

# Data-**Driven** or Data-**Blinded**?

The Uses and Abuses of Analytics in Games



# Data is Awesome

---



# Data is a Hot Mess





# So who am I to talk?



Kongregate.com

Cross-Platform Publisher

Coming Summer 2018



# How Did I Get **Here?**



so much data

# Don't Be **Intimidated**

---

You don't need an advanced degree in statistics to get data analysis right.



You can easily get it wrong even if you do have one.

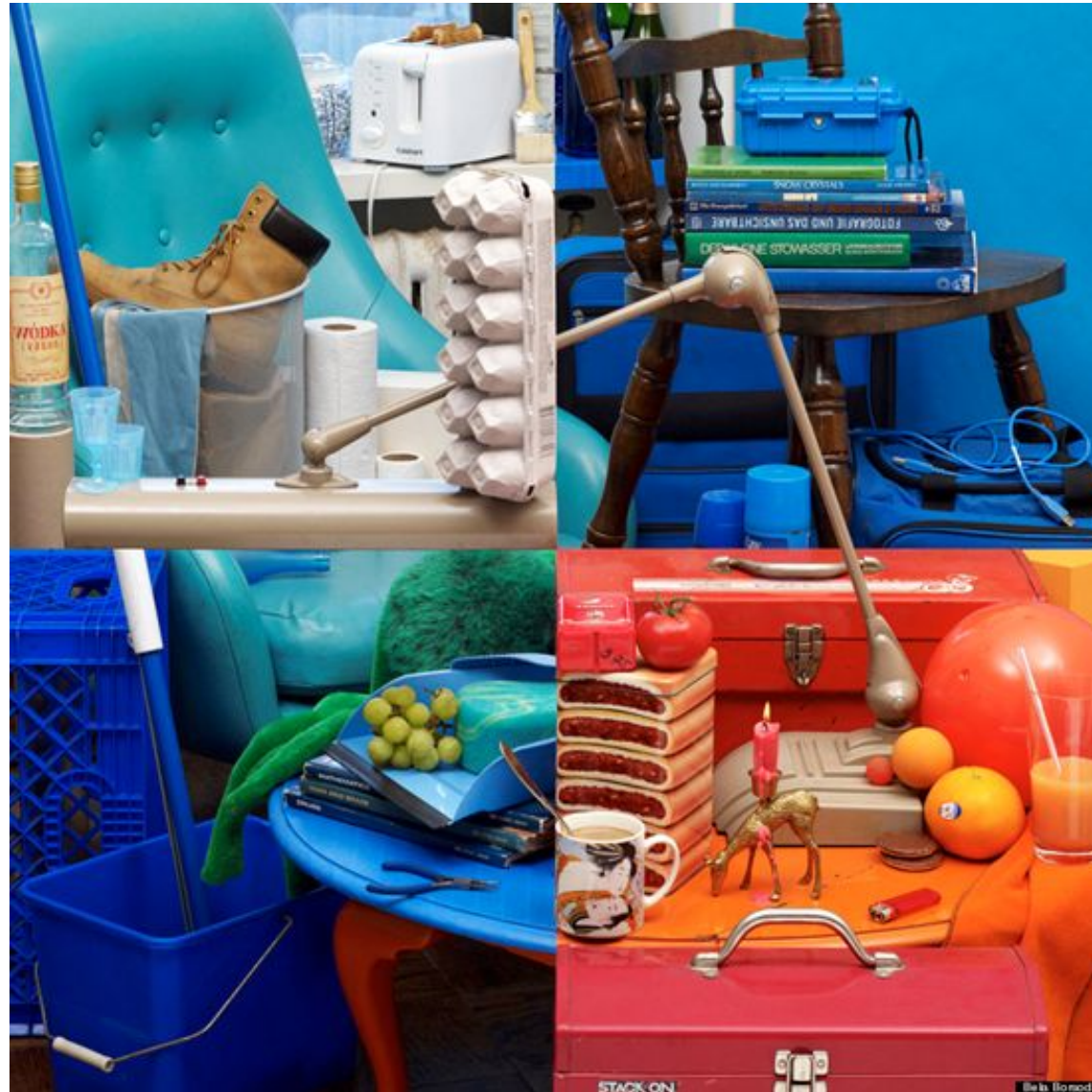
(I'm not saying it doesn't help, just that it's not essential.)

# Get Paranoid

---



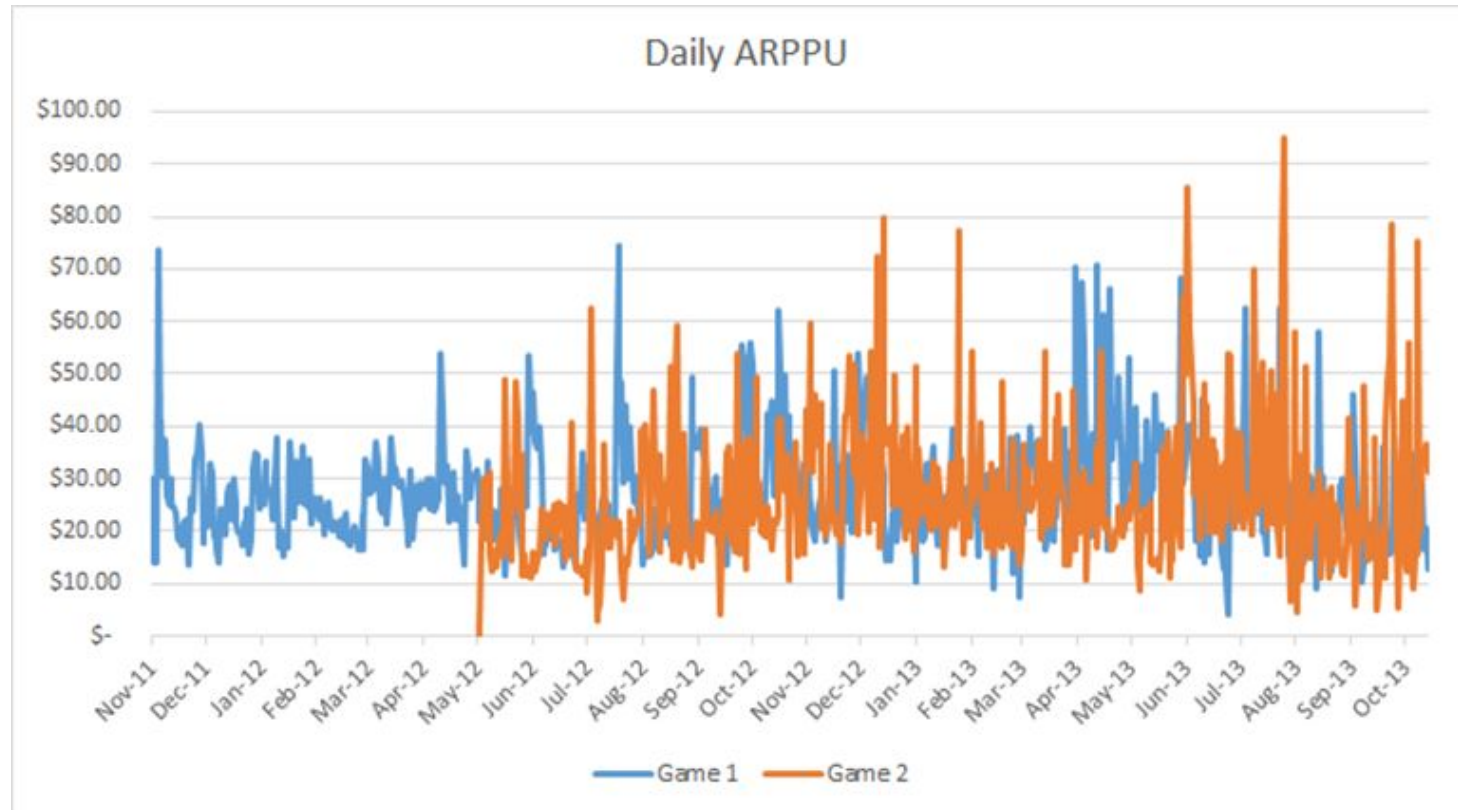
Under an apparently solid surface data is often a steaming pile of shit





Bela Borsodi

# A Tale of **Two Games**



Two games on Kongregate.com:

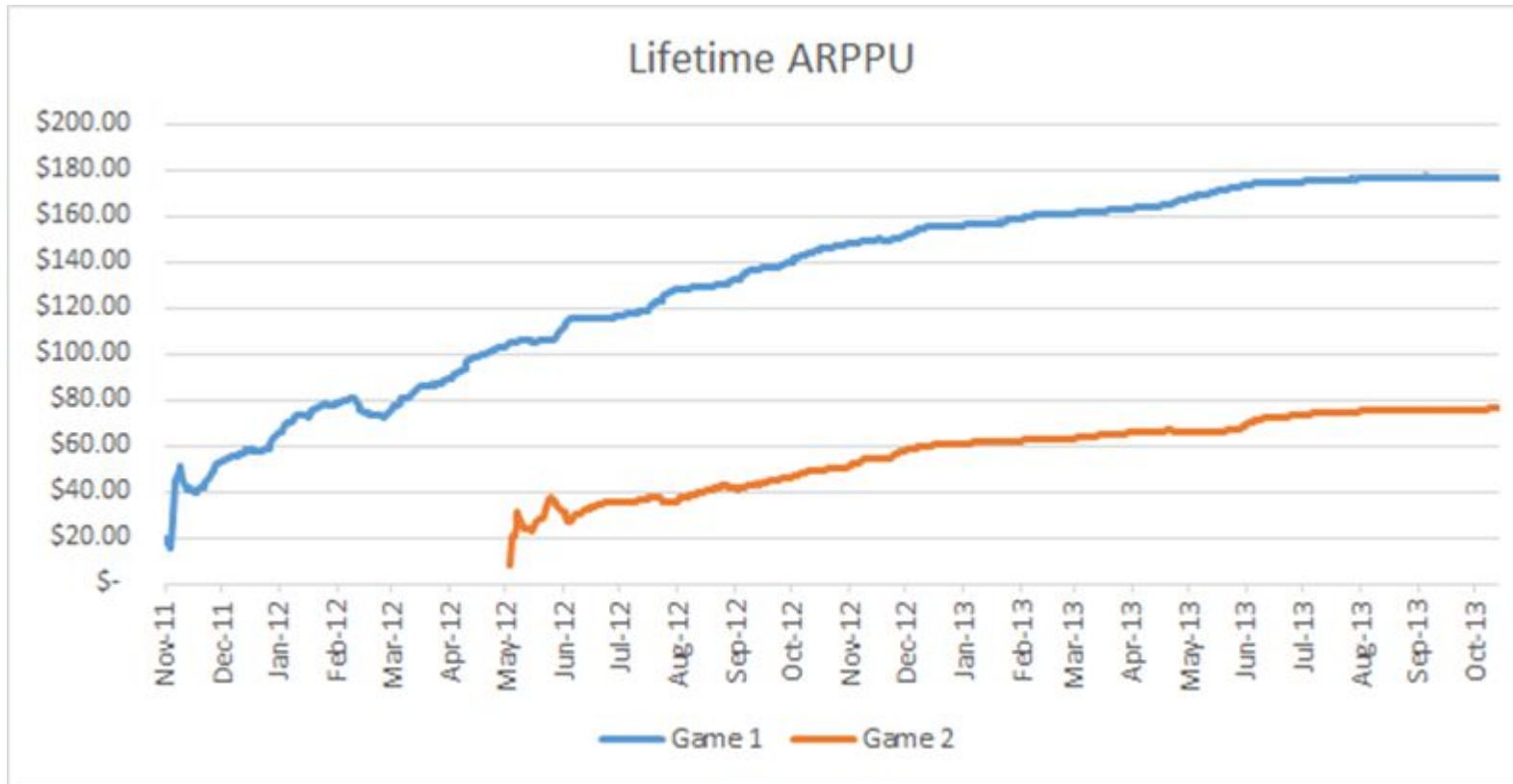
- Same genre
- Similar Day 1 and Day 7 retention (Game 1 slightly higher)
- Similar lifetime buyer % (Game 2 slightly higher)
- Similar ARPPU\* (chart)

$$\text{ARPU}^{**} = \text{Buyer \%} \times \text{ARPPU}$$

So they'll have similar ARPUs, too, right?

\*ARPPU = Average Revenue per Paying User    \*\*ARPU = Average Revenue per Paying User

# A Different Tale of **Two Games**



Game 1 ARPU = \$2.27

Game 2 ARPU = \$0.84

Game 1 has much higher:

- D30
- Transactions/Buyer
- Lifetime ARPPU (chart)

# You're a **Detective**

---



# Triangulating Truth

## Client Data



KONGREGATE



## Server Data



KONGREGATE

## Platform Data



ANDROID

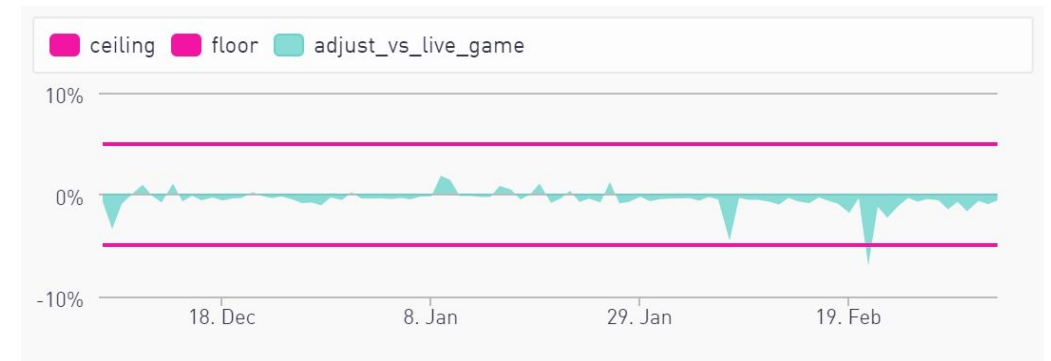


## Benchmark Data



KONGREGATE

OS TYPE	ADJUST REV	LIVE GAME REV	APPANNIE REV	ADJUST VS LIVE GAME	LIVE GAME VS AA
ios	\$113,297	\$113,132	\$124,287	0.1%	-9.0%
android	\$192,000	\$194,767	\$198,187	-1.4%	-1.7%

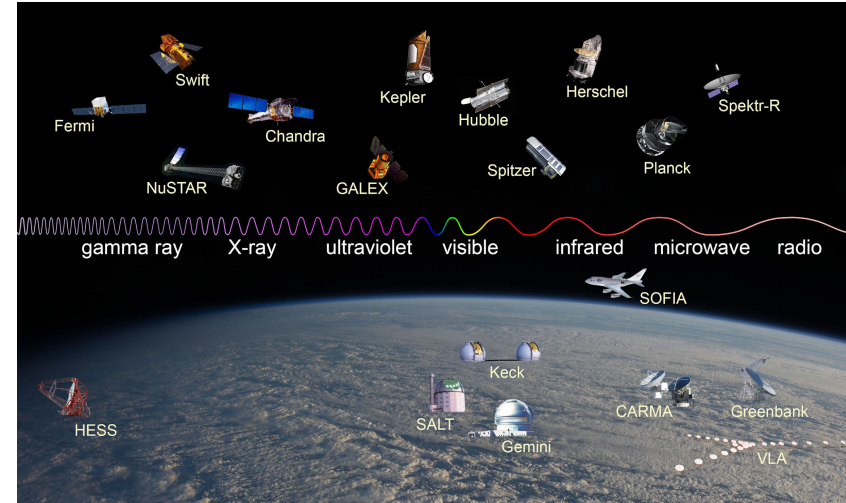


# You're also a **CSI**

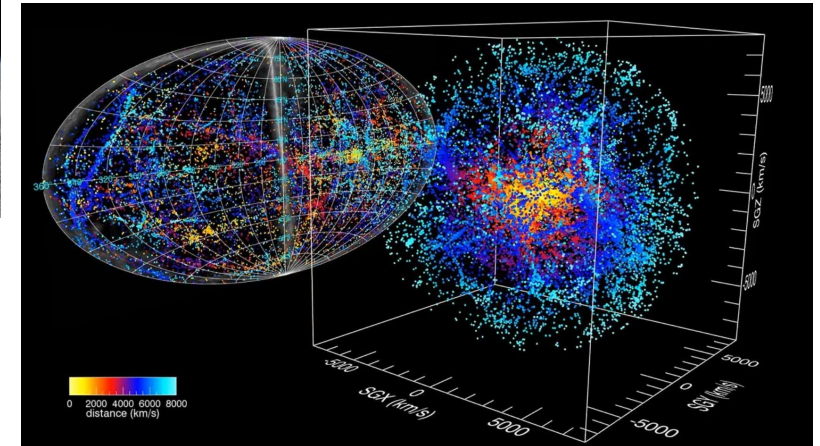


# Building Worlds

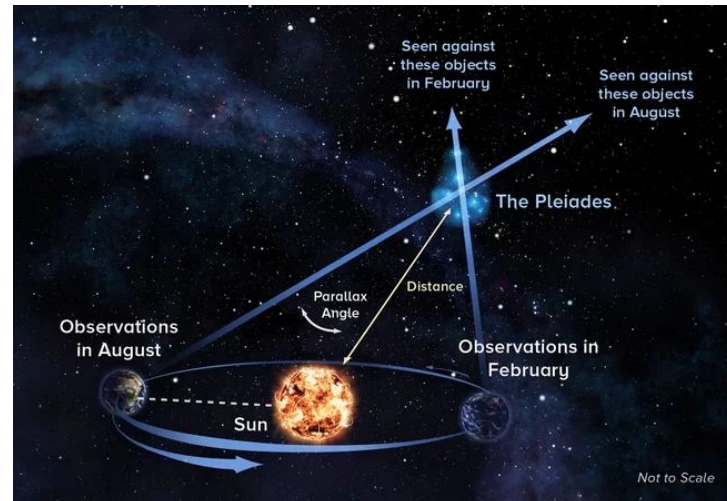
Limited view of enormous,  
changing systems



Deploy every method we  
can invent



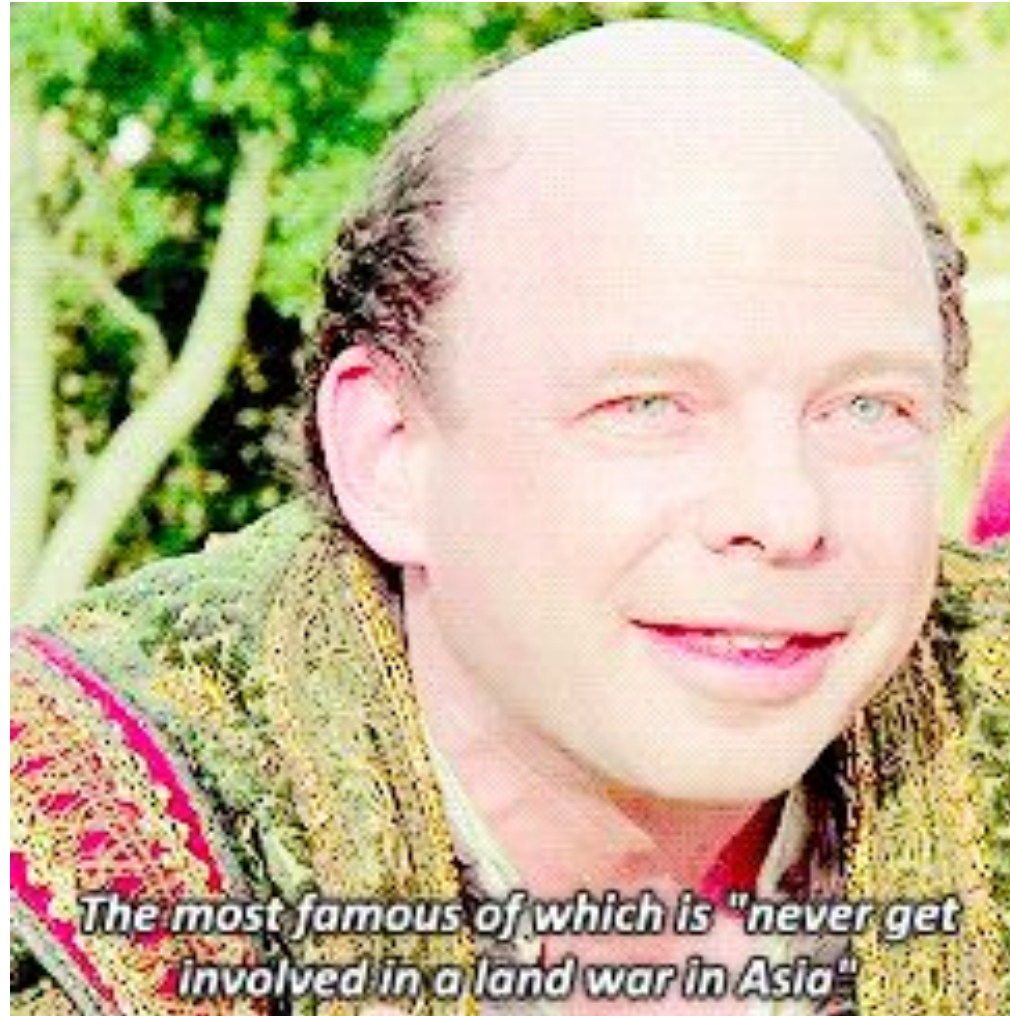
Compare observations over from  
different times and angles



Map of the Universe

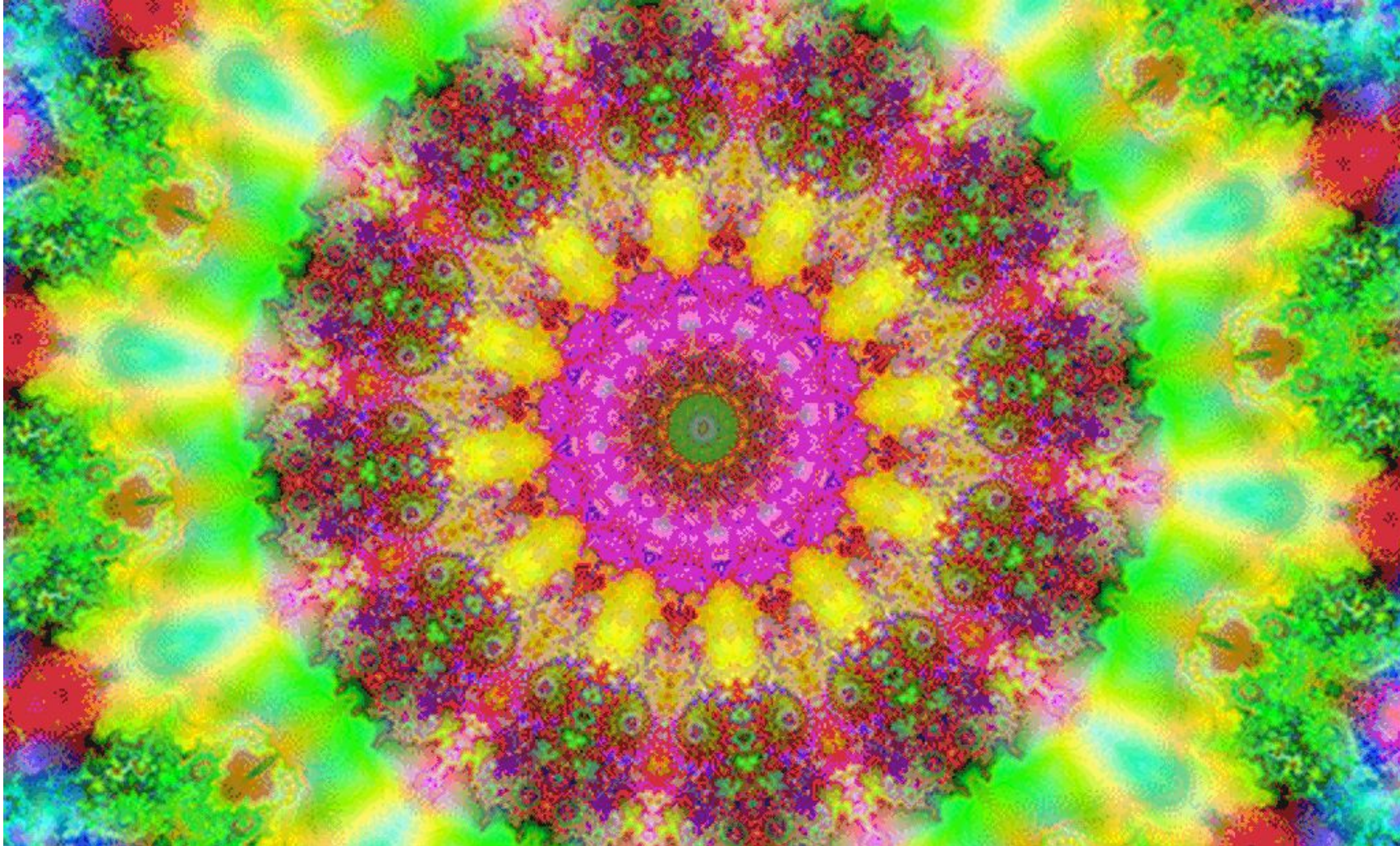
# Classic **Blunders**

---



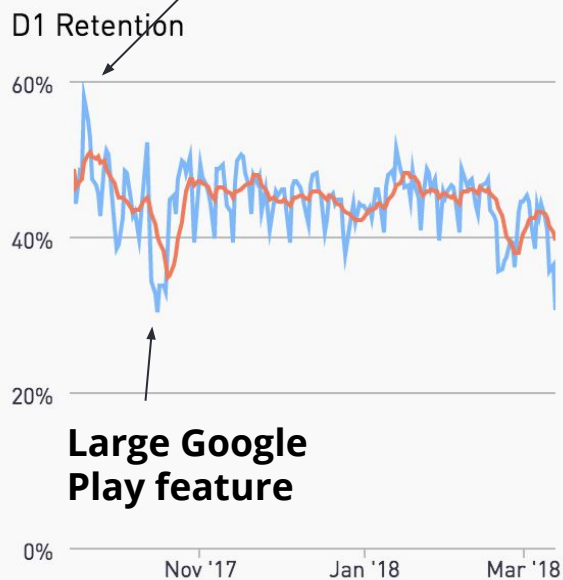
# Audience **Mix**

---



# Audience Mix

**US iOS-only UA spike**

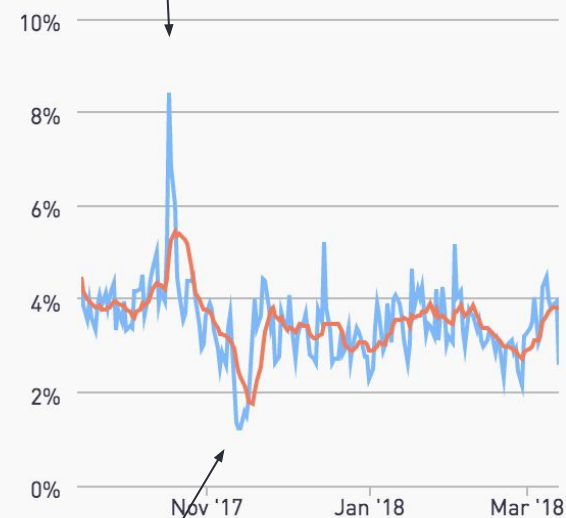


**UA spike cohort**

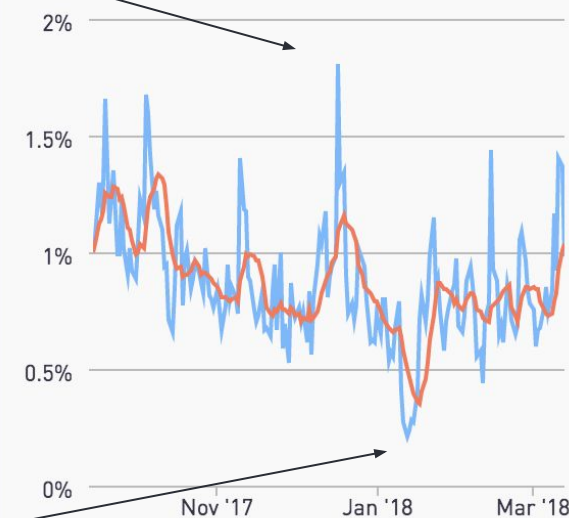
D7 Retention



D30 Retention

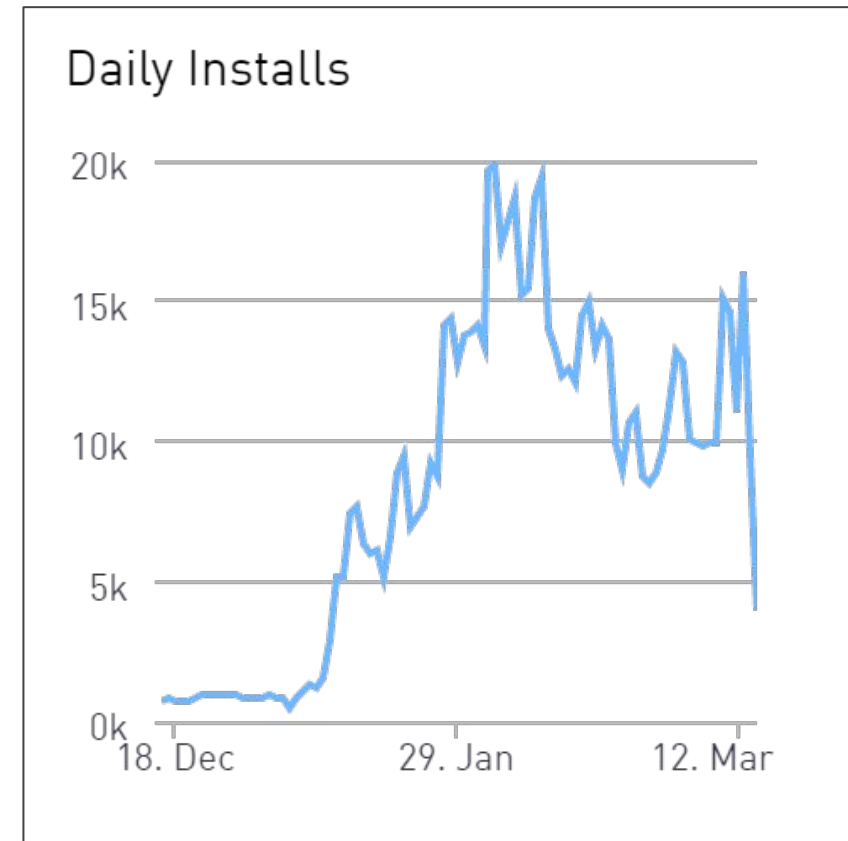
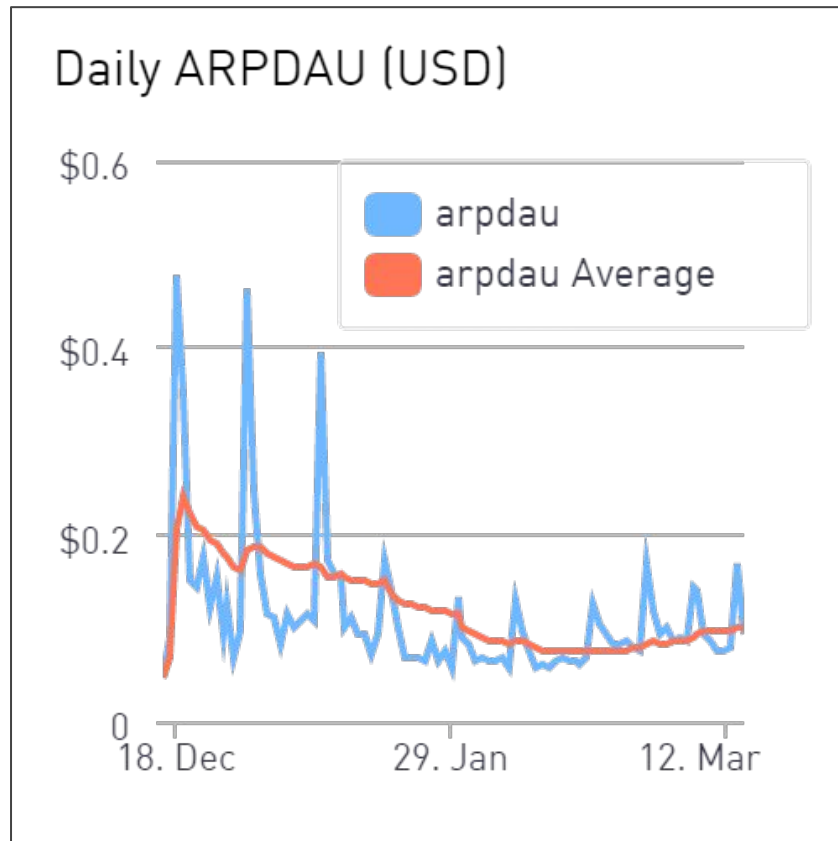


D90 Retention

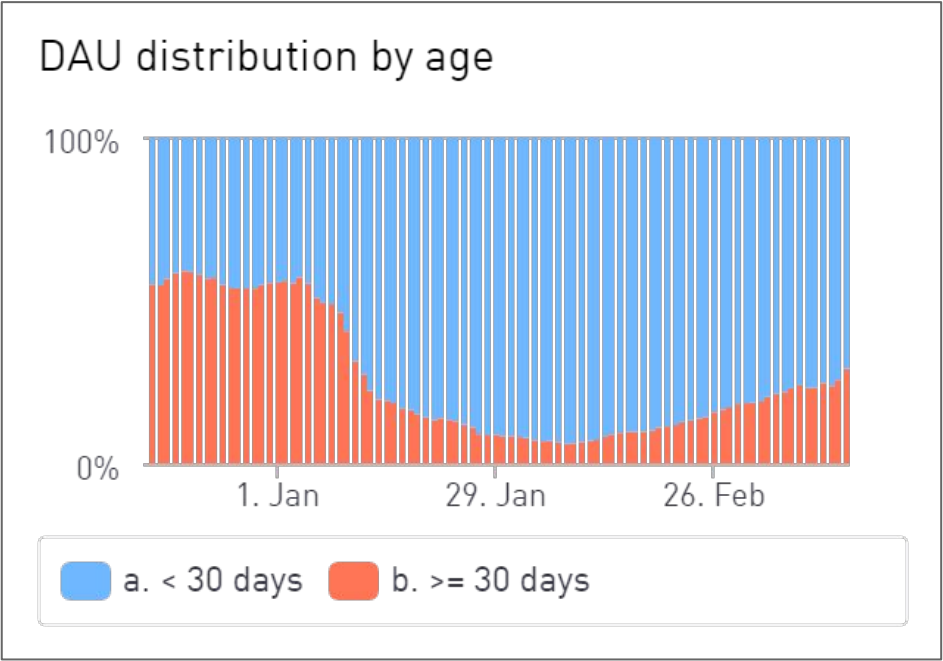
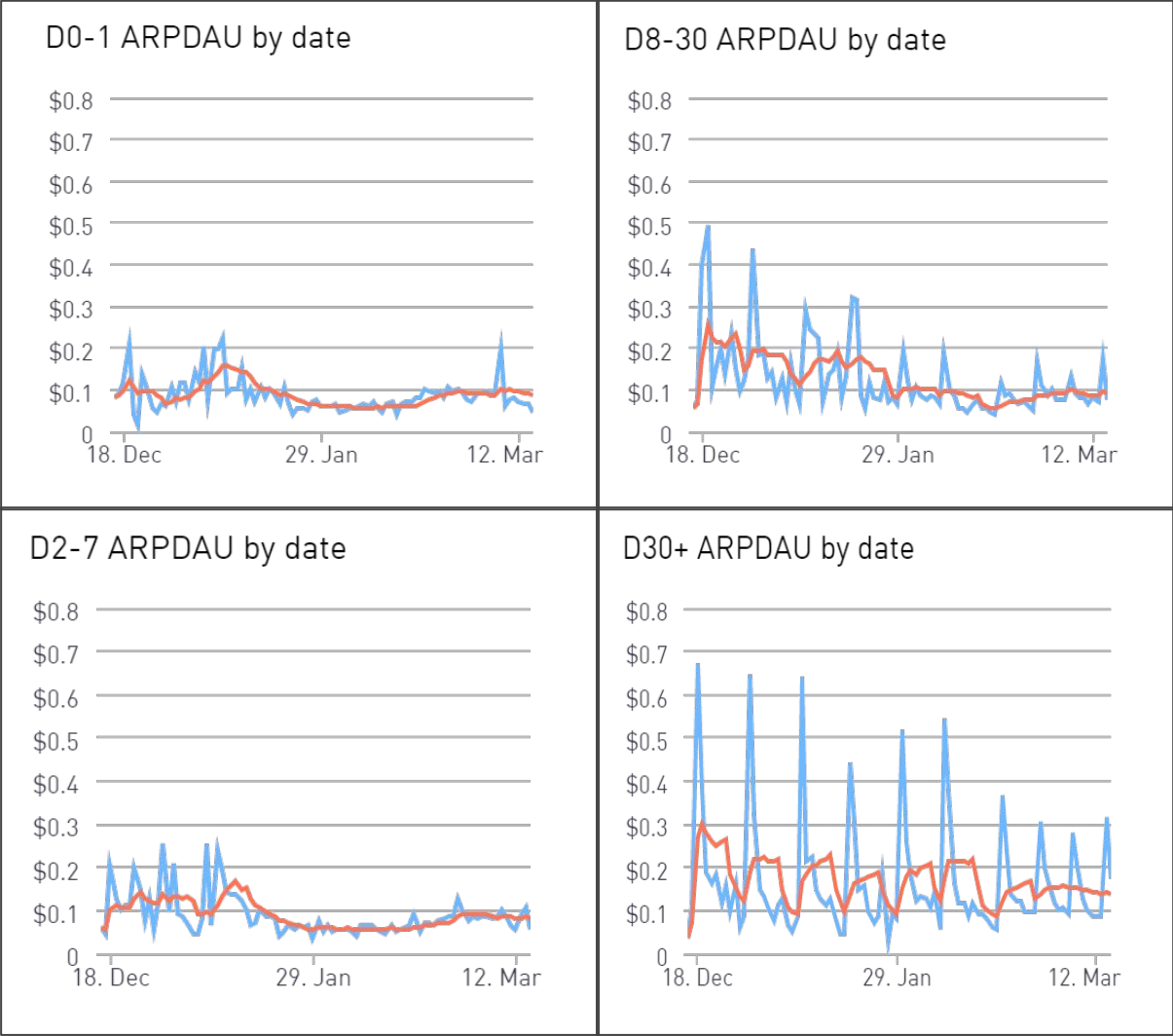


**Feature cohort**

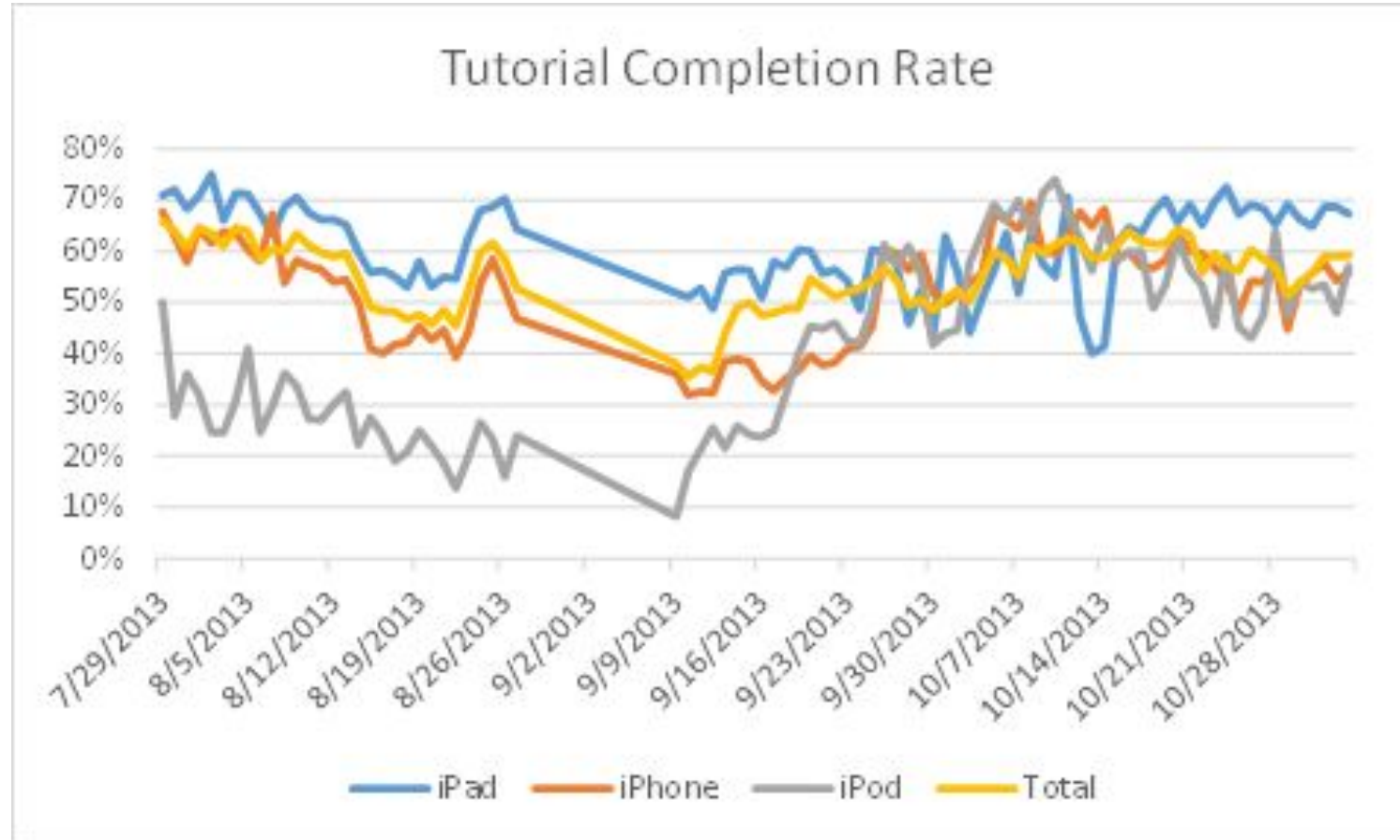
# Audience **Age**



# Audience Age



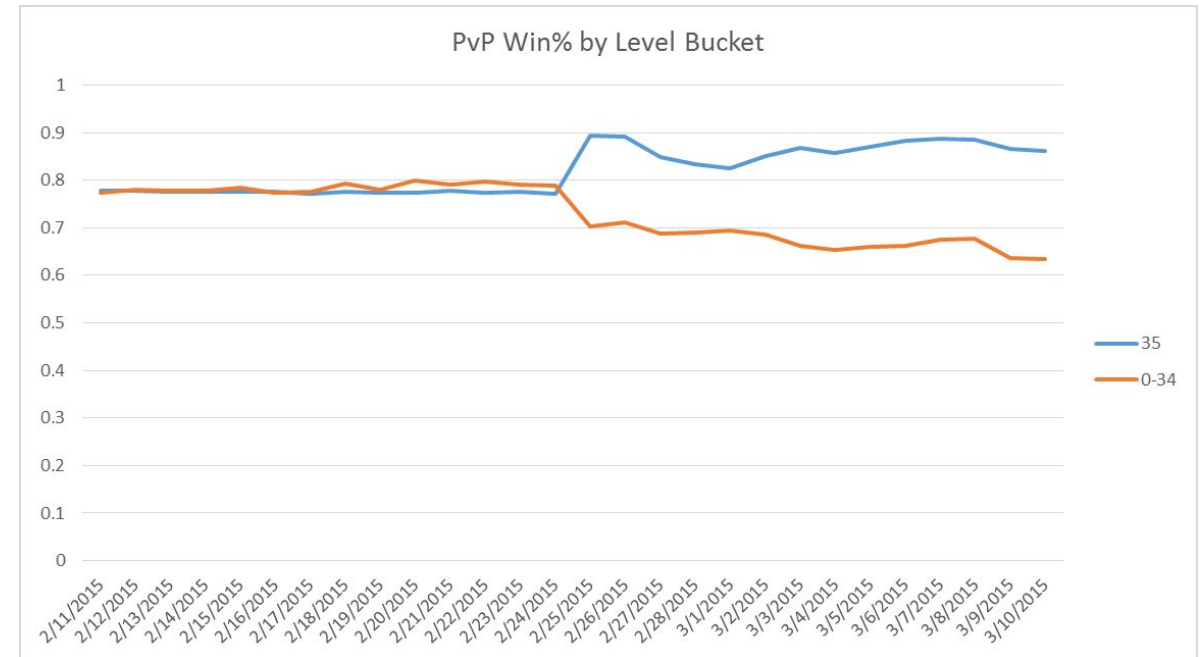
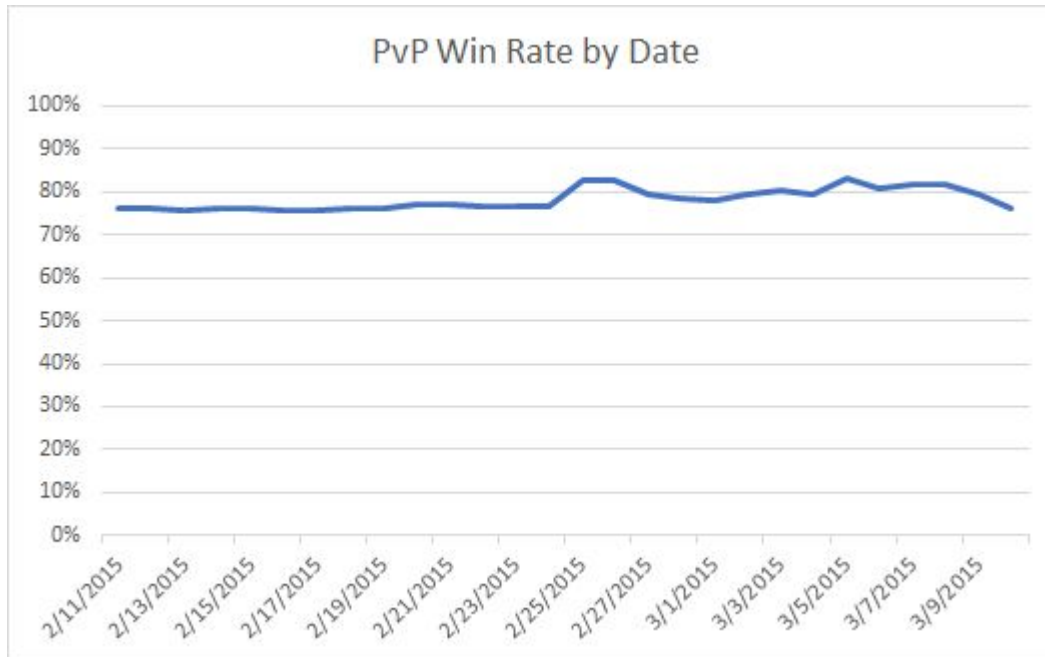
# Averages are **Average**



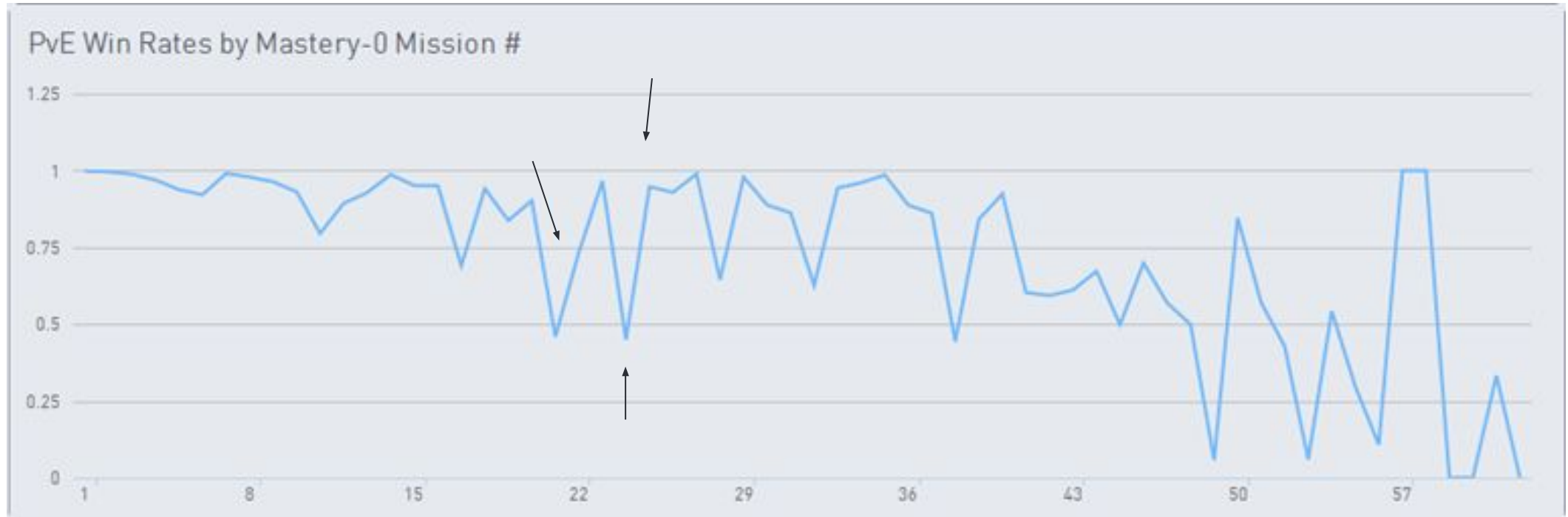
Numbers from a dominant subgroup can mask problems (or successes) in smaller subgroups.

# Players are **Not** Average

Changes can have drastically different impacts on new and late-game players.

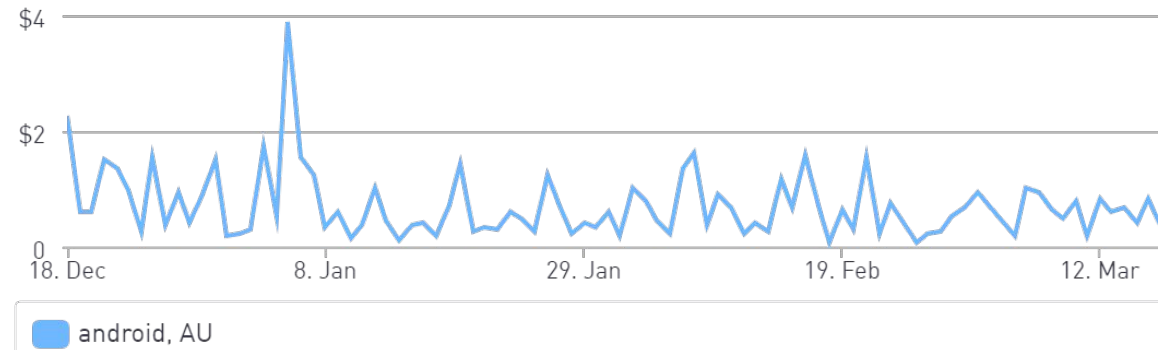


# Funnel Subgroups

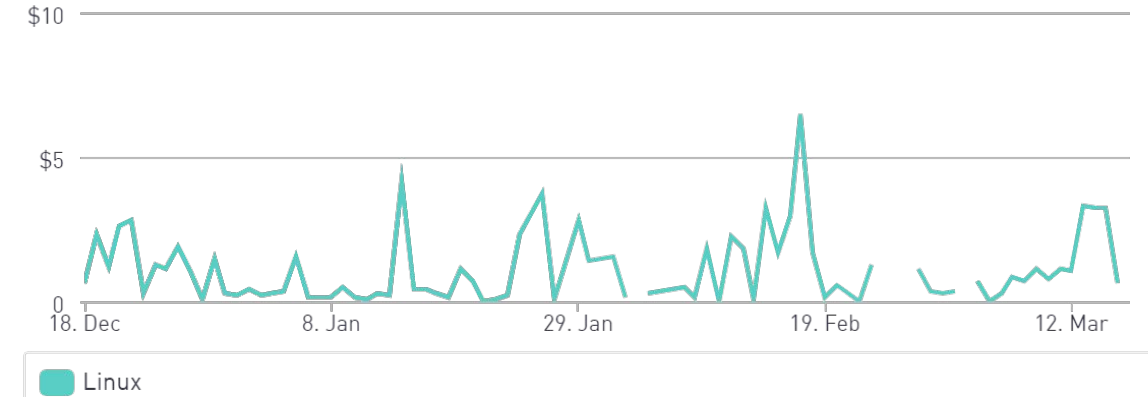


# Small Sample Sizes

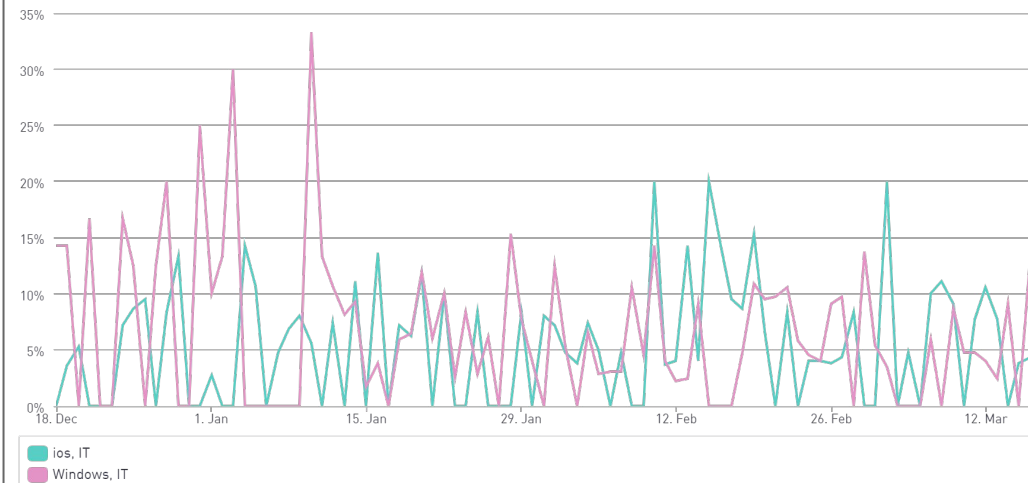
Daily ARPDAU



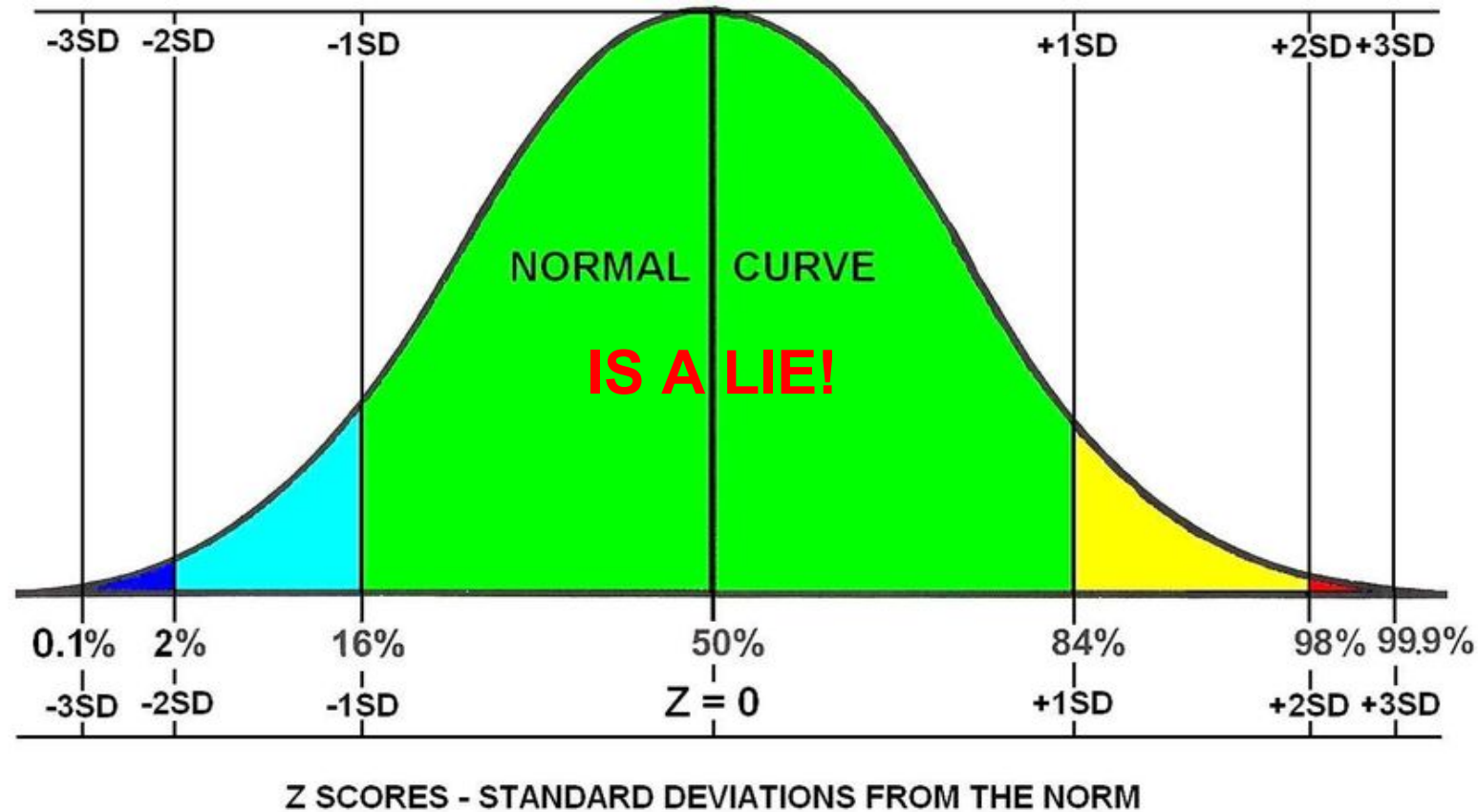
Daily ARPDAU by Platform



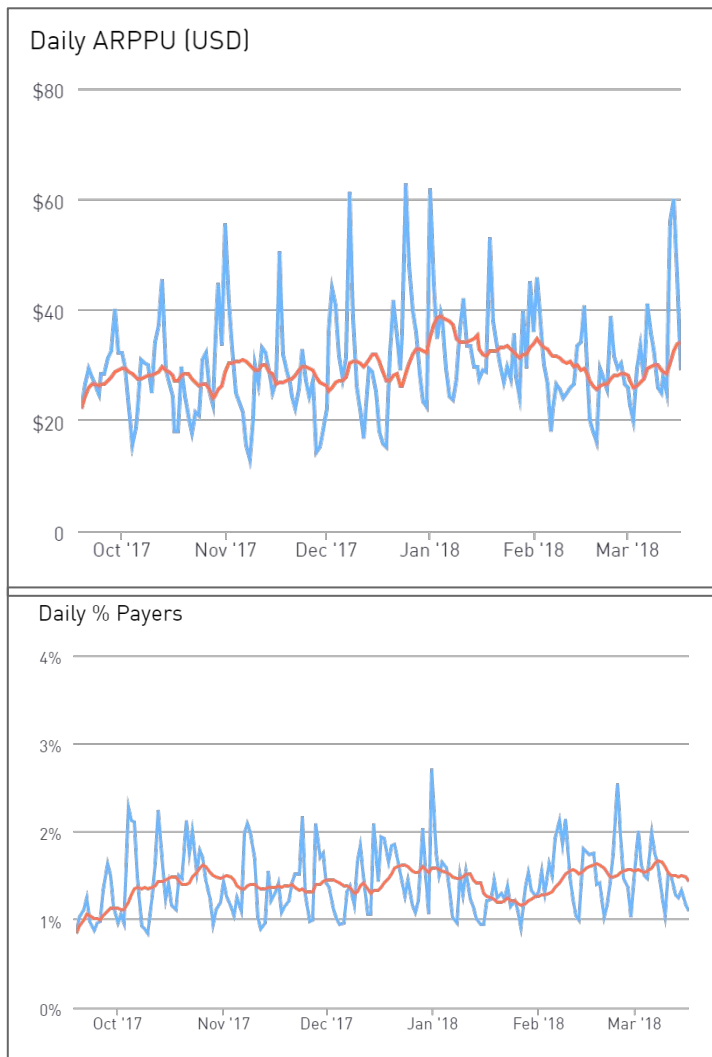
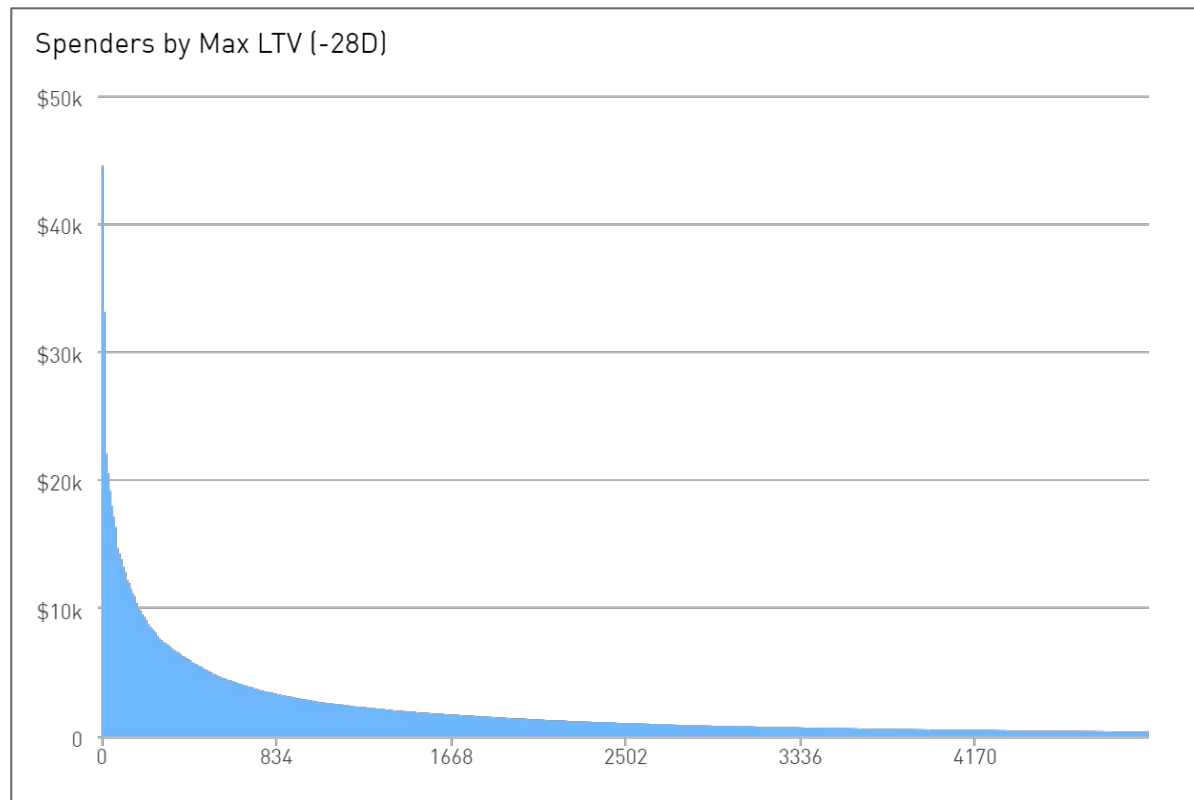
D30 Retention



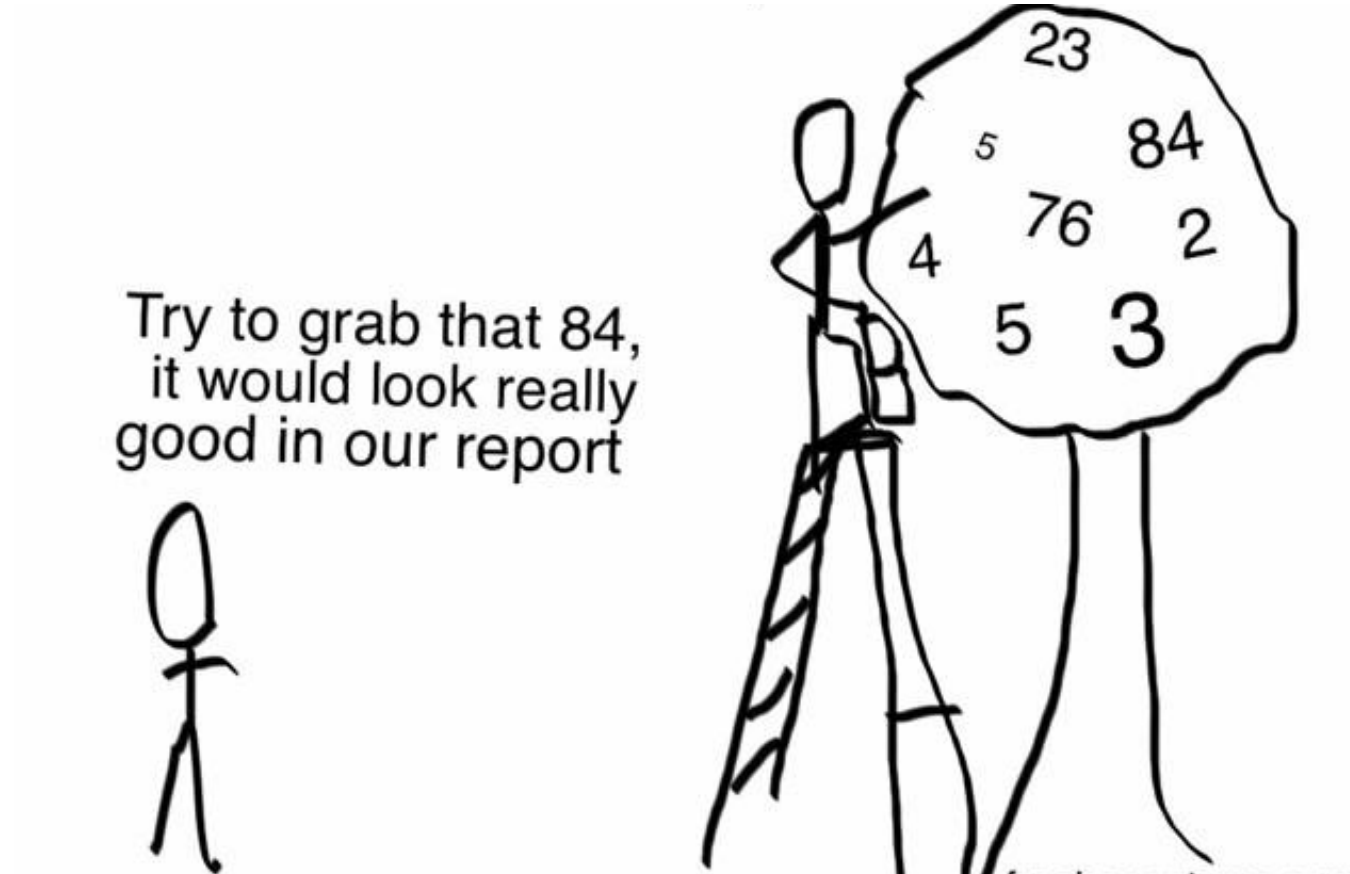
# Nothing is **Normal**



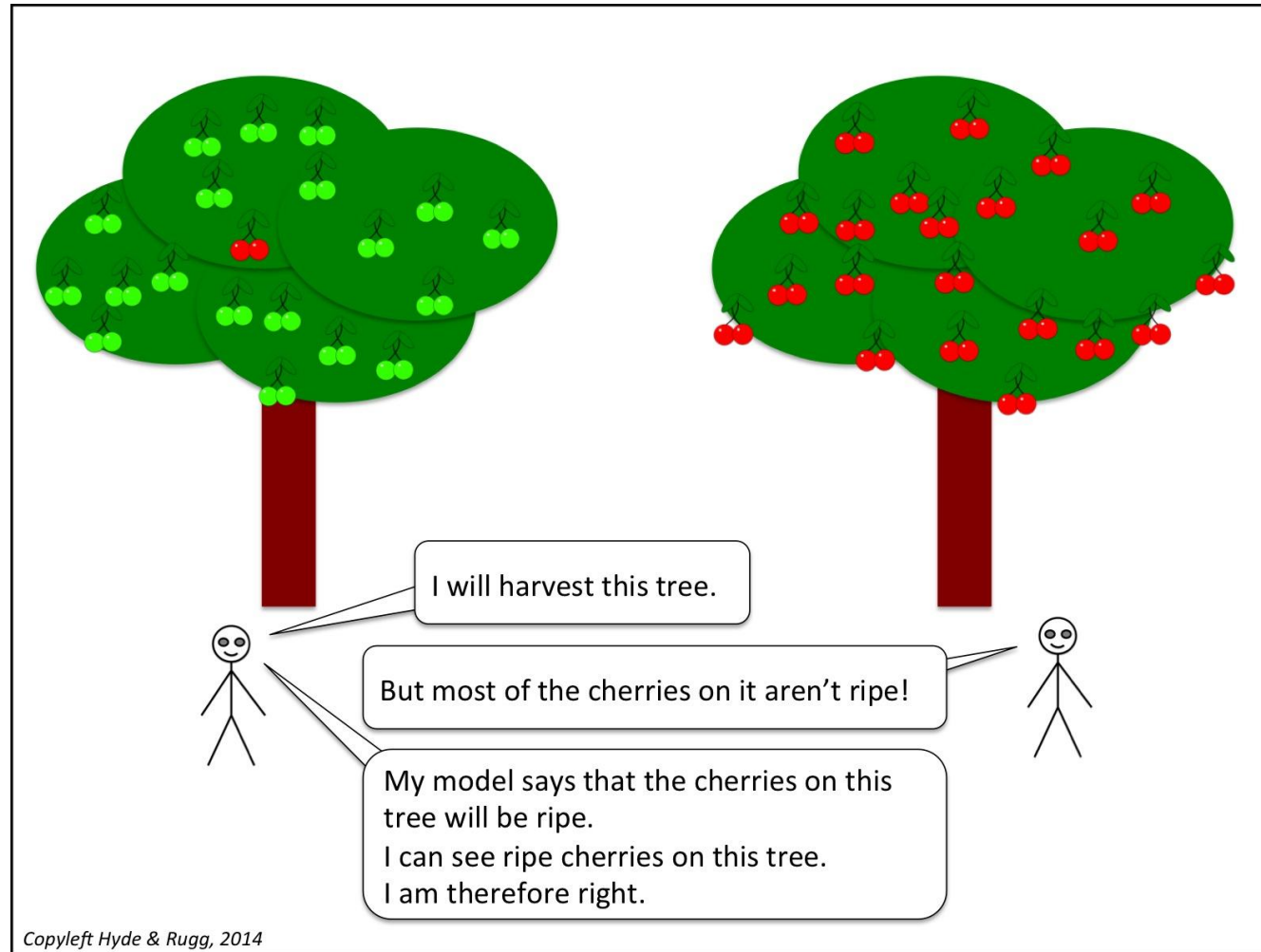
# Outliers Rule



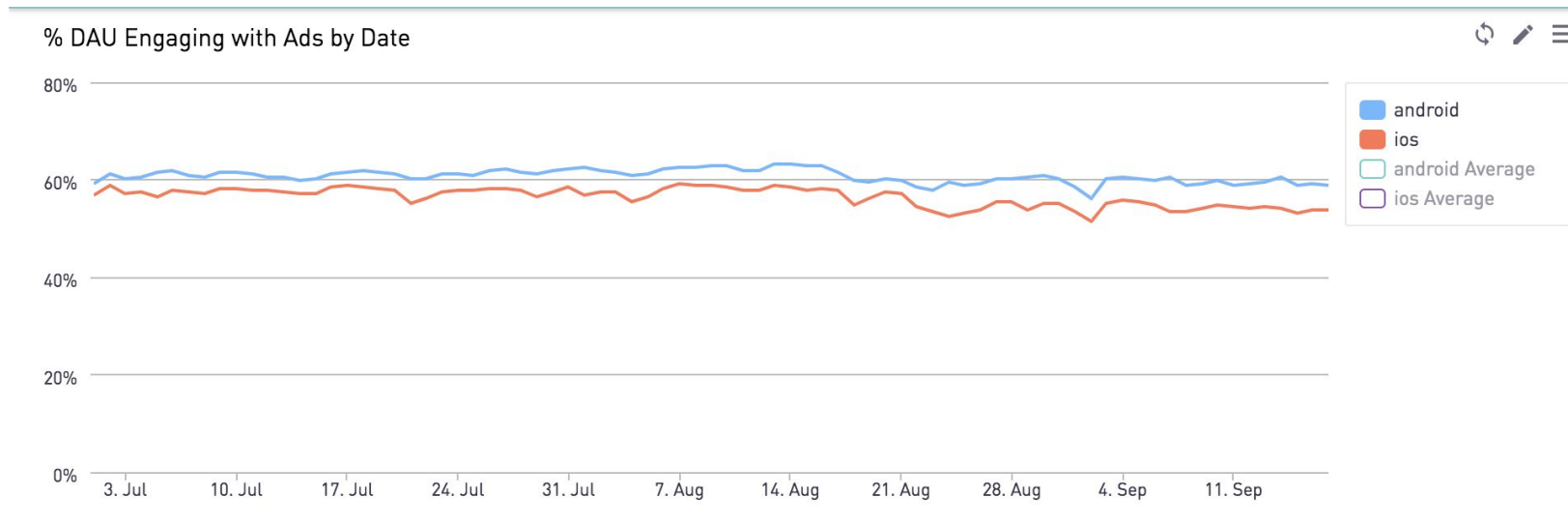
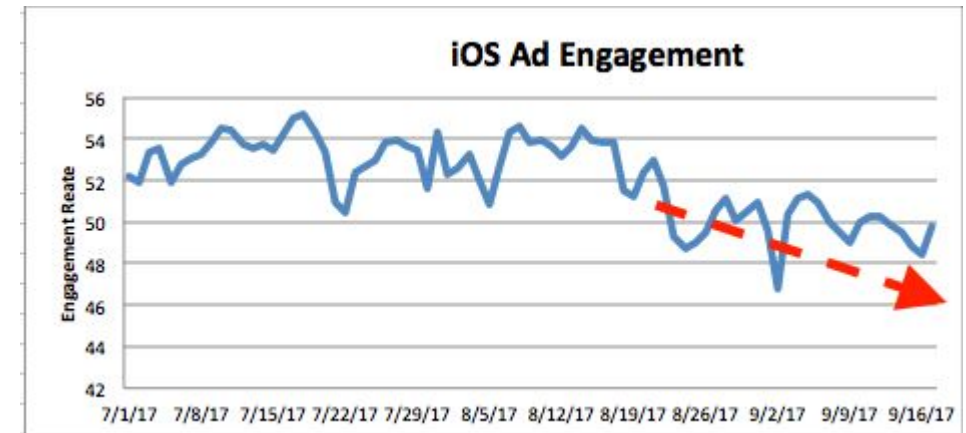
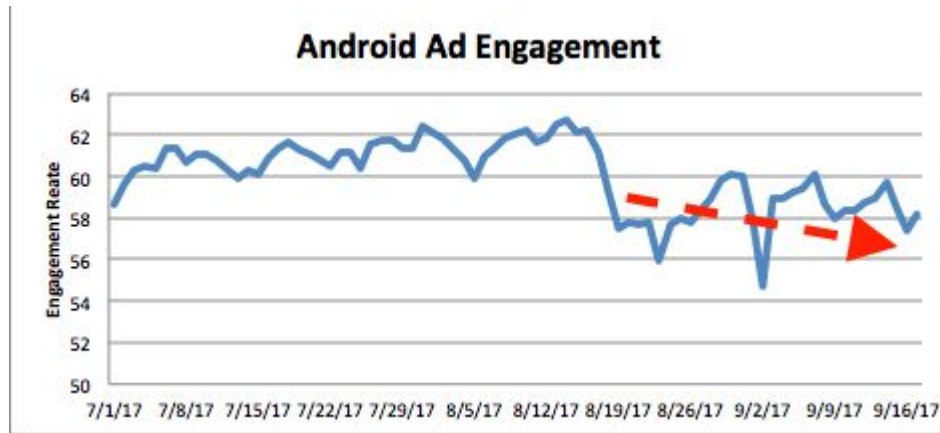
# Cherry Picking



# Confirmation Bias



# Axis of Evil



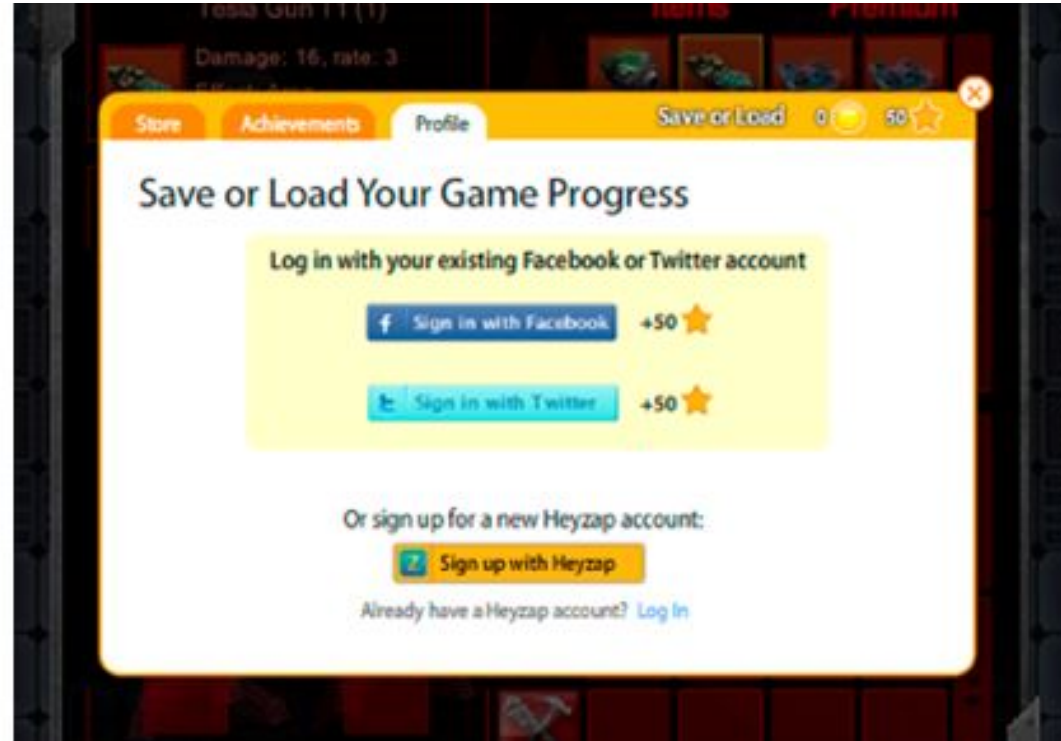
# Correlation **!=** Causation

---



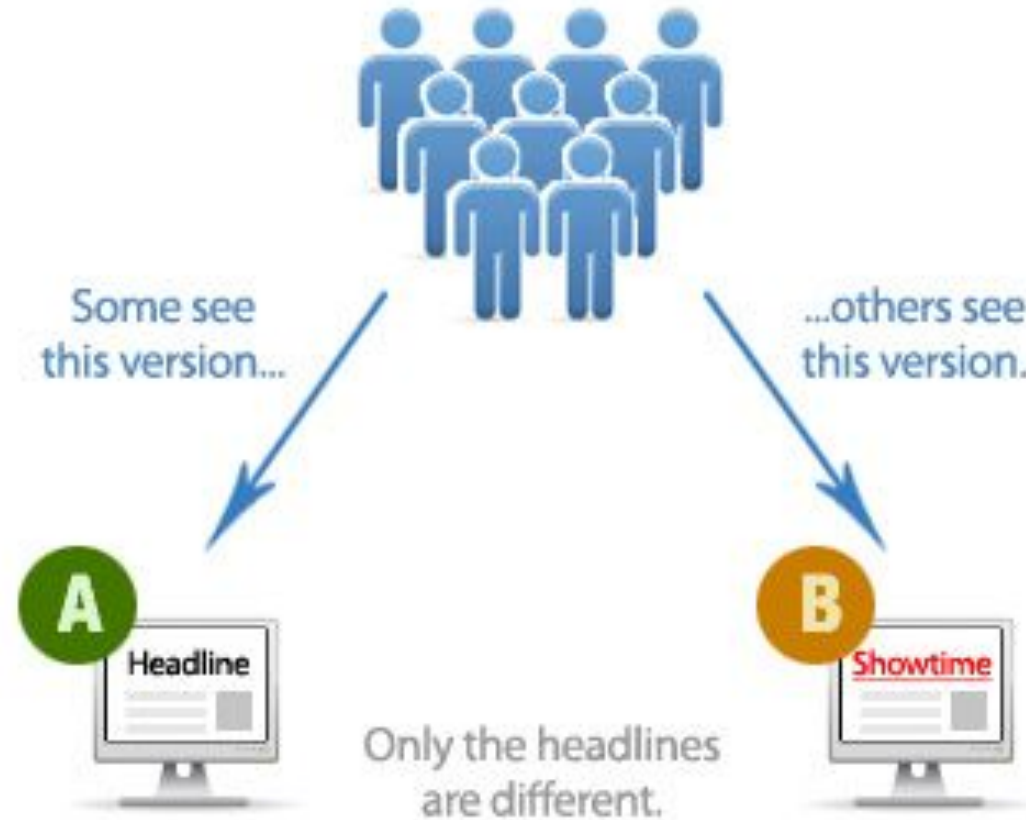
# Everything Correlates with **Engagement**

*“Players who connect to Facebook are 2x as engaged, let’s incentivize Facebook connections!”*



# A/B Tests

---



# A/B Test Traps

---

## **Assignment Issues**

**Sample Sizes**

**Audience Mix**



**Cherry Picking**

**Power Distributions**

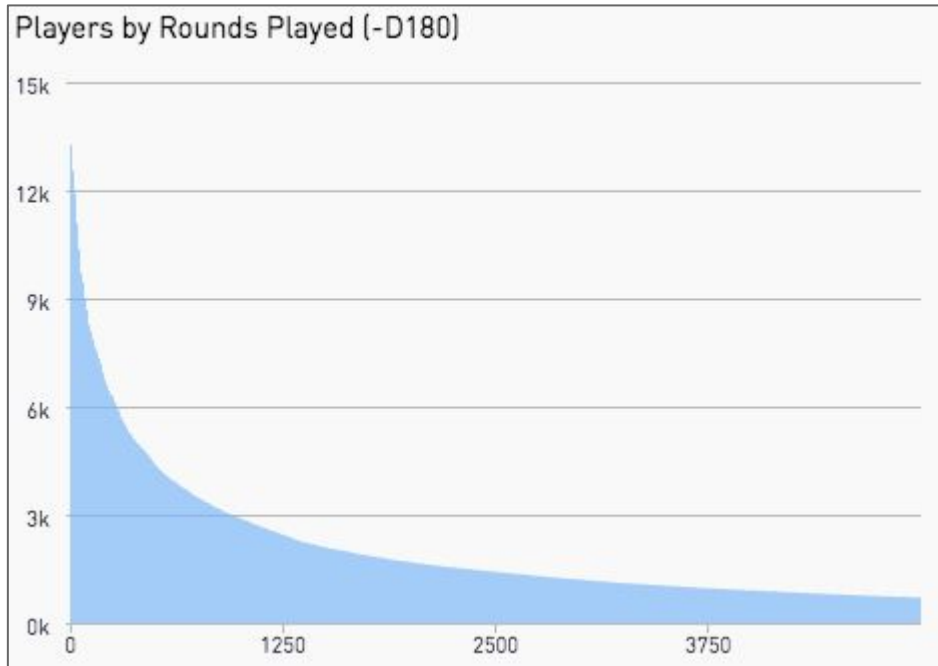
**Confirmation Bias**

# (Miss) Assignment

Tip: Make sure to split on what you're actually testing



# (Huge Miss) **Assignment**



Your best players ALWAYS show up first.  
Any test that doesn't take that into account  
is inherently flawed.

Even if you're testing to new players only  
the people who start playing a game on  
Saturday are not the same people who start  
on Tuesday.

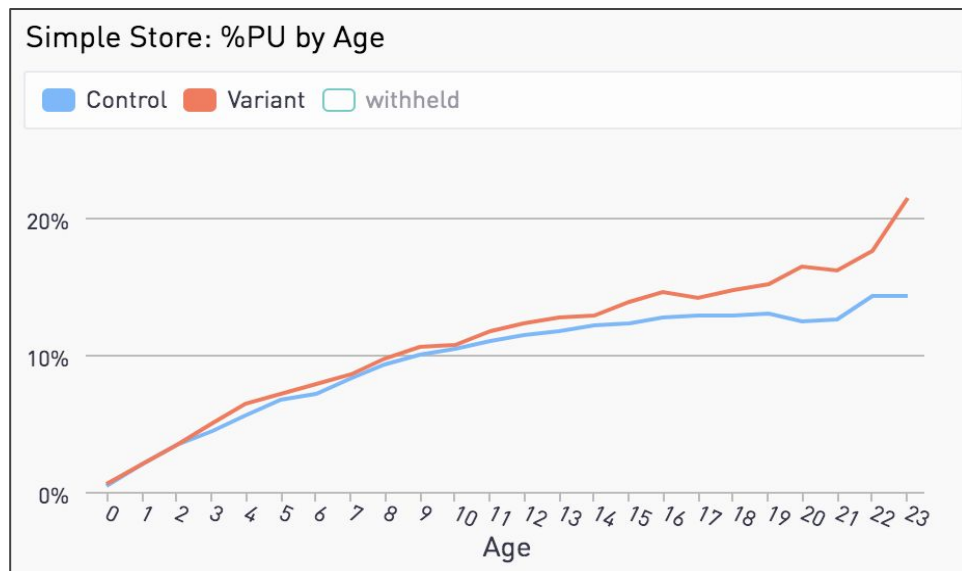
# Downstream Impacts

---

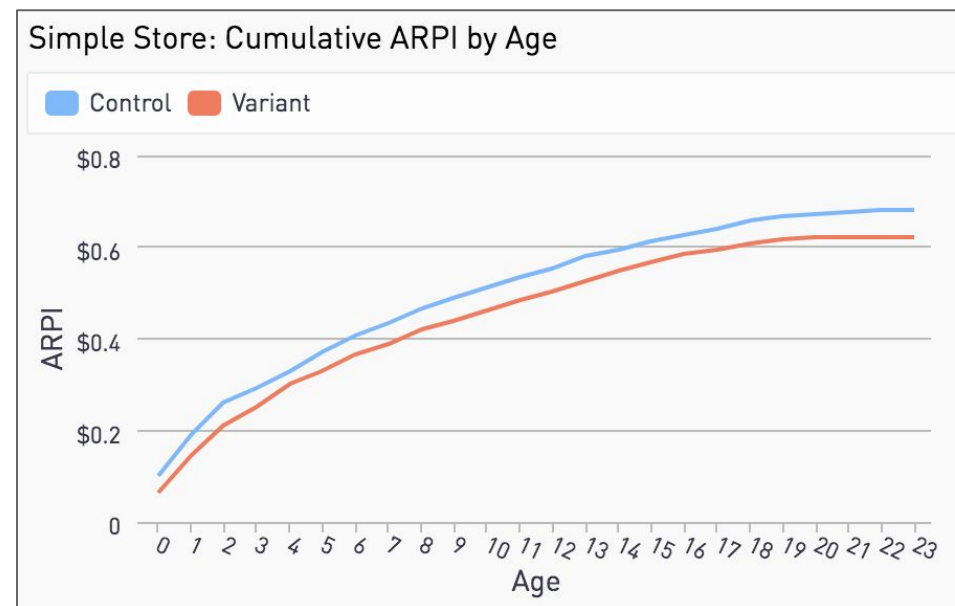


# Downstream Example

Simple New Buyer Store: Conversion Rate D0-D7 by Variation			
VARIATION	TOTAL INSTALLS	TOTAL CONVERTS	CONVERSION
C	38039	569.0	1.49%
V1	37545	610.0	1.62%



Simpler store increased conversion by **9%**!



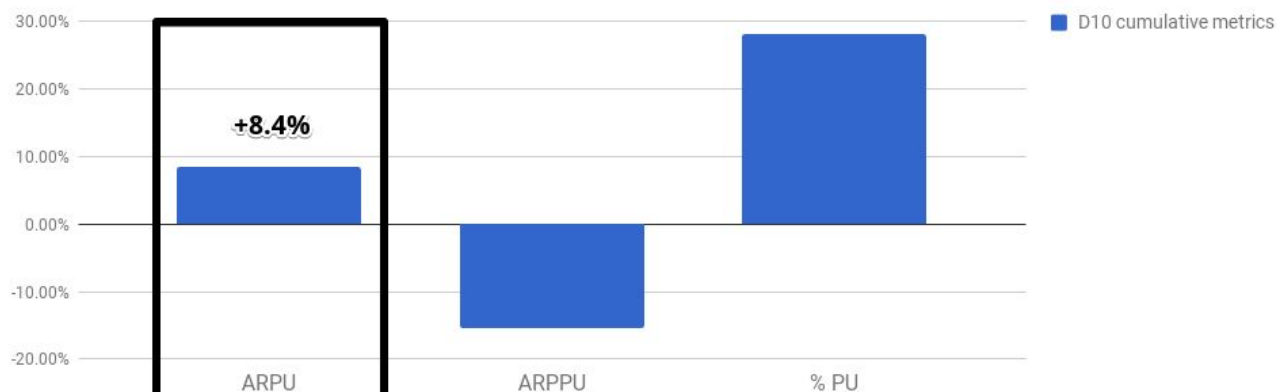
But drastically reduced ARPPU meant revenue actually decreased by **11%**. Oops.

# Premature Analysis

## Pop-Up Flash Item Sales A/B Test



Change in monetization KPIs: Control versus Variant



Change in monetization KPIs: Control versus Variant

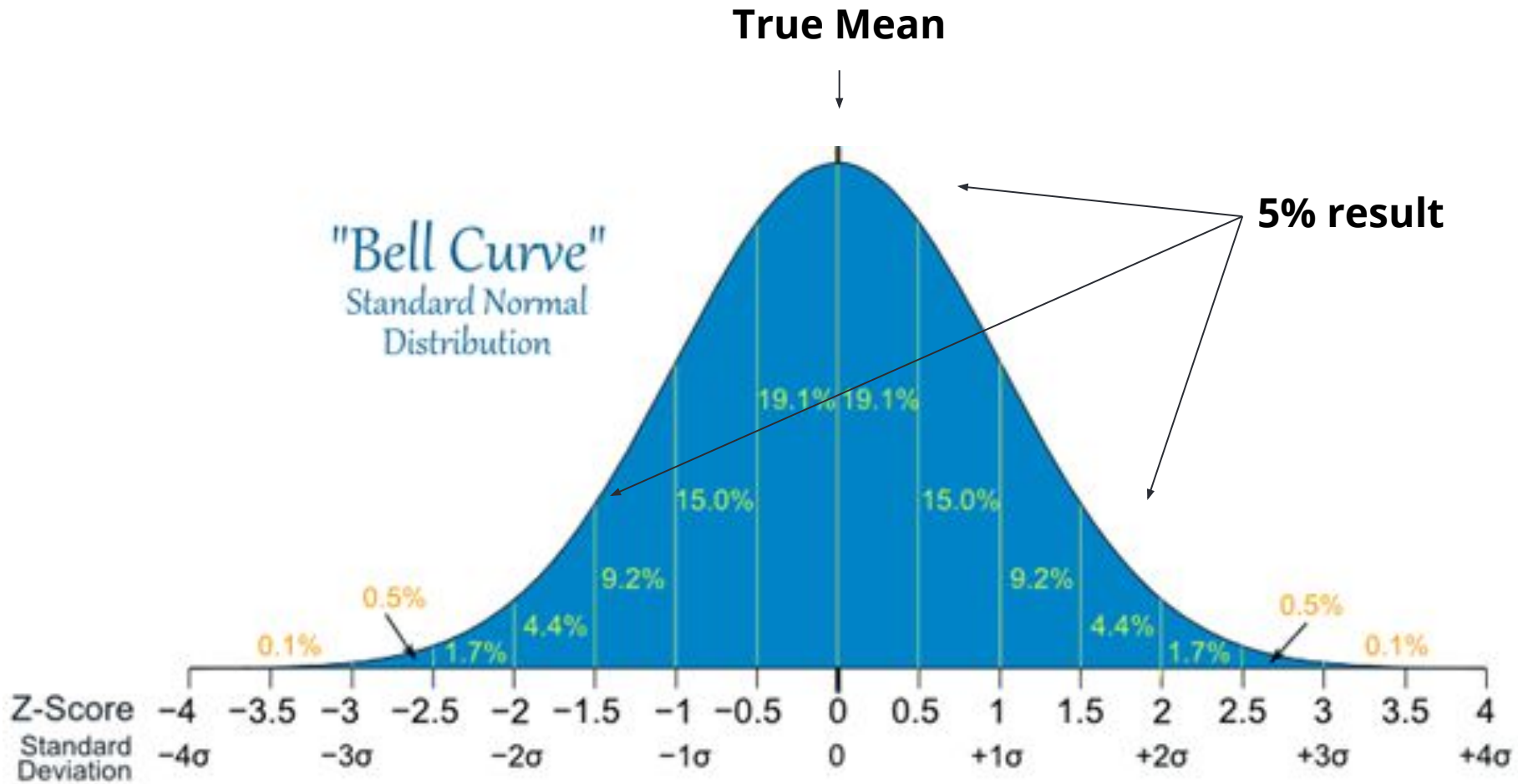


# Statistical **Significance**

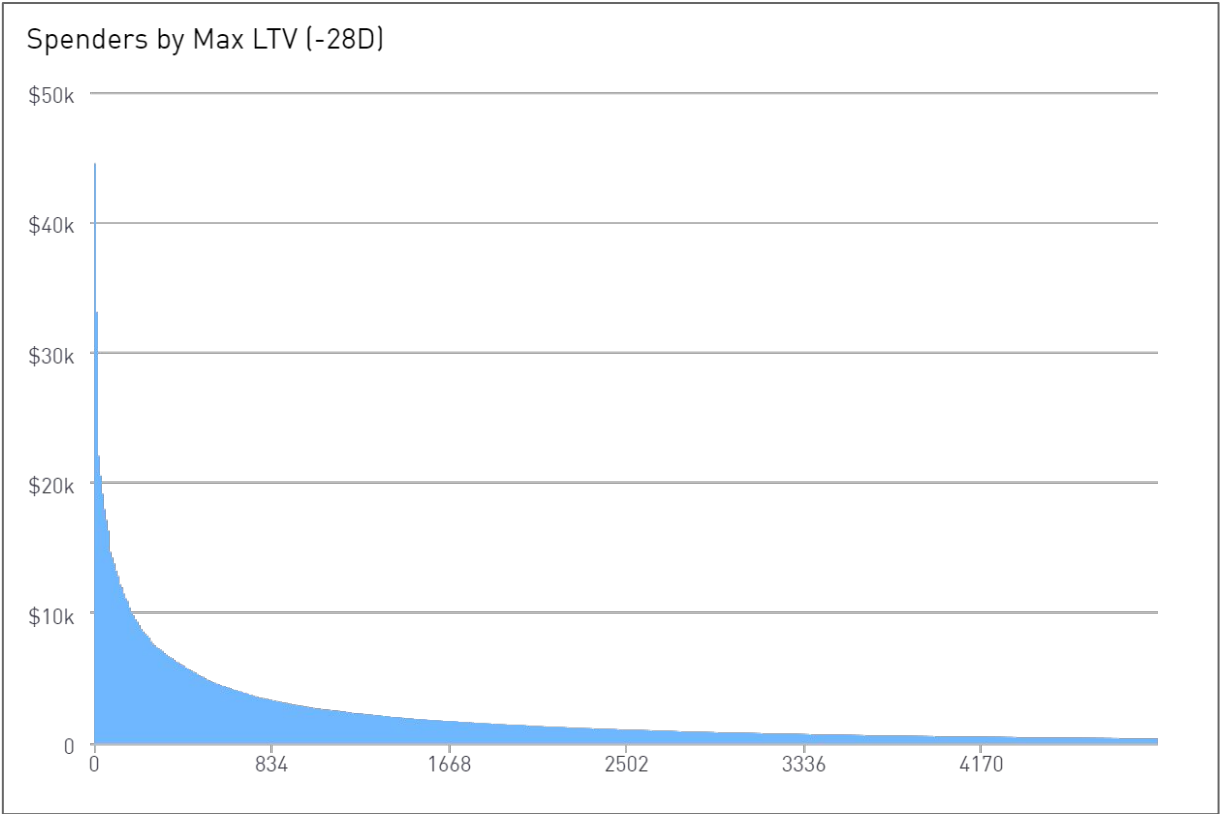
---



# Statistical Significance



# Abnormal Testing

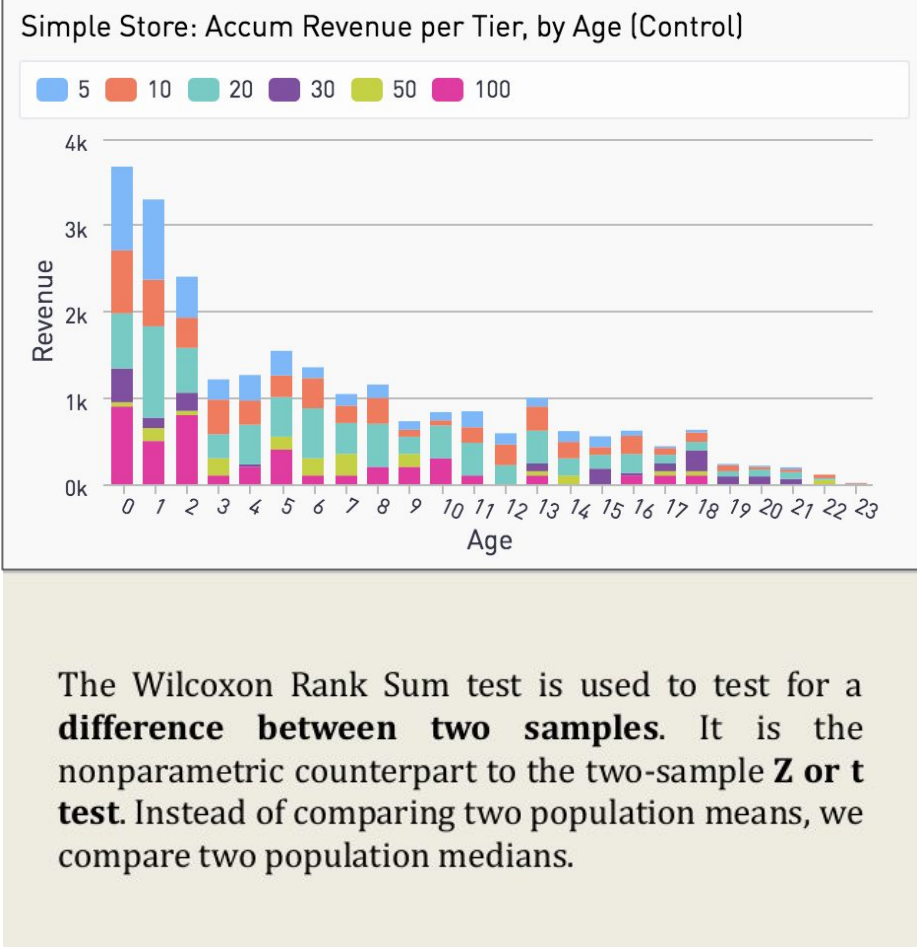


A	B	
\$1,000	\$500	
\$100	\$100	
\$40	\$40	
\$20	\$20	
\$10	\$10	
\$5	\$5	
\$5	\$5	
\$5	\$5	
Avg \$148	Avg \$86	-42%

# Medians and Distribution FTW

A	B
\$1,000	\$500
\$100	\$100
\$40	\$40
\$20	\$20
\$10	\$10
\$5	\$5
\$5	\$5
\$5	\$5
Avg \$148	Avg \$86
Median \$15	Median \$15

A	B
\$1,000	\$825
\$100	\$200
\$40	\$80
\$20	\$40
\$10	\$20
\$5	\$10
\$5	\$10
\$5	\$10
Avg \$148	Avg \$149
Median \$15	Median \$30



# Testing's **Dirty Secret**

---

Most tests don't show a significant change.

# Extreme Description Testing

"...a design masterpiece." 5/5 - TouchArcade  
2016 Game of the Year - TouchArcade, Gamezebo  
2016 Action Game of the Year - Pocket Tactics  
DICE 2016 Mobile Game of the Year Nominee  
Time's Top 10 Games / Top 50 Apps of 2016

Craft, battle, and quest your way through Crashlands, an outlandish story overflowing with sass!

Become Flux Dabes, a galactic trucker whose latest shipment gets derailed by a chin-strapped alien menace named Hewgodooko, leaving you stranded on an alien planet. As you hustle to retrieve your packages you'll become enmeshed in a nefarious plot of world domination, which will require all of your wits and both of your glutes to overcome. Learn recipes from the local sentient life, make new friends, uncover ancient secrets and deadly bosses, tame everything and build yourself a home-away-from-home as you learn to thrive on planet Woanope.

## ▼▼ Key Features ▼▼

### ● Expansive Crafting System ●

Unlock over 500 craftable items as you explore the world and learn its secrets!

### ● Self-managing, Infinite Inventory ●

In Crashlands, your inventory is infinite, manages itself, and retrieves your tools when you need them, so you can focus on adventuring, questing, and building. You'll never have dig through your bag or return to your base to free up inventory space!

### ● RPG-Style Character Progression ●

Become more powerful through creating ever-more-amazing items! As you grow in power, you can venture to new regions of the world, meet strange characters, discover new stories, and encounter new and interesting enemies.

### ● Skill-Based Combat ●

Learn the attacks of the enemies you encounter, and use your skill, agility, and wits to defeat them! You can even augment your fighting prowess with the power of the dozens of gadgets you can craft. Set your enemies on fire, stun them, slow down time, and more!

### ● Intuitive Base Building ●

Building a base in Crashlands is so simple it feels like fingerpainting. You can create beautiful, sprawling bases in minutes!

### ● Tameable Creatures ●

Every creature in Crashlands can become a trusty combat sidekick. Find an egg, incubate it, and hatch your very own adorable or hideous bundle of joy. You can even craft special items to grow and empower them!

### ● Huge World... with Huge Problems ●

Four sentient races, three continents, an epic bid for the future of the planet, and you - trapped in the middle, trying to deliver your freakin' packages. Take your time to dive into the sidestories of the characters you meet or just rush headlong into making that special delivery. With hundreds upon hundreds of quests, there's a lot to do and discover on planet Woanope!

### ● Effortless Cloud Saving ●

Just because your battery died or you accidentally dropped your device into a bottomless chasm, doesn't mean your save has to die with it. With BscotchID, you can easily store and retrieve your save from the cloud, and move it between your devices!

### ● Controller Support ●

Tired of rubbing your sweaty hands all over your beautiful touchscreen? No problem! We've got support for most mobile-compatible controllers, so you can rub your sweaty hands on some joysticks instead!

## ----- Recommended Hardware & OS:

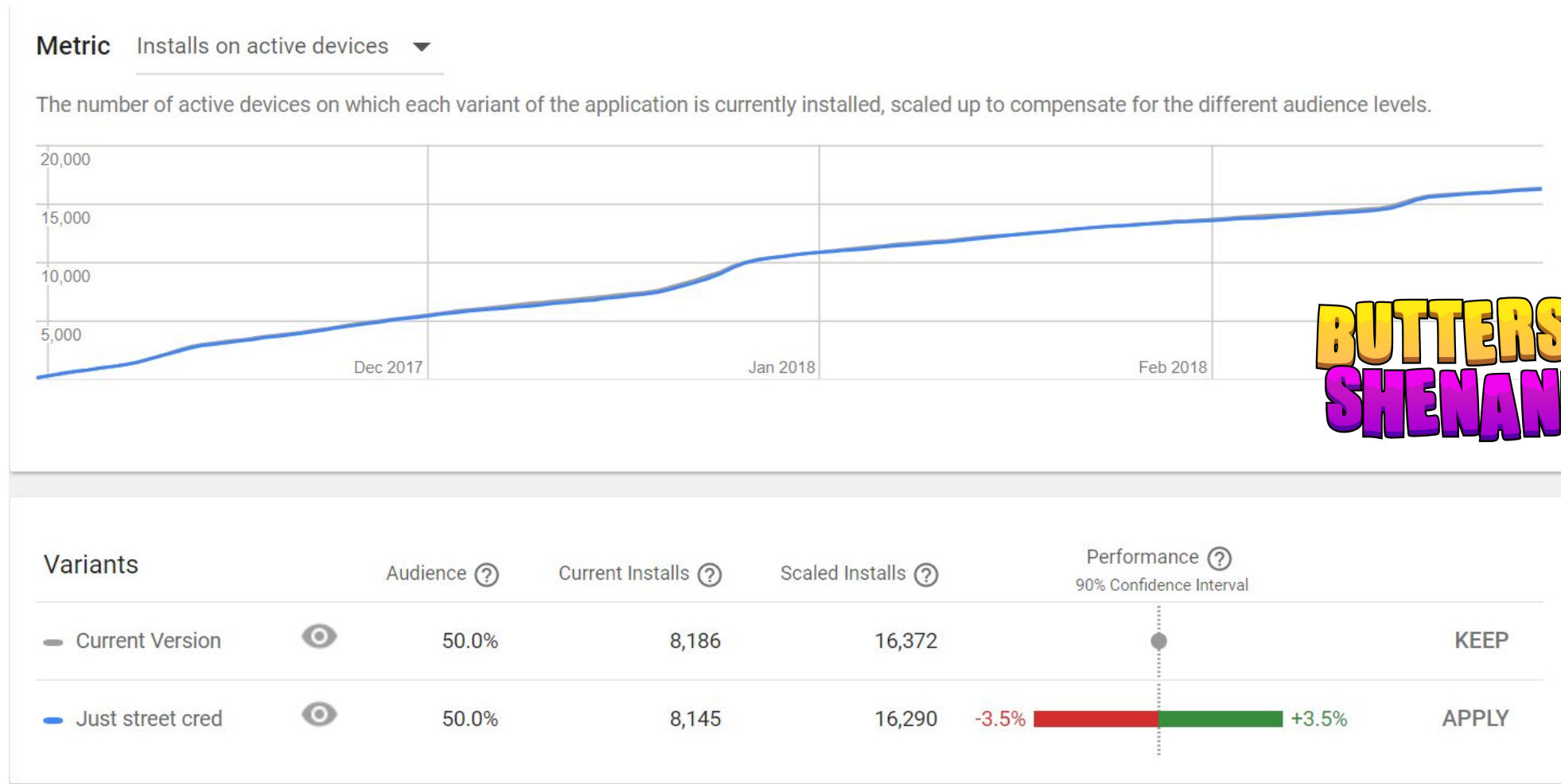
- Android 4.1 or newer
- At least 1GB RAM
- At least 960x540px screen resolution

"...a design masterpiece." 5/5 - TouchArcade  
2016 Game of the Year - TouchArcade, Gamezebo  
2016 Action Game of the Year - Pocket Tactics  
DICE 2016 Mobile Game of the Year Nominee  
Time's Top 10 Games / Top 50 Apps of 2016

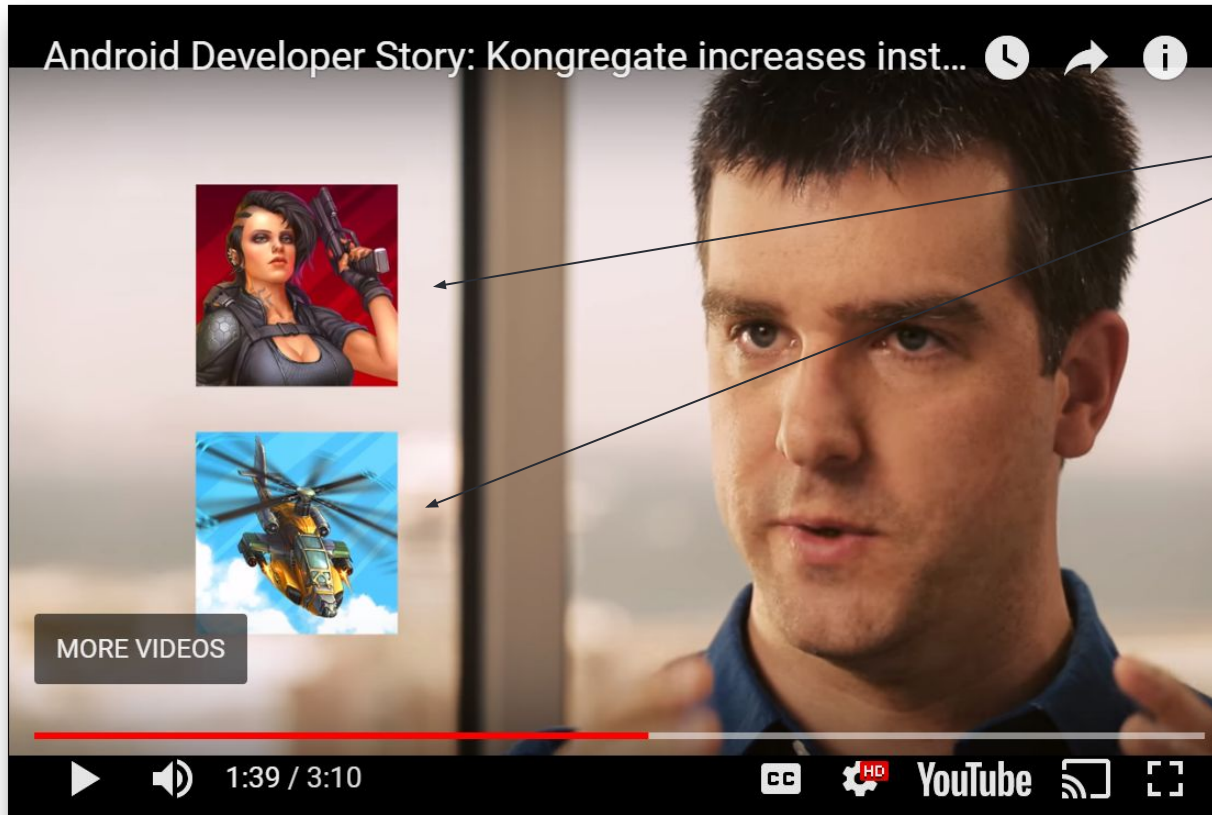
VS



# Meaningful No Difference



# A/B/C(context) Testing



On Google Play, Helicopter beat Girl with Gun by **92%**

...but we were using Girl with Gun because it beat Helicopter by **47%** on Kongregate.com

creative	impressions	clicks	ctr
5 - girl with gun	511,009	1,446	0.28%
11 - girl, short burnette	514,333	1,321	0.26%
5b - logo	525,426	1,221	0.2%
10 - girl, pony tail	522,983	1,188	0.2%
3 - guy with eye patch	538,999	1,222	0.2%
9 - girl, red streak	538,946	1,165	0.22%
3b - logo	544,249	1,122	0.21%
8 - blonde girl	577,632	1,125	0.19%
7 - helicopter	599,334	1,113	0.19%
flame car	555,663	1,025	0.18%
10noborder - girl, pony tail	547,888	1,003	0.18%
09noborder, girl, red streak	581,934	1,054	0.18%
11noborder - girl, short burnette	596,356	1,005	0.17%

Kongregate's tips for success with Store Listing Experiments

# Not Everything is **Testable**

---



**Daniel Cook** @danctheduck · 20h



A key premise of many data-driven approaches is you are mapping out static properties of a 'natural structure'. In this case, a set of human interests that does not change.

The mistake is not realizing you've created a feedback loop. The changes the audience.

**game**

# Hierarchy of Testing

Critical

Difficult

Advertising

Conversion

Initial  
Experience

First Weeks

Late Game

Test everything, all the time, everywhere.  
Tools abound.

As much testing of visual assets as possible. Tools more limited.

Significant testing possible but tests will often have only minor effects.

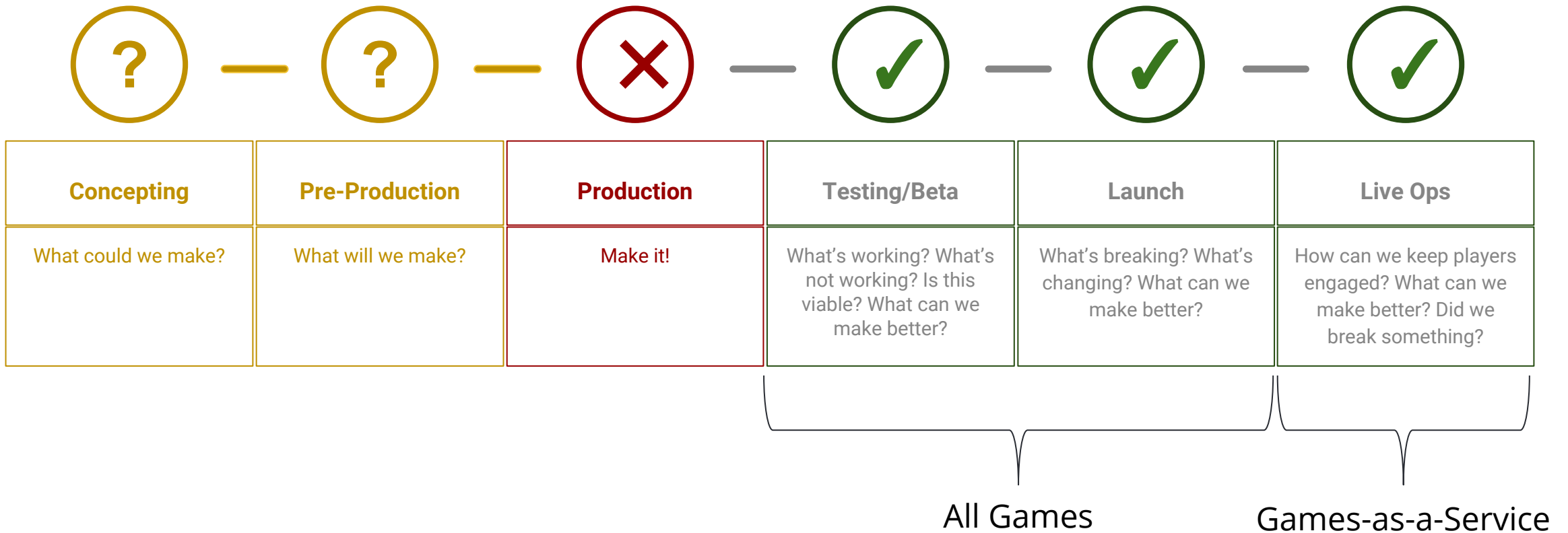
Still possible, especially around store, offers, feature unlocks.

Tread carefully. Sample size, audience expectations and player fairness become a challenge

# Game Data Lifecycle

## Creation

## Optimization



# What Should We **Make**?

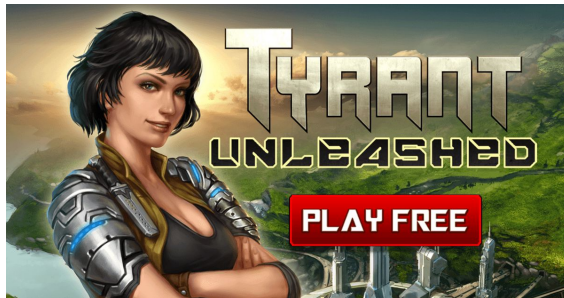


**THE** decision, always, for game companies

# Creative Iteration

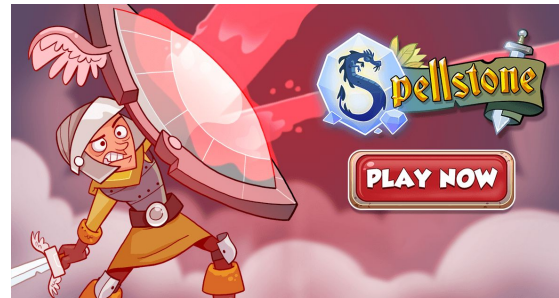
## Tyrant Unleashed

Launched August 2013  
\$7 CPIs



## Spellstone

Launched September 2015  
\$5 CPIs



## Animation Throwdown

Launched September 2016  
\$2.50 CPIs



# Pre-Production Testing



Source Totals	Impressions	Clicks	CTR
Relic Raiders	39,194	949	2.42%
Castaway Cove	40,450	1,068	2.64%
Eon Heroes	26,028	503	1.93%

# Live Disappointment



Source	Impressions	Clicks	CTR	Conversion
Castaway Cove Art Test Round 1	53,929	1,157	2.15%	n/a
Castaway Cove Art Test Round 2	40,450	1,068	2.64%	n/a
Castaway Cove Art Test Round 3	175,762	3,323	1.89%	n/a
<b>Castaway Cove Test Markets</b>	<b>3,912,062</b>	<b>43,765</b>	<b>1.11%</b>	<b>22.39%</b>

← Target = 30%

CPIs for live version of Castaway Cove are okay, but much higher than we'd been targeting

# Data Maps the **Known** World



# But What Don't We **Know**?



# Known World vs Unknown World



# Determination Driven

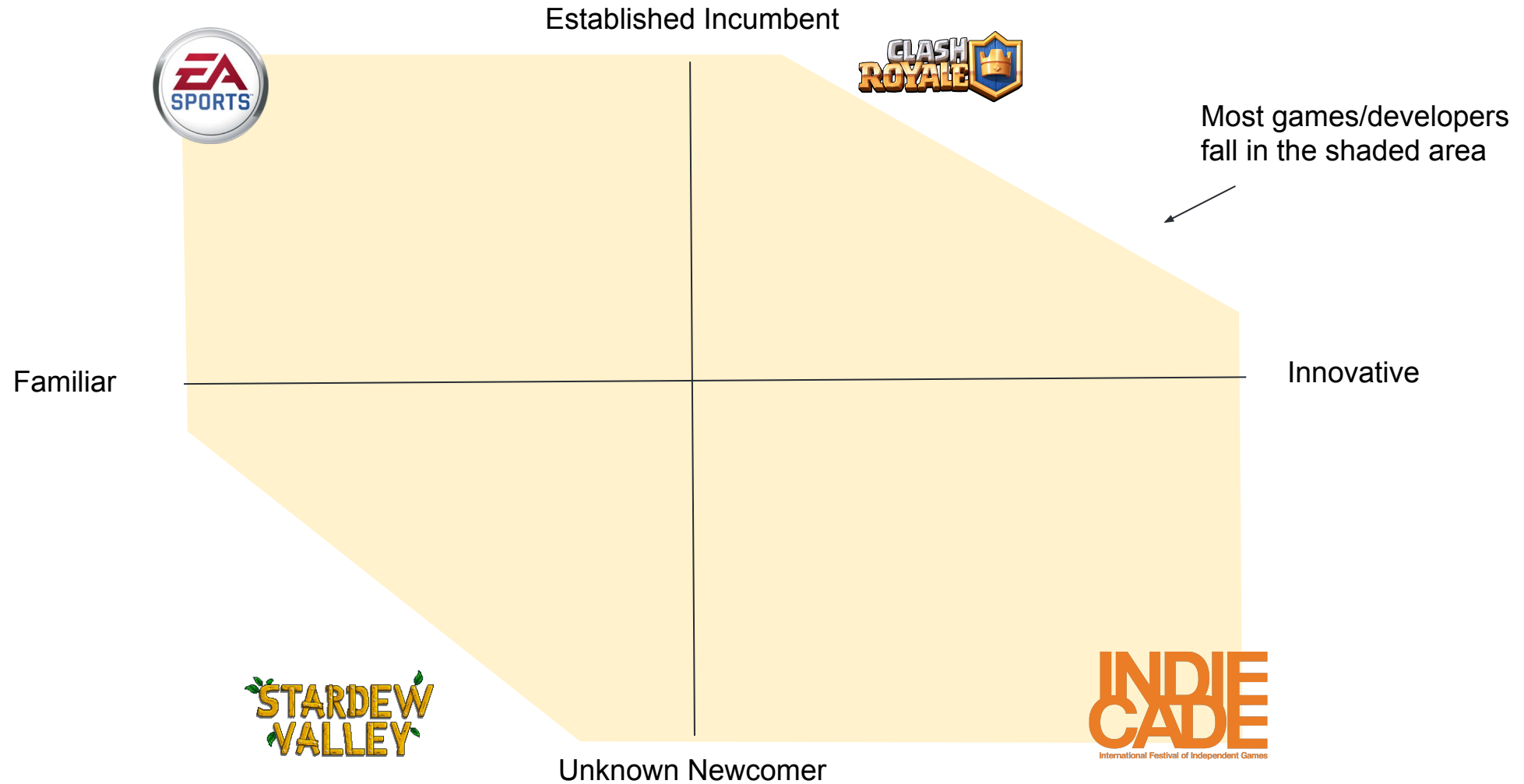
*Supersonic Acrobatic  
Rocket-Powered Battle Cars*



*Rocket League*



# There's No **Right** Answer



# Good **Data**, Good Luck

---



Keep your eyes and minds open

# Ask Me Anything

Special thanks to **Tammy Levy**, **Drew Levin**, **Zebulon Reynolds**, **Heather Gainer** and **Butterscotch Shenanigans** for help with data examples!

More great data & talks from the whole team on our blog  
<https://blog.kongregate.com>

Or follow us on Twitter:  
**@EmilyG**  
**@KongregateDevs**

And finally a good explanation of Wilcoxon Rank Sum test can be found here:

<https://www.slideshare.net/KrysselMaeCabili/wilcoxon-ranksum-mann-whitney-u-kolmogorovsmirnov-12>

## We're hiring for analytics!

### Open Roles

Director of Analytics  
Product Manager  
Data Analyst

### Locations

Portland, OR  
San Francisco, CA  
Chicago, IL  
Montreal, Canada