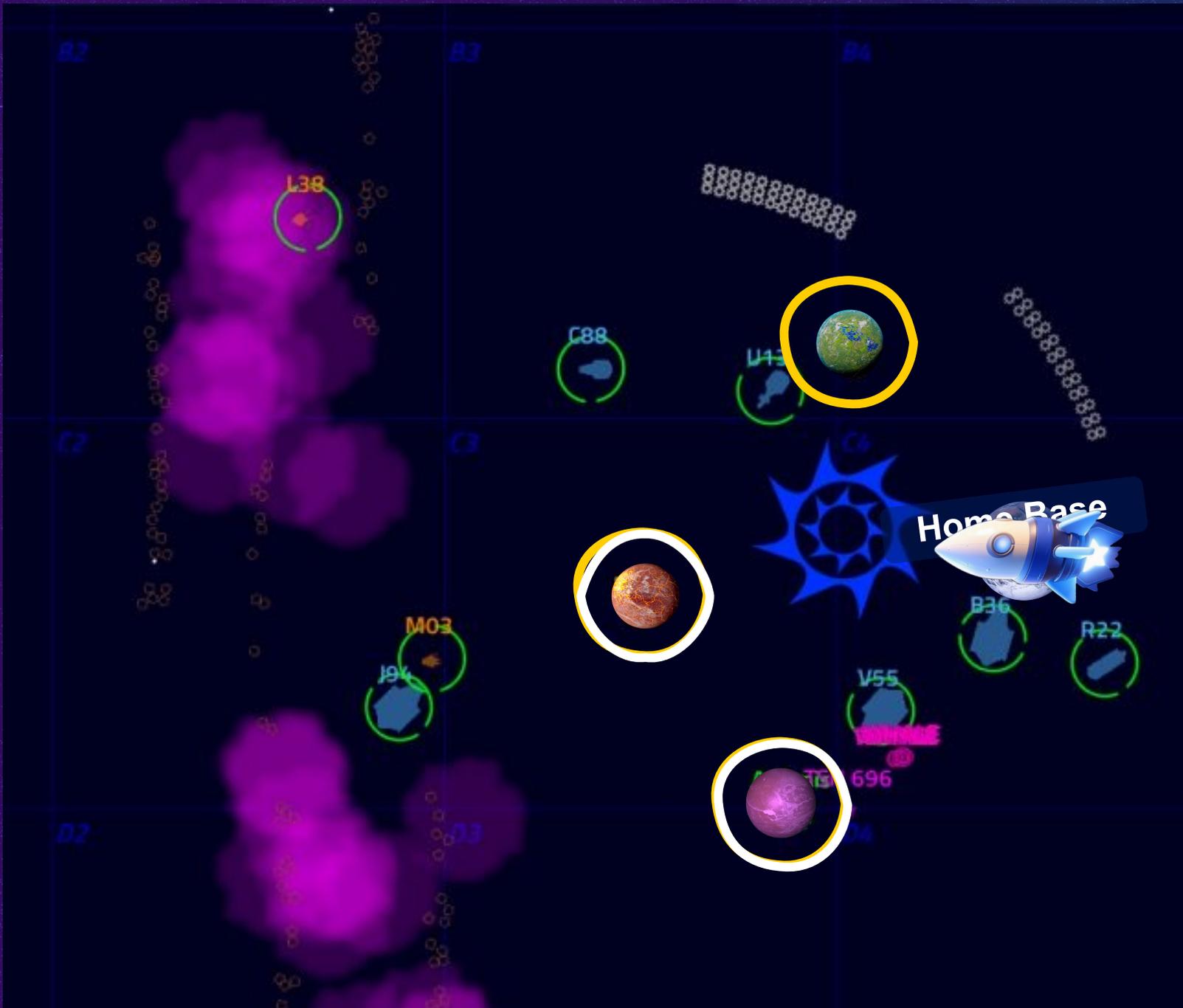
THUNDERDOME



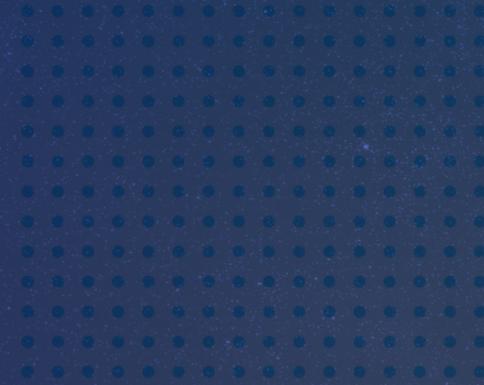
THUNDERDOME

First: Explore your options with **Stats Engine** and discover a few almost-winners or winners

Next: Throw 'em all in a battle royale with Stats Accelerator to discover the **winningest winner**



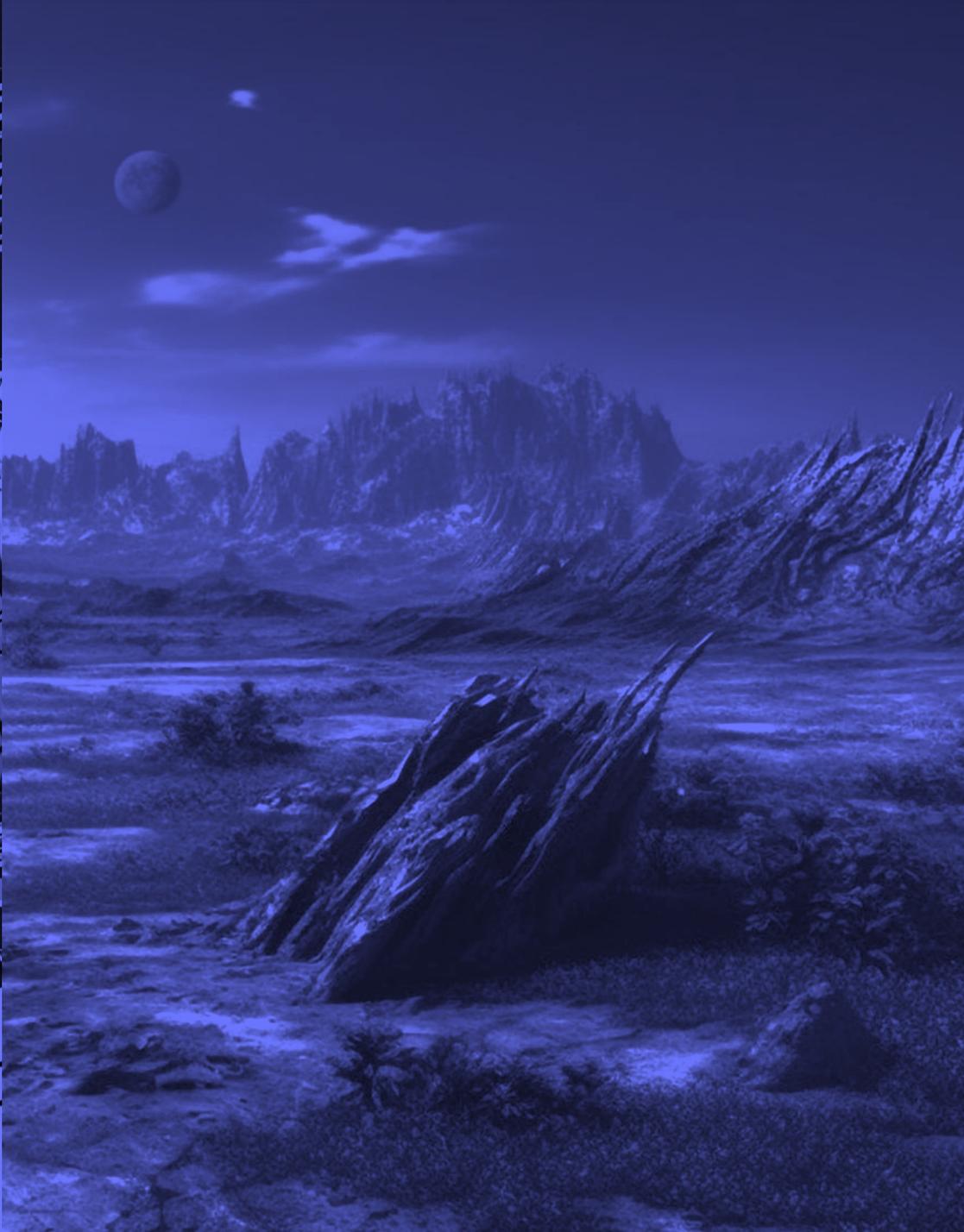
**Let's talk about
mishaps with MAB**





**IT'S TIME TO
PARTY**

**MISSION
ACCOMPLISHED**





MAB SMARTS

MABs are for improving

- Add to Cart
- CTAs
- Form fills

MAB results do not
have stat sig...

and that's OK!

04

THE HERO'S RETURN

... and gearing up for the next mission with a pre-flight checklist



WELCOME

BACK!



START

EXPLORATIONS
INTO BANDITS
WITH THE **MAB**

STATS ACCELERATOR

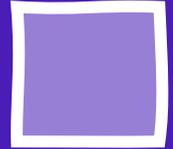


THUNDERDOME

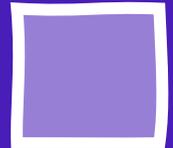


PREFLIGHT CHECKLIST

MAB



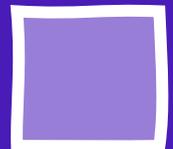
At least one test variation plus a baseline



You want to maximize improvement from a temporary experience



like headlines, sales, limited offers

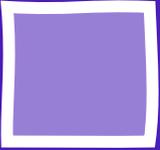
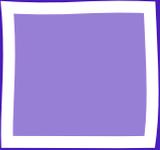
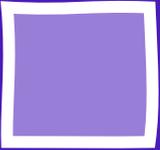


You don't need stat sig as your guide



PREFLIGHT CHECKLIST

Stats Accelerator

-  At least two test variations plus a baseline
-  You know how your users respond to the test variations
-  Speed and learning from the confidence interval and statistical significance values are a priority



Go forth and **boldly** experiment, intrepid heroes!



Inspired? Reach infinity and beyond
with your experiment program when
you partner with Optimizely's
strategy consultants!

alek.toumert@
optimizely.com (USA)

isabel.meijaard@
optimizely.com (EU/APJ)

Connect with us

Rebecca R. Carter
[linkedin.com/in/rebeccarcarter/](https://www.linkedin.com/in/rebeccarcarter/)

Cam Hayes
[linkedin.com/in/camhayes3/](https://www.linkedin.com/in/camhayes3/)

Slides will be available **after** Opticon