StoreIQ

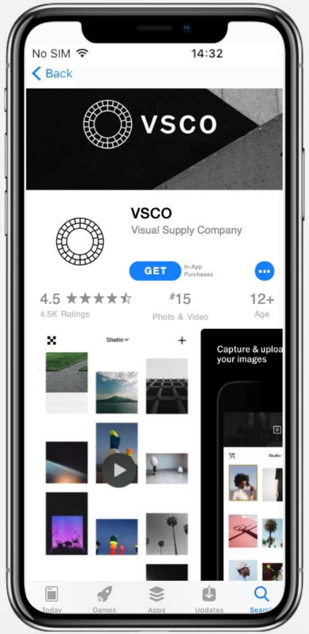
The Multi-Armed Bandit
for App Store Optimization

Introduction

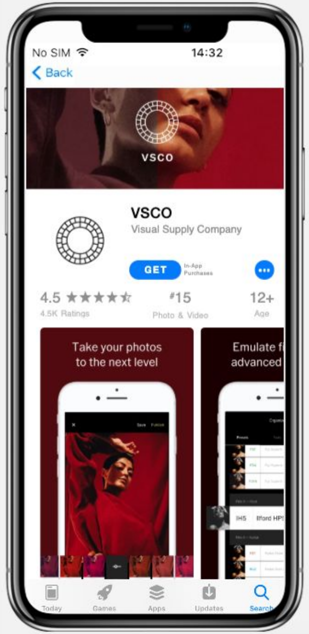
100% of users pass through the App Store essentially making it your new "homepage" and impacting your business success.

Introduction

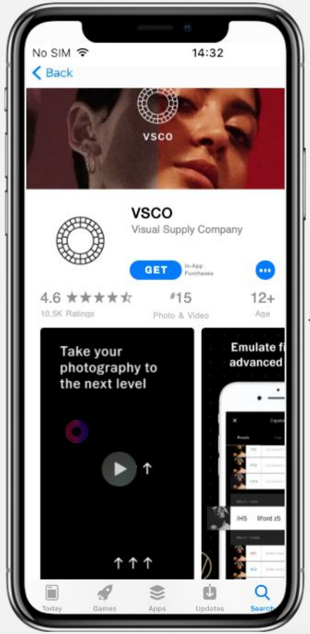
A



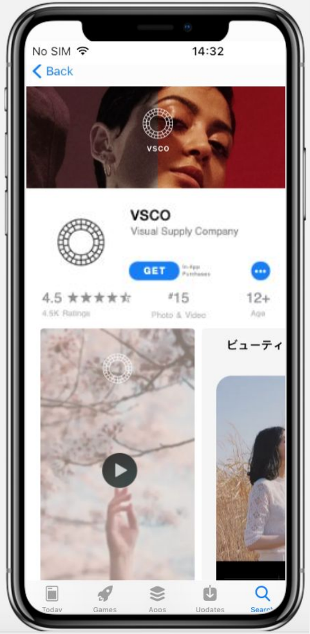
B



C



D

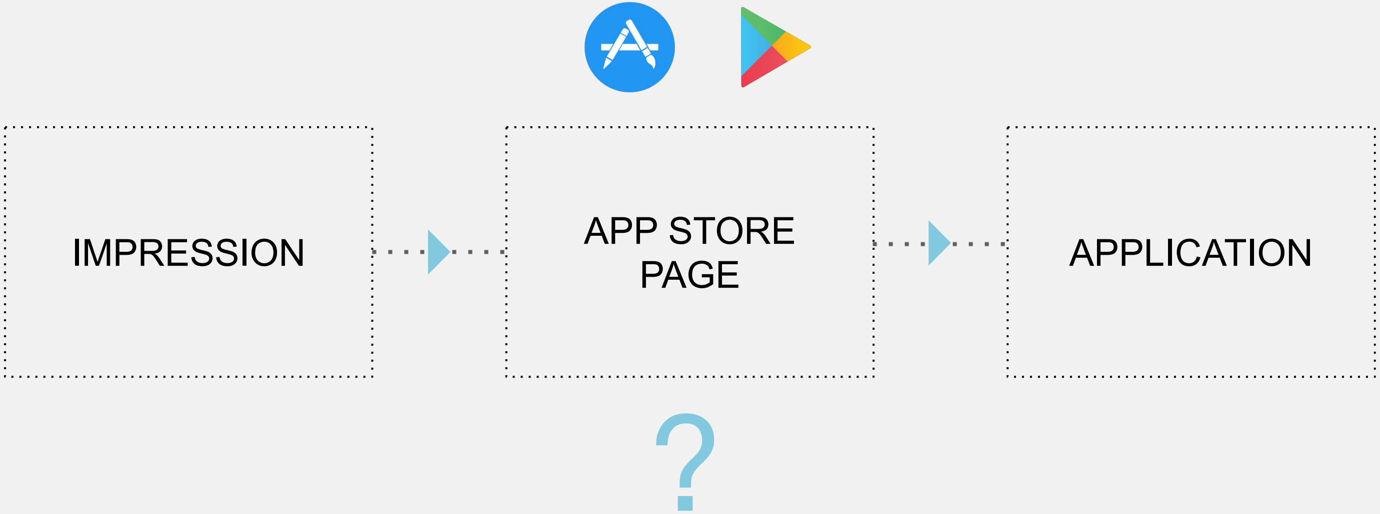


Goal

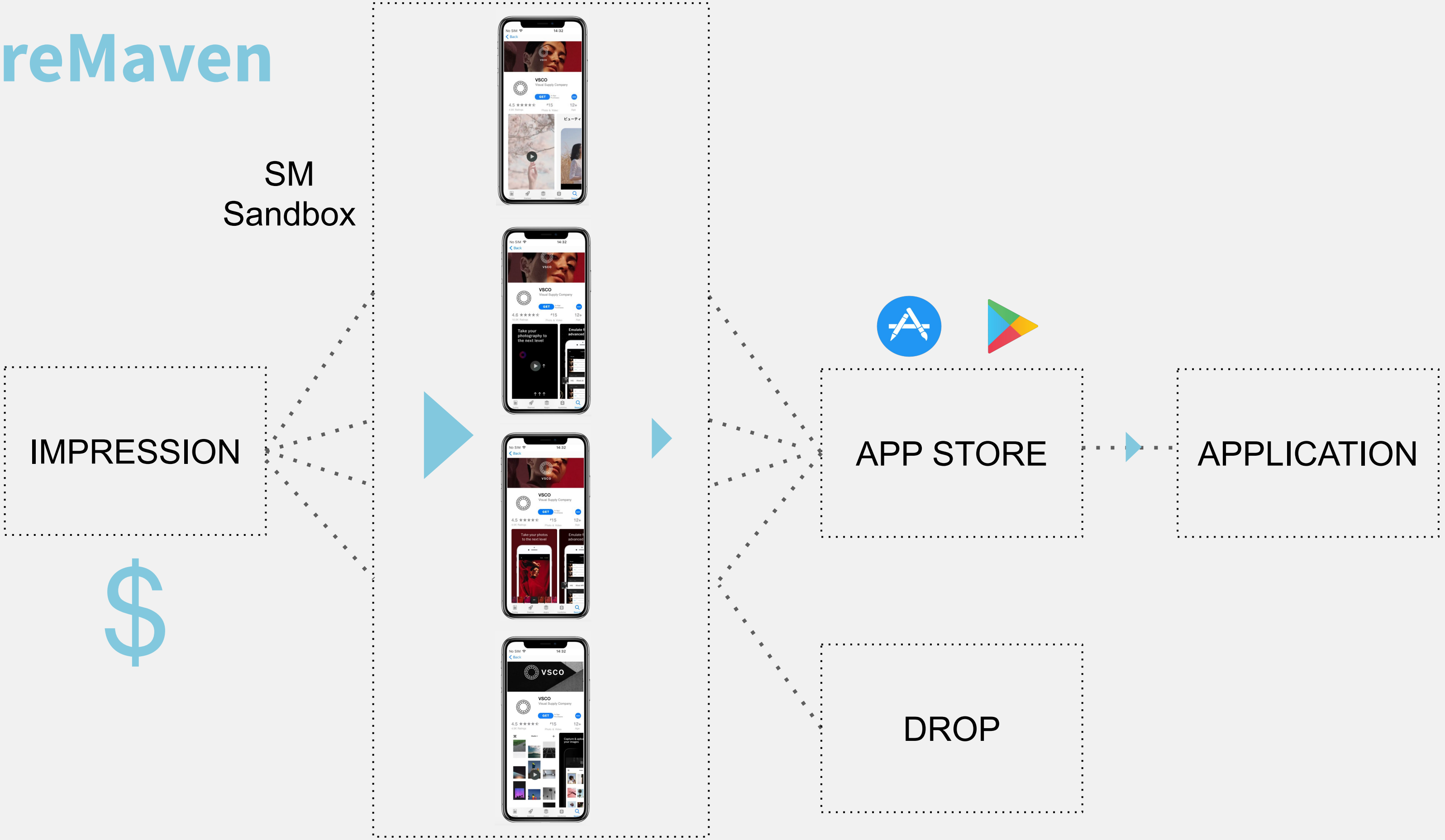
Conversion Rate

Cost Per Install

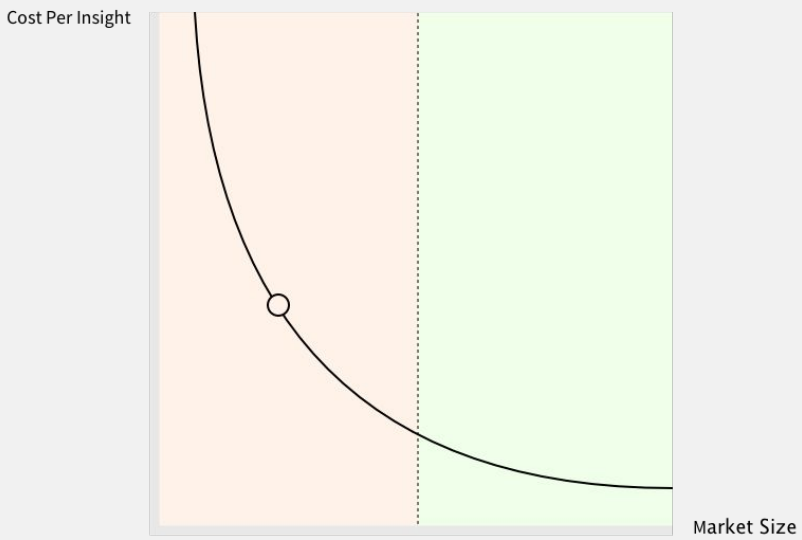
App Installation Flow

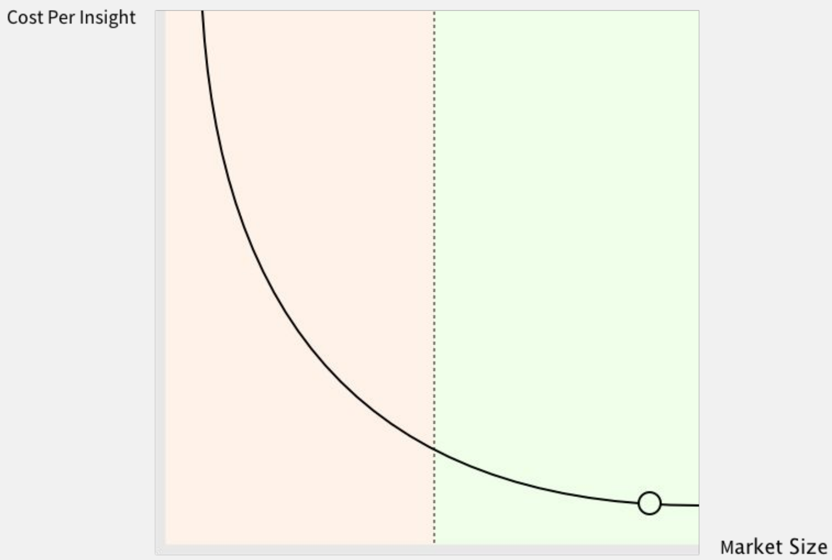


StoreMaven



StoreMaven





The Challenge

Minimize the cost per insight
without compromising accuracy

Today's Talk

1	Why not Proportion Testing ?
2	Multi-Armed Bandit – consideration & rejection
3	Cracking our business challenge with **StoreIQ**
4	Q & A

Proportion Testing – Why not ?

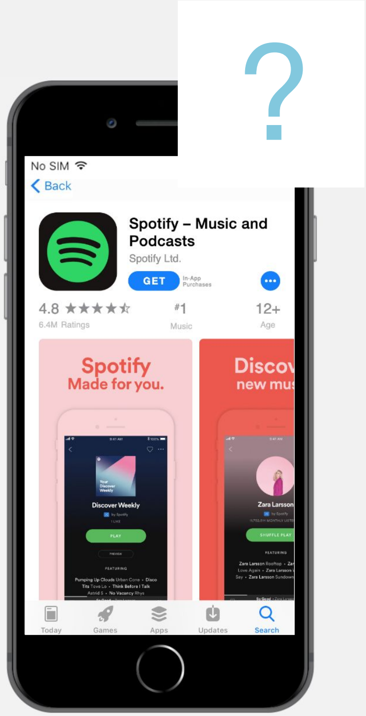
High cost per experiment: collecting large number of samples to achieve statistical validity in classical A/B test

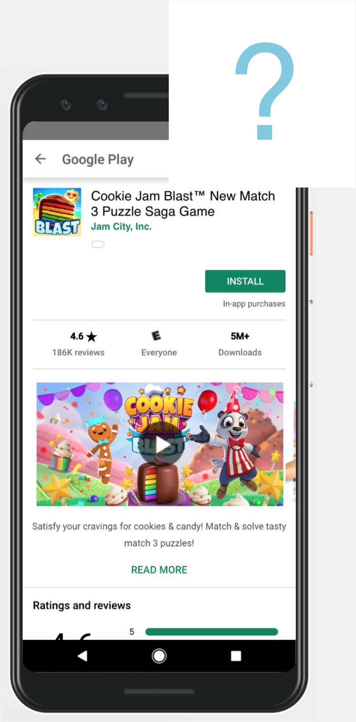
Multiple Hypotheses problem: test that $H_0: P_i = P_j, i \neq j$ when there are more than 2 variations, significance is 'harder' to reach

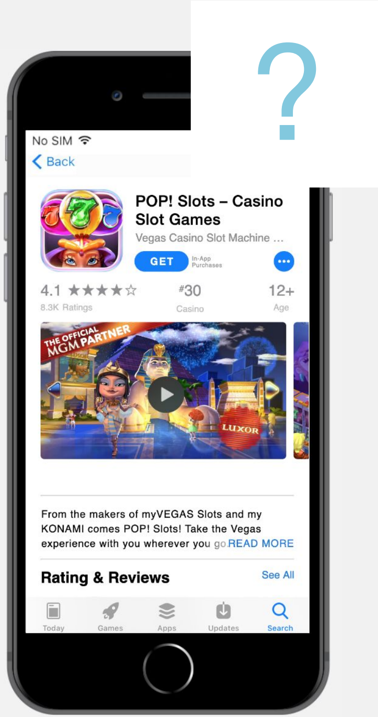
Accuracy issue: assumes all observations are IID (Independent Identical Distribution)

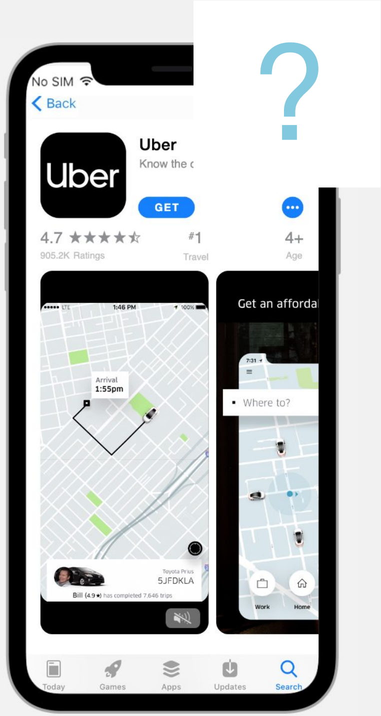
Robust experimental design is barely achievable pre-test though crucial for valid frequentist setting in the dynamic ASO ecosystem

How?


?
No SIM
Back
Spotify – Music and Podcasts
Spotify Ltd.
GET
In-App Purchases
4.8 ★★★★★
6.4M Ratings
#1
Music
12+
Age
Spotify
Made for you.
Discov
new mus
Today
Games
Apps
Updates
Search


?
Google Play
Cookie Jam Blast™ New Match 3 Puzzle Saga Game
Jam City, Inc.
INSTALL
In-app purchases
4.6 ★
186K reviews
Everyone
5M+
Downloads
Satisfy your cravings for cookies & candy! Match & solve tasty match 3 puzzles!
READ MORE
Ratings and reviews


?
No SIM
Back
POP! Slots – Casino Slot Games
Vegas Casino Slot Machine ...
GET
In-App Purchases
4.1 ★★★★☆
8.3K Ratings
#30
Casino
12+
Age
THE OFFICIAL MGM PARTNER
LUXOR
From the makers of myVEGAS Slots and my KONAMI comes POP! Slots! Take the Vegas experience with you wherever you go READ MORE
Rating & Reviews
See All
Today
Games
Apps
Updates
Search

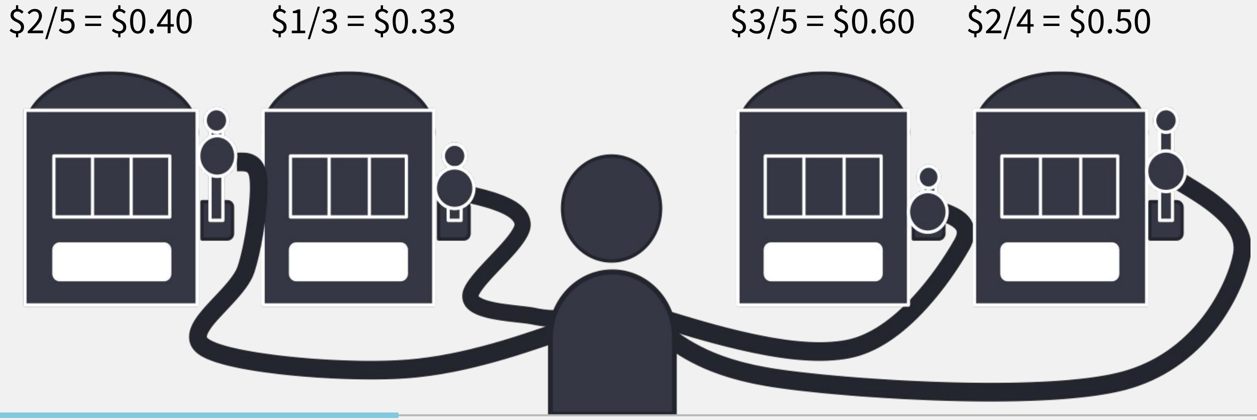

?
No SIM
Back
Uber
Know the c
GET
4.7 ★★★★★
905.2K Ratings
#1
Travel
4+
Age
Get an afforda
Arrival
1:55pm
Where to?
Toyota Prius
5JFDKLA
Work
Home
Today
Games
Apps
Updates
Search

Multi-Armed Bandit

Exploitation vs Exploration dilemma

Bandits try to balance the trade-off to maximize total rewards

$$R_T = \Sigma_{t=1}^{T}\left(P^* - PA_{(t)}\right)$$



Multi-Armed Bandit – Why Not?

Experiments are conducted in our ‘sandbox’ testing environment, no real-time optimizations are done (our regret is defined differently)

Doesn’t necessarily reduce the cost of testing. Usually it is not a testing model it is an optimization method

Many ways to determine a definitive winner and conclude the experiment

StoreIQ

Wikipedia ©

Backed by

4 years of Mobile App Stores user engagement events

More than 3 billion data points

Hundreds of millions of unique mobile users examined

20+ GEOs around the world

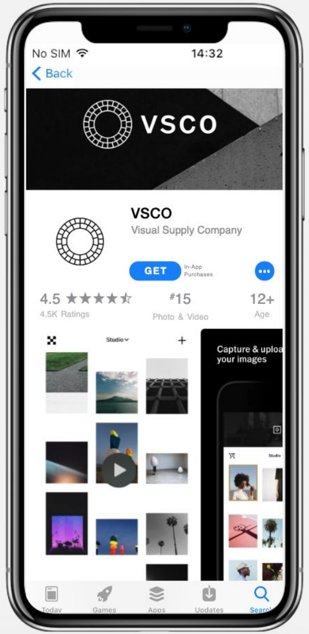
Bayesian approach

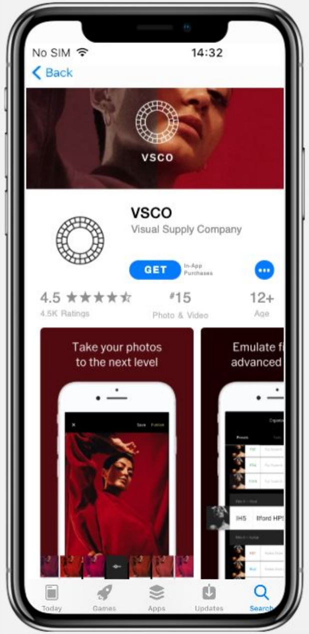
We use prior knowledge when calculating statistics

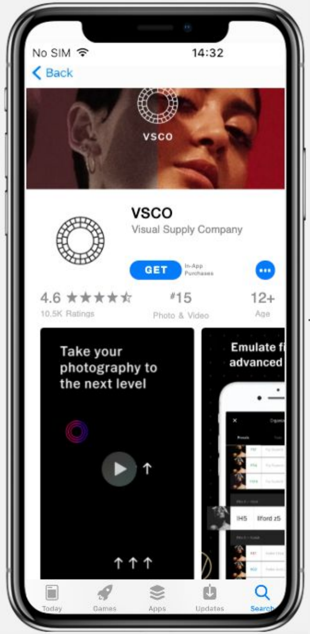
Bayes provides starting point from which we update our knowledge

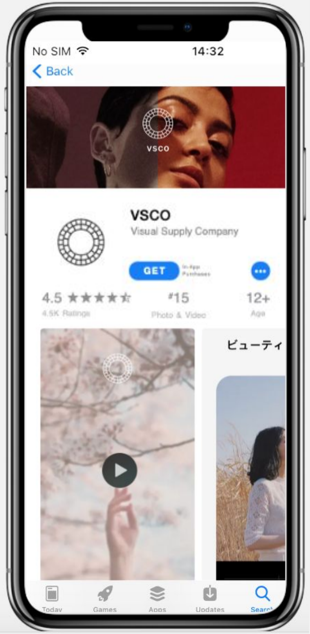
StoreIQ

An Experiment









StoreIQ

Each experiment starts with a warm-up period in which the algorithm gathers initial information on the competing variations

Defining **warm-up thresholds**:

T ; days of traffic (to control for the time of day)

O ; minimum observations per variation (users who started a session within the page)

C ; minimum conversions per variation (user who clicked through / installed)

StoreIQ

Prior knowledge is comprised of:

Learning the daily volatility of the traffic per variation

Learning weights per traffic sources & app categories

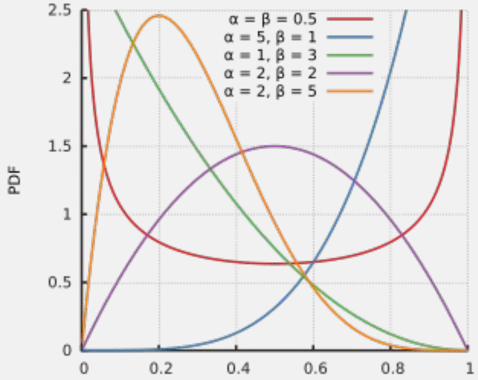
Learning ‘high quality’ user behaviour using app’s experiments history

StoreIQ

StoreIQ uses beta distribution to model its belief regarding competing variant conversion rates

The beta distribution is a suitable model for the random behavior of percentages and proportions.

The beta distribution is the conjugate prior probability distribution for the Bernoulli distribution



Wikipedia ©

$$IF\ CVR \sim Beta(\alpha, \beta)$$
$$THEN\ CVR \in [0, 1]$$

$$Round\ 1:$$

$$x \sim Bernoulli(\theta) \quad IID$$
$$x \in \{0,1\}$$

$$X - vector\ of\ x$$

$$f_\theta = ?$$
$$\theta = ?$$

In round 1, our prior is a derivation of the Beta distribution and is not dependent on the posterior a,b

$$f_{\theta|X} \propto L(X|\theta) \cdot prior = \theta^{\Sigma x} \cdot (1-\theta)^{N-\Sigma x} \cdot prior$$
$$\sim$$
$$Beta(\Sigma x + 1,\ N - \Sigma x + 1)$$

StoreIQ

$$Round\ 2:$$

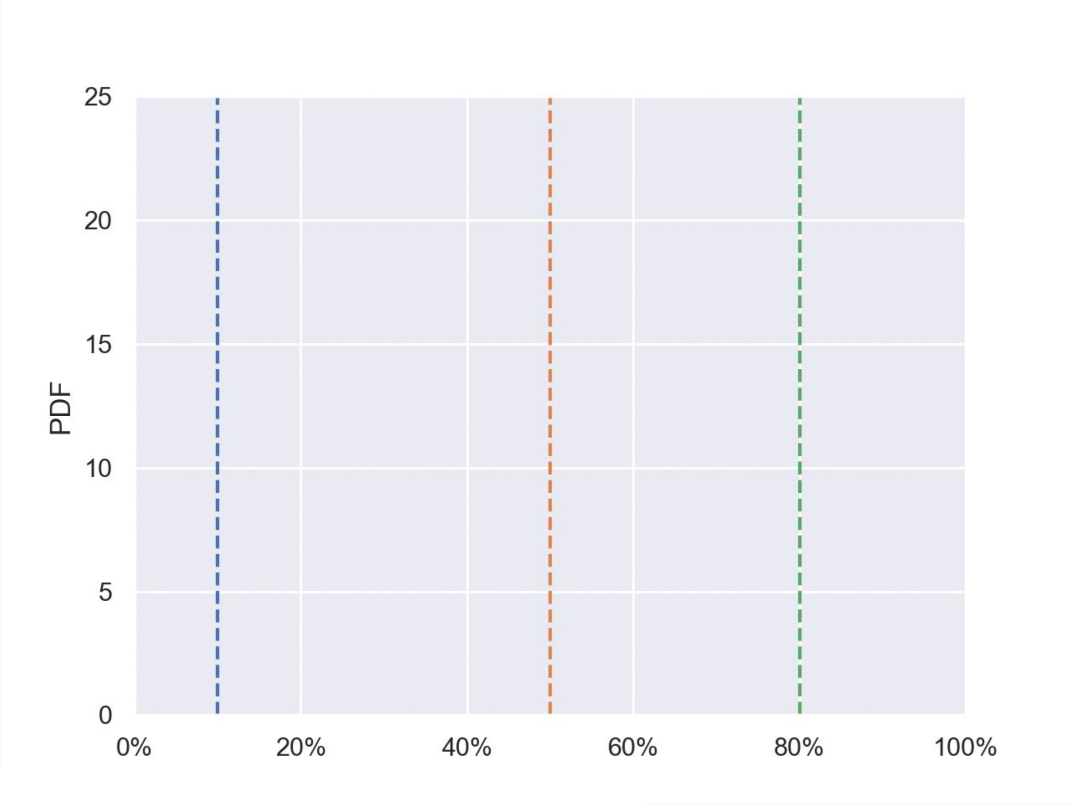
$$y \sim Bernoulli(\theta) \quad IID$$
$$y \in \{0,1\}$$

$$X - vector\ of\ y$$

$$f_\theta = ?$$
$$\theta = ?$$

$$f_{\theta|X} \propto L(X|\theta) \cdot prior = \theta^{\Sigma x + \Sigma y} \cdot (1-\theta)^{M+N-\Sigma x - \Sigma y} \cdot prior$$
$$\sim$$
$$Beta(\Sigma x + \Sigma y + 1,\ M + N - \Sigma x + \Sigma y - \ + 1)$$

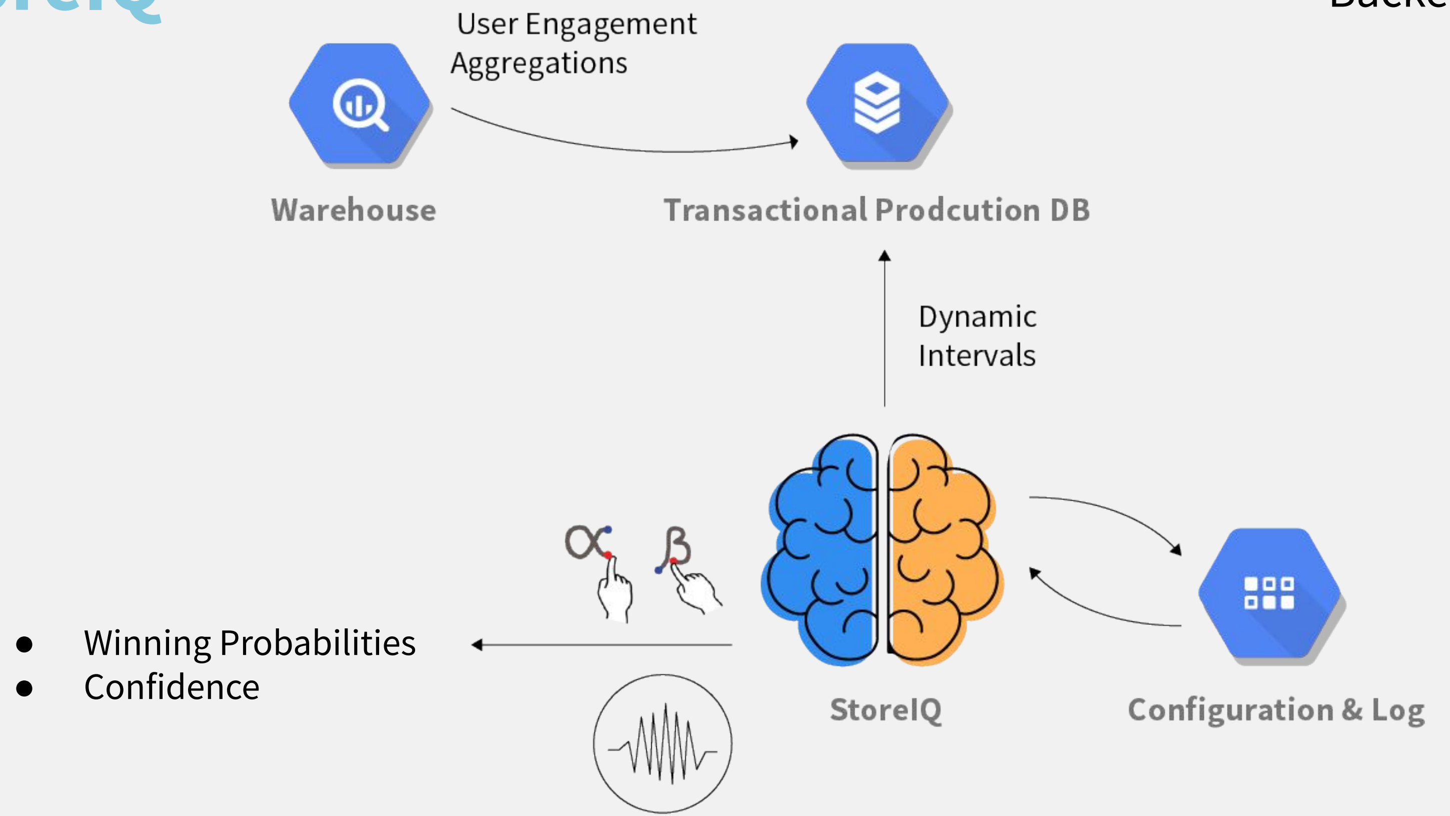
StoreIQ



Spotx.tv ©

StoreIQ

Backend Peek

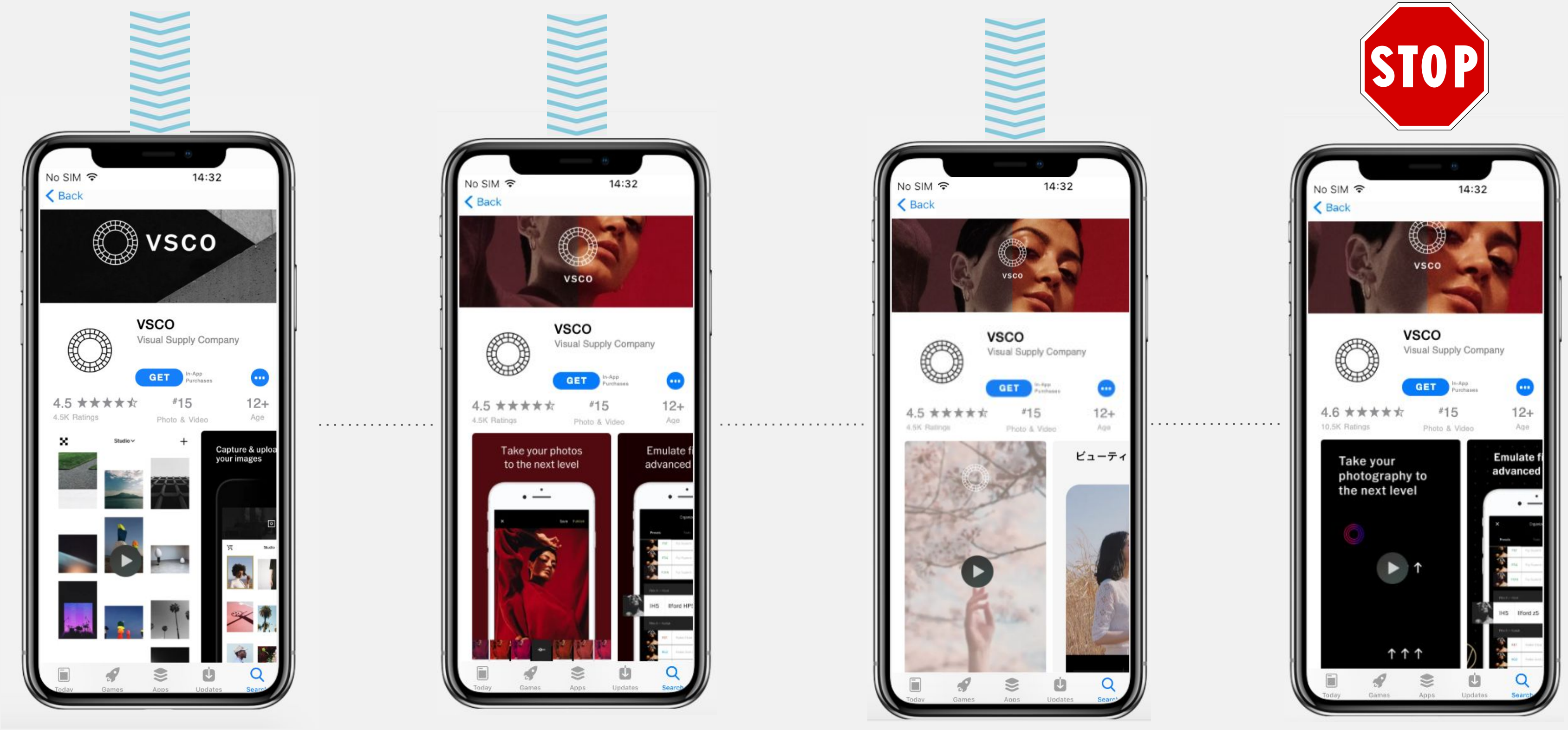


StoreIQ

Winning Probabilities

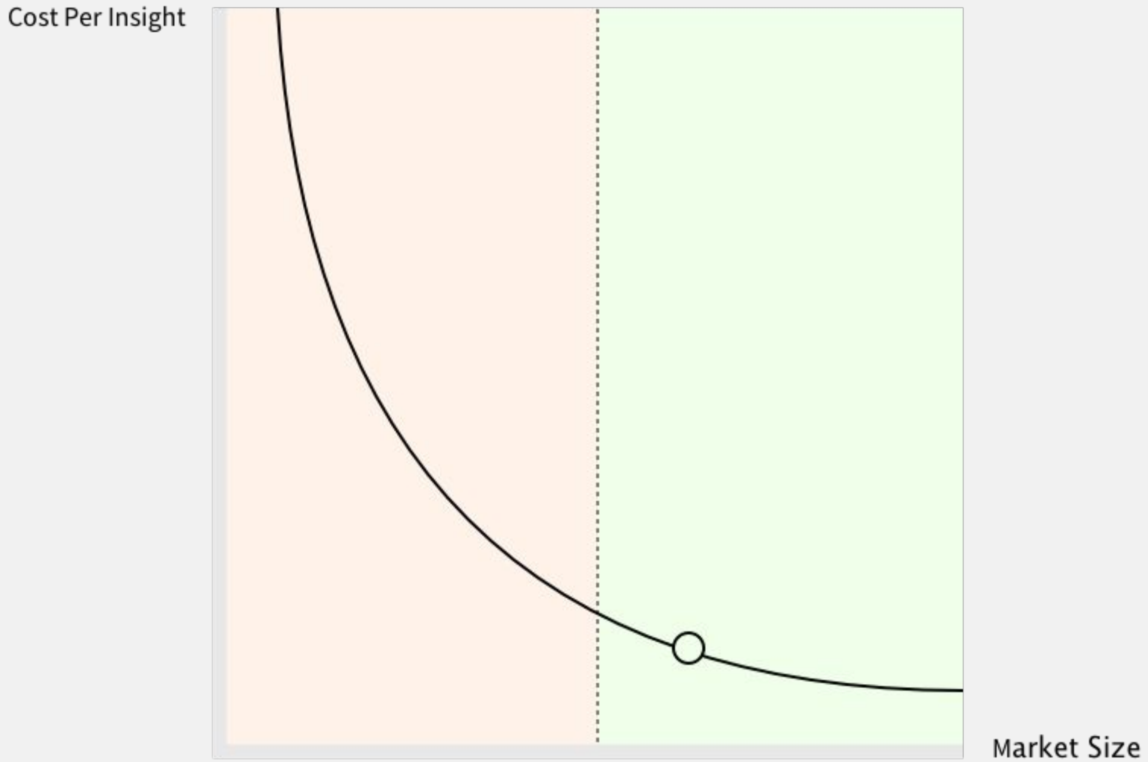
Iteration	Variation 1	Variation 2	Variation 3	Variation 4
1	0.251654	0.201544	0.4481	0.1043
2	0.355897	0.353784	0.156574	0.13
3	0.484654	0.5464651	0.564546	0.234
. . .	. . .	. . .	. . .	. . .
N	0.3684615	0.2654684	0.68474	0.1593
Winning Probabilities	0.122	0.0005	0.8745 (Leader!)	0.003

StoreIQ



STOP
No SIM
14:32
Back
VSCO
Visual Supply Company
GET
In-App Purchases
4.5
4.5K Ratings
#15
Photo & Video
12+
Age
Capture & upload your images
Take your photos to the next level
Emulate film advanced
ビューティ
4.6
10.5K Ratings
Take your photography to the next level
Today
Games
Apps
Updates
Search

StoreIQ



StoreIQ

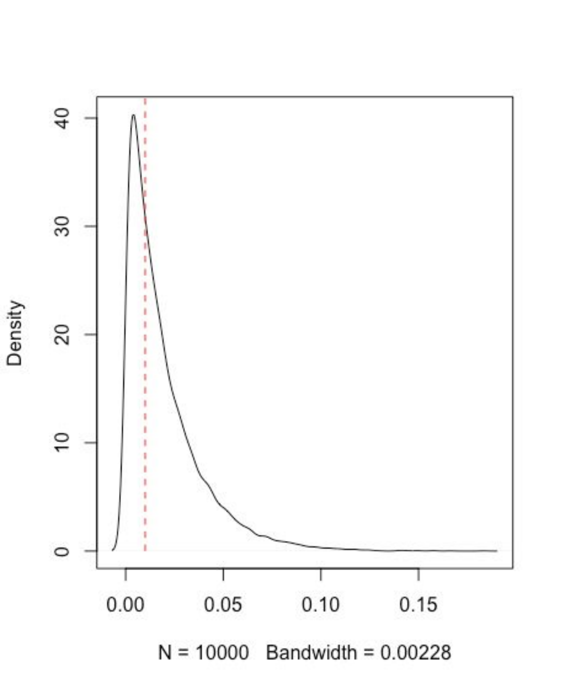
Confidence

Iteration	Variation 1	Variation 2	Variation 3 - Current Leader	Lift function of iteration - Winner from Leader	Lift below threshold
1	0.251654	0.201544	0.5481	0	1
2	0.355897	0.354884	0.33566	0.06029	0
3	0.484654	0.5464651	0.564546	0	1
. . .	. . .	. . .	. . .	. . .	. . .
N	0.4684615	0.2654684	0.68474	0	1
Confidence	-	-	-	-	**0.95**

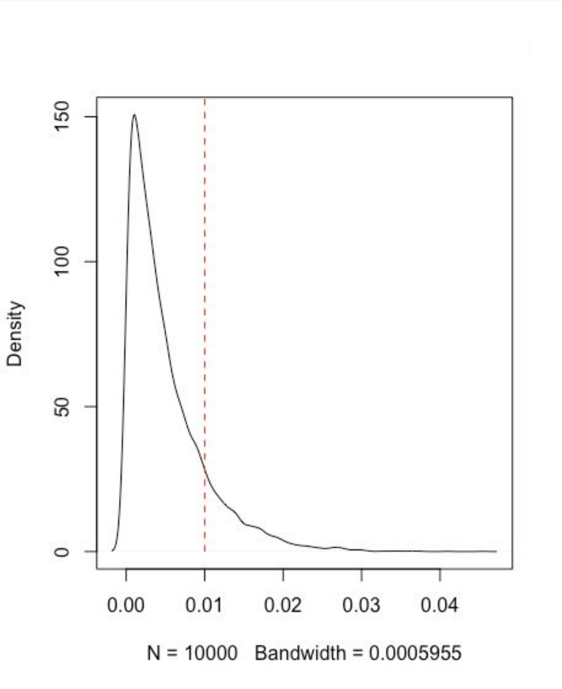
StoreIQ

Confidence

Low confidence

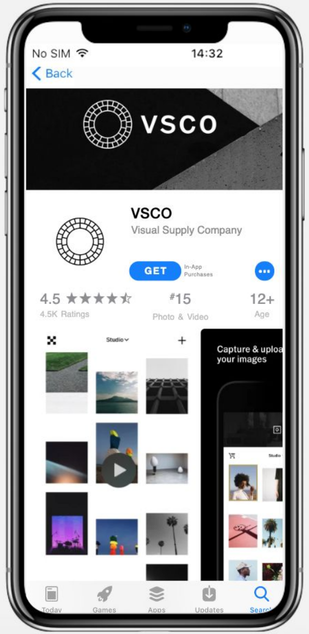


High confidence

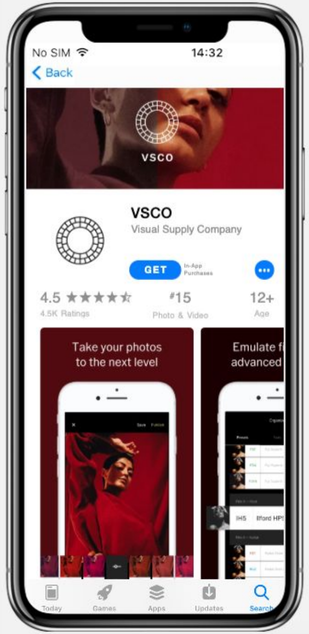


StoreIQ

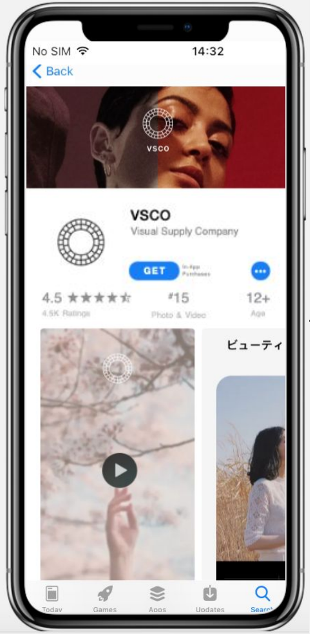
A



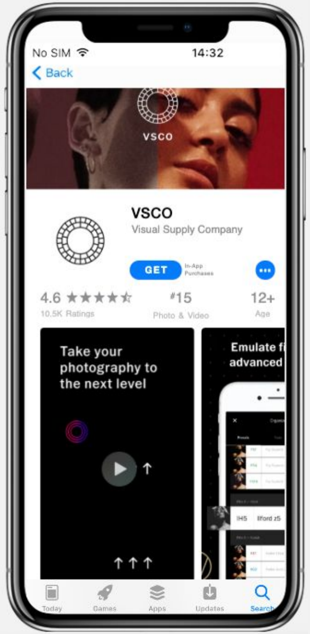
B



C - Winner!

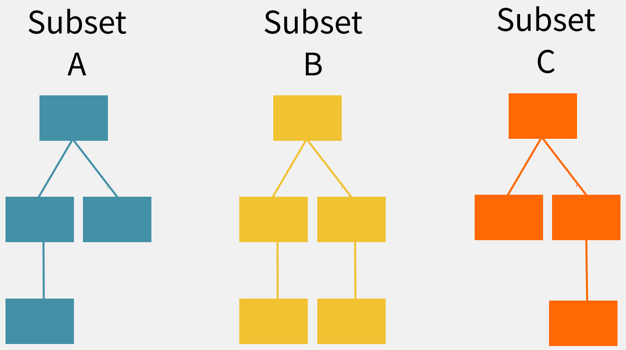


D



StoreIQ Booster – Bagging Trees (RF)

Train Dataset = user behavior from **app's** experiment history + user activity from a live experiments interval 0 to t-1



C1 = Compute the correlation between predicted and empirical CVR per variation

C2 = Compute the correlation between the per store CVR of the explorers and non-explorers

C3 = Compute the accuracy of the current model on Test Dataset (users from interval t in a live experiment)

Boost = baseline + (# of boosts given so far * $\mathcal{E}$)

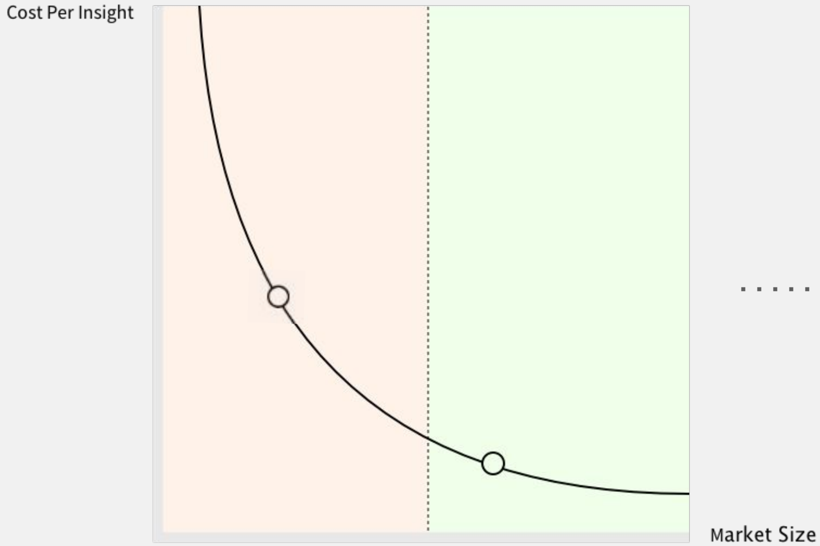
Result = a model that predicts for a given user whether he will Install or not

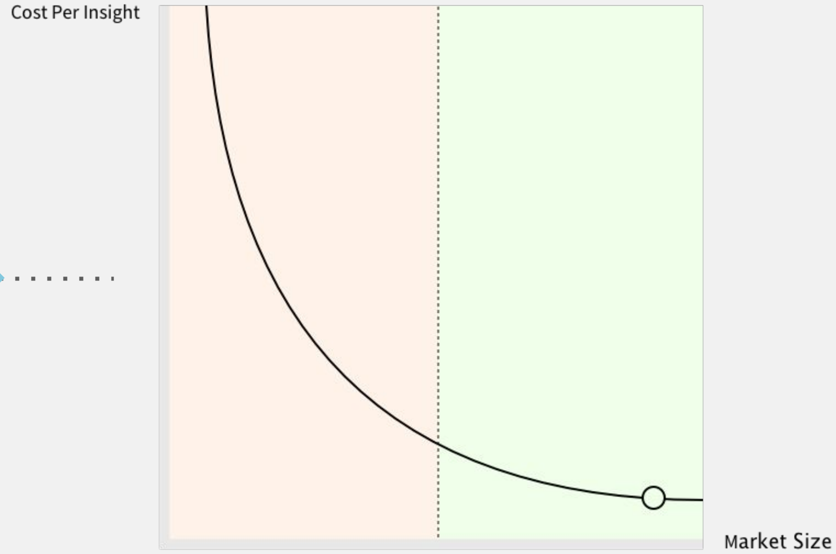
If (C3 >= **accuracyThreshold** and C1*C2 >= **correlationThreshold**):

observationsToAdd_i = round(w_i*C3*C2*C1*boost*N)
conversionsToAdd_i = round(cvr_i * observationsToAdd_i)

STOREMAVEN

StoreIQ – Boosting





Q & A

Thank You!

STOREMAVEN