MS&E 233
Game Theory, Data Science and AI
Lecture 16

Vasilis Syrgkanis
Assistant Professor
Management Science and Engineering
(by courtesy) Computer Science and Electrical Engineering
Institute for Computational and Mathematical Engineering

Computational Game Theory for Complex Games

1
- Basics of game theory and zero-sum games (T)
- Basics of online learning theory (T)
- Solving zero-sum games via online learning (T)
- *HW1: implement simple algorithms to solve zero-sum games*
- Applications to ML and AI (T+A)
- *HW2: implement boosting as solving a zero-sum game*

2
- Basics of extensive-form games
- Solving extensive-form games via online learning (T)
- *HW3: implement agents to solve very simple variants of poker*

3
- General games, equilibria and online learning (T)
- Online learning in general games
- *HW4: implement no-regret algorithms that converge to correlated equilibria in general games*

Data Science for Auctions and Mechanisms

4
- Basics and applications of auction theory (T+A)
- Basic Auctions and Learning to bid in auctions (T)
- *HW5: implement bandit algorithms to bid in ad auctions*

5
- Optimal auctions and mechanisms (T)
- Simple vs optimal mechanisms (T)
- *HW6: implement simple and optimal auctions, analyze revenue empirically*

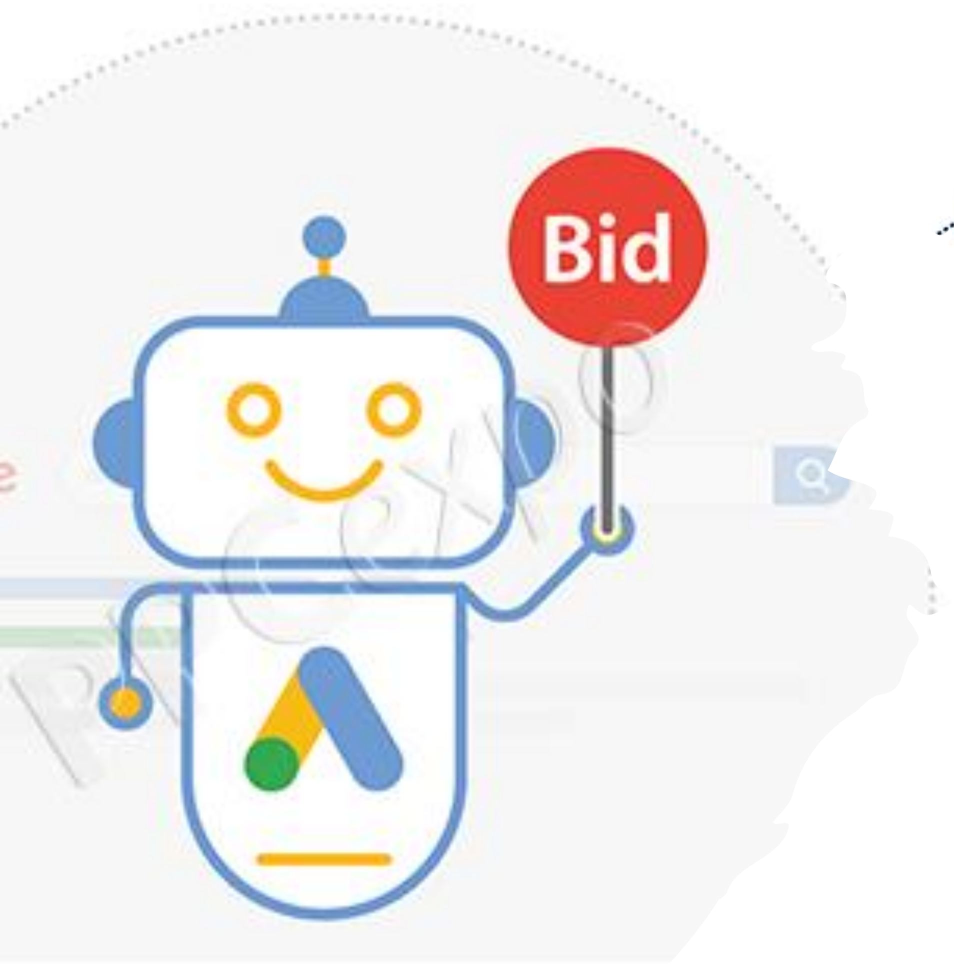
6
- Basics of Statistical Learning Theory (T)
- Optimizing Mechanisms from Samples (T)
- *HW7: implement procedures to learn approximately optimal auctions from historical samples*

Further Topics

7
- Econometrics in games and auctions (T+A)
- **A/B testing in markets (T+A)**
- *HW8: implement procedure to estimate values from bids in an auction*

Guest Lectures

- Mechanism Design for LLMs, Renato Paes Leme, Google Research
- Auto-bidding in Sponsored Search Auctions, Kshipra Bhawalkar, Google Research



Econometrics in Games and Auctions

Econometrics in Games and Auctions

- We are given data from actions of players in a game (and potentially auxiliary contextual information about the game)

- Multiple instances were players played the same type of game

- We don’t know the exact utilities of the players in the game

- We want to use the data to learn the parameters of the utilities of the players in the game or the distribution of these parameters

Why useful?

Scientific: economically meaningful quantities

Perform counter-factual analysis: what would happen if we change the game?

Performance measures: welfare, revenue

Testing game-theoretic models: if theory on estimated quantities predicts different behavior, then in trouble

If I know the equilibrium bid distribution G, then whenever *I see a bid b_i*, I can *reverse engineer* and *uniquely determine the value* that led to such a bid

unobserved value

$$v_i = b_i + \frac{1}{(n-1)\frac{g(b_i)}{G(b_i)}}$$

Reverse hazard ratio
of distribution of bids
"Probability that opponent bid is immediately below b_i given that it is below b_i"

observed equilibrium bid

More competition ⇒ less "value reduction"

Side Note (Asymmetric Bidders): If I know the equilibrium bid distributions G_i, then whenever *I see a bid b_i*, I can *reverse engineer* and *uniquely determine the value v_i* that led to such a bid

unobserved value

$$v_i = b_i + \frac{1}{\sum_{k \neq i} \frac{g_k(b_i)}{G_k(b_i)}}$$

observed equilibrium bid

Reverse hazard ratio of distribution of bids of k-th opponent

"Probability that opponent bid is immediately below b_i given that it is below b_i"

More competition ⇒ less "value reduction"

Estimating CDFs from Truthful Samples

Given truthful bids $v_1, \ldots, v_m$ of players in instances of Second Price Auction the CDF of the distribution can be approximated by the empirical CDF to an error of $\approx \frac{1}{\sqrt{n}}$

$$F(z) \stackrel{\text{def}}{=} \Pr(v < z) \approx \frac{1}{n \cdot m} \sum_{i,j} 1\{v_{ij} < z\} \stackrel{\text{def}}{=} \hat{F}(z)$$

Estimating CDFs and PDFs of Bids from FPA Bid Samples

Given bids $b_1, \dots, b_m$ of players in instances of First Price Auction the CDF and PDF of the *bid distribution* can be approximated by empirical CDF and a Kernel Density Estimate

$$G(z) \stackrel{\text{def}}{=} \Pr(b < z) \approx \frac{1}{n \cdot m} \sum_{i,j} 1\{b_{ij} < z\} \stackrel{\text{def}}{=} \hat{G}(z)$$

$$g(z) = \partial_z G(z), \qquad \boxed{\hat{g}(z) = \frac{1}{n \cdot m} \sum_{i,j} \frac{1}{h_n} K\left(\frac{b_{ij} - z}{h_n}\right)}$$

Fraction of samples that ≈lie within h from z, divided by region length

Estimating CDFs and PDFs of Values from FPA Bid Samples

Given bids $b_1, \ldots, b_m$ of players in instances of First Price Auction the CDF and PDF of the *value distribution* can be approximated using the plug-in approach, by approximately "inverting the bid" and using the "recovered value as a truthful sample"

$$\hat{v}_{ij} = b_{ij} + \frac{\hat{G}(b_{ij})}{(n-1)\,\hat{g}(b_{ij})}$$

$$\hat{F}(z) \stackrel{\text{def}}{=} \frac{1}{n \cdot m} \sum_{i,j} 1\{\hat{v}_{ij} < z\}, \qquad \hat{f}(z) = \frac{1}{n \cdot m} \sum_{i,j} \frac{1}{h_n} K\left(\frac{\hat{v}_{ij} - z}{h_n}\right)$$

Formal Guarantees

- Suppose pdf f has R uniformly bounded continuous derivatives
- If we observed values then error rate of $\left(\frac{nm}{\log(nm)}\right)^{-\frac{R}{2R+1}}$ [Stone'82]
- Now that only bids are observed, [GPV'00] show that best achievable is: $\left(\frac{nm}{\log(nm)}\right)^{-\frac{R}{2R+3}}$
- The density f depends on the derivative of g

Why useful?

Scientific: economically meaningful quantities

Perform counter-factual analysis: what would happen if we change the game?

Performance measures: welfare, revenue

Testing game-theoretic models: if theory on estimated quantities predicts different behavior, then in trouble

What if all we want is to compare between auctions A and B in terms of revenue?

What I could potentially do is:
For each auction flip a coin;
If heads, then run auction A else run auction B

After many auctions compare average revenue from A auctions, vs., average revenue from B auctions

Is this correct?

We will see that it can be problematic and needs thought of how to analyze such data or structure such A/B tests!

Experimentation (aka A/B Testing)

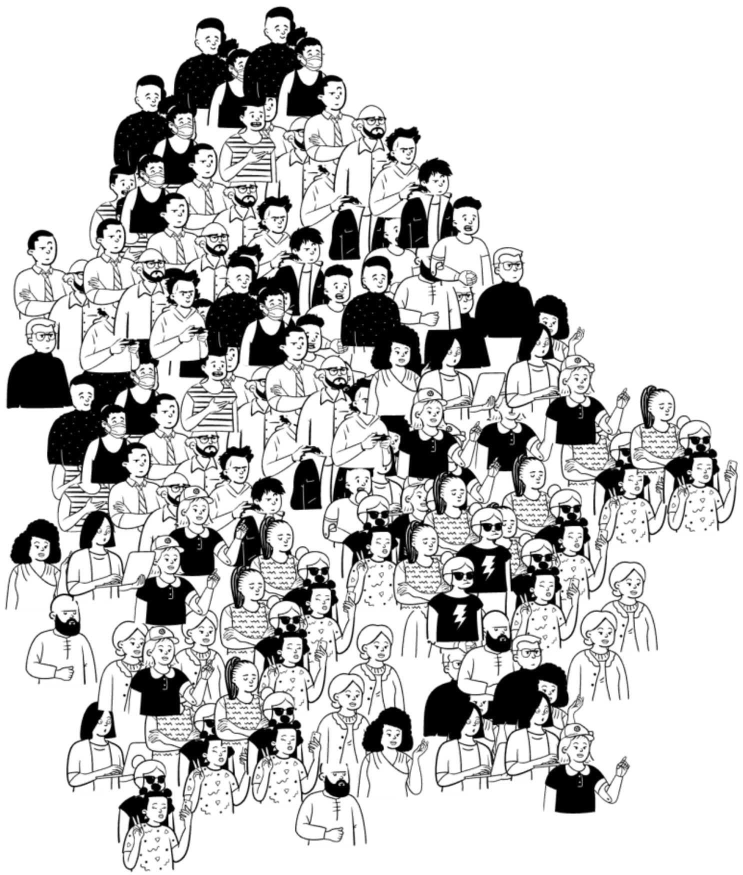
The Basics of A/B Testing

Randomization, Causality, Statistical Inference

The Mechanics

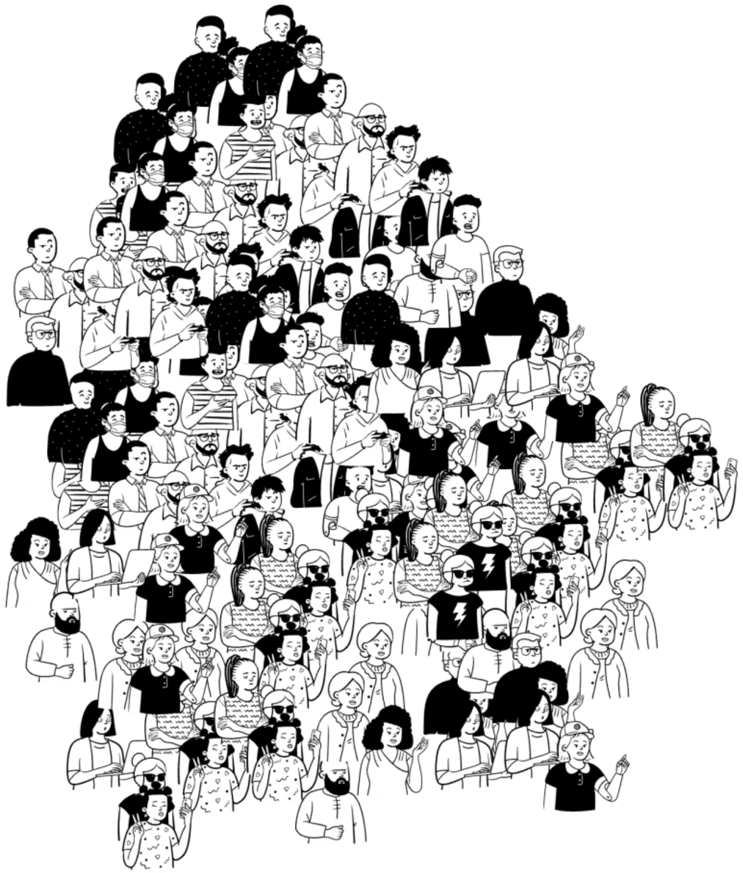
A/B Testing

user base



A/B Testing

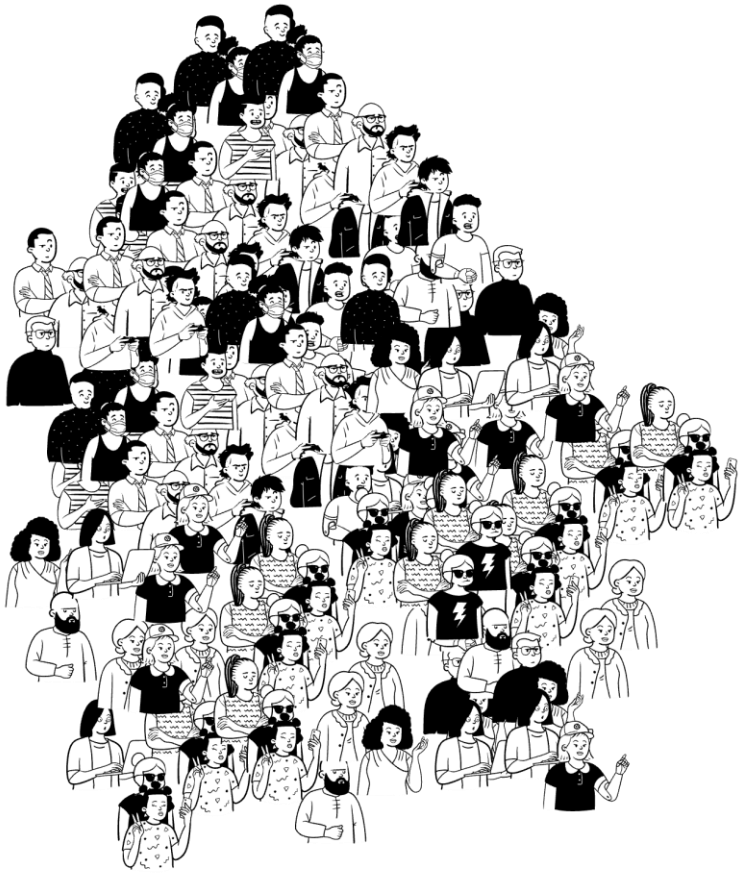
user base



sample

A/B Testing

user base

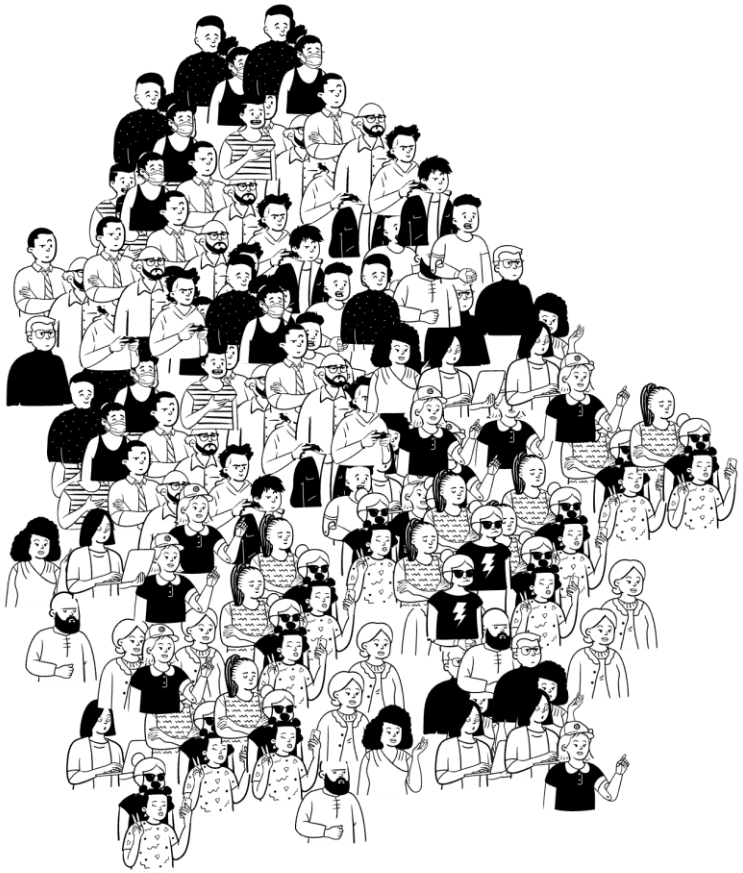


sample

flip a coin for each user

A/B Testing

user base

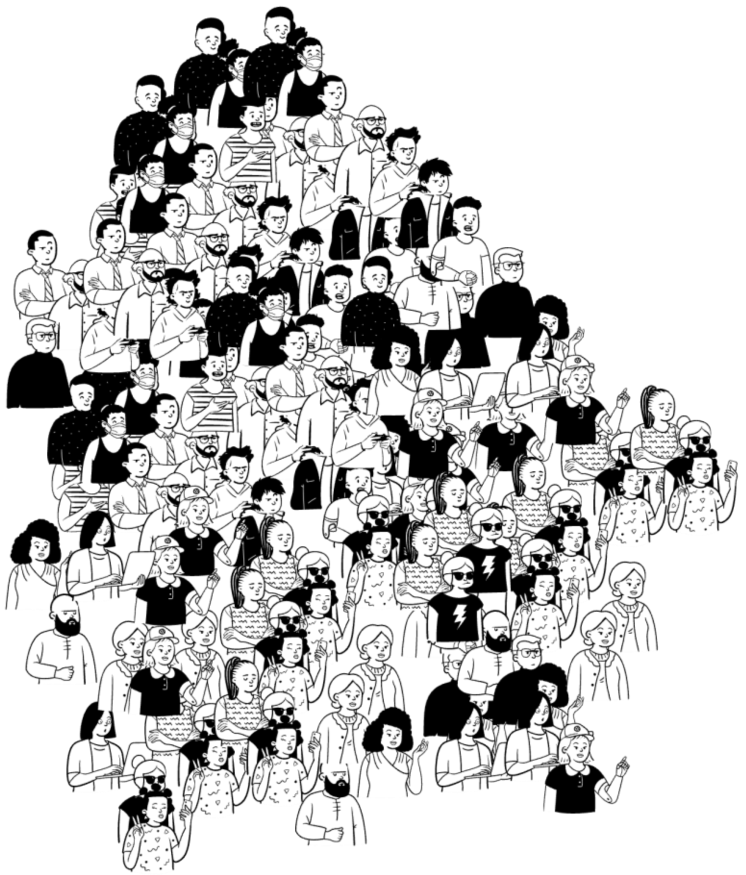


sample

split into groups based on coin

A/B Testing

user base



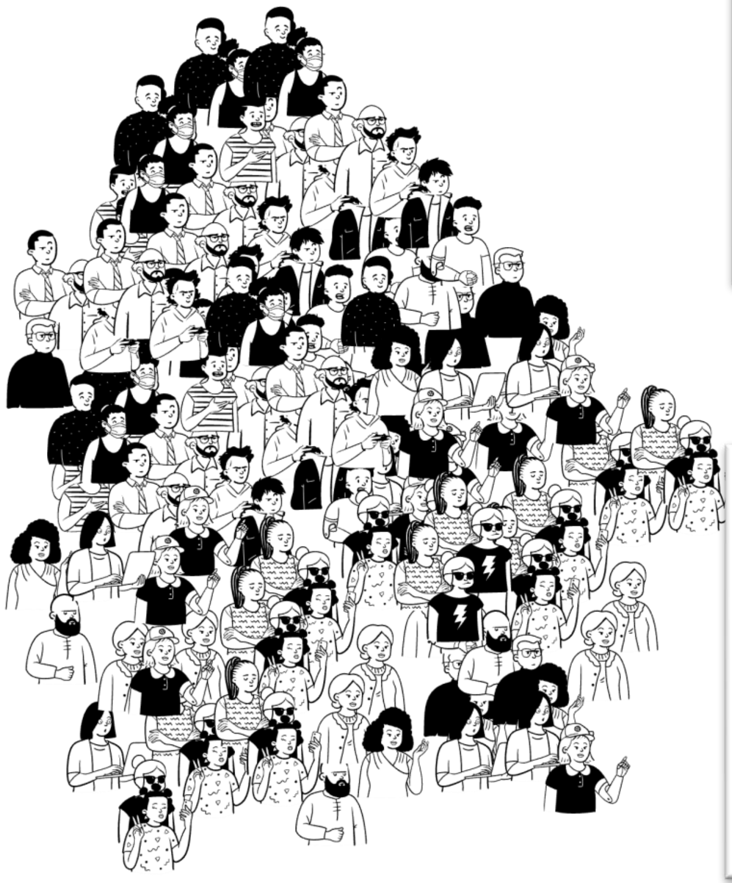
Group A

Group B

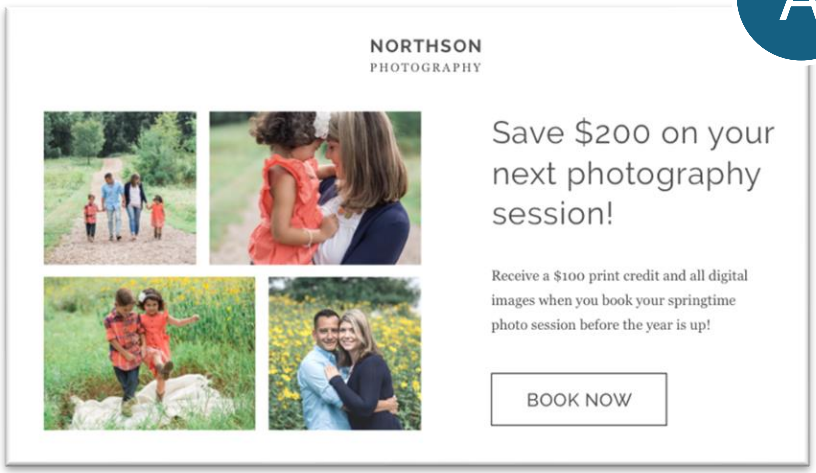
Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

A/B Testing

user base

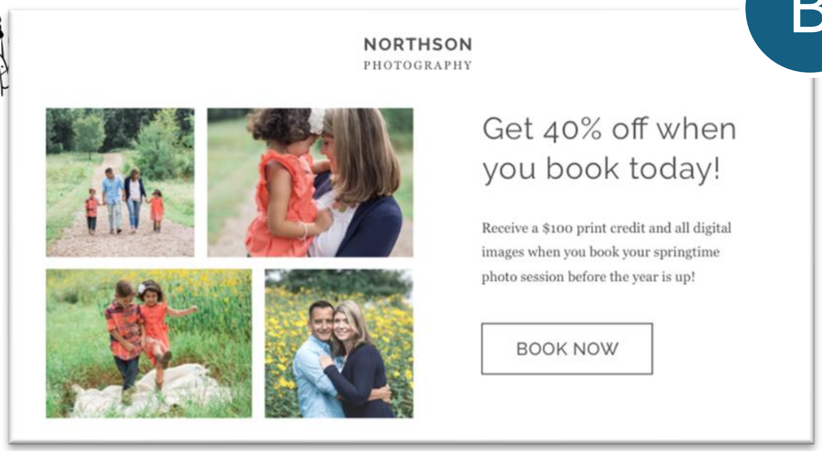


A



Group A

B

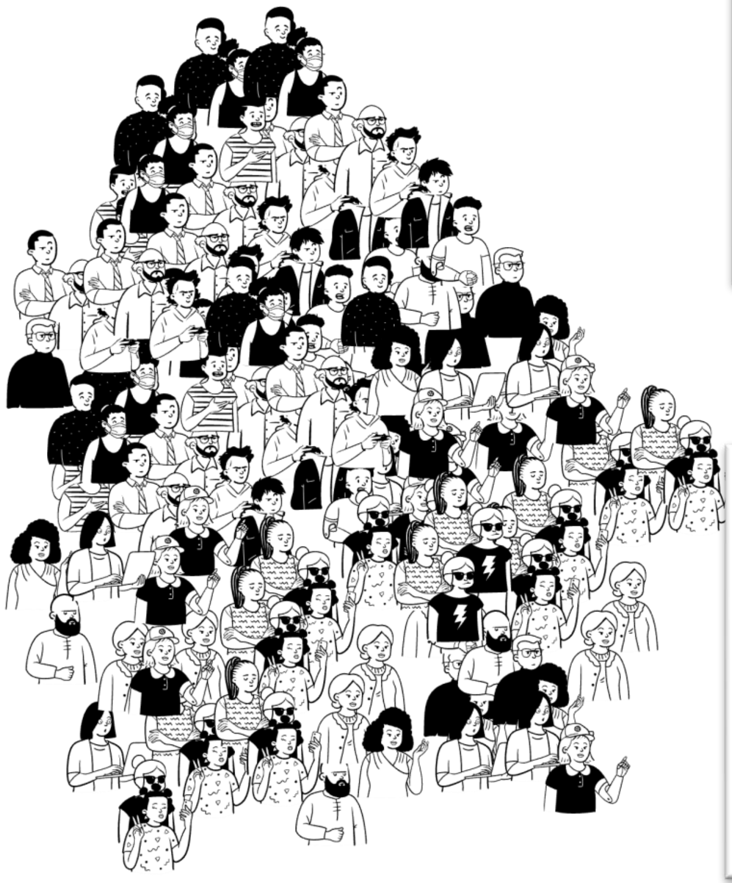


Group B

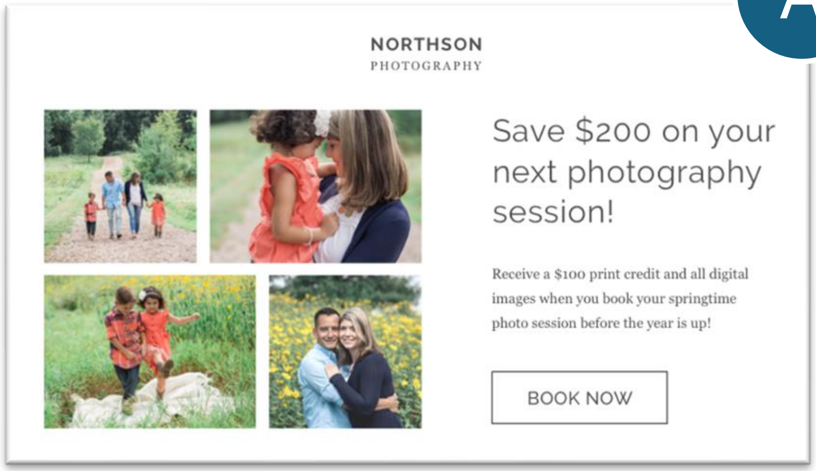
Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

A/B Testing

user base

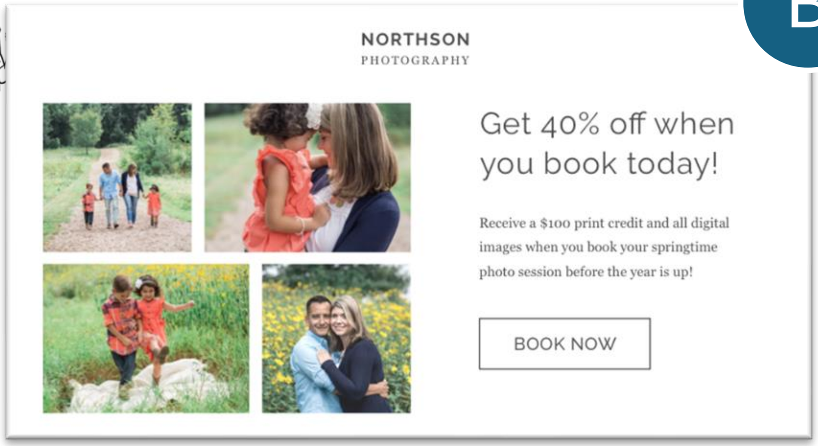


A



Group A

B

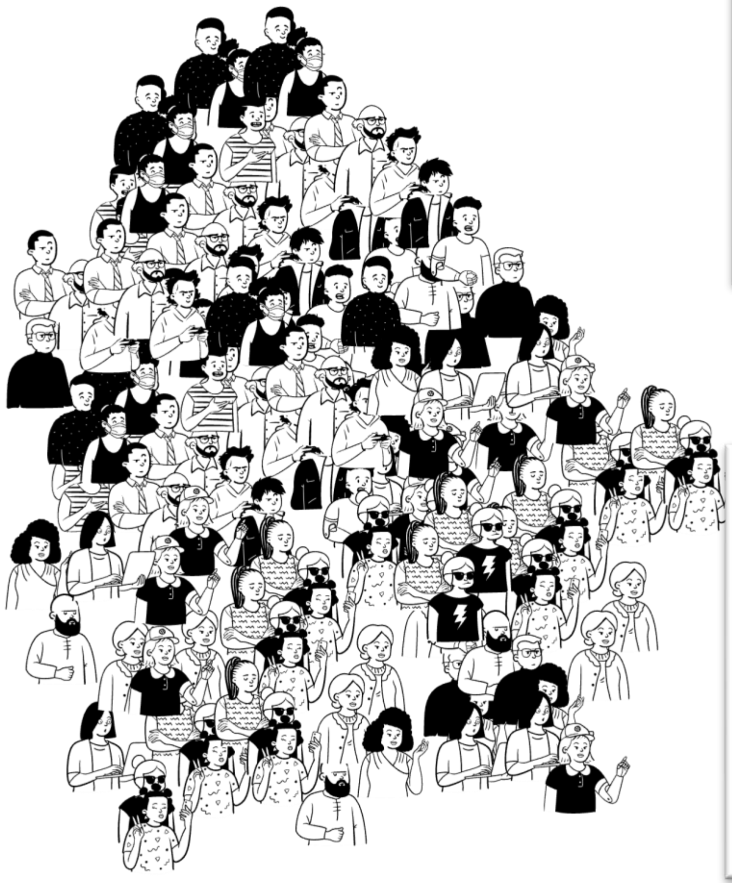


Group B

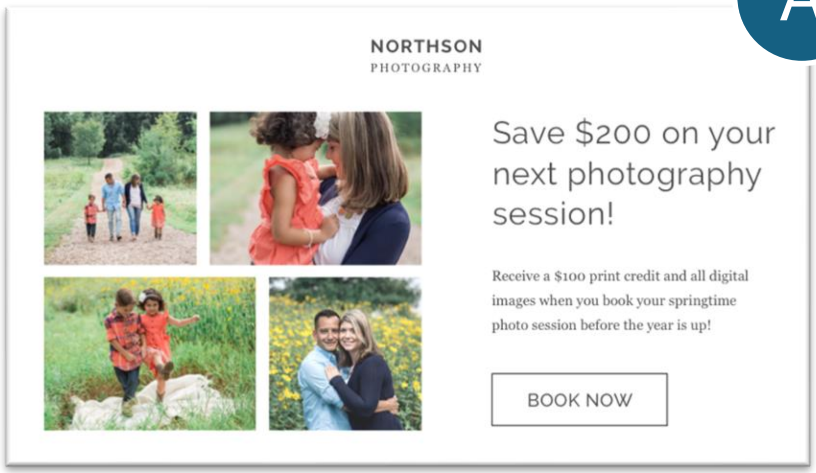
Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

A/B Testing

user base

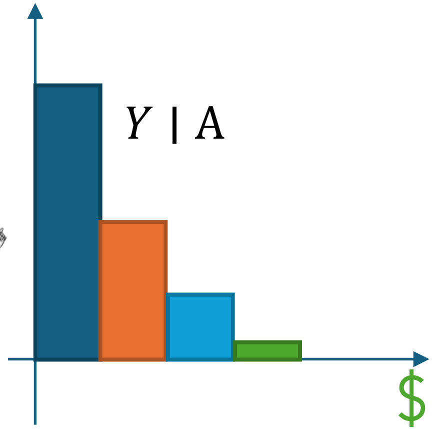


A



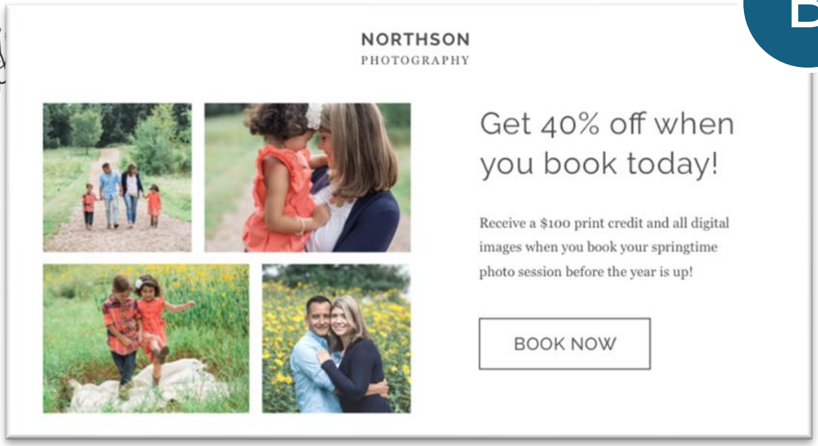
Group A

% of people



$\mu_A = 10\$$ (average spend)

B

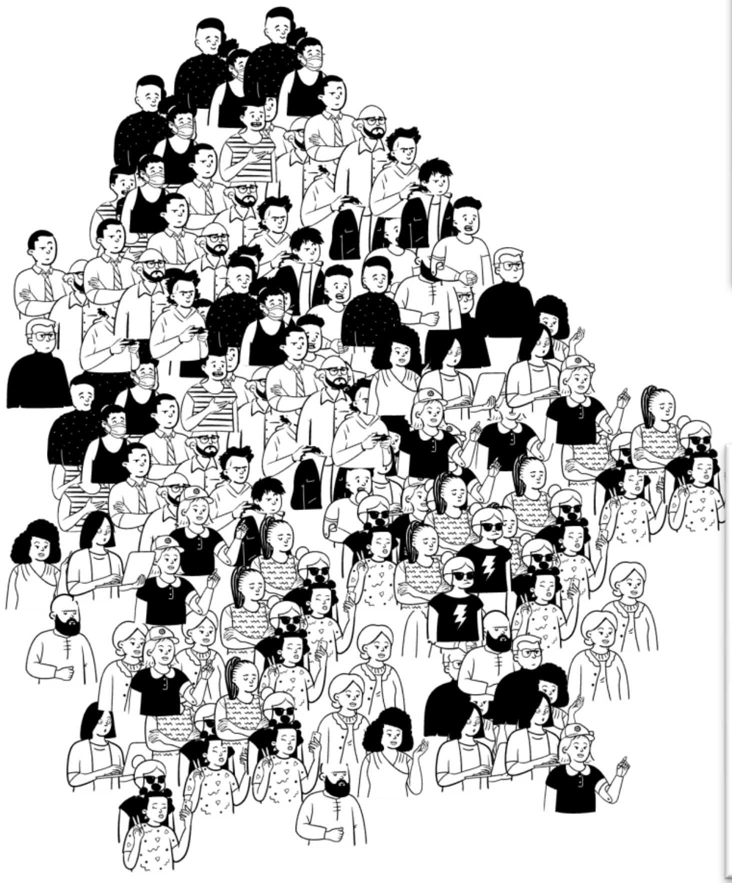


Group B

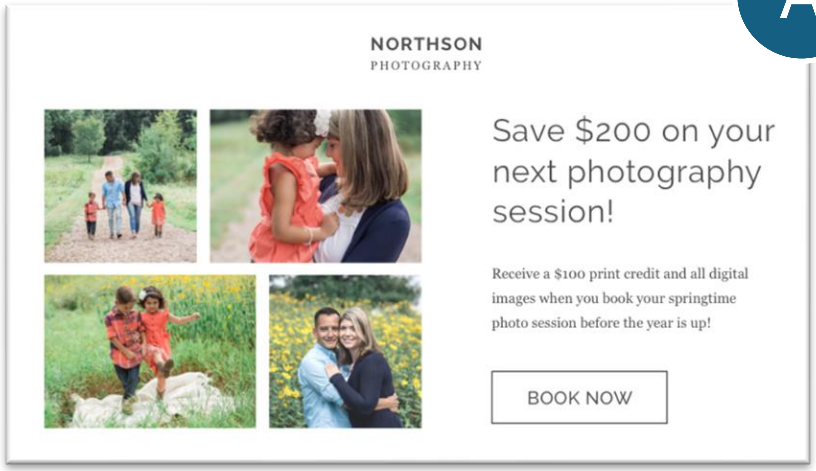
Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

A/B Testing

user base

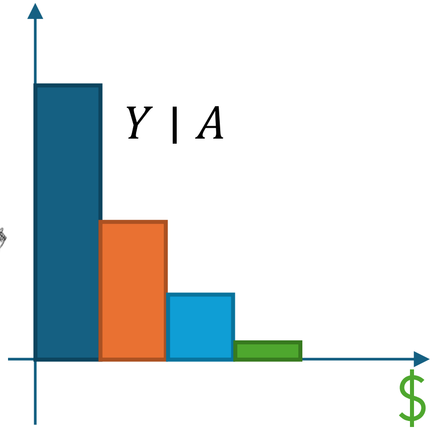


A



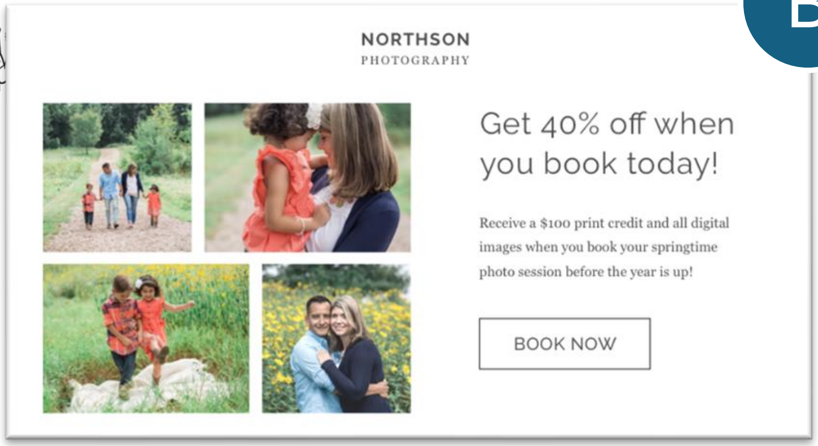
Group A

% of people



$\mu_A = 10\$$ (average spend)

B

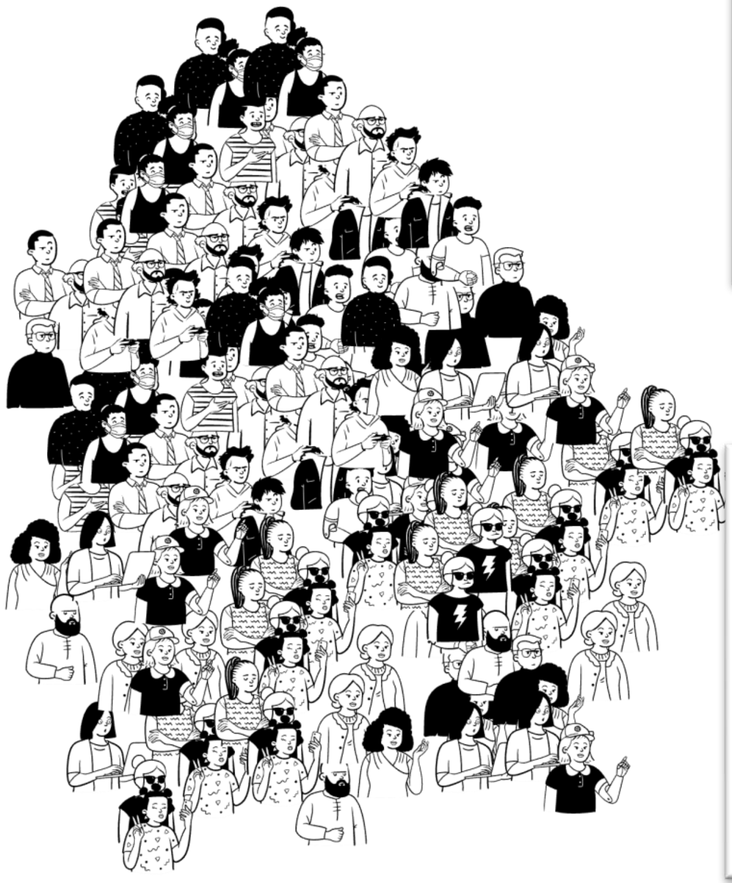


Group B

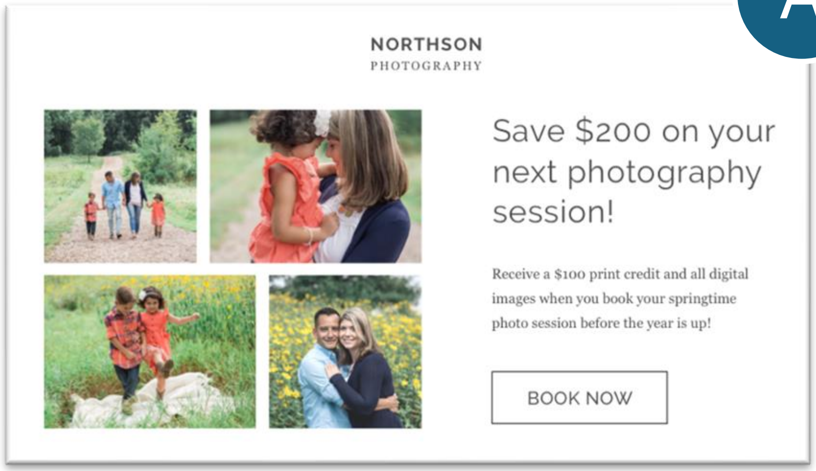
Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

A/B Testing

user base

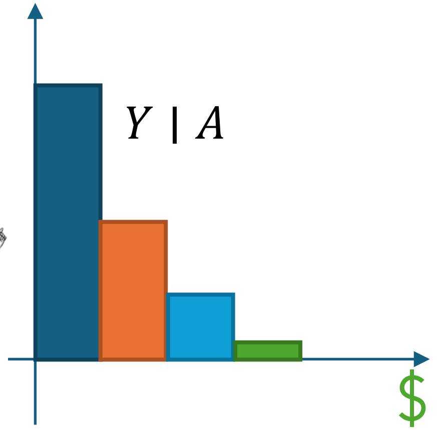


A



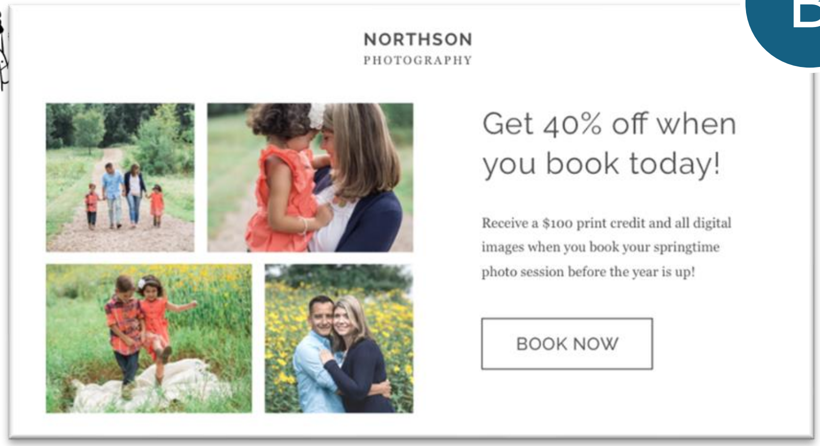
Group A

% of people



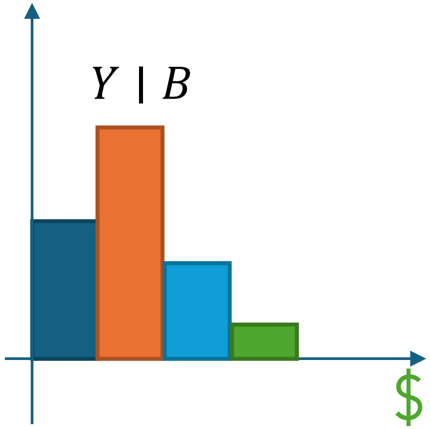
$\mu_A = 10\$$ (average spend)

B



Group B

% of people

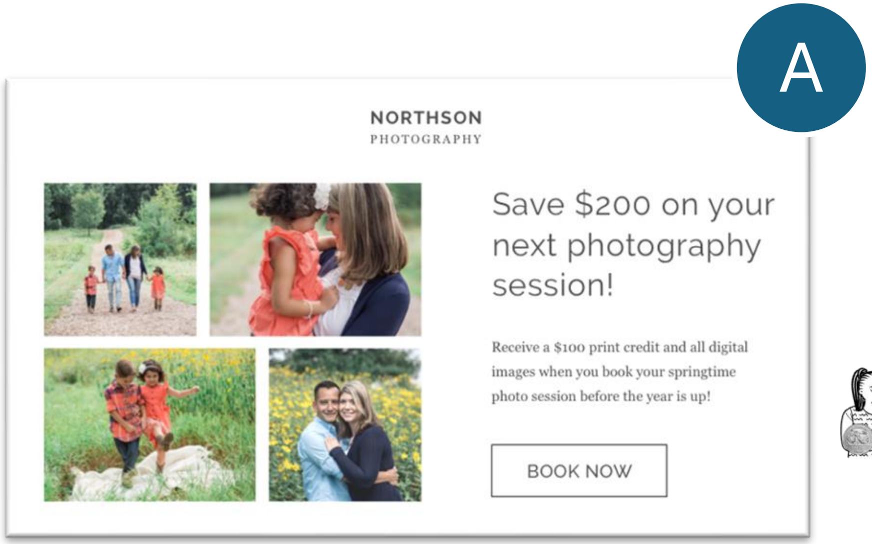


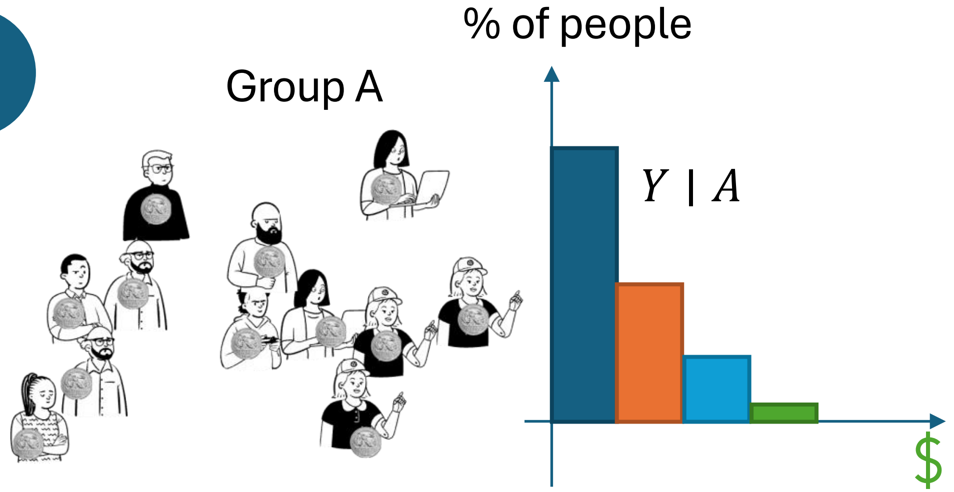
$\mu_B = 20\$$ (average spend)

Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

A/B Testing

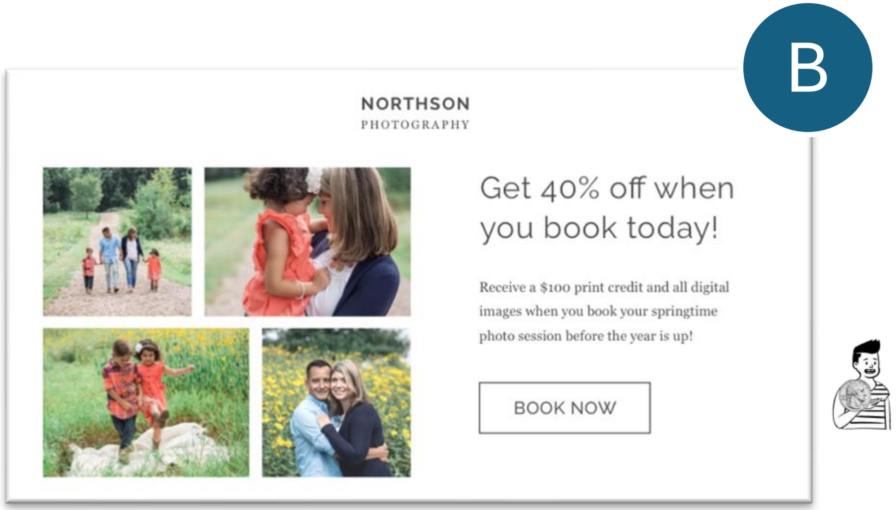
Control
Baseline
Status quo

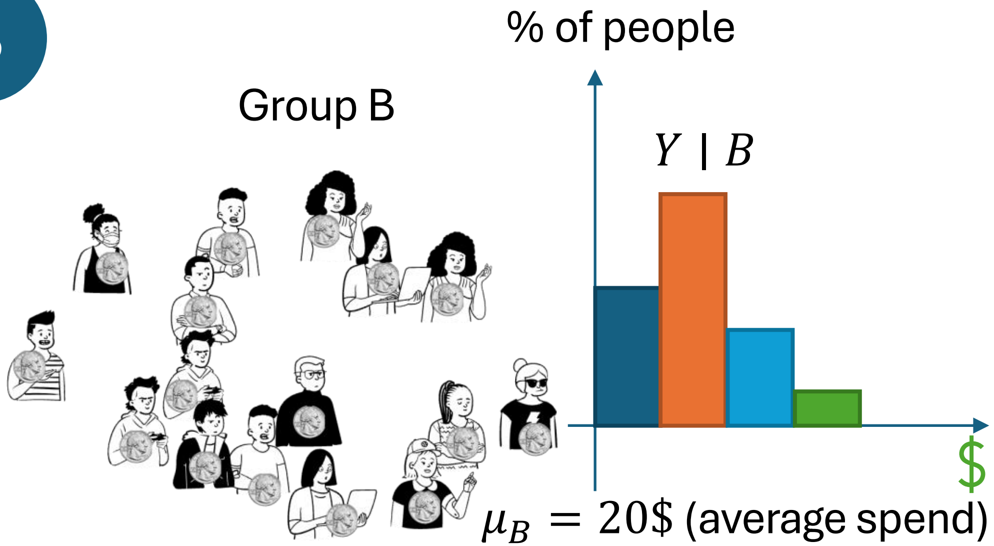




$\mu_A = 10\$$ (average spend)

Treatment
Innovation





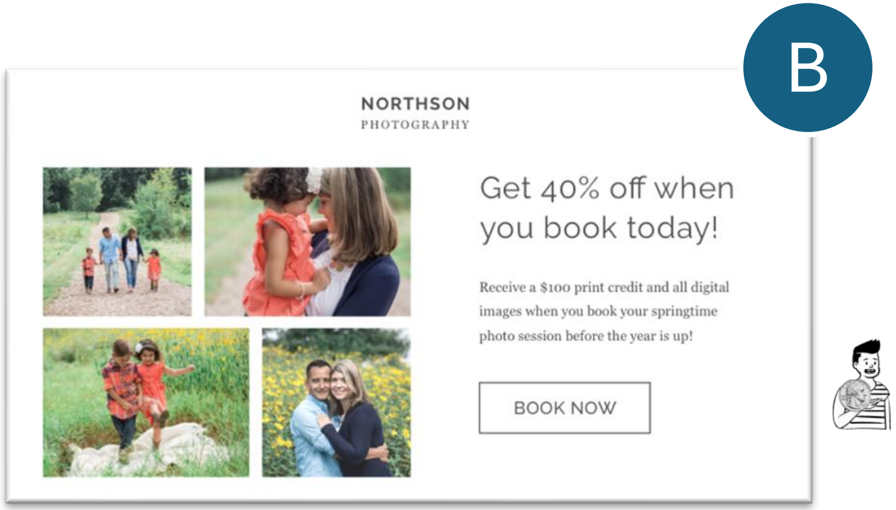
$\mu_B = 20\$$ (average spend)

Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

A/B Testing

Control
Baseline
Status quo

Treatment
Innovation



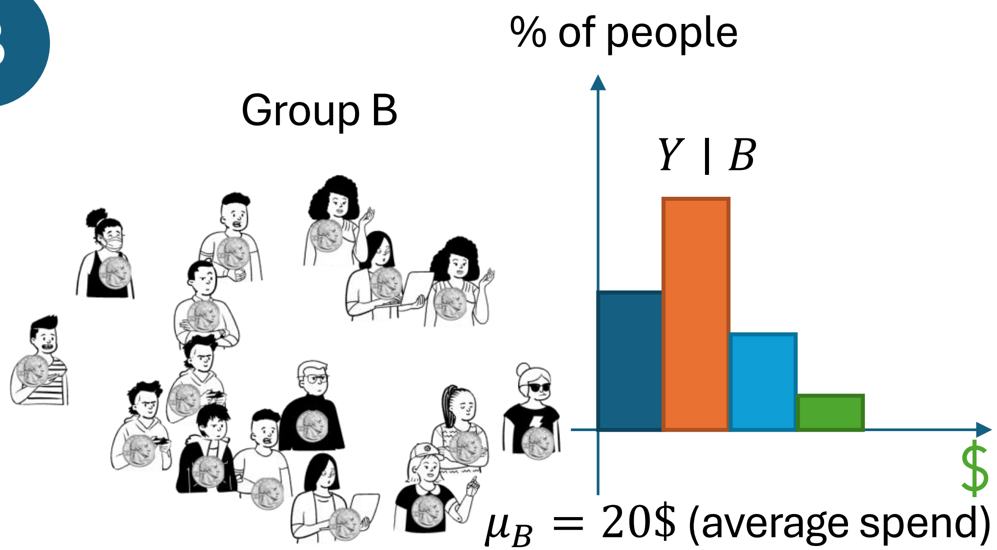
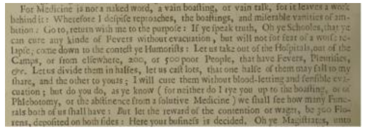


Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

A Brief History of Experimentation

1662

John Baptista Van Helmont proposes a randomized experiment for the effect of bloodletting


For Medicine is not a naked word, a vain boaſting, or vain talk, for it leaves a work behind it: Wherefore I deſpiſe reproaches, the boaſtings, and miſerable vanities of ambition: Go to, return with me to the purpoſe: If ye ſpeak truth, Oh ye Schooles, that ye can cure any kinde of Fevers without evacuation, but will not for fear of a worſe relapſe; come down to the conteſt ye Humoriſts: Let us take out of the Hoſpitals, out of the Camps, or from elſewhere, 200, or 500 poor People, that have Fevers, Pleuriſies, &c. Let us divide them in halfes, let us caſt lots, that one halfe of them may fall to my ſhare, and the other to yours; I will cure them without blood-letting and ſenſible evacuation; but do you do, as ye know (for neither do I tye you up to the boaſting, or of Phlebotomy, or the abſtinence from a ſolutive Medicine) we ſhall ſee how many Funerals both of us ſhall have: But let the reward of the contention or wager, be 300 Florens, depoſited on both ſides: Here your buſineſs is decided. Oh ye Magiſtrates, unto

1828

Louis conducts a non-randomized controlled experiment showing strong evidence against bloodletting

1923

Neyman publishes “On the Application of Probability Theory to Agricultural Experiments”

1925

Fisher publishes Statistical Methods for Research Workers

1935

Michigan Supreme Court authorizes RCTs as part of medical practice

Fisher publishes The Design of Experiments

1947

First published medical RCT in Great Britain

1962

FDA to prohibit the marketing of any new drug in the absence of “substantial evidence.”

1969

FDA: “…partially controlled studies are not acceptable evidence to support claims of effectiveness”

2000-

Abhijit Banerjee, Esther Duflo, and Michael Kremer, popularize the use of RCTs in economics (esp. for policy problems related to poverty).

10s of thousands of RCTs run annually by companies like Airbnb, Amazon, Booking.com eBay, Facebook, Google, LinkedIn, Lyft, Microsoft, Netflix, Twitter, Uber, Yahoo! and Yandex

Credit: Ron Kohavi, [History of Controlled Experiments]

RCTs are the gold standard for measuring the “causal effect” of a “treatment” on an “outcome”

Causality

Causal Effect

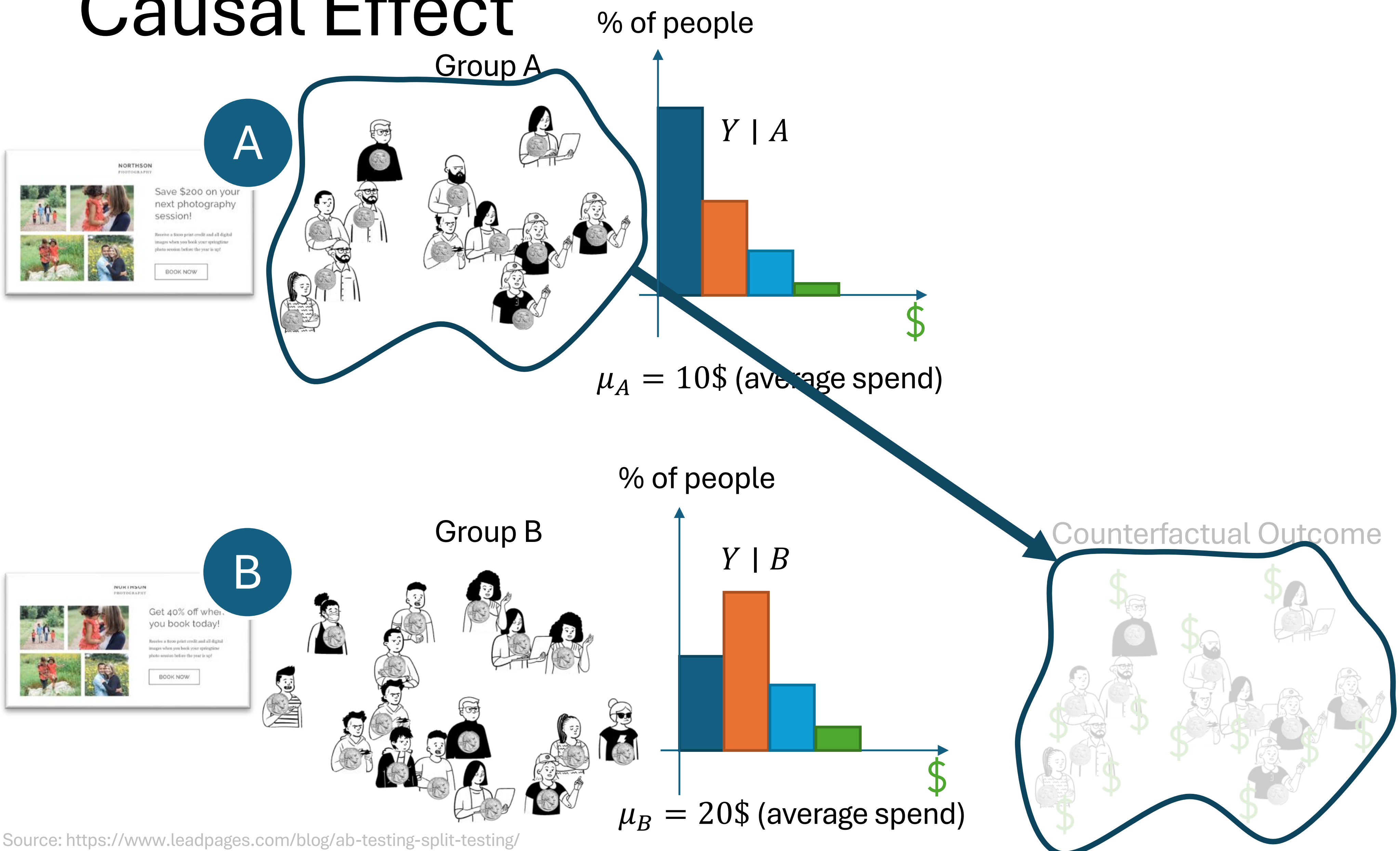


Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

Causal Effect

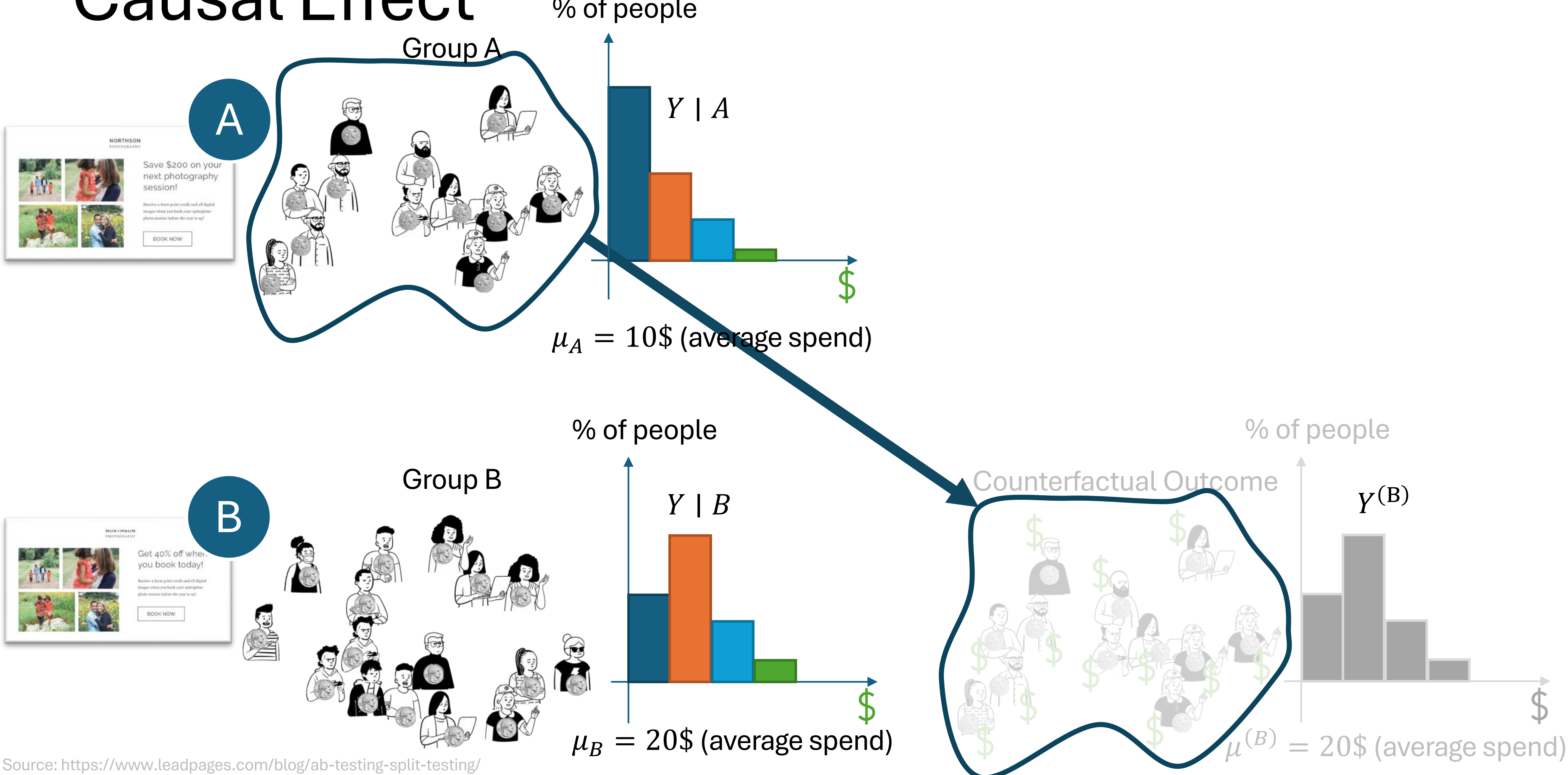


Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

Causal Effect

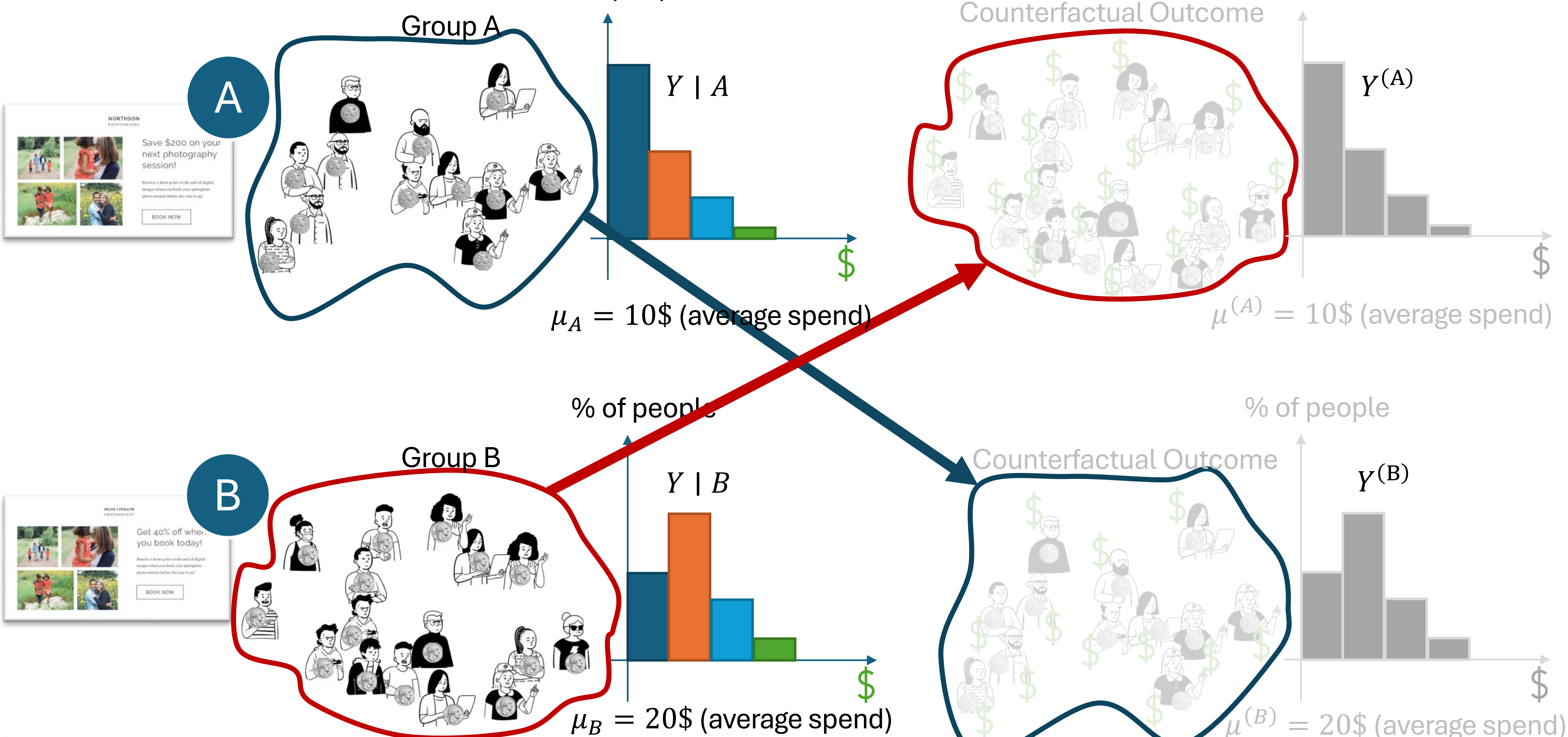


Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

Randomization implies

$$Y|A \sim Y^{(\mathrm{A})}$$

$$Y|B \sim Y^{(\mathrm{B})}$$

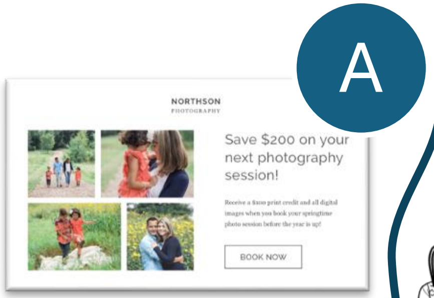
Aggregate differences between groups

$$\mathrm{E}[Y|A] - \mathrm{E}[\mathrm{Y}|B]$$

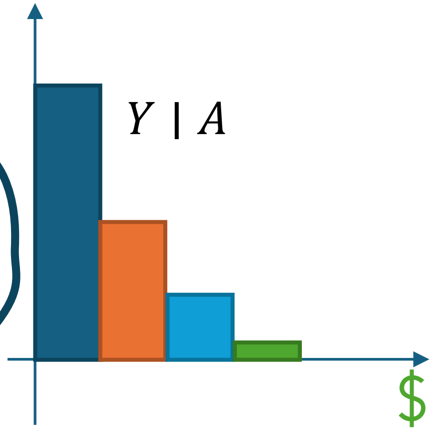
Equal aggregate causal effects

$$E\left[Y^{(\mathrm{A})} - Y^{(\mathrm{B})}\right]$$

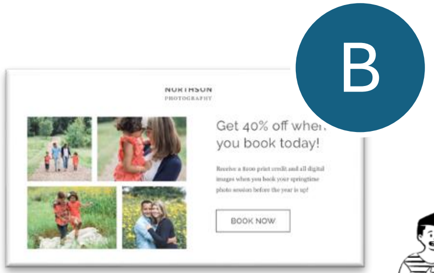
Historical Data



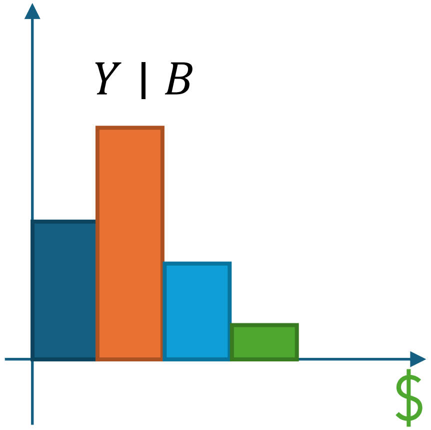
Group A



$\mu_A = 10\$$ (average spend)



Group B



$\mu_B = 20\$$ (average spend)

Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

Historical Data

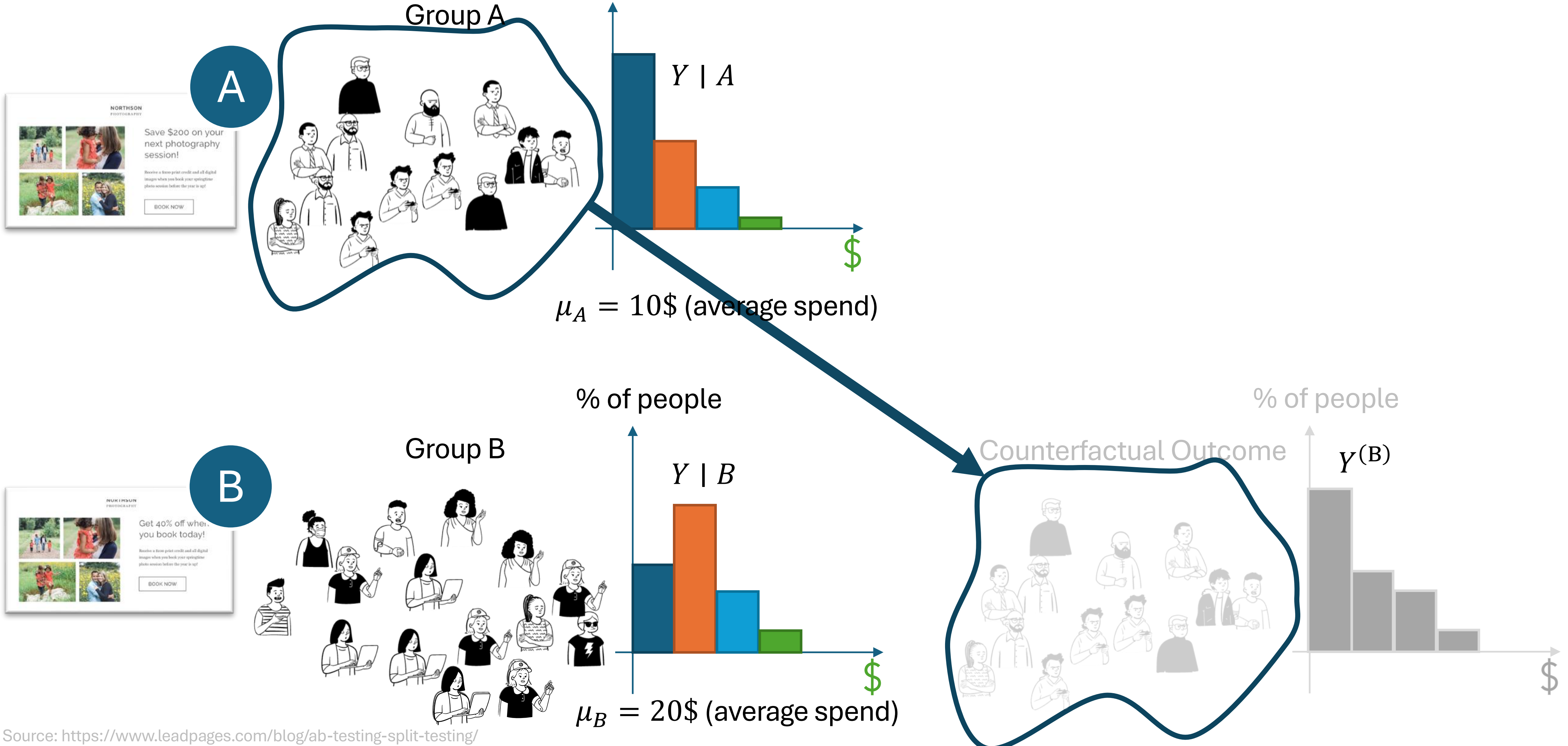


Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

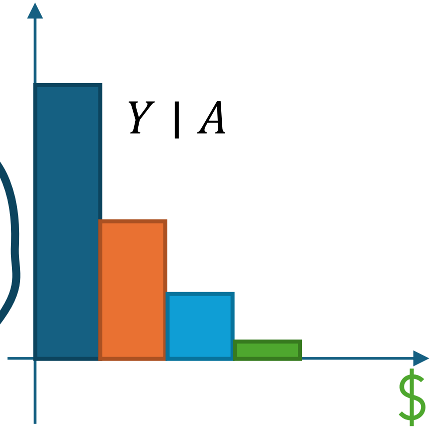
Historical Data

No cloud specialist

A

Group A

% of companies



$\mu_A = 10\$$ (average spend)

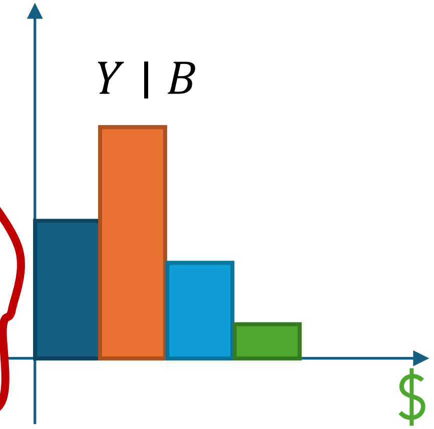
Assign Cloud Specialist

B

Group B

Fortune500

% of companies



$\mu_B = 20\$$ (average spend)

Image source: https://www.linkedin.com/pulse/why-90-fortune-500-companies-now-using-microsoft-cloud-natasha-spurr/

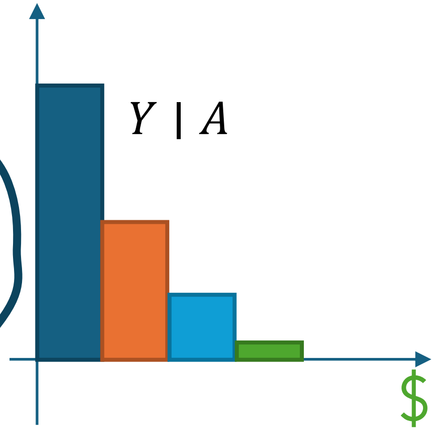
Historical Data

No cloud specialist

A

Group A

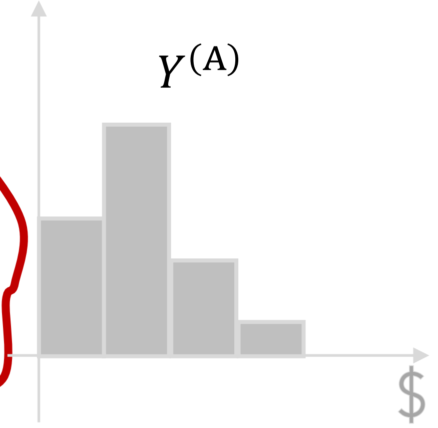
% of companies



$\mu_A = 10\$$ (average spend)

Counterfactual outcome

% of companies



$\mu^{(A)} = 20\$$ (average spend)

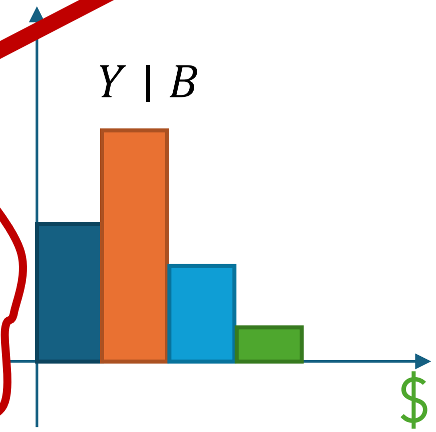
Assign Cloud Specialist

B

Group B

Fortune500

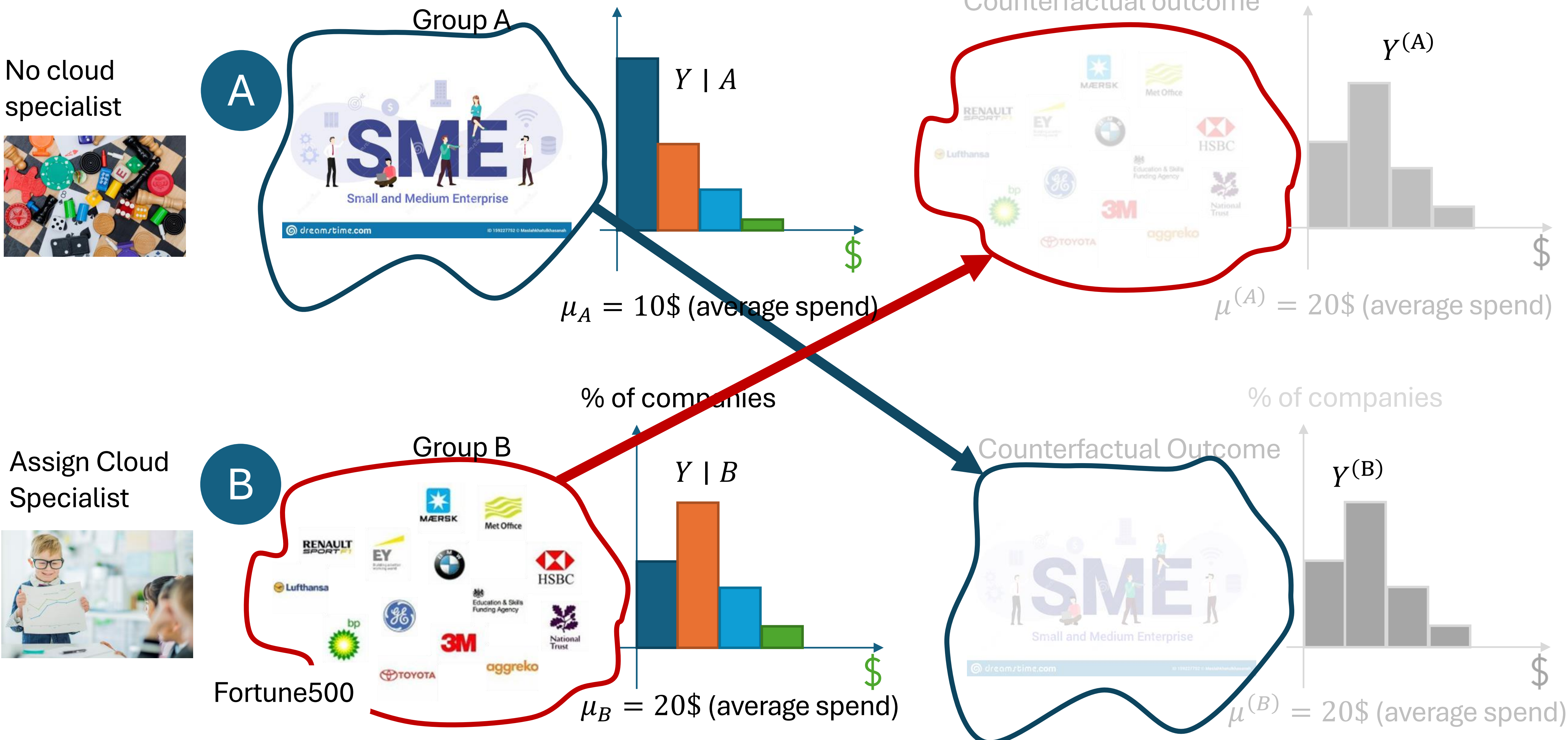
% of companies



$\mu_B = 20\$$ (average spend)

Image source: https://www.linkedin.com/pulse/why-90-fortune-500-companies-now-using-microsoft-cloud-natasha-spurr/

Historical Data

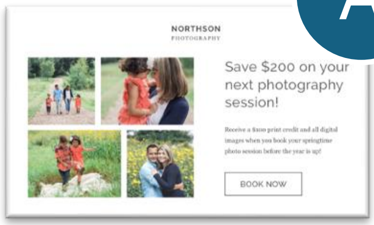


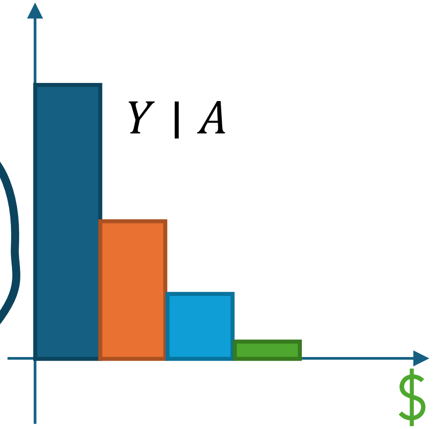
Statistics

Statistics

Sampled A

A

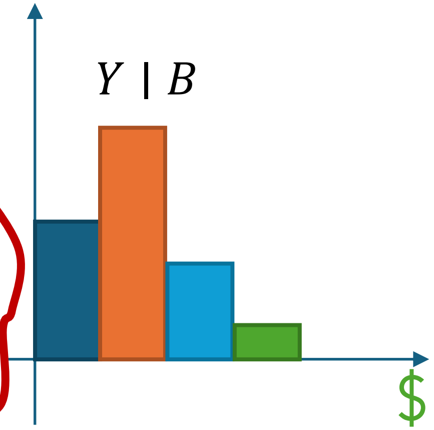




$\mu_A = 10\$$ (average spend)

Sampled B

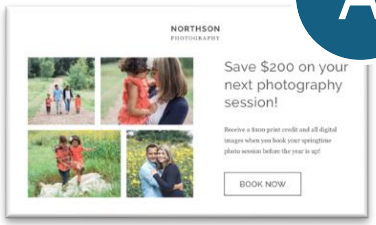
B



$\mu_B = 20\$$ (average spend)

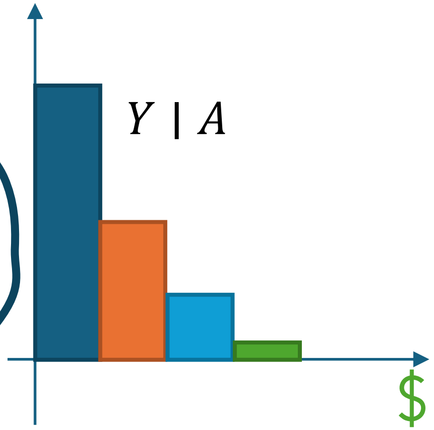
Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

Statistics



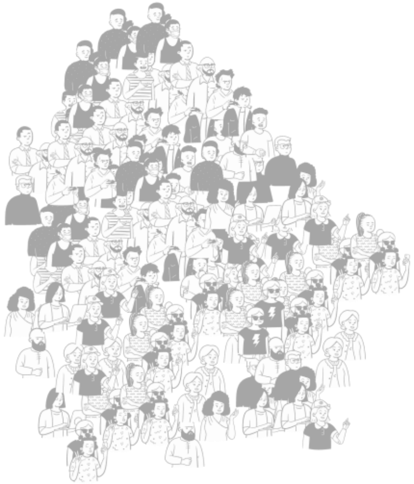
Sampled A

% of people

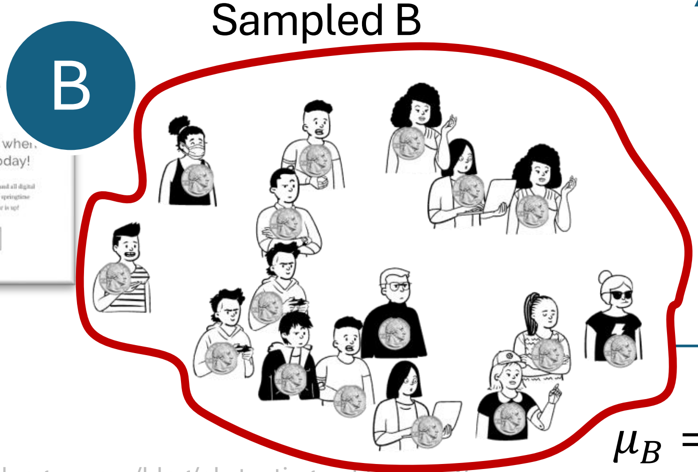


$\mu_A = 10\$$ (average spend)

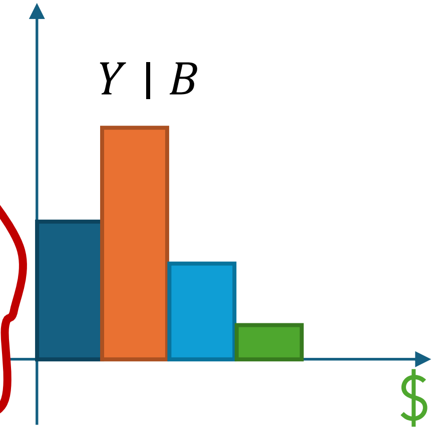
all user base



Sampled B



% of people



$\mu_B = 20\$$ (average spend)

all user base

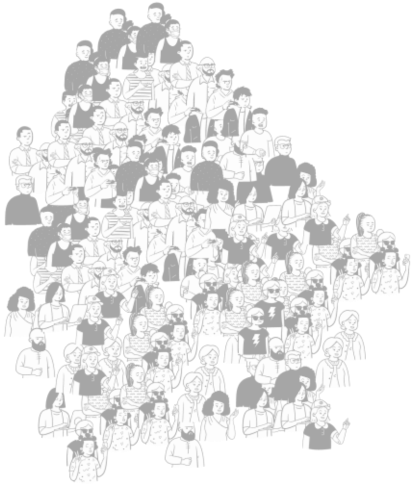
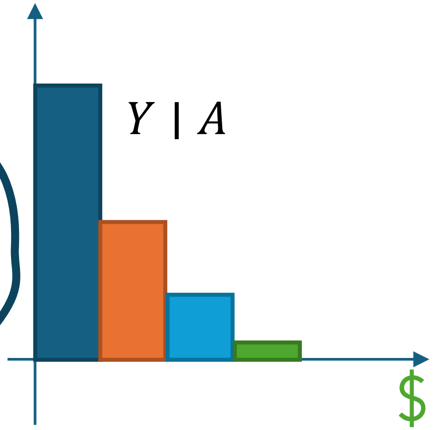


Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

Statistics

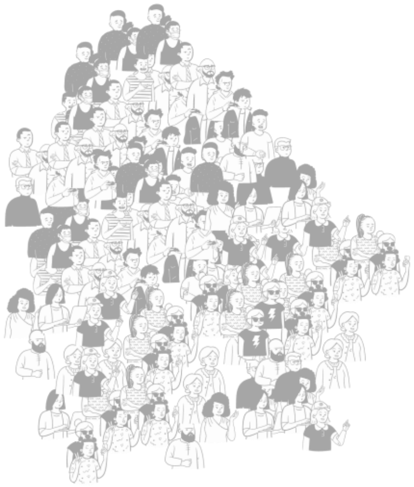
Sampled A

% of people

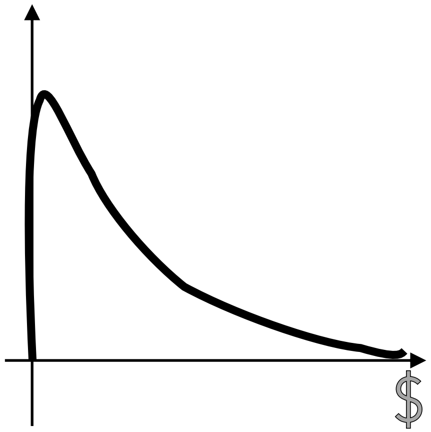


$\mu_A = 10\$$ (average spend)

all user base

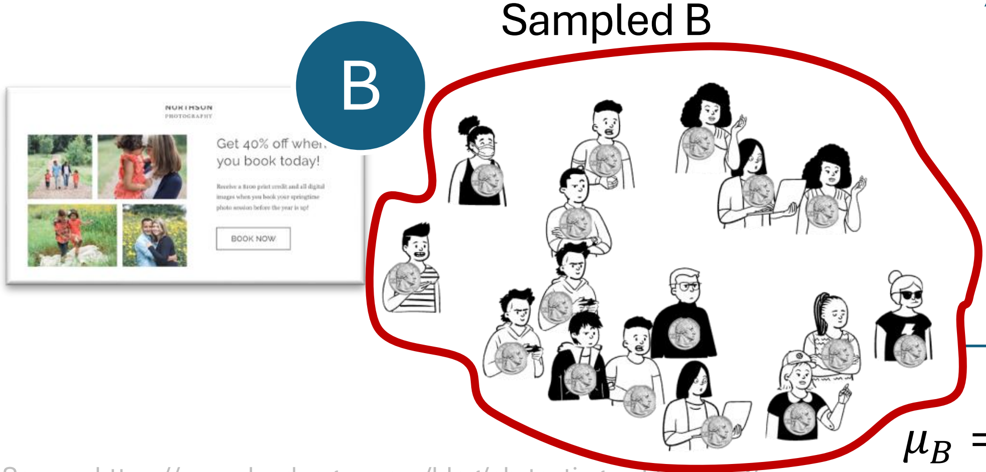


% of people

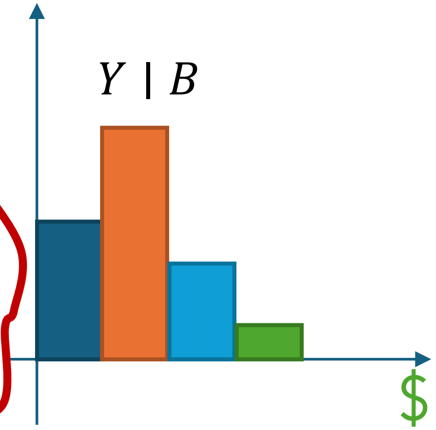


$\mu_A^0 = 12\$$ (average spend)

Sampled B

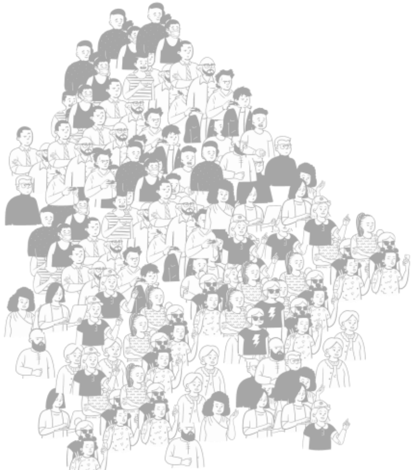


% of people

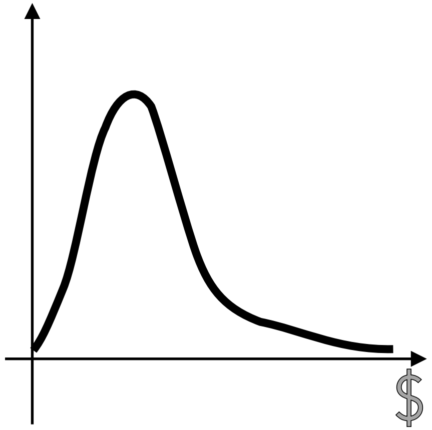


$\mu_B = 20\$$ (average spend)

all user base



% of people



$\mu_B^0 = 24\$$ (average spend)

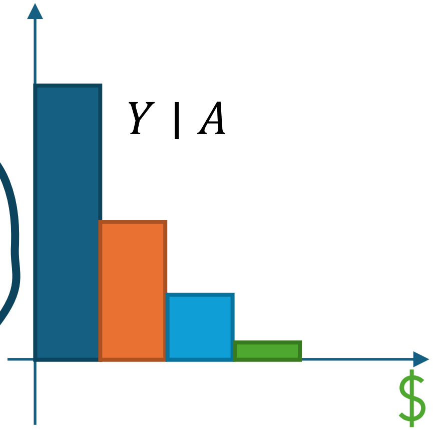
Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

Statistics

with probability 95%

Sampled A

% of people



$\mu_A = 10\$$ (average spend)

population mean

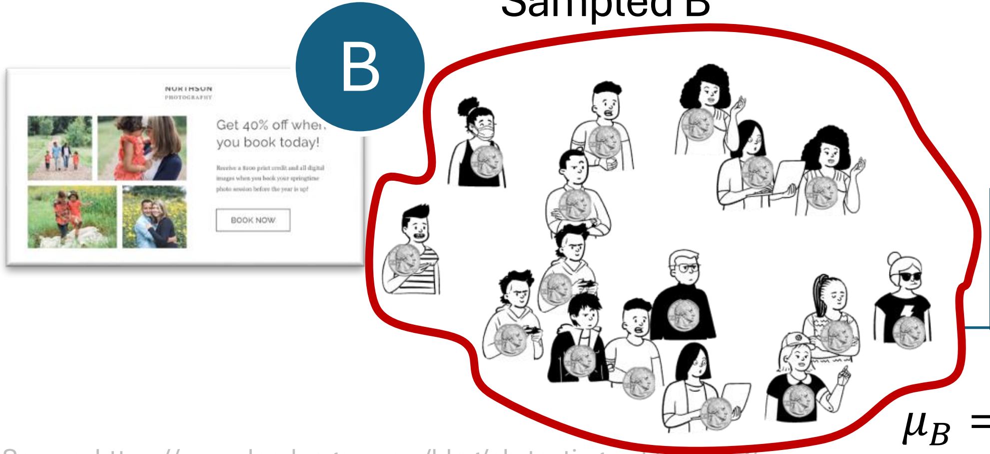
$$\boxed{\mu_A^0} \approx \boxed{\mu_A} \pm \boxed{2\sqrt{\frac{\mathrm{Var}(Y|A)}{N_\mathrm{A}}}}$$

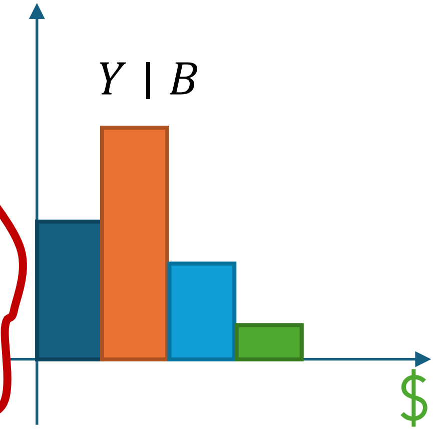
sample mean

sampling error

% of people

Sampled B





$\mu_B = 20\$$ (average spend)

population mean

$$\boxed{\mu_B^0} \approx \boxed{\mu_B} \pm \boxed{2\sqrt{\frac{\mathrm{Var}(Y|B)}{N_\mathrm{B}}}}$$

sample mean

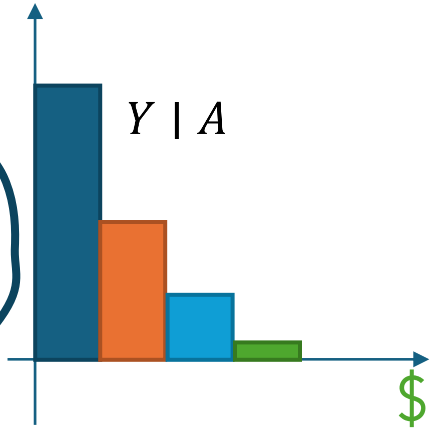
sampling error

Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

Statistics

with probability 95%

Sampled A

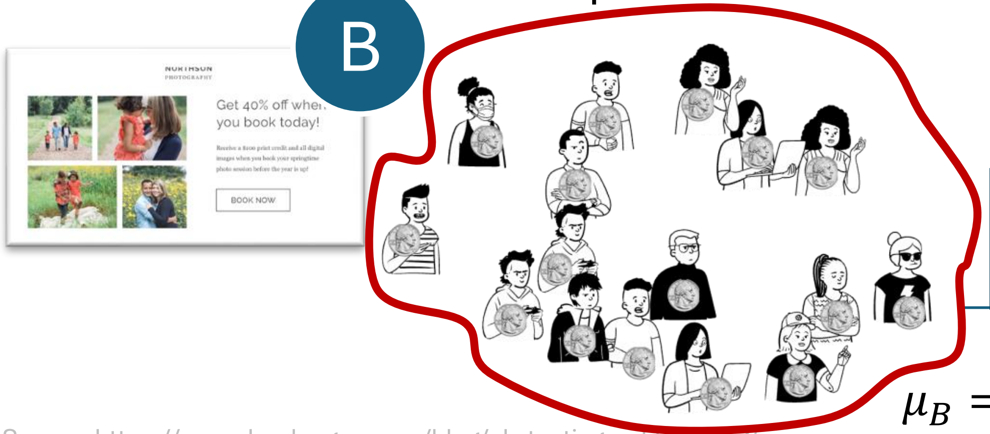


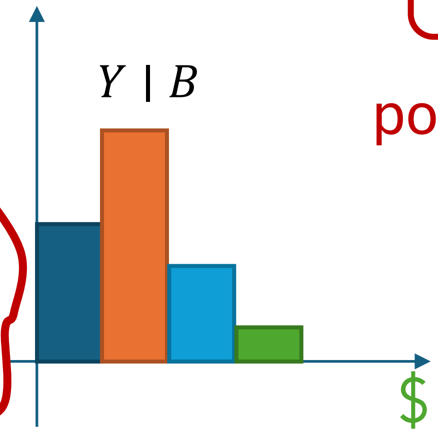
$\mu_A = 10\$$ (average spend)

$$\boxed{\mu_B^0 - \mu_A^0} \approx \boxed{\mu_B - \mu_A} \pm \boxed{2\sqrt{\frac{\mathrm{Var}(Y|A)}{N_\mathrm{A}} + \frac{\mathrm{Var}(Y|B)}{N_\mathrm{B}}}}$$

population effect | sample effect | sampling error

Sampled B





$\mu_B = 20\$$ (average spend)

Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

Experimentation Platform

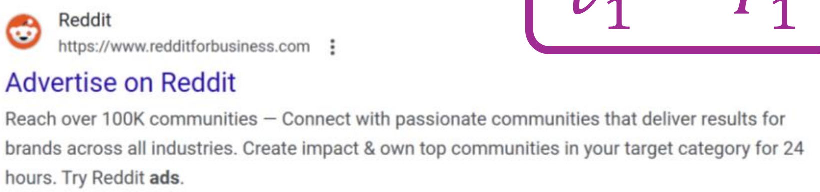
Experimentation Platform (ExP) is a team of 60+ Data Scientists, Software Engineers and Program Managers. Our mission is to accelerate innovation through trustworthy experimentation. Most major products such as Bing, Cortana, Edge, Exchange, Identity, MSN, Office client, Office online, Photos, Skype, Speech, Store, Teams, Visual Studio Code, Windows, Xbox use our platform ExP to run trustworthy Online Controlled Experiments – aka A/B tests.

Interference!
The Big Challenge of A/B Testing in Markets and Platforms

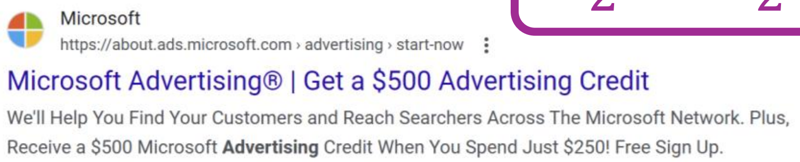
Interference

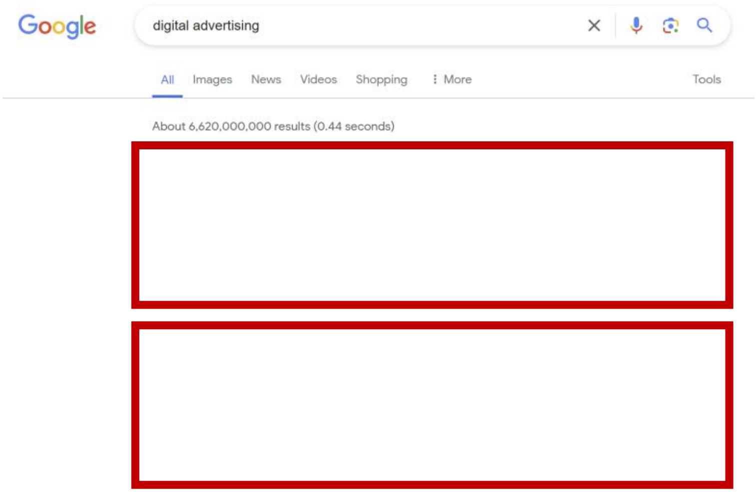
- Social Network interference
- Equilibrium effects
- Stateful systems and time effects

$v_1 \sim F_1$

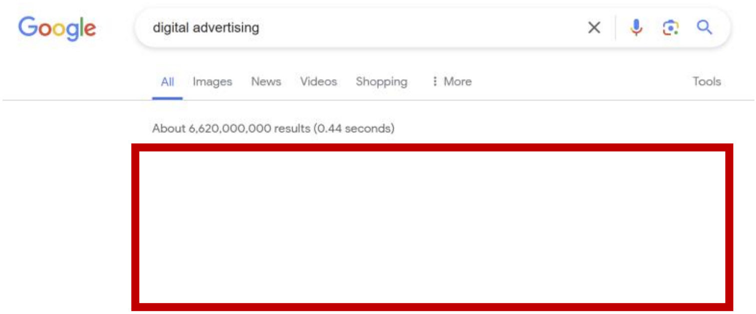


$v_2 \sim F_2$

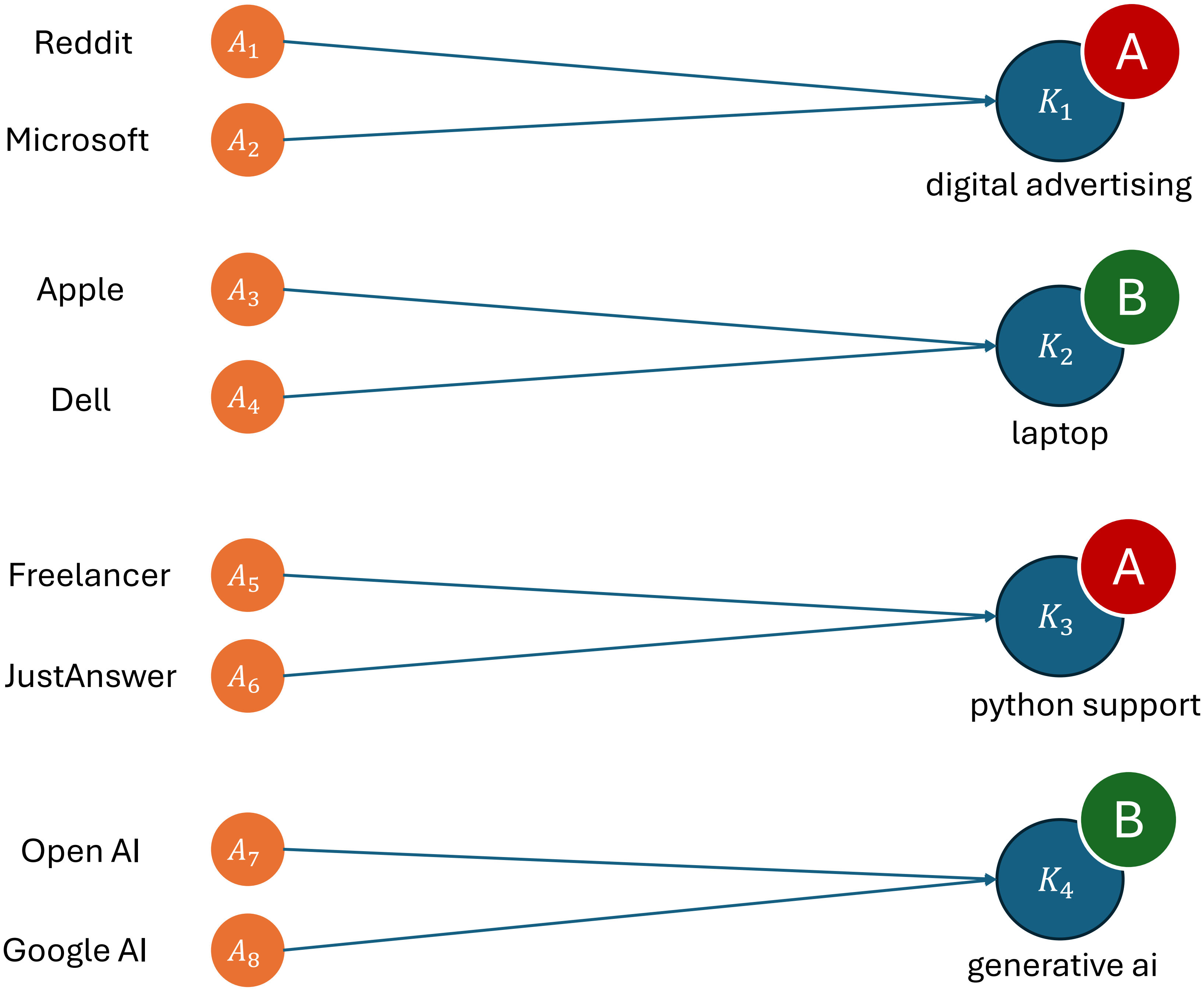


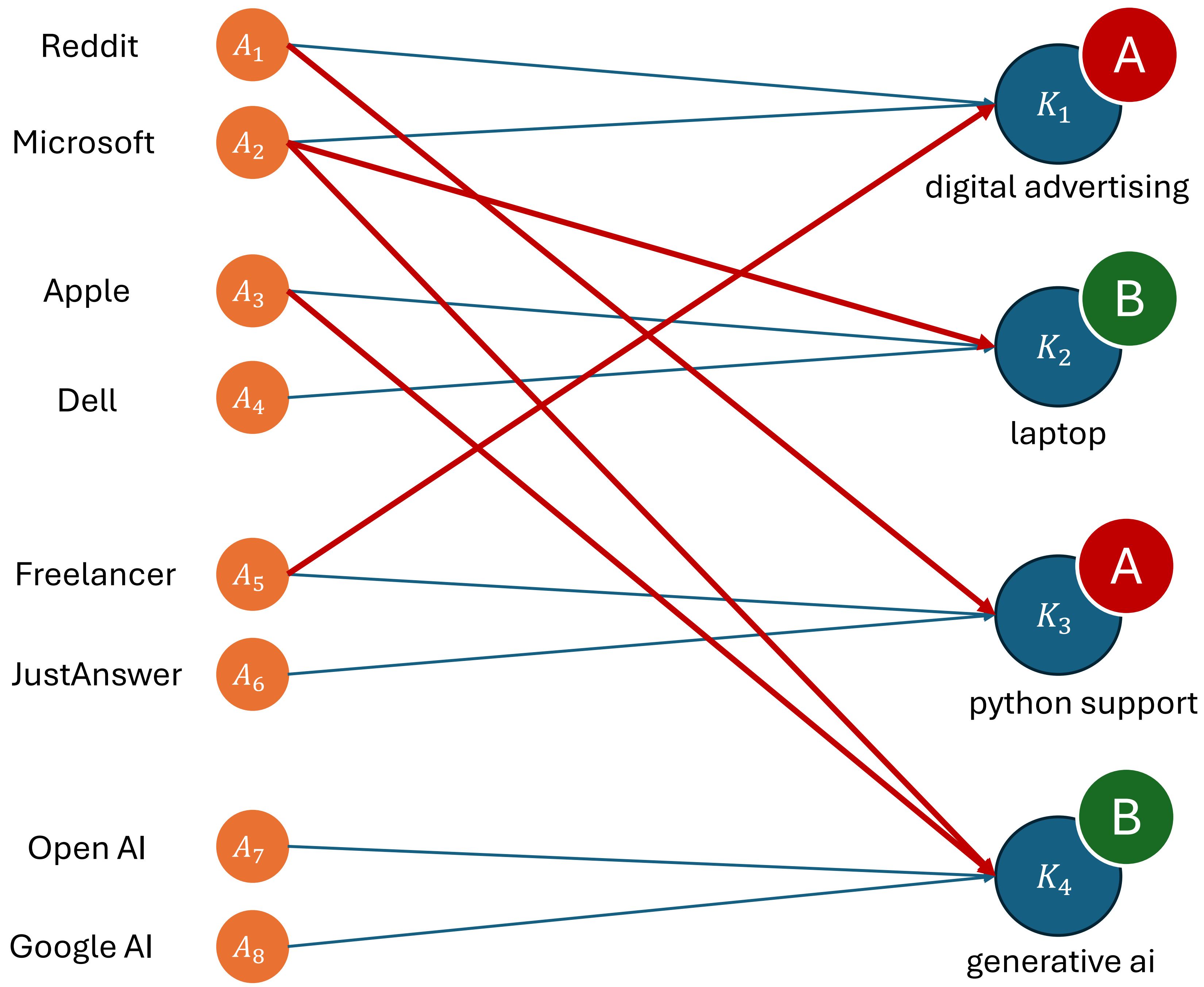


A



B


Reddit
A_1
Microsoft
A_2
Apple
A_3
Dell
A_4
Freelancer
A_5
JustAnswer
A_6
Open AI
A_7
Google AI
A_8
A
K_1
digital advertising
B
K_2
laptop
A
K_3
python support
B
K_4
generative ai



Reddit
Microsoft
Apple
Dell
Freelancer
JustAnswer
Open AI
Google AI
A_1
A_2
A_3
A_4
A_5
A_6
A_7
A_8
K_1
A
digital advertising
K_2
B
laptop
K_3
A
python support
K_4
B
generative ai

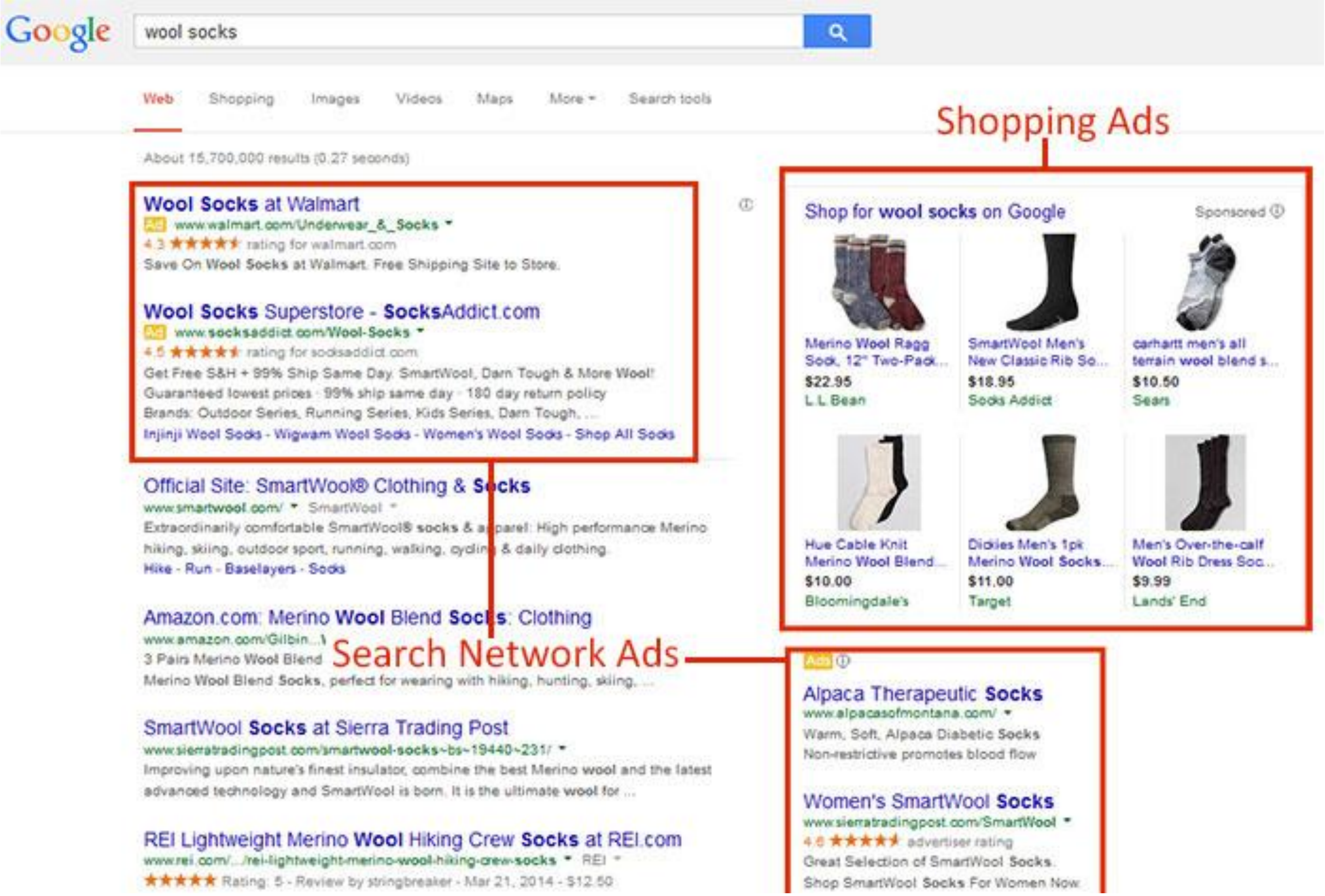
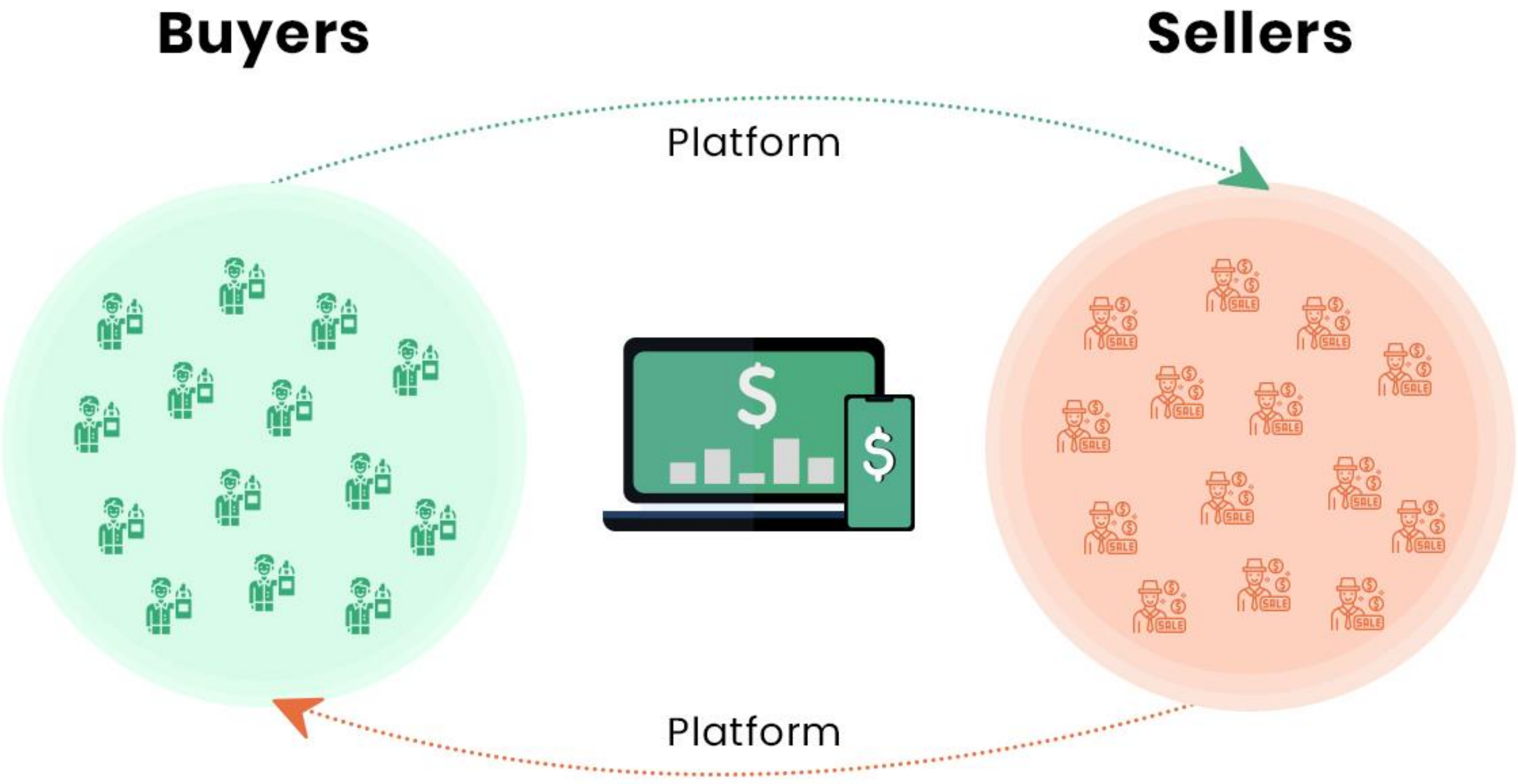
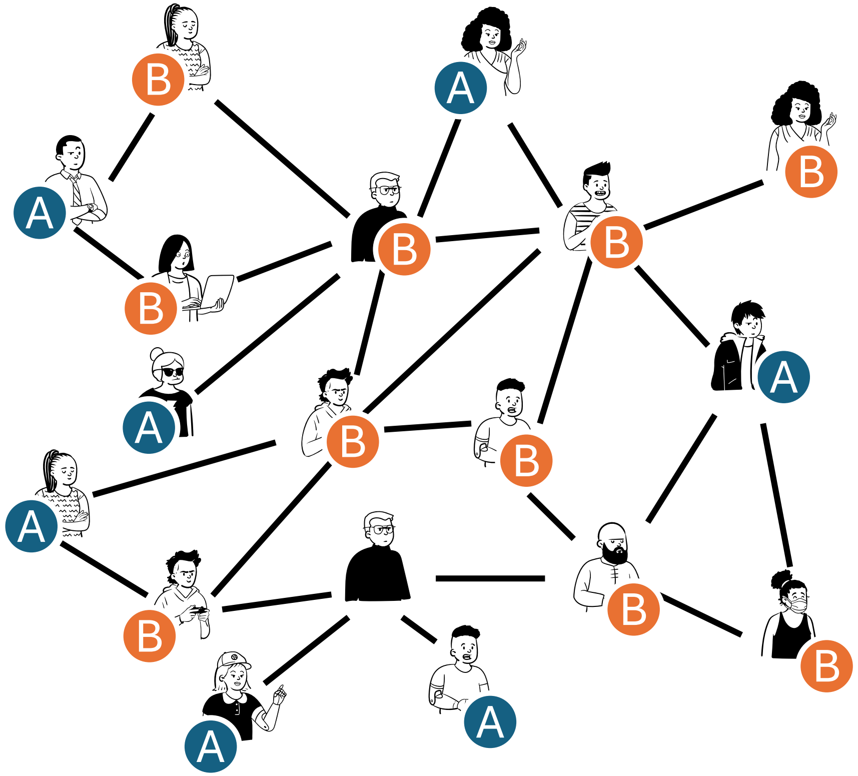


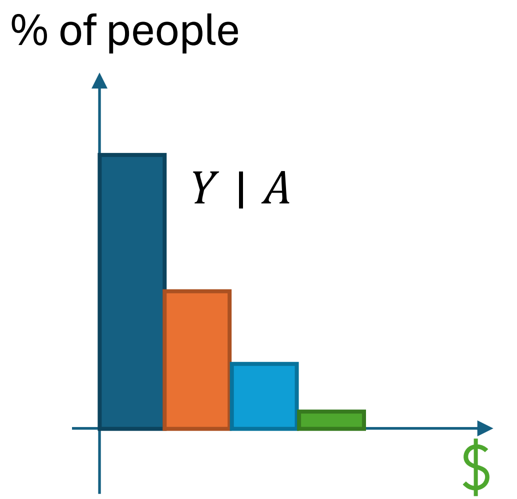
Image source: https://googleadsstrategy.com/google-adwords-search-network-vs-display-network/

Two-Sided Matching Markets

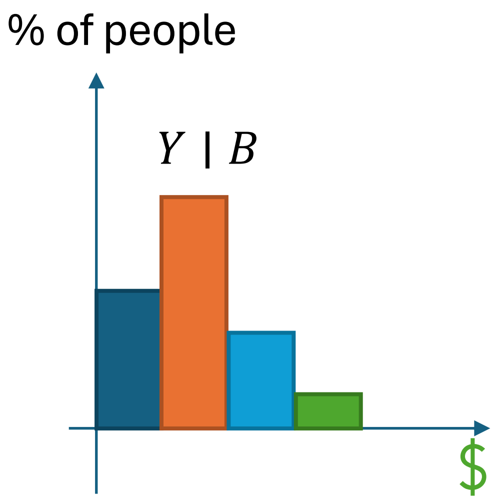


Social Network Interference



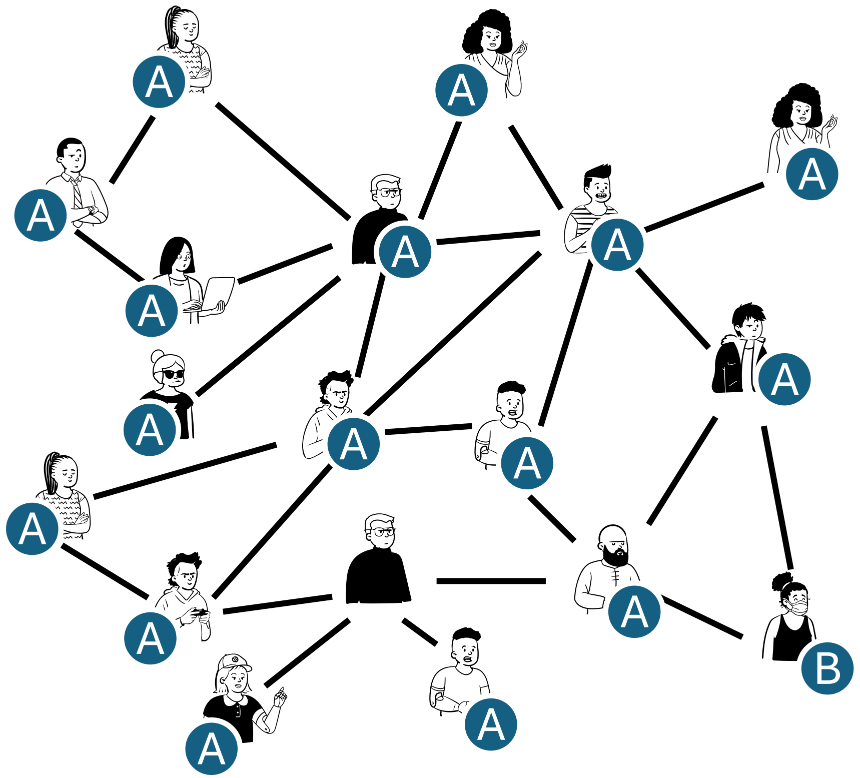


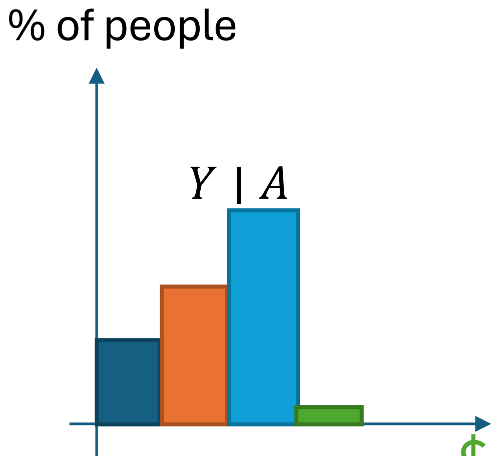
$\mu_A = 10\$$ (average spend)



$\mu_B = 20\$$ (average spend)

Counterfactual world





$\mu_A = 30\$$ (average spend)

Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

IKEA
@IKEA

Big challenges call for big solutions. Tune in to #OHOP21 on 9 November to hear thinkers, doers and leaders discuss the global response to climate change. Watch the event, get inspired and discover how we can take action ➡️ingka.com/one-home-one-p...
#AssembleABetterFuture #COP26

Source: @IKEA

IKEA
@IKEA

The moment is now. Climate action can't wait any longer. Join global thinkers, doers & leaders at #OHOP21 on 9 Nov – where they'll discuss the need for urgent change & action to help create a better future. Learn more: ingka.com/one-home-one-p...
#COP26 #AssembleABetterFuture

Source: @IKEA

Image source: https://blog.hootsuite.com/social-media-ab-testing/

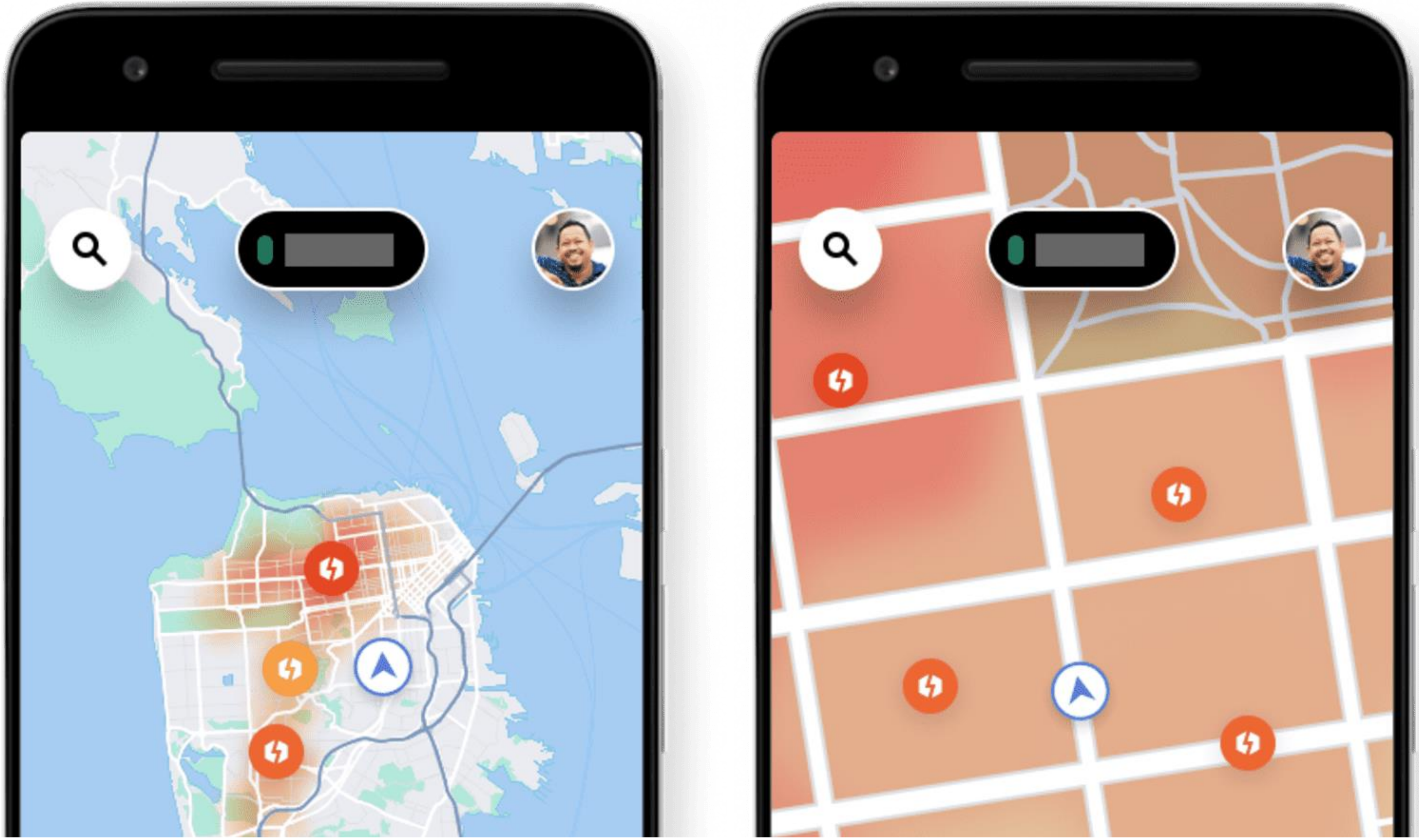
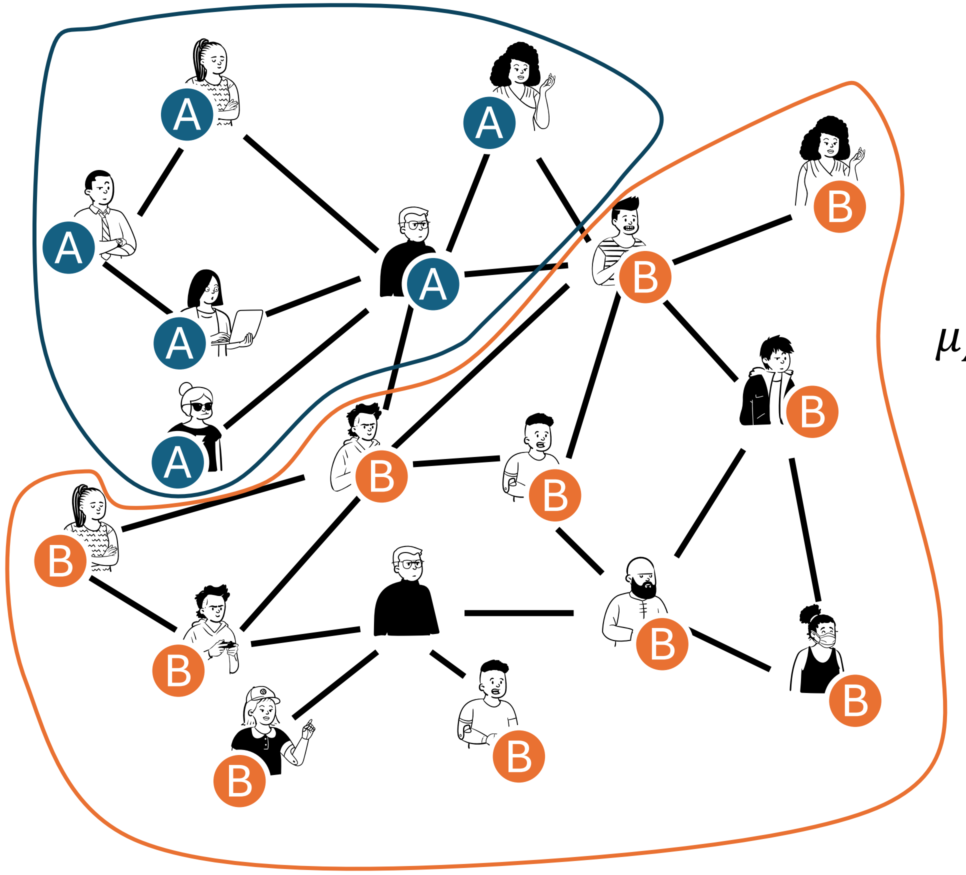
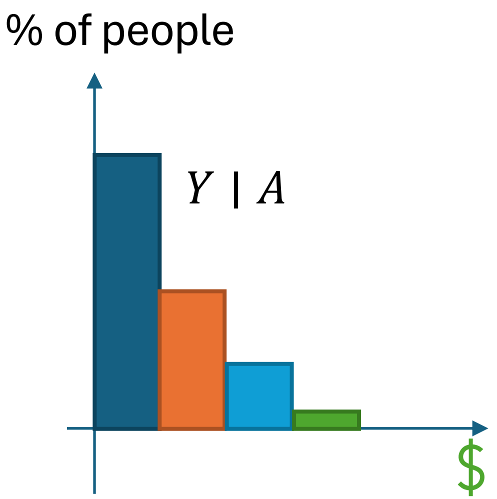


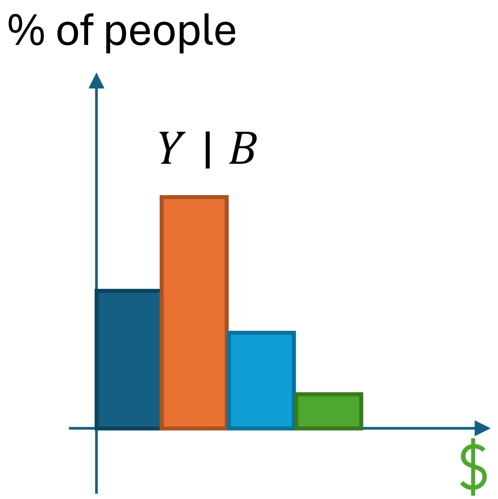
Image source: https://www.uber.com/us/en/drive/driver-app/how-surge-works/

Approach: Clustering





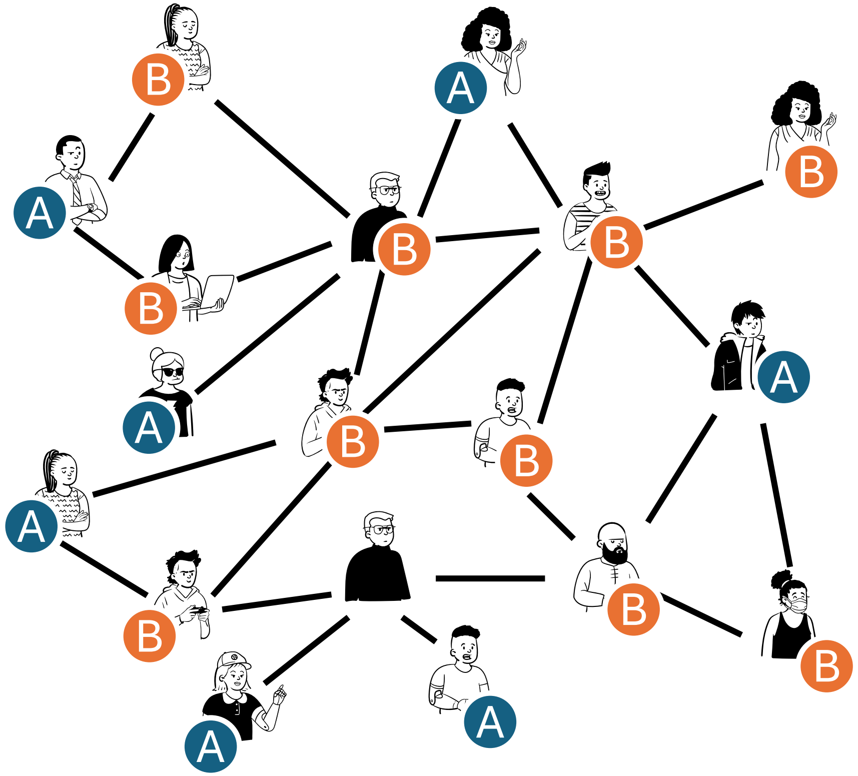
$\mu_A = 10\$$ (average spend)

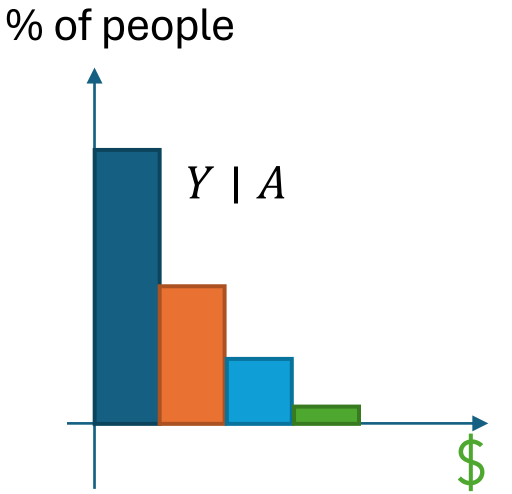


$\mu_B = 20\$$ (average spend)

Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

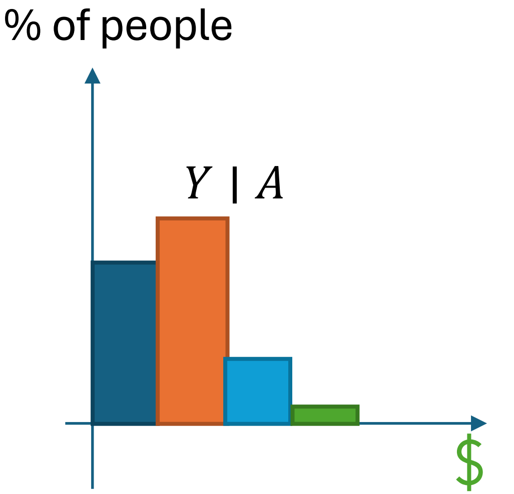
Approach: Structural Bias Correction





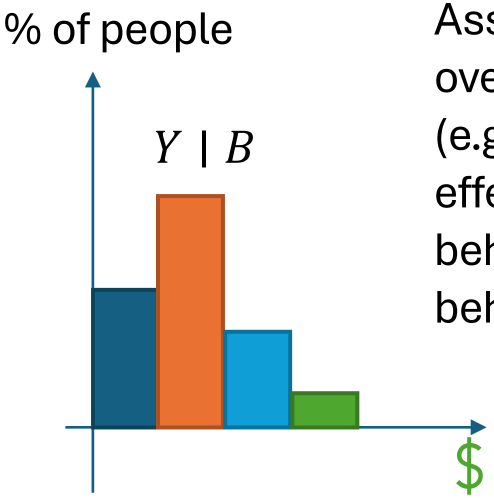
$\mu_A = 10\$$ (average spend)

Correct Spill-Over Bias

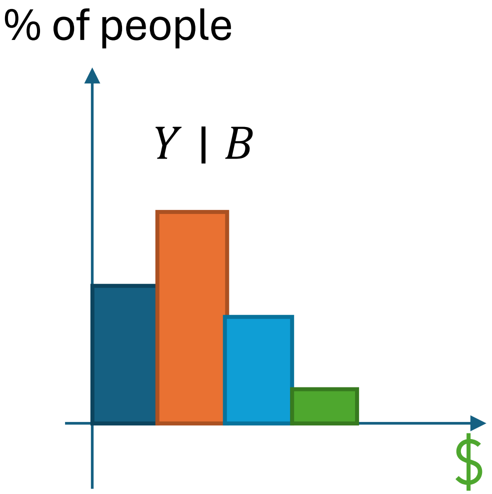


$\mu_A = 10\$$ (average spend)

Use Network Information + Assumptions on how spill-overs change outcome (e.g. additive homophily effects, market equilibrium behavior, Nash equilibrium behavior)



$\mu_B = 20\$$ (average spend)



$\mu_B = 20\$$ (average spend)

Image Source: https://www.leadpages.com/blog/ab-testing-split-testing/

A/B Testing in Auctions

A/B Testing over Position Auction Formats

Context A/B Testing for Position Auctions

- We want to optimize over the space of position auctions

- We are allowed to play with the click probabilities of slots
  - By reducing or increasing the space allocated to each slot
  - Changing the probability that the slot appears on the impression
  - Randomizing which slot the k-th highest bidder gets

High-Level Idea

- We will see that we can run a single randomized auction

- Using data that contain m samples of bids from that single randomized auction, we can estimate the revenue for every other auction in the design space at an estimation rate of $\frac{1}{\sqrt{m}}$
- Hence, we can choose the best auction in the space, with only a few rounds of experimentation!

- To do that we will need to use optimal auction theory!

Formal Setting

- We have N bidders and N slots (wlog) with CTRs $a_1 \geq \cdots \geq a_N \geq a_{N+1} = 0$
- Bidders are charged their bid-per-click (GFP)
- k-th highest bidder assigned with some distribution to one of the slots
- Slot distributions are solely determined by bid rank
- k-th highest bidder gets an implicit expected CTR of x_k

$$x_k = p_{k1}a_1 + \cdots + p_{kN}a_N$$

- These expected CTRs are monotone decreasing, $x_1 \geq x_2 \geq \cdots \geq x_N$
- No bidder is over-assigned $\sum_j p_{kj} \leq 1$
- No slot is over-assigned $\sum_k p_{kj} \leq 1$

Feasibility Characterization

- They must be feasible: for each prefix, $x_1, \ldots, x_k$ I cannot allocate a total probability more than the cumulative top k highest slots

$$\sum_{i=1}^{k} x_i \leq \sum_{i=1}^{k} a_i$$

Feasibility Characterization

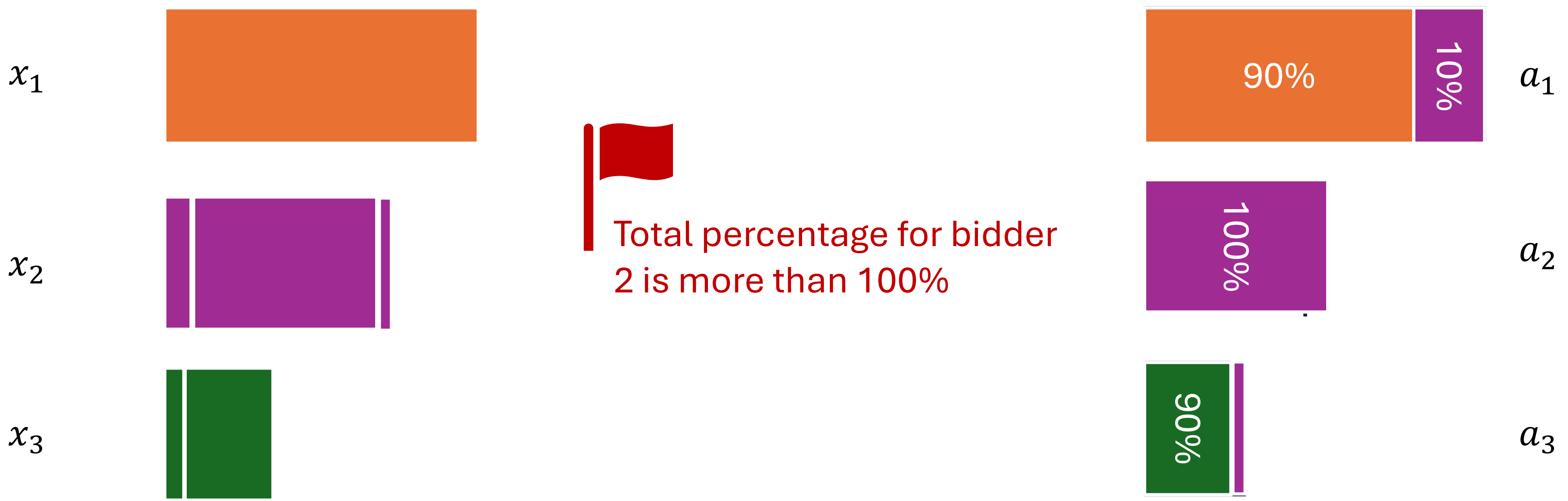
- They must be feasible: for each prefix, $x_1, \ldots, x_k$ I cannot allocate a total probability more than the cumulative top k highest slots

$$\sum_{i=1}^{k} x_i \leq \sum_{i=1}^{k} a_i$$

Feasibility Characterization

- They must be feasible: for each prefix, $x_1, \dots, x_k$ I cannot allocate a total probability more than the cumulative top k highest slots

$$\sum_{i=1}^{k} x_i \leq \sum_{i=1}^{k} a_i$$



Equivalently: Position Auction with Flexible CTRs

- We have N bidders and N slots
- Bidders are charged their bid-per-click (GFP)
- Slots are assigned in decreasing order of bidders
- k-th slot has CTR x_k. CTR of k-th slot is part of the design space
- Can choose the CTRs in any manner that satisfies $\forall k \leq N$:

$$\sum_{j=1}^{k} x_j \leq \sum_{j=1}^{k} a_j$$

for some set of predefined quantities $a_1 \geq \cdots \geq a_N \geq a_{N+1} = 0$

Equivalently: Distribution over k-Unit Auctions

- In a k-unit auction we are selling k-units of the same good
- The top-k bidders win a unit and pay their bid

Theorem. Position auction with $x_1 \geq \cdots \geq x_N \geq x_{N+1} = 0$, equivalent to distribution over k-unit auctions. k-th unit auction chosen w.p.

$$w_k = x_k - x_{k+1}, \qquad k \geq 1, \qquad \text{and}, \qquad w_0 = 1 - x_1$$

Proof. If you are the j-th bidder in position auction, you win w.p. x_j
If you are the j-th bidder in random k-unit auction, you win if $k \geq j$

$$\Pr(k \geq j) = \sum_{k \geq j} w_j = \sum_{k \geq j} x_k - x_{k+1} = x_j$$

Equivalently: Distribution over k-Unit Auctions

- In a k-unit auction we are selling k-units of the same good
- The top-k bidders win a unit and pay their bid

Theorem. Position auction with $x_1 \geq \cdots \geq x_N \geq x_{N+1} = 0$, equivalent to distribution over k-unit auctions. k-th unit auction chosen w.p.

$$w_k = x_k - x_{k+1}, \qquad k \geq 1, \qquad \text{and,} \qquad w_0 = 1 - x_1$$

Proof. If you are the j-th bidder in position auction, you win w.p. x_j
If you are the j-th bidder in random k-unit auction, you win if $k \geq j$

$$\Pr(k \geq j) = \sum_{k \geq j} w_j = \sum_{k \geq j} x_k - x_{k+1} = x_j$$

Equivalently: Distribution over k-Unit Auctions

- In a k-unit auction we are selling k-units of the same good
- The top-k bidders win a unit and pay their bid

- We run k-unit auction with probability w_k

- When bidders are symmetric, every such auction has a symmetric monotone equilibrium (in fact it has a unique equilibrium that is symmetric and monotone)

Revenue of Randomized k-unit Auction

- By Myerson, revenue of any auction is expected virtual welfare

$$\text{Rev} = \sum_i E[\phi_i(v_i) \cdot x_i(v_i)] = \sum_i \sum_k w_k E\left[\phi_i(v_i) \cdot x_{i,k}(v_i)\right]$$

- Allocation function is solely determined by rank

$$x_{i,k}(v) = \Pr(\leq \mathrm{k} - 1 \text{ bidders above you})$$

$$= \sum_{t=1}^{k-1} \binom{n-1}{t} \left(1 - F(v)\right)^t F(v)^{n-1-t}$$

- Expected allocation only depends on quantile $q(v) = F(v)$

Revenue of Randomized k-unit Auction

- By Myerson, revenue of any auction is expected virtual welfare

$$\text{Rev} = \sum_i E[\phi_i(v_i) \cdot x_i(v_i)] = \sum_i \sum_k w_k E\big[\phi_i(v_i) \cdot x_{i,k}(v_i)\big]$$

- Allocation function is solely determined by rank

$$x_{i,k}(v) = \Pr(\leq \mathrm{k} - 1 \text{ bidders above you})$$

$$= \sum_{t=1}^{k-1} \binom{n-1}{t} \big(1 - F(v)\big)^t F(v)^{n-1-t}$$

- Expected allocation only depends on quantile $q(v) = F(v)$
- Convenient to re-express everything in quantiles instead of values

Revenue of Randomized k-unit Auction

- By Myerson, revenue of any auction is expected virtual welfare

$$\text{Rev} = \sum_i E[\phi_i(q_i) \cdot x_i(q_i)] = \sum_i \sum_k w_k E\big[\phi_i(q_i) \cdot x_{i,k}(q_i)\big]$$

- Allocation function is solely determined by rank

$$x_{i,k}(q) = \sum_{t=1}^{k-1} \binom{n-1}{t} (1-q)^t \, q^{n-1-t}$$

- Quantiles q are uniformly distributed in $[0,1]$:

$$v(q) = F^{-1}(q), \qquad \Pr(Q \le q) = \Pr\big(v \le v(q)\big) = F\big(v(q)\big) = q$$

- Virtual values simplify, since by derivative of inverse $v'(q) = \big(F^{-1}(q)\big)' = 1/f\big(v(q)\big)$

$$\phi_i(q) = v(q) - \frac{1 - F\big(v(q)\big)}{f\big(v(q)\big)} = v(q) - (1-q) \cdot v'(q) = -\big(v(q) \cdot (1-q)\big)'$$

Suffices to Analyze Estimation of Revenue of k-th unit Auction

- The revenue is the weighted sum of terms (using also symmetry)

$$R_k = E[\phi(q) \cdot x_k(q)]$$

- The function $x_k(q)$ is known in closed form
- The function $\phi(q)$ is negative derivative of the **revenue function**

$$\phi(q) = -R'(q), \qquad R(q) = v(q) \cdot (1 - q)$$

- Integration-by-Parts yields

$$E[\phi(q) \cdot x_k(q)] = -\int_0^1 R'(q) \cdot x_k(q) dq = \int_0^1 R(q) \cdot x_k'(q) dq = E\left[R(q) \cdot x_k'(q)\right]$$

- It suffices that we estimate terms

$$R_k := E\left[v(q) \cdot (1 - q) \cdot x_k'(q)\right]$$

For any randomized k-unit first-price auction among symmetric bidders, we have that:

$$\mathrm{Rev} = n \sum_{k \leq N} w_k \, E\left[v(q) \cdot (1 - q) \cdot x_k'(q)\right]$$

Estimating $R_k = E\left[v(q) \cdot (1-q) \cdot x_k'(q)\right]$

- The value function $v(q) = F^{-1}(q)$ relates to distribution of values
- Only observed from data distribution of bids with CDF G and pdf g
- Define the bid function $b(q) = G^{-1}(q)$: *what is my bid if I'm at the bottom q-th percentile of the distribution of values, equivalently, if I'm at the q-th percentile of the distribution of bids*

- Want to relate value of quantile q to bid of quantile q

- Similar to bid inversion question in last lecture

Estimating $R_k = E\left[v(q) \cdot (1 - q) \cdot x_k'(q)\right]$

- At symmetric equilibrium

$$b(q) = \mathrm{argmax}_z\ (v(q) - z) \cdot x\left(b^{-1}(z)\right)$$

- The first order condition (using derivative of inverse):

$$\left(v(q) - b(q)\right) \cdot x'(q) \frac{1}{b'(q)} - x(q) = 0$$

- We can write a similar bid inversion formula

$$v(q) = b(q) + \frac{b'(q)x(q)}{x'(q)}$$

- ***Reminder:*** The functions $x(q)$ and $x'(q)$ are known in closed form

Estimating $R_k = E\left[v(q) \cdot (1-q) \cdot x_k'(q)\right]$

- We can write a similar bid inversion formula
$$v(q) = b(q) + \frac{b'(q)x(q)}{x'(q)}$$
- Need to estimate $\mathrm{b}(\mathrm{q})$ and $\mathrm{b}'(\mathrm{q})$ from data
- ***Reminder:*** $b(q) = G^{-1}(q), \quad b'(q) = \frac{1}{g\left(G^{-1}(q)\right)}$
- Estimating $b(q)$ and $b'(q)$ is the same as estimating G, g

- **Main message.** The quantity R_k for any k depends only on $b(q)$ and not on $b'(q)$ because it is an integral over q! Leads to much faster rates.

Estimating $R_k = E\left[v(q) \cdot (1-q) \cdot x_k'(q)\right]$

- We can write

$$R_k = E\left[b(q) \cdot (1-q) \cdot x_k'(q)\right] + E\left[\frac{b'(q)x(q)}{x'(q)} \cdot (1-q) \cdot x_k'(q)\right]$$

- First part only depends on $b(q)$. Analogous to estimating a CDF
- Second part seemingly problematic. But integration-by-parts

$$E\left[\frac{b'(q)x(q)}{x'(q)} \cdot (1-q) \cdot x_k'(q)\right] = -E\left[b(q)\left(\frac{x(q)(1-q) \cdot x_k'(q)}{x'(q)}\right)'\right]$$

- This only depends on $b(q)$ and known quantities

For any randomized k-unit first-price auction among symmetric bidders, we have that:

$$\text{Rev} = n \sum_{k \leq N} w_k \, E[b(q) \cdot f(q)]$$

for a function $f(q)$ known in closed form

We can estimate Rev by estimating the CDF of bids using the empirical CDF $\hat{G}$. Then use $\hat{b} = \hat{G}^{-1}$ and

$$\widehat{\mathrm{Rev}} = n \sum_{k \leq N} w_k \int_0^1 \hat{b}(q) \cdot f(q) dq$$

for a function $f(q)$ known in closed form

Assuming $f(q)$ is bounded (e.g. holds if $w_k \geq \epsilon$), then

$$\left|\widehat{\mathrm{Rev}} - \mathrm{Rev}\right| \lesssim 1/\sqrt{m}$$

Conclusion

- Run a single randomized auction as our experimentation strategy
- Using data that contain m samples of bids from that single randomized auction, we can estimate the revenue for every other auction in the design space at an estimation rate of $\frac{1}{\sqrt{m}}$
- Hence, we can choose the best auction in the space, with only a few rounds of experimentation!
- To do that we used optimal auction theory!

A/B Testing across Many Keywords with Budgets

Budgets!

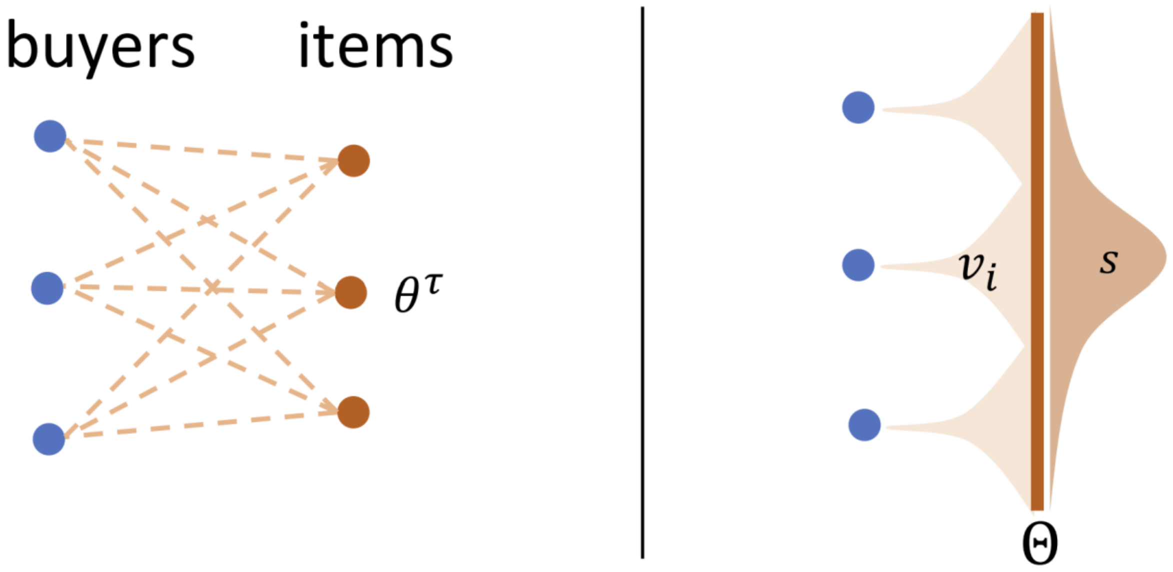
- So far we did not place any budget constraints on bidders
- In practice, budget constraints are very important
- Bidders participate in many auctions and have a budget limit
- Can only spend at most B_i in total across all the auctions

- This couples the bidding strategy across auctions
- Makes learning (e.g. no-regret learning hard)
- In its full generality a stochastic dynamic program

Simplified Budgets: Pacing Equilibria

Interference Among First-Price Pacing Equilibria: A Bias and Variance Analysis (arxiv.org)

- In practice, people use the following simplification
- We have n bidders and a continuum of items
- Items have type θ which follows some distribution with measure s
- $v_i(\theta)$ is bidder i's value for an item of type θ



Simplified Budgets: Pacing Equilibria

The multipliers $\beta = (\beta_1, \ldots, \beta_n)$ and price function $p(\theta)$ are a *pacing equilibrium* if there exists and allocation function $x(\theta)$ such that

- First-price payment: $p(\theta) = \max_i \beta_i v_i(\theta)$
- Highest-bidder wins: $x_i(\theta) \geq 0 \Rightarrow \beta_i v_i(\theta) = \max_k \beta_k v_k(\theta)$
- Budgets are respected
$$\int_\theta x_i(\theta) p(\theta) s(\theta) d\theta \leq B_i$$
- No-overselling: $\sum_i x_i(\theta) \leq 1$
- Full-allocation of competitive items: $p(\theta) > 0 \Rightarrow \sum_i x_i(\theta) = 1$
- No un-necessary pacing: $\int_\theta x_i(\theta) p(\theta) s(\theta) d\theta < B_i \Rightarrow \beta_i = 1$

Characterization of Pacing Equilibria

Multipliers in pacing equilibrium are characterized as solutions to a convex optimization problem (related to market equilibrium)

$$\beta_* = \operatorname*{argmin}_{\beta \in (0,1]^n} E\left[\max_i \beta_i v_i(\theta)\right] - \sum_i B_i \log(\beta_i)$$

Clustered Experiment Designs and Debiasing

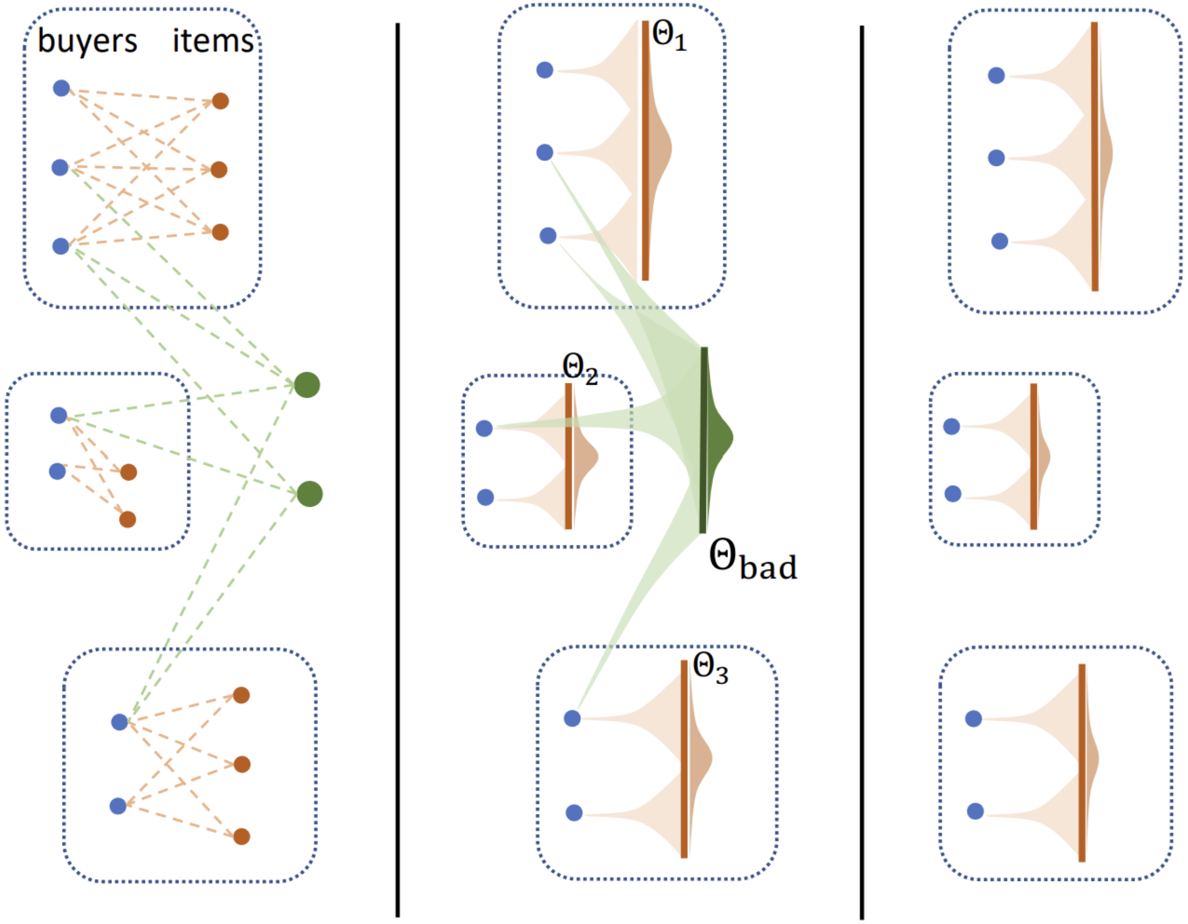
[Interference Among First-Price Pacing Equilibria: A Bias and Variance Analysis (arxiv.org)](https://arxiv.org)

1. For each sub-market want pacing multipliers as if the bad items don't exist

2. With such multipliers, can estimate idealized revenue for each sub-market, as if isolated

3. Characterization of multipliers as minimizers of market equilibrium program ⇒ closed form first-order bias that bad items introduce

4. Subtract bias and measure revenue of A and B clusters using debiased multipliers



A/B Testing in Two-Sided Matching Markets

Two-Sided Randomized Designs

[Experimental Design in Two-Sided Platforms: An Analysis of Bias | Management Science (informs.org)]