



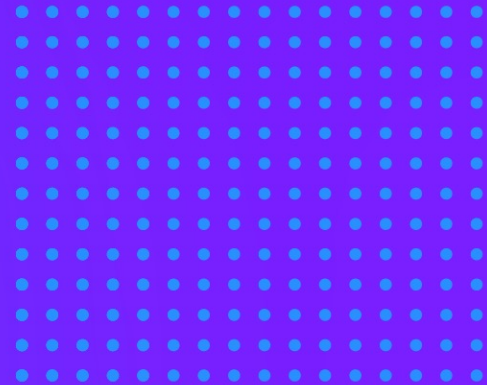
# To Stat Sig and Beyond: Expand your experiment horizons today with bandit tests

**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 crowded exhibition hall. In the foreground, several people are standing and talking. In the background, there are large curved structures, possibly part of a stage or display, and a large screen showing a presentation. The word 'OPTICON' is overlaid in large white letters across the center of the image.





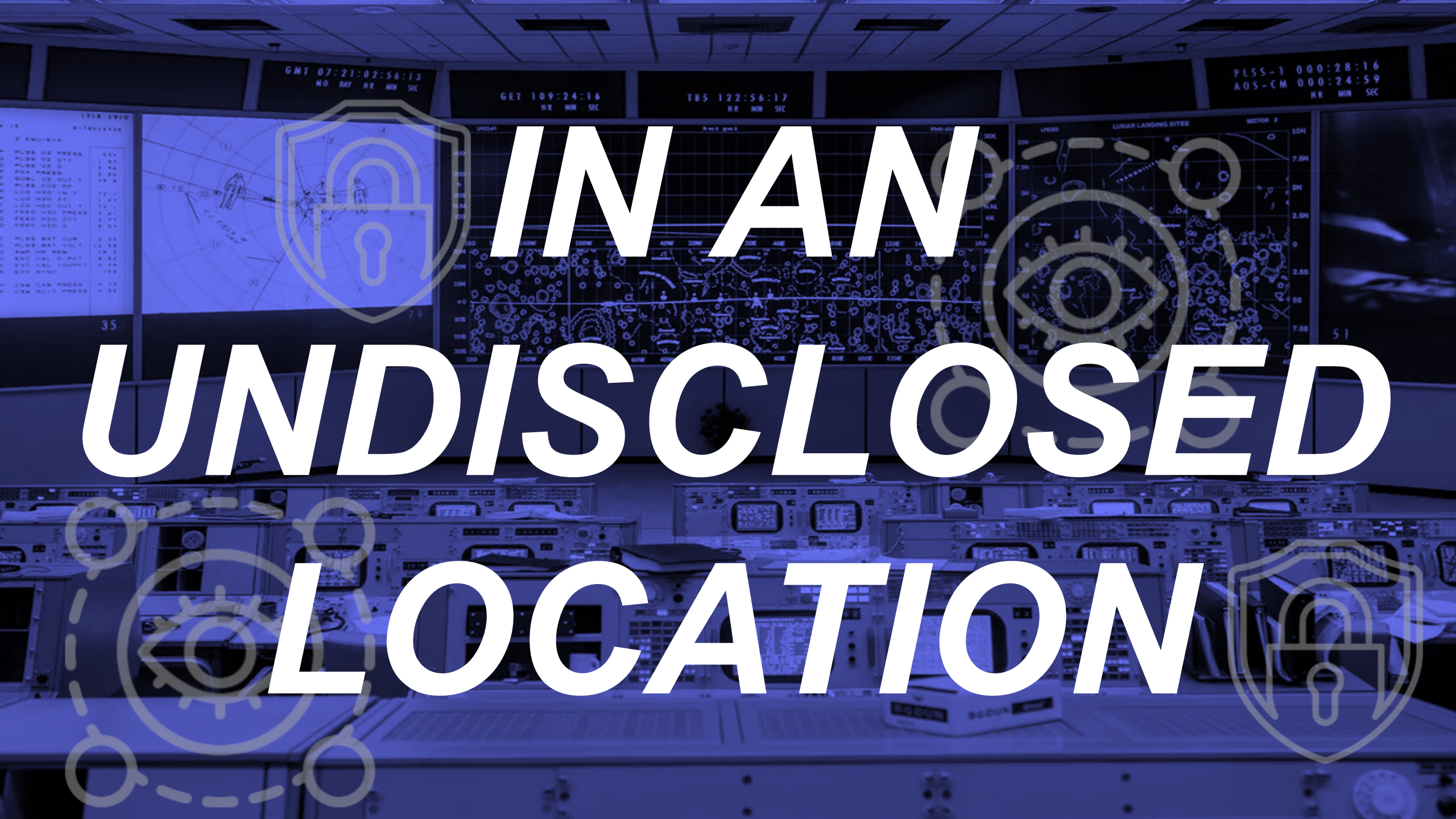
# SANDIEGO





# MISSION CONTROL





# *INAN UNDISCLOSED LOCATION*



**WELCOME  
INTREPID  
HEROS**

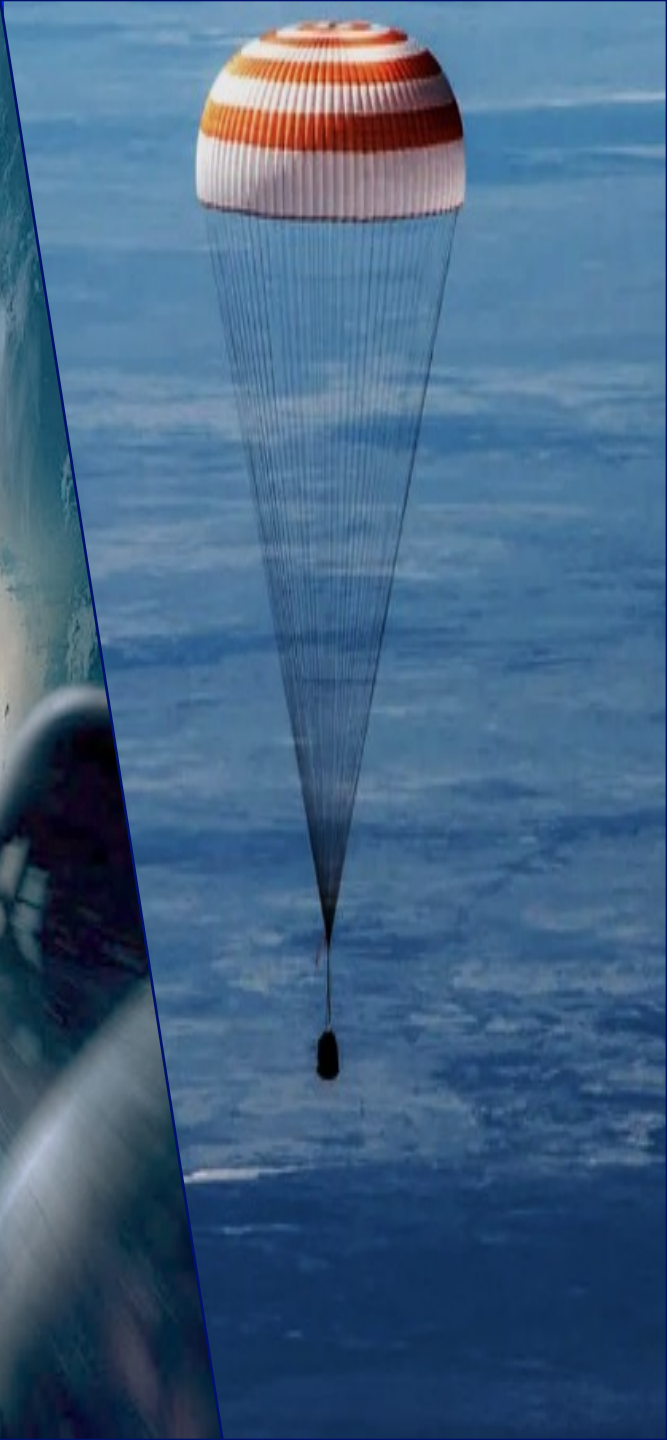
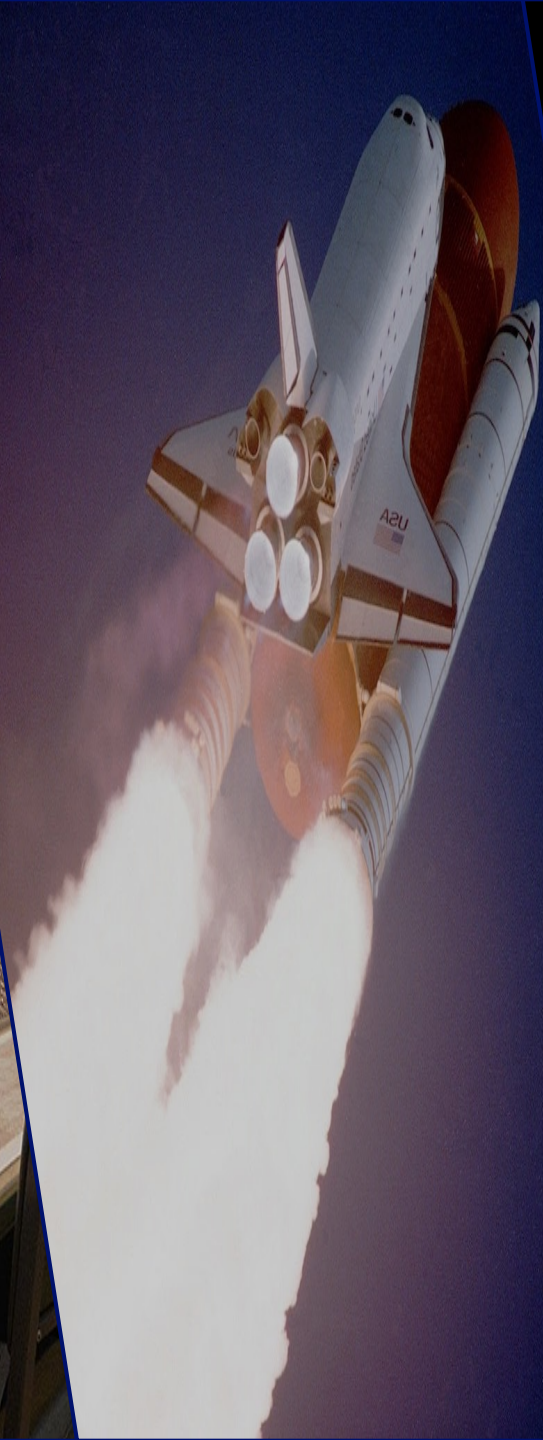






THIS IS YOU







# 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**



**IT'S 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?!**





Optimize WWWW - Web

← Variations

[Michiel]

Experimentation vs A/...

A/B Test • Archived



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

experimentati

testing to e

web and f

people to



Reset Results

## Experiment Results

Date Range

Oct 13, 2021 - Oct 25, 2021

Segment

All Visitors

Baseline

A/B Testing and Feature Managa...

Print View

Export CSV

Share

Manage Metrics

Edit Experiment

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

### Summary

Variations	Visitors	Visit Page: Visited ...	Web clicks	Full stack clicks
A/B Testing	10,307	--	--	--
Feature Management	49.64%	--	0.99%	0.25%
Experimentation	10,373	--	-55.19%	-23.57%
	50.16%	--	0.44%	0.19%

Link & Experimentation Edit

Visit Page: Visited Full Stack



Statistical  
Significance

--

Baseline

> 99%

Loser

Improvement Over Time



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 features two astronauts in white space suits standing on a reddish-brown, rocky surface that resembles Mars. In the background, a large, white, cylindrical spacecraft component lies on the ground. The sky is a hazy, orange-brown color. A large, semi-transparent white speech bubble with a black outline is positioned in the upper right, containing the text 'THESE RESULTS ARE STRANGE'. Another smaller, similar speech bubble is on the left, containing the text 'IS THIS A BUG?'. The right side of the image transitions into a dark blue space background with stars and a grid of small dots.

**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

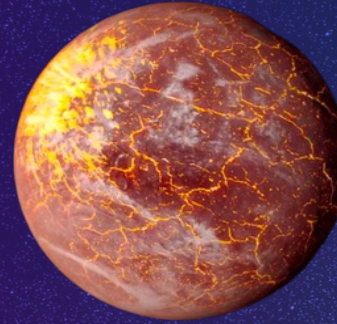




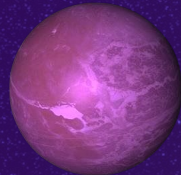
# How Stats Accelerator navigates



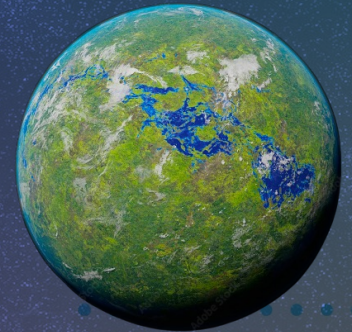
Baseline is  
first comparison



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



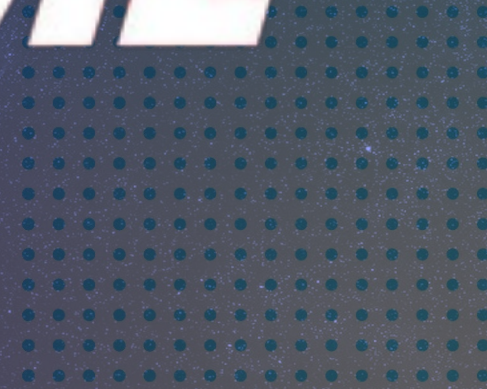
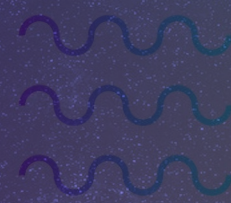
Here the performance estimates are kinda **similar** to the baseline



STATS ACCELERATOR



**THUNDERDOME**

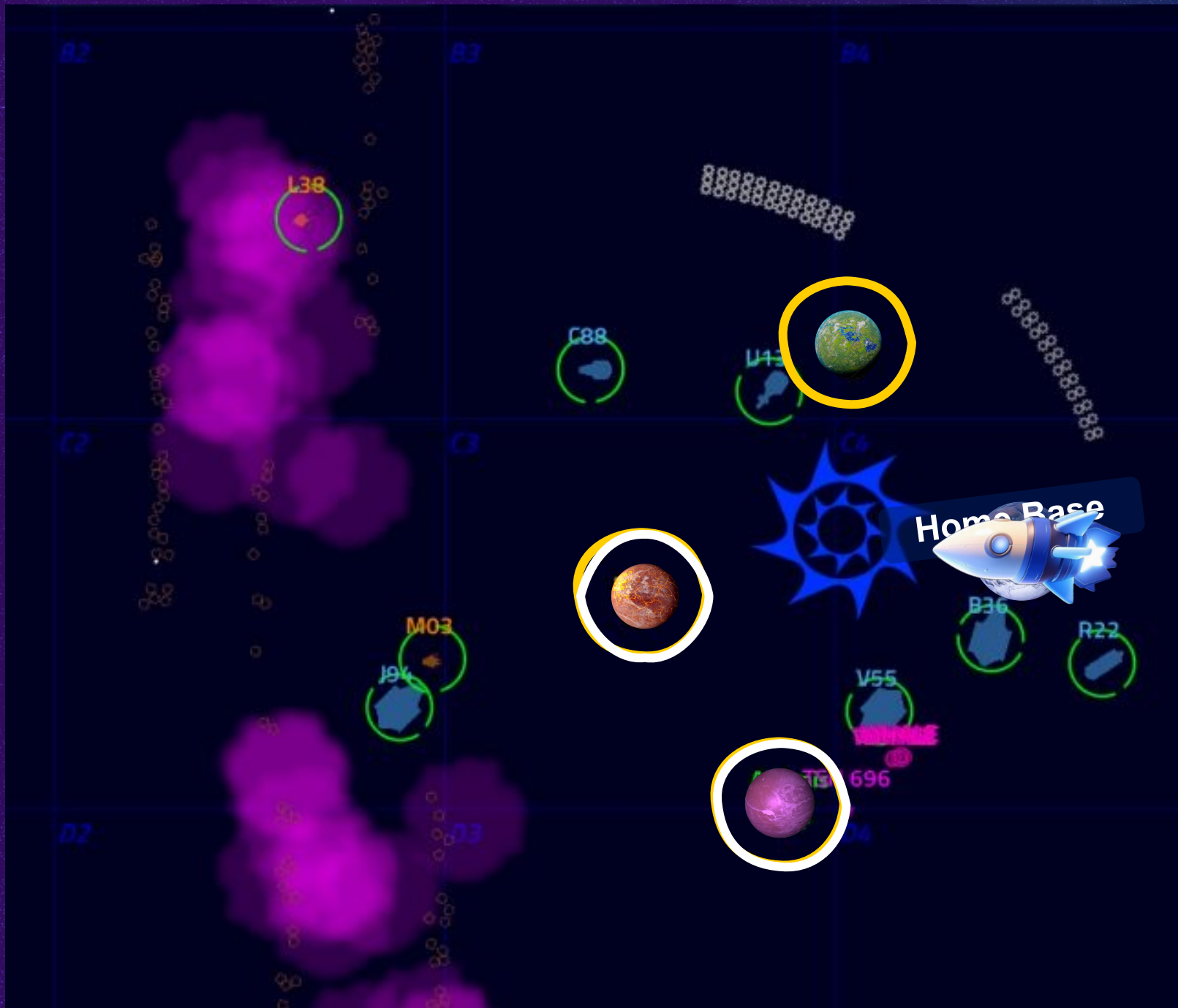
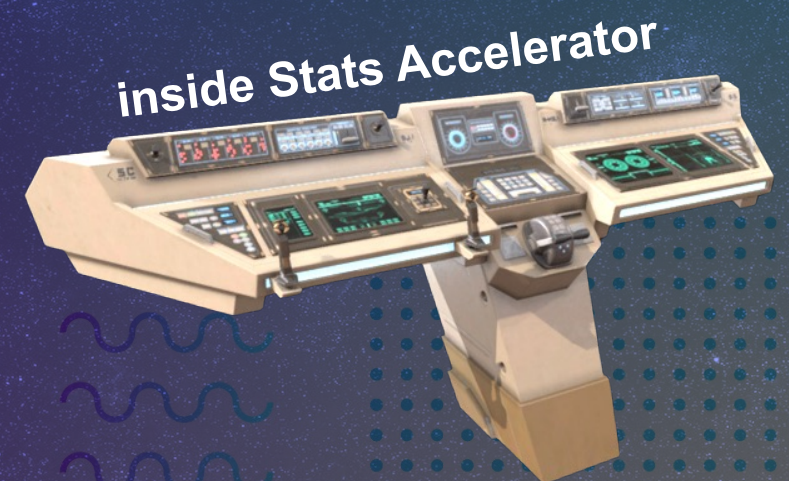




# THUNDERDOME

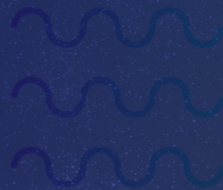
First: Explore your options with **Stats Engine** and discover a few almost-winners or winners

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





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



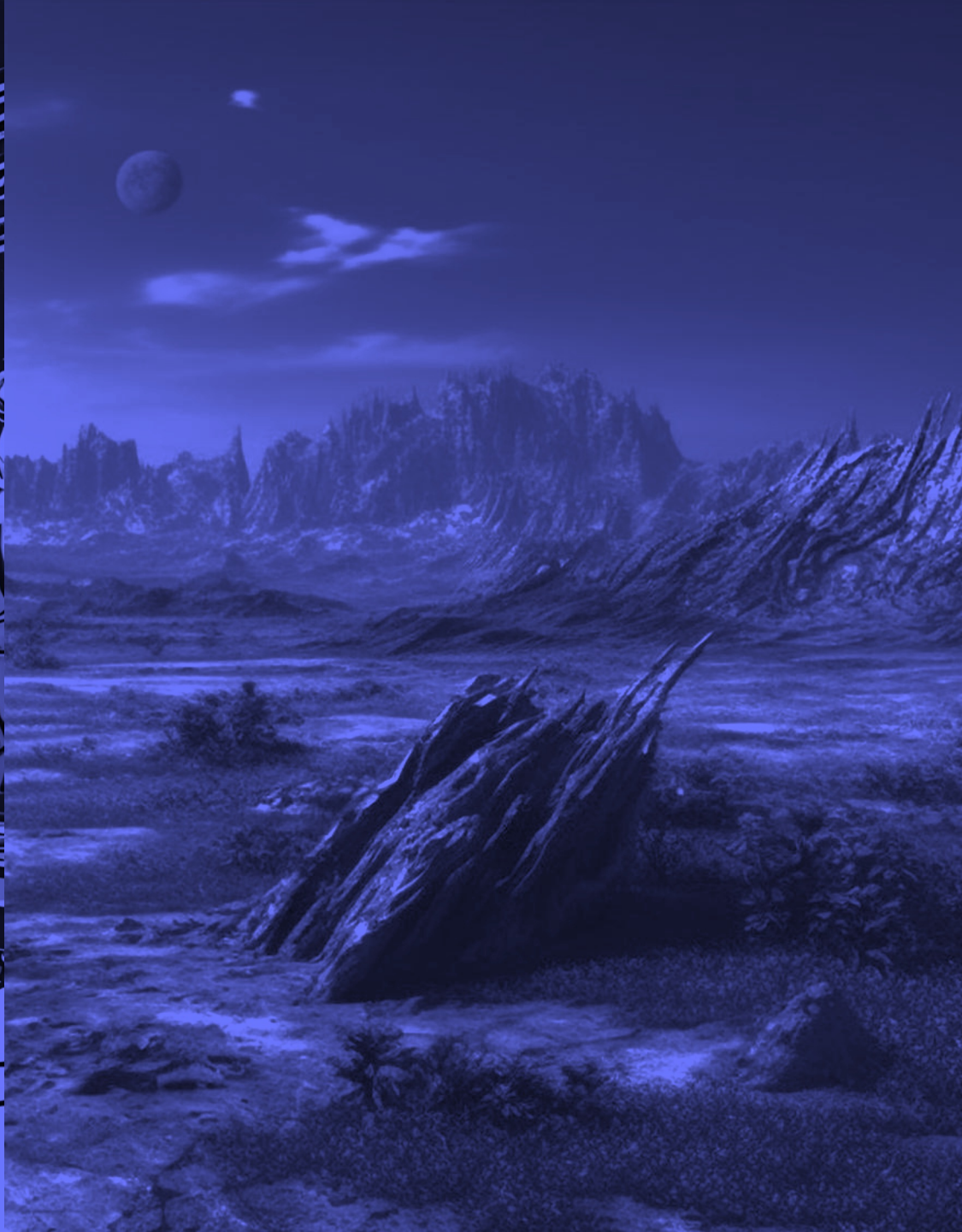




**IT'S TIME TO  
PARTY**

**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

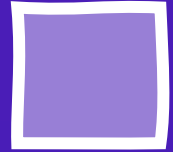
**THUNDERDOME**



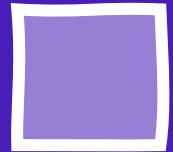


# PREFLIGHT CHECKLIST

## MAB



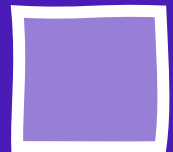
At least one test variation plus a baseline



You want to maximize improvement from a temporary experience



*like headlines, sales, limited offers*




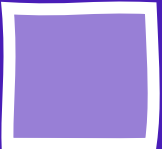
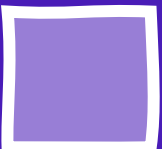
You don't need stat sig as your guide





# 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)

Slides will be available **after** Opticon

**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/)