

B9651 – Marketing Analytics

Session 6: Choice Models

Professor Hortense Fong

Logistics

- Individual Assignment 1 due Wed, Oct 16 at 8PM
- No classes next week
- Midterm Oct 22 + 23
 - Weeks 1-5, closed-note, calculator allowed

Last Time

- Recommender Systems + Modeling Customer Churn
- Today: Modeling Customer Choices

Today: Modeling Choice

1. How can we model the choice process of customers?
2. What are the different types of choice models?
 1. Binary logit
 2. Multinomial Logit
3. How can we estimate the Logit Models?
4. What are the limitations of Multinomial Logit Models?

Today's Goals

Understand

- What are choice models
- How to construct a utility function
- How to use Maximum Likelihood Estimation to model consumer choices
- The limitations of Multinomial Logit Models

Be able to

- Construct a statistical model of choice
- Estimate a choice model in Excel and Python
- Analyze own- and cross-price elasticities

Course Roadmap

STP Analytics (Identify Value)	Customer Analytics (Deliver Value)	4P Analytics (Capture Value)
Module 1	Module 2	Module 3
What datasets can we use? How can we segment and target our customers? How should we position our products/services?	How much are our customers worth? Are our customers leaving? How do our customers make choices?	How do we build a new product? How should we price our products? How do we distribute them? How do we quantify the impact of our promotions?

Choice Models

Motivation

Modeling Choice

- Imagine that you work at Megabus
 - Provide regular intercity bus routes at a low cost
- Problem: You don't understand when people choose to take a bus or another transportation mode!
 - Alternatives: Train, Plane, Car
- What do you do?
- What if you want to know how price sensitive people are or how important it is compared to travel time?
- Today: we will build a (choice) model to address this type of question



Choice Models – Other Applications

- Choice Models
 - Describe how people make choices
 - Predict choices under different conditions
- Common Choice Problems
 - Purchase or Not (Purchase incidence)
 - Brand Choice
 - Credit Default Prediction

Choice Models

Assumptions

Choice Problem

- Choices are made by **decision makers**
 - Individual Consumers, Households, Firms
- **Choice sets** contain
 - A finite number of alternatives
 - Alternatives are mutually exclusive
 - Collectively exhaustive
- Choose one **alternative** from a **choice set**

Decision maker



Choice set

Bus

Train

Plane

Car

Bus

Consumers

- Consumer choices are modeled in terms of
 - **Characteristics** of consumers (age, income, etc.)
 - **Attributes** of alternatives (price, travel time, etc.)
- Consumers have **preferences** over the attributes
 - Preferences for attributes are represented by attribute weights
 - **Preferences for attributes** translate into **preferences for alternatives**

Decision maker



25 y.o
Female
\$80,000

Choice set

Bus

\$20
5 hours

Train

\$100
4 hours

Plane

\$200
1 hour

Car

\$50
4.5 hours

$$\beta_1 \text{Price} + \beta_2 \text{Time}$$

Bus

Consumer Decision Rule

- Consumers preferences for alternatives are represented by **utility functions**
- Utility functions assign one scalar numerical value to each alternative
- Utilities are functions of
 - Attributes of alternatives
 - Characteristics of consumers

} E.g., Transportation choice is function of price, time, and income
- **Rational consumers choose the alternative with the highest utility**
 - Utility Maximizing Decision Rule

Example: Maximum Utility Choice Rule

- t = time, c = cost, y = income in \$K
- Suppose utility is given by $U(t, c, y) = -t - \frac{5c}{y}$

Mode	Time (t) (Hours)	Cost (c) (\$)	Utility $y = \$40k$
Drive Alone	0.50	2.00	
Carpool	0.75	1.00	
Bus	1.00	0.75	

Which mode of transportation maximizes utility?

0

Drive Alone

0

Carpool

0

Bus

0

Example: Maximum Utility Choice Rule

- t = time, c = cost, y = income in \$K
- Suppose utility is given by $U(t, c, y) = -t - \frac{5c}{y}$

Mode	Time (t) (Hours)	Cost (c) (\$)	Utility $y = \$40k$	Utility $y = \$10k$
Drive Alone	0.50	2.00	-0.75	-1.50
Carpool	0.75	1.00	-0.88	-1.25
Bus	1.00	0.75	-1.09	-1.38

- How can you explain the above utilities for the two travelers?

Effect of Reduction in Bus Travel Time by a Quarter of an Hour

- t = time, c = cost, y = income in \$K
- Suppose utility is given by $U(t, c, y) = -t - \frac{5c}{y}$

Mode	Time (t) (Hours)	Cost (c) (\$)	Utility $y = \$40k$	Utility $y = \$10k$
Drive Alone	0.50	2.00	-0.75	-1.50
Carpool	0.75	1.00	-0.88	-1.25
Bus	0.75	0.75	-0.84	-1.13

Choice Models

From Assumptions To Mathematics

Consumer Preferences: Mathematical Representation

- Let \mathcal{J} represent the choice set
- Let \mathbf{w}_{ij} represent **all** the attributes of alternative j that consumer i faces
 - E.g., Time, Cost
- Let \mathbf{r}_i be the vector of **all** consumer characteristics that are relevant for choice
 - E.g., Income
- Utility is a function $U_{ij} = U(\mathbf{w}_{ij}, \mathbf{r}_i), \forall j \in \mathcal{J}$
(for all alternatives in the choice set)
- Consumer decision rule is deterministic:
 - Choose alternative k if $U_{ik} > U_{ij}$ for all $j \neq k, j \in \mathcal{J}$ (Choose the alternative with the highest utility)

The Problem with Deterministic Utility

- The previous travel example used a deterministic utility function
- The function implies **consistency** in behavior
 - For each consumer, each time the consumer faces the same task
 - For identical consumers
- Real datasets show a lot of inconsistency

Inconsistency in Choice

- What inconsistencies do you see in the example below?

Customer	Age	Price_A	Price_B	Choice
1	23	1.25	1.15	B
1	23	1.25	1.35	A
1	23	1.25	1.15	A
2	25	1.15	1.25	B
2	25	1.25	1.15	A
2	25	1.15	1.25	A
3	31	1	1.15	A
4	23	1.25	1.35	B

Inconsistency in Choice

- Potential sources of inconsistency:
 1. Consumers have incomplete or incorrect information about the attributes
 2. Analyst has incomplete or incorrect information about the attributes or circumstances of the customers
- To account for incomplete information, we use a random utility model to rationalize the observed data

Choice Models

Assumptions (again!) and Random Utility

Random Utility

- Researcher does not know utilities exactly
 - Only few consumer characteristics, \mathbf{z}_i out of \mathbf{r}_i are known
 - Only few of the attributes, \mathbf{x}_{ij} out of \mathbf{w}_{ij} are known
- Solution: treat the utility, U_{ij} , as random with additive errors



$$U_{ij} = V_{ij} + \epsilon_{ij}$$

- $V_{ij}(\mathbf{x}_{ij}, \mathbf{z}_i; \boldsymbol{\beta})$ is the **systematic** part of the utility
- ϵ_{ij} is the **stochastic** part
 - Represents total impact of all unobserved attributes and demographics relevant to a given choice occasion

Probabilistic Choice

- Given the stochastic part, we can only model choice probabilistically
- Probability of choosing alternative $j \in \mathcal{J}$ by customer i is given by

$$P_{ij} = \text{Prob}(U_{ij} > U_{ik}, \forall k \neq j, k \in \mathcal{J})$$

$$P_{ij} = \text{Prob}(V_{ij} + \epsilon_{ij} > V_{ik} + \epsilon_{ik}, \forall k \neq j, k \in \mathcal{J})$$

$$P_{ij} = \text{Prob}(\underbrace{\epsilon_{ik} - \epsilon_{ij}}_{\text{Random}} < \underbrace{V_{ij} - V_{ik}}_{\text{Systematic}}, \forall k \neq j, k \in \mathcal{J})$$

Example: Two Brands

- Let the utilities be given by
 - $U_{i1} = V_{i1} + \epsilon_{i1}$
 - $U_{i2} = V_{i2} + \epsilon_{i2}$
- Probability of choosing brand 2 for individual i is

$$\begin{aligned} P_{i2} &= \text{Prob}(U_{i1} < U_{i2}) \\ &= \text{Prob}(\epsilon_{i1} - \epsilon_{i2} < V_{i2} - V_{i1}) \end{aligned}$$

Only differences in utilities matter!

Example: Three Brands

- Let the utilities be given by
 - $U_{i1} = V_{i1} + \epsilon_{i1}$
 - $U_{i2} = V_{i2} + \epsilon_{i2}$
 - $U_{i3} = V_{i3} + \epsilon_{i3}$
- What is the probability of choosing brand 3 in terms of differences in errors?

$$\begin{aligned}P_{i3} &= \text{Prob}(U_{i1} < U_{i3}, U_{i2} < U_{i3}) \\&= \text{Prob}(V_{i1} + \epsilon_{i1} < V_{i3} + \epsilon_{i3}, V_{i2} + \epsilon_{i2} < V_{i3} + \epsilon_{i3}) \\&= \text{Prob}(\epsilon_{i1} - \epsilon_{i3} < V_{i3} - V_{i1}, \epsilon_{i2} - \epsilon_{i3} < V_{i3} - V_{i2})\end{aligned}$$

Only differences in utilities matter!

Implications

- Only differences in utilities matter
- Implications on the systematic part of the utility
 - What happens when we add the same constant to all utilities?
 - Probabilities don't change: $V_1 = 4, V_2 = 5$ is the same as $V_1 = 2, V_2 = 3$
 - Probability of choosing brand 2 for individual i is

$$\begin{aligned} P_{i2} &= \text{Prob}(\epsilon_{i1} - \epsilon_{i2} < V_{i2} - V_{i1}) \\ &= \text{Prob}(\epsilon_{i1} - \epsilon_{i2} < 5 - 4) \\ &= \text{Prob}(\epsilon_{i1} - \epsilon_{i2} < 3 - 2) \end{aligned}$$

Choice Models

Systemic Utility

Systematic Part of the Utility

- The systematic part $V_{ij}(\cdot)$ can be written as

$$V_{ij} = V_j(x_{ij}) + V_j(z_i) + V_j(x_{ij}, z_i) + bias_j$$

- consumer i
- alternative j

- $V_j(x_{ij})$ is the portion that is associated with the **attributes** of alternative j faced by consumer i
- $V_j(z_i)$ is the portion of utility associated with the **characteristics** of the consumer
- $V_j(x_{ij}, z_i)$ contains the interactions between the **attributes and characteristics**
- $bias_j$ is the alternative specific constant
- Our entire objective will be to estimate the various components

Portion 1: Systematic Utility Composed of Attributes

$$V_j(x_{ij})$$

- The attribute portion of the systematic utility, $V_j(x_{ij})$ can be written as

$$V_j(x_{ij}) = \beta_1 x_{ij1} + \beta_2 x_{ij2} + \cdots + \beta_k x_{ijk}$$

- β_k is the **coefficient** of the k th attribute (e.g., **weight** of price)
 - This is what we want to know/estimate!**
- x_{ijk} is the **value** of the k th attribute (e.g., price faced by consumer i for alternative j)

Portion 1: Systematic Utility Composed of Attributes

$$V_j(x_{ij})$$

- Consider two modes A (airplane) and B (bus) described on two attributes: Price and TravelTime
- The attribute-specific component of the systematic utility is

$$V_A(x_{iA}) = \beta_1 Price_{iA} + \beta_2 TravelTime_{iA}$$

$$V_B(x_{iB}) = \beta_1 Price_{iB} + \beta_2 TravelTime_{iB}$$

- Notice that the coefficients are the same across alternatives
 - Why?
 - Assumption is that Price (and TravelTime) sensitivity is the same across modes

Portion 2: Systematic Component for Individual Characteristics $V_j(z_i)$

- Suppose we have two demographics for consumer i : **Income** and **Family-size**
- The individual-specific component can be written as

$$V_A(z_i) = \beta_{1A}Income_i + \beta_{2A}FamilySize_i$$

$$V_B(z_i) = \beta_{1B}Income_i + \beta_{2B}FamilySize_i$$

- Notice that each mode has a **different** coefficient for each variable.
 - Why? Only differences in utility matter: if it was the same, $V_A(z_i) - V_B(z_i) = \beta_1Income_i + \beta_2FamilySize_i - (\beta_1Income_i + \beta_2FamilySize_i) = 0$

Portion 3: Systematic Component with Interactions

$$V_j(x_{ij}, z_i)$$

- Different individuals may evaluate attributes differently
- We can interact the attributes with the demographics

$$V_A(x_{iA}, z_i) = \beta_1 Price_{iA} + \beta_2 Price_{iA} \times Income_i$$

$$V_B(x_{iB}, z_i) = \beta_1 Price_{iB} + \beta_2 Price_{iB} \times Income_i$$

- How do we interpret the coefficients?
 - Total effect of a unit increase in price: $\beta_1 + \beta_2 \times Income_i$
 - How much price matters depends on income
 - High income people will be less price sensitive compared to low-income people (when $\beta_1 < 0$ and $\beta_2 > 0$)

Portion 4: Alternative Specific Constants *bias_j*

- The systematic utilities also contain intercepts (or biases) that are alternative specific

$$\begin{aligned}V_{iA} &= \beta_{0A} + \dots \\V_{iB} &= \beta_{0B} + \dots\end{aligned}$$

- These represent the mean of all the unobserved variables ϵ_{ij}
 - Ex. Comfort, safety, privacy (difficult to measure variables)
- They capture the **baseline utilities**
 - What is **unique** about each alternative

Putting It All Together: Systematic Utility V_{ij}

- We can assemble all components to get

$$V_{iA} = \beta_{0A} + \beta_{1A}Income_i + \beta_2Price_{iA} + \beta_3Price_{iA} \times Income_i$$

For illustration, we include only one attribute (price) and one consumer characteristic (income)

$$V_{iB} = \beta_{0B} + \beta_{1B}Income_i + \beta_2Price_{iB} + \beta_3Price_{iB} \times Income_i$$

- Because only differences in utilities matter, not all these coefficients are identifiable (i.e., have a unique value)

$$\begin{aligned} V_{iA} - V_{iB} = & (\beta_{0A} - \beta_{0B}) \\ & + (\beta_{1A} - \beta_{1B}) Income_i \\ & + \beta_2(Price_{iA} - Price_{iB}) \\ & + \beta_3Income_i \times (Price_{iA} - Price_{iB}) \end{aligned}$$

← unidentifiable

Many values of the estimates would lead to the same difference in utility

Putting It All Together: Systematic Utility V_{ij}

$$\begin{aligned} V_{iA} - V_{iB} = & (\beta_{0A} - \beta_{0B}) \quad \leftarrow \text{unidentifiable} \\ & + (\beta_{1A} - \beta_{1B}) \text{Income}_i \\ & + \beta_2(\text{Price}_{iA} - \text{Price}_{iB}) \\ & + \beta_3 \text{Income}_i \times (\text{Price}_{iA} - \text{Price}_{iB}) \end{aligned}$$

- Because only differences in utility matter, the two intercepts cannot be estimated separately. Only their difference can be estimated.
- Many values of the estimates would lead to the same difference in utility
 - E.g., $\beta_{0A} = 2, \beta_{0B} = 1$ same as $\beta_{0A} = 3, \beta_{0B} = 2$

Putting It All Together: Systematic Utility V_{ij}

- We can set some parameters to zero to obtain identification

$$V_{iA} = 0 + 0 \times \text{Income}_i + \beta_2 \text{Price}_{iA} + \beta_3 \text{Price}_{iA} \times \text{Income}_i$$

$$V_{iB} = \beta_{0B} + \beta_{1B} \text{Income}_i + \beta_2 \text{Price}_{iB} + \beta_3 \text{Price}_{iB} \times \text{Income}_i$$

$$\begin{aligned} V_{iA} - V_{iB} &= (0 - \beta_{0B}) \\ &\quad + (0 - \beta_{1B}) \text{Income}_i \\ &\quad + \beta_2 (\text{Price}_{iA} - \text{Price}_{iB}) \\ &\quad + \beta_3 \text{Income}_i \times (\text{Price}_{iA} - \text{Price}_{iB}) \end{aligned}$$

Putting It All Together: Systematic Utility V_{ij}

- We can set some parameters to zero to obtain identification

$$V_{iA} = 0 + 0 \times \text{Income}_i + \beta_2 \text{Price}_{iA} + \beta_3 \text{Price}_{iA} \times \text{Income}_i$$

$$V_{iB} = \beta_{0B} + \beta_{1B} \text{Income}_i + \beta_2 \text{Price}_{iB} + \beta_3 \text{Price}_{iB} \times \text{Income}_i$$

- How do we interpret the remaining intercept for Brand B?
 - Relative to the intercept of Brand A
 - If positive, baseline utility of B higher than baseline utility of A
- How do we interpret the income coefficient for Brand B?

Putting It All Together

- How can we model the choice process of customers?
- How to construct a utility function

$$U_{ij} = V_{ij} + \epsilon_{ij}$$

$$V_{ij} = V_j(x_{ij}) + V_j(z_i) + V_j(x_{ij}, z_i) + bias_j$$

- consumer i
- alternative j

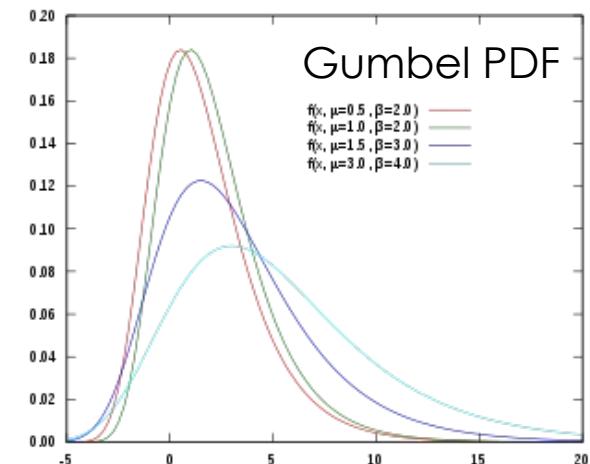
- $V_j(x_{ij})$ is the portion that is associated with the attributes of alternative j faced by consumer i
- $V_j(z_i)$ is the portion of utility associated with the characteristics of the consumer
- $V_j(x_{ij}, z_i)$ contains the interactions between the attributes and characteristics
- $bias_j$ is the alternative specific constant

Choice Models

Stochastic Error Term

Stochastic Part

- We assume that the stochastic part ϵ_{ij} varies across alternatives j and across consumers i
- As the errors are not known, we assume that these come from a probability distribution
- Different assumption on probability distributions of errors leads to different discrete choice model
 - Errors are Gumbel means we get logit models
 - Computational advantages
 - Closed-form choice model
 - Closely approximates normal distribution
 - Gumbel PDF: $f(x) = e^{-(x+e^{-x})}$
 - Errors are normal means we get probit models



Multinomial Logit (MNL) Model

- Since errors ϵ_{ij} are i.i.d. extreme value (Gumbel), we have **logit** models

Cumulative density function: $F(x) = e^{-e^{-x}}$ ($\Pr(\epsilon < x) = e^{-e^{-x}}$)

- i.i.d means that the errors are independent
 - Across utility equations for a given consumer: $\epsilon_{ij} \perp \epsilon_{ik}$
 - Across different consumers: $\epsilon_{ij} \perp \epsilon_{mj}$ & $\epsilon_{ij} \perp \epsilon_{mk}$
- After estimation, the probability of consumer i choosing brand j is given by

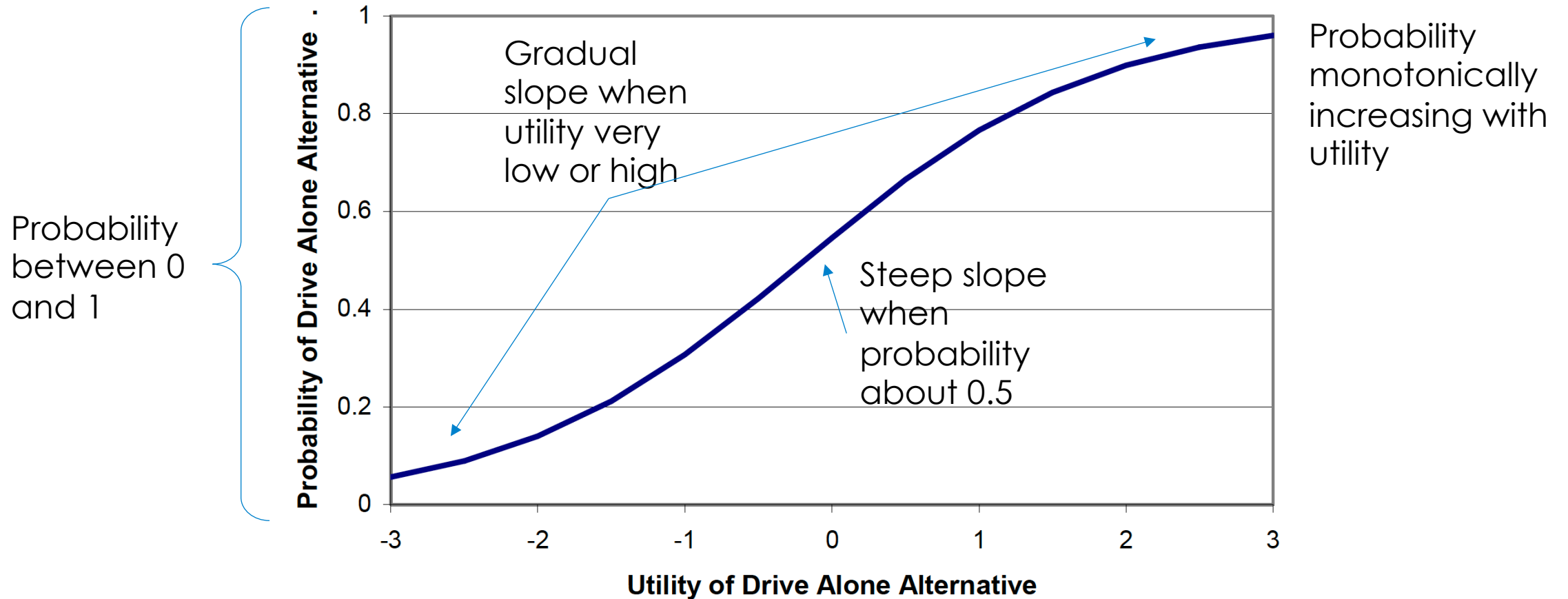
$$P_{ij} = \frac{\exp(V_{ij})}{\sum_k \exp(V_{ik})}$$

- V_{ij} = systematic part of utility
- k captures all possible alternatives in choice set including j

- Does the above equation for probability make sense?

Multinomial Logit (MNL) Model

Assuming choice is drive alone or not drive alone (i.e., 2 choices)



Extra – Multinomial Logit Derivation

Don't worry. This will not be tested on the exam.

Probability consumer purchases k vs j:

$$\begin{aligned}\Pr(U_{ik} > U_{ij}) &= \Pr(V_{ik} + \epsilon_{ik} > V_{ij} + \epsilon_{ij}) \\ &= \Pr(V_{ik} - V_{ij} + \epsilon_{ik} > \epsilon_{ij})\end{aligned}$$

Imposing CDF of Gumbel distribution and conditioning on ϵ_{ik} , we get:

$$\Pr(U_{ik} > U_{ij} | \epsilon_{ik}) = F(V_{ik} - V_{ij} + \epsilon_{ik})$$

Cumulative density function: $\Pr(\epsilon < x) = F(x)$

What about when we have more than alternatives j and k? We can multiply the above probability since the unobserved utility is independent across goods. For all alternatives $j \neq k$,

$$\Pr(U_{ik} > U_{ij} \forall j \neq k | \epsilon_{ik}) = \prod_{j \neq k} F(V_{ik} - V_{ij} + \epsilon_{ik})$$

The unconditional probability (integrating over ϵ_{ik}) that k is chosen is:

$$P_{ik} = \Pr(U_{ik} > U_{ij} \forall j \neq k) = \int_{-\infty}^{\infty} \prod_{j \neq k} F(V_{ik} - V_{ij} + \epsilon_{ik}) f(\epsilon_{ik}) d\epsilon_{ik}$$

From last slide (unconditional probability of choosing k):

$$P_{ik} = \Pr(U_{ik} > U_{ij} \forall j \neq k) = \int_{-\infty}^{\infty} \prod_{j \neq k} F(V_{ik} - V_{ij} + \epsilon_{ik}) f(\epsilon_{ik}) d\epsilon_{ik}$$

Imposing Gumbel distribution, we get

Cumulative density function: $F(x) = e^{-e^{-x}}$

$$P_{ik} = \int_{-\infty}^{\infty} \prod_{j \neq k} \exp\left(-\exp\left(-(V_{ik} - V_{ij} + \epsilon_{ik})\right)\right) \underbrace{\exp(-\epsilon_{ik}) \exp(-\exp(-\epsilon_{ik}))}_{f(\epsilon_{ik})} d\epsilon_{ik}$$

Probability density function: $f(x) = e^{-(x+e^{-x})}$

Since $\exp(-\exp(-\epsilon_{ik})) = \exp(-\exp(-(V_{ik} - V_{ik} + \epsilon_{ik})))$, the above simplifies to:

$$P_{ik} = \int_{-\infty}^{\infty} \prod_j \exp\left(-\exp\left(-(V_{ik} - V_{ij} + \epsilon_{ik})\right)\right) \exp(-\epsilon_{ik}) d\epsilon_{ik} \quad \text{(Product now includes k)}$$

Since product of exponentials is the exponential of the sum of the exponents:

$$P_{ik} = \int_{-\infty}^{\infty} \exp\left(-\sum_j \exp\left(-(V_{ik} - V_{ij} + \epsilon_{ik})\right)\right) \exp(-\epsilon_{ik}) d\epsilon_{ik} = \int_{-\infty}^{\infty} \exp\left(-\exp(-\epsilon_{ik}) \sum_j \exp\left(-(V_{ik} - V_{ij})\right)\right) \exp(-\epsilon_{ik}) d\epsilon_{ik}$$

We need to use a change of variables:

$$t = -\exp(-\epsilon_{ik})$$
$$dt = \exp(-\epsilon_{ik}) d\epsilon_{ik} \text{ where } t \in (-\infty, 0)$$

Then,

$$P_{ik} = \int_{-\infty}^{\infty} \exp\left(-\exp(-\epsilon_{ik}) \sum_j \exp\left(-(V_{ik} - V_{ij})\right)\right) \exp(-\epsilon_{ik}) d\epsilon_{ik} = \int_{-\infty}^0 \exp\left(t \sum_j \exp\left(-(V_{ik} - V_{ij})\right)\right) dt$$

Completing the integral:

$$P_{ik} = \left(\frac{\exp\left(t \sum_j \exp\left(-(V_{ik} - V_{ij})\right)\right)}{\sum_j \exp\left(-(V_{ik} - V_{ij})\right)} \right) \Big|_{-\infty}^0$$

$$P_{ik} = \frac{1}{\sum_j \exp\left(-(V_{ik} - V_{ij})\right)} = \frac{1}{\sum_j \exp(-V_{ik}) \exp(V_{ij})} = \frac{1}{\exp(-V_{ik}) \sum_j \exp(V_{ij})} = \frac{\exp(V_{ik})}{\sum_j \exp(V_{ij})}$$

Break

10-minutes

Choice Models

Estimation

So What? From Theory to Practice!

- We established a framework for how people make choices **BUT** how do we use it?
- In practice, we want to **estimate** the **coefficients** in systematic part of the utility
 - Why? For interpretation!
 - How price sensitive are consumers?
 - How time sensitive are consumers?
 - When buying a car, how sensitive are consumers to mileage per gallon?

So What? From Theory to Practice!

- Consider data with 200 consumer choices:
 - Two brands, named 1 and 2
 - One alternative specific variable named Price
- What are the systematic utility functions for the two brands?
 - Systematic utility for consumer i choosing alternative 1: $V_{i1} = \beta Price_{i1} + \beta_{01}$
 - Systematic utility for consumer i choosing alternative 2: $V_{i2} = \beta Price_{i2} + \beta_{02}$
- Based on the theory, how many parameters can we estimate?
 - 2
 - Systematic utility for consumer i choosing alternative 1: $V_{i1} = \beta Price_{i1}$
 - Intercept $(\beta_{01}) = 0$
 - Systematic utility for consumer i choosing alternative 2: $V_{i2} = \beta Price_{i2} + \beta_{02}$

Data Extract

- Here is an extract from 9 consumers

Consumer	Choice	Price_1	Price_2
1	1	1.15	1.2
2	1	1.15	1.24
3	2	1.1	1.09
4	2	1.15	1.2
5	2	1.1	0.9
6	2	1.1	1.24
7	2	1.1	1.2
8	2	1.15	0.9
9	1	1.25	1.2

- We can use data to estimate the **intercept** for brand 2 and the **price coefficient**
- How?

Maximum Likelihood Estimation

Maximum Likelihood Estimation (MLE) – Reminder

- Maximum Likelihood Estimation
 - Use the data to find values of the model parameters (θ) that maximize the likelihood of observing the data that we have
- We estimate the model parameters by maximizing the likelihood function $L(\theta)$
- The resulting parameter estimates, θ_{ML} are called “maximum likelihood estimates”

Likelihood – Consumer Choice

- Let y_i be the observed choice for customer i , which takes the values 1 or 2
- Let δ_{i1} and δ_{i2} be two binary variables
 - $\delta_{ij} = 1$, if $y_{ij} = j$, and zero otherwise
- The likelihood for observation i is given by
$$\mathcal{L}_i(\theta | y_i) = \text{Prob}(y_i = 1)^{\delta_{i1}} \text{Prob}(y_i = 2)^{\delta_{i2}}$$
 - Why? Note that $\delta_{i2} = 1 - \delta_{i1}$ so δ_{ij} serves as an indicator variable for probability
- The likelihood for the **entire data** is the product of the observation-level likelihoods

$$\mathcal{L}(\theta | D) = \prod_{i=1}^N \mathcal{L}_i(\theta | y_i)$$

Log-Likelihood

- In practice, we maximize the log-likelihood
- The log-likelihood for consumer i is given by

Recall probability of consumer i choosing brand j ,

$$P_{ij} = \frac{\exp(V_{ij})}{\sum_k \exp(V_{ik})}$$

$$\begin{aligned} \mathcal{LL}_i(\theta|y_i) &= \delta_{i1} \log(\text{Prob}(y_i = 1|\theta)) + \delta_{i2} \log(\text{Prob}(y_i = 2|\theta)) = \\ &\delta_{i1} \log\left(\frac{\exp(V_{i1})}{\sum_k \exp(V_{ik})}\right) + \delta_{i2} \log\left(\frac{\exp(V_{i2})}{\sum_k \exp(V_{ik})}\right) \end{aligned}$$

- The overall log-likelihood is the sum of the consumer-specific log-likelihoods

$$\mathcal{LL}(\theta|D) = \sum_{i=1}^N \mathcal{LL}_i(\theta|y_i)$$

Your Turn!

- Open dataBinaryAnalysisProblems.xlsx
- Start with an intercept and a β_{price} of 0.5
- For each individual (each row), add columns that compute
 - The systematic utilities for each brand
 - The probabilities of purchasing each brand
 - The loglikelihood (you can use more columns if you prefer)
 - $\mathcal{LL}_i(\theta|y_i) = \delta_{i1} \log\left(\frac{\exp(V_{i1})}{\sum_k \exp(V_{ik})}\right) + \delta_{i2} \log\left(\frac{\exp(V_{i2})}{\sum_k \exp(V_{ik})}\right)$
- Add a cell that contains the total log-likelihood and use solver to find the coefficients
 - $\mathcal{LL}(\theta|D) = \sum_{i=1}^N \mathcal{LL}_i(\theta|y_i)$
- You have 20 minutes!

In-Class Exercise Solution

- Systematic utility for consumer i choosing alternative 1: $V_{i1} = \beta Price_1$
 - Intercept (β_{01}) = 0
- Systematic utility for consumer i choosing alternative 2: $V_{i2} = \beta Price_2 + \beta_{02}$
- $\mathcal{LL}(\theta|D) = \sum_{i=1}^N \delta_{i1} \log \left(\frac{\exp(V_{i1})}{\exp(V_{i1}) + \exp(V_{i2})} \right) + \delta_{i2} \log \left(\frac{\exp(V_{i2})}{\exp(V_{i1}) + \exp(V_{i2})} \right)$

β_{02}	β
0.434886855	-1.2880701
Total LogLikelihood	-131.201885

Choice Models

Model Comparison

MNL Application - Megabus

- Data: 210 travelers stated their choice among 4 travel modes (1=Air, 2=Bus, 3=Car, 4=Train)
- Independent variables
 - Time: terminal waiting time
 - Invc: In-vehicle cost
 - Invt: In-vehicle time
 - Hinc: Household income in thousands
- We normalize by dividing each of these variables by 100 in the data, before estimation
 - Note: We could have standardized instead
- You will work on this dataset for the in-class concept check

MNL Travel Data Estimates – Model 1

Parameter	Estimate	Std Error	2.5%	97.5%
intBus	-1.434	0.681	-2.768	-0.099
intCar	-4.740	0.868	-6.440	-3.040
intTrain	-0.787	0.603	-1.968	0.394
Time	-9.689	1.034	-11.716	-7.662
Invc	-1.391	0.665	-2.695	-0.088
Invt	-0.400	0.085	-0.566	-0.233

- Intercept for plane = 0
- Only attributes of options included
- No consumer characteristics

- Time: terminal waiting time,
- Invc: In-vehicle cost
- Invt: In-vehicle time

95% Confidence Interval

MNL Travel Data Estimates – Model 2

Parameter	Estimate	Std Error	2.5%	97.5%
intBus	-0.184	0.897	-1.942	1.573
intCar	-4.247	1.007	-6.220	-2.275
intTrain	1.242	0.817	-0.359	2.843
Time	-9.528	1.036	-11.558	-7.499
InvC	-0.450	0.721	-1.863	0.964
InvT	-0.366	0.087	-0.537	-0.196
Bus_Hinc	-2.311	1.646	-5.537	0.914
Car_Hinc	0.210	1.210	-2.160	2.581
Train_Hinc	-5.590	1.536	-8.600	-2.580

- Intercept for plane = 0
- Plane_Hinc = 0
- Time: terminal waiting time,
- InvC: In-vehicle cost
- InvT: In-vehicle time
- Hinc: Household income in thousands

Model Comparison

- Models can be compared using the Bayes Information Criterion (BIC)

- BIC is given by

$$\text{BIC} = -2 * LL(\theta_{ML}) + K * \ln(N)$$

- K is the total number of parameters estimated
 - BIC penalizes having more parameters
- N is the total number of observations
- Lower BIC = better. Why?
- Model 1: $\text{BIC}_1 = -2 * (-192.89) + 6 * \ln(210) = 417.86$
- Model 2: $\text{BIC}_2 = -2 * (-182.22) + 9 * \ln(210) = 412.564$

Elasticities & IIA

Choice Probability Derivatives

- Probabilities are functions of observed variables
- By **varying variables**, we can analyze how **probabilities vary**
 - How does an increase in price impact purchase probability?
 - How? **derivatives**
- Own Derivatives
 - To what extent will the probability of choosing **Bus** change when we decrease **Bus's** cost?
- Cross Derivatives
 - To what extent will the probability of choosing **Bus** change when we decrease **Train's** price?

Own Derivatives

- Let i be consumer, j be alternative, and m be an attribute (e.g., price)
- P_{ij} is the probability that consumer i chooses j
- Own Derivative: Impact on probability of alternative j when attribute of j is changed

$$\frac{\partial P_{ij}}{\partial x_{ijm}} = \beta_m P_{ij} (1 - P_{ij})$$

- Notice that P_{ij} is function of x_{ijm} . How?
- When is the derivative null?
 - When $P_{ij} = 0$ or $P_{ij} = 1$: no uncertainty in purchase choice
- When is the derivative the highest?
 - When $P_{ij} = 0.5$: uncertainty is the highest

Own Elasticities

- Own choice elasticity is given by

$$\frac{\% \text{ Change in Probability of } j}{\% \text{ Change in an attribute of } j}$$

- Own elasticity is

$$E_{j x_{ijm}} = \frac{\frac{\partial P_{ij}}{P_{ij}}}{\frac{\partial x_{ijm}}{x_{ijm}}} = \frac{\partial P_{ij}}{\partial x_{ijm}} \frac{x_{ijm}}{P_{ij}} = \beta_m P_{ij} (1 - P_{ij}) \left(\frac{x_{ijm}}{P_{ij}} \right) = \beta_m x_{ijm} (1 - P_{ij})$$

How important the attribute is

Value of the attribute

Probability of choosing the alternative

Cross Derivatives

- Let i be consumer, j and k be alternatives, and m be an attribute (e.g., price)
- Cross Derivative: Impact on probability of alternative j when attribute of k is changed

$$\frac{\partial P_{ij}}{\partial x_{ikm}} = -\beta_m P_{ij} P_{ik}$$

- When is the above derivative the highest?
 - Unclear

Cross Elasticities

- Cross Choice elasticities is given by

$$\frac{\% \text{ Change in Probability of } j}{\% \text{ Change in an attribute of } k}$$

- Cross elasticity for j is

$$E_{j x_{ikm}} = \frac{\partial P_{ij}}{\partial x_{ikm}} \frac{x_{ikm}}{P_{ij}} = -\beta_m x_{ikm} P_{ik}$$

- Cross elasticities are the same for all j
 - When k changes its attribute value by 1 percent, it impacts the probabilities of all other alternatives by the same percentage

Elasticities

- Own and Cross elasticities with respect to Invt (In Vehicle Time)
- Effect on the choice probability of the row alternative when the time of the column alternative changes

	Air	Bus	Car	Train
Air	-0.252	0.019	0.023	0.039
Bus	0.013	-0.094	0.023	0.039
Car	0.013	0.019	-0.022	0.039
Train	0.013	0.019	0.023	-0.100

- Notice cross elasticities in each column
 - This pattern is due to Independence of Irrelevant alternatives

Independence of Irrelevant Alternatives

- Independence of Irrelevant Alternatives (IIA)
 - Ratio of choice probabilities between pairs of alternatives is independent of availability or attributes of other alternatives

$$\frac{P_{ij}}{P_{ik}} = \frac{\exp(V_{ij})}{\exp(V_{ik})}$$

- Characteristics of one particular choice alternative do not impact the relative probabilities of choosing other alternatives
- Why?
- Denominator is the same for all probabilities and numerator only depends on alternative

Independence of Irrelevant Attributes

- Suppose consumers are indifferent between a Car and a Red Bus.
- Then $P(car) = 0.5$, $P(RedBus) = 0.5$

$$\frac{P(car)}{P(RedBus)} = 1$$

- The company introduces a Blue Bus: identical to the Red Bus, except for its color (irrelevant attribute)
- → Utilities for car and Red bus don't change
- According to MNL $\frac{P(car)}{P(RedBus)} = 1$ and we expect that $\frac{P(RedBus)}{P(BlueBus)} = 1$
- What are probabilities of all the alternatives?

What are probabilities of all the alternatives?



$P(car)=0.33$, $P(RedBus)=0.33$, and $P(BlueBus)=0.33$

0

$P(car)=0.5$, $P(RedBus)=0.25$, and $P(BlueBus)=0.25$

0

Independence of Irrelevant Attributes

- Suppose consumers are indifferent between a Car and a Red Bus.
- Then $P(car) = 0.5, P(RedBus) = 0.5$

$$\frac{P(car)}{P(RedBus)} = 1$$

- The company introduces a Blue Bus: identical to the Red Bus, except for its color (irrelevant attribute) → Utilities for car and Red bus don't change
- According to MNL $\frac{P(car)}{P(RedBus)} = 1$ and we expect that $\frac{P(RedBus)}{P(BlueBus)} = 1$
- What are probabilities of all the alternatives?
 - $P(car) = 0.33, P(RedBus) = 0.33, \text{ and } P(BlueBus) = 0.33$

Blue Bus - Red Bus

- MNL implies the following choice probabilities

	Car	Red Bus	Blue Bus
Two Alternatives	0.5	0.5	NA
MNL: Three Alternatives	0.33	0.33	0.33
Expected: Three Alternatives	0.5	0.25	0.25

- The new alternative draws proportionally from each of the existing alternatives

IIA

- IIA is beneficial in modeling when choice sets differ across observations
 - Allows addition or removal of an alternative from choice set
 - Why? Structure and parameters of the model won't be impacted
- But
 - Can be problematic in predicting choice shares when new brands are introduced
 - Can give misleading elasticities
 - Alternatives?
 - Nested Logit; Multinomial Probit...

Concept Check In-Class

Choice Modeling

Model 1 – Systematic Utilities

- Elements of model: intercept, time, in-vehicle cost, in-vehicle time
- How many parameters can be identified?
 - 6 (3 intercepts, time, invc, invt)

- Systematic utility for consumer i choosing air:

$$V_{i,air} = \beta_{time}Time_{i,air} + \beta_{invc}Invc_{i,air} + \beta_{invt}Invt_{i,air}$$

- Intercept ($\beta_{0,air}$) = 0
- Systematic utility for consumer i choosing bus:
$$V_{i,bus} = \beta_{bus} + \beta_{time}Time_{i,bus} + \beta_{invc}Invc_{i,bus} + \beta_{invt}Invt_{i,bus}$$
- Similar to above for choosing car or train

Model 1 – Log Likelihood

$$\mathcal{LL}(\theta|D) = \sum_{i=1}^N \delta_{i,air} \log \left(\frac{\exp(V_{i,air})}{\exp(V_{i,air}) + \exp(V_{i,bus}) + \exp(V_{i,car}) + \exp(V_{i,train})} \right) + \\ \delta_{i,bus} \log \left(\frac{\exp(V_{i,bus})}{\exp(V_{i,air}) + \exp(V_{i,bus}) + \exp(V_{i,car}) + \exp(V_{i,train})} \right) + \dots$$

β_{bus}	β_{car}	β_{train}	β_{time}	β_{invc}	β_{invt}
-1.43371	-4.73997	-0.78674	-9.68878	-1.3912	-0.39947

Total LogLike

-192.889

Model 1 – BIC

- BIC is given by

$$\text{BIC} = -2 * LL(\theta_{ML}) + K * \ln(N)$$

- K is the total number of parameters estimated – 6
 - N is the total number of observations – 210
- **BIC = 417.86**

Model 2 – Systematic Utilities

- Elements of model: intercept, time, in-vehicle cost, in-vehicle time, **household income**
- How many parameters can be estimated?
 - 9 (3 intercepts, 3 household income, time, invc, invt)

- Systematic utility for consumer i choosing air:

$$V_{i,air} = \beta_{time}Time_{i,air} + \beta_{invc}Invc_{i,air} + \beta_{invt}Invt_{i,air}$$

- Intercept ($\beta_{0,air}$) = 0, $\beta_{air,hinc} = 0$
- Systematic utility for consumer i choosing bus:
$$V_{i,bus} = \beta_{bus} + \beta_{time}Time_{i,bus} + \beta_{invc}Invc_{i,bus} + \beta_{invt}Invt_{i,bus} + \beta_{bus,hinc}Hinc_i$$
 - Similar to above for choosing car or train

Model 2 – Prediction for Individual 1

β_{bus}	β_{car}	β_{train}	β_{time}	β_{invc}	β_{invt}	$\beta_{bus,hinc}$	$\beta_{car,hinc}$	$\beta_{train,hinc}$
-0.1844	-4.2476	1.2421	-9.5285	-0.4499	-0.3665	-2.3111	0.2103	-5.5896

Id	Choice	Time.air	Invc.air	Inv.t.air	Time.bus	Invc.bus	Inv.t.bus
1	car	0.69	0.59	1	0.35	0.25	4.17

Time.car	Invc.car	Inv.t.car	Time.train	Invc.train	Inv.t.train	Hinc
0	0.1	1.8	0.34	0.31	3.72	0.35

Id	Vair	Vbus	Vcar	Vtrain	Den	Prob(air)	Prob(bus)	Prob(car)	Prob(Train)
1	-7.20663	-5.96901	-4.87866	-5.45675	0.015173	0.04888	0.168508	0.501364	0.281248

Highest probability = car

Model 2 – Elasticities

- Own choice elasticity is given by

$$\frac{\% \text{ Change in Probability of } j}{\% \text{ Change in an attribute of } j}$$

- Own elasticity is

How important the attribute is

Value of the attribute

$$E_{air,inv} = \beta_{inv} inv_{1,air} (1 - P_{1,air}) = -0.36647 * 1(1 - 0.04888) = -0.349$$

Probability of choosing the alternative

Interpretation

A 10% increase in in-vehicle time for air travel reduces the choice probability of flying by 3.49% for Individual 1

Let's Go to Python

Choice Modeling

Takeaways

- Multinomial Logit is the most widely used choice model
- Identification: Only differences in utility matter
 - Need to set one alternative specific constant to zero
 - Need to set the coefficients of the individual characteristics to zero for one alternative
- Can be used to predict brands bought on different purchase occasions
- Can be used to compute own and cross-elasticities
- Beware of IIA

Break

5-minutes

Midterm Review

Week 1 – Marketing Datasets

- Each dataset has pros and cons
 - What are they?
 - What type of question can I answer?
- Important to quickly know what is possible or not with your data

Data Taxonomy

	Primary Data <i>Data that is gathered by the researcher for the purpose of answering a specific question.</i>	Secondary Data <i>Data that was gathered for a purpose other than answering the specific question.</i>
Structured <i>Data that can be easily and meaningfully represented and manipulated in a traditional database (spreadsheet). Typically numeric or “choice” data.</i>	Surveys (ratings, choice) Experiments	Transaction logs Scanner panel data Ad tracking Product usage data
Unstructured <i>Data that cannot be meaningfully stored in a traditional data structure (spreadsheet) without further processing. Examples include text, images, video, and voice.</i>	Focus groups Interviews Surveys (free response) Observation Eye tracking Physiological/neural	Online reviews Social media Most digital content Call logs

Types of Marketing Research

Exploratory Research

(Ambiguous Problem)

“Our sales are declining and we do not know why.”

Descriptive Research

(Aware of Problem)

“What kinds of people are buying our products?”

“Who buys our competitors’ products?”

Causal Research

(Problem Clearly Defined)

“Will buyers purchase more of our product in a new package?”

Week 2 – Segmentation and Targeting

- What is STP?

Deliver the right products, to the right people, in the right way
Targeting Segmentation Positioning

- What type of data can we use for segmentation?
 - Geodemographics, psychographics, behavioral, benefits and needs
- How to implement and interpret results from hierarchical clustering and k-means
 - Basic idea: use similarity in columns to group rows in segments
 - Hierarchical clustering: sequentially join individuals together based on distance until we get one large unique cluster then select number of segments
 - K-means: find groups of data that are the same within and distinct across groups
- You should be able to determine the number of segments and interpret them
- Segments are **L**arge, **I**dentifiable, **D**istinctive, **S**table and **actionable!**
- How to choose a target segment? Opportunity + Competition + Customer + Company “fit”

Week 3 – Segmentation and Positioning

- Dimension reduction techniques – Factor Analysis (PCA)
 - Assume that independent variables are derived from underlying “concepts”
 - Uncover underlying structure between many variables
- Steps to PCA: determining the number of factors and interpreting them
 - Good factors: uncorrelated, capture as much of the original variance as possible
 - Factors are often intuitive, easier to use, and managerially interesting
- Understand the difference between loading, communalities and scores
 - Loadings = how the original variables relate to the factors
 - Communalities = how much variability in the original variables is explained by the factors
 - Scores = translation of original data into factors

Week 4 – Perceptual Maps + Ford Ka + Customer Lifetime Value (CLV)

- Be able to build, interpret and use a perceptual map
- How to conduct an end-to-end marketing strategy
 - Be aware of common potential problems: how to reach target, data limitations,...

- Margin m ; Retention rate r ; Discount rate i ; Acquisition cost AC

$$CLV = m \left(\frac{r}{1 + i - r} \right) - AC$$

- Understand the impact of each parameter on CLV
 - E.g., How much should a company spend to acquire a new account?

Week 5 – CRM + Churn

- Managing the CLV
 - Customer acquisition
 - Acquisitions, affiliation network...
 - Customer expansion
 - Bundling, recommendation (matrix factorization),...
 - Customer retention
 - Causes of churn, double effect of high retention, impact on market share, firm value...
- Collaborative filtering for recommendation systems
- Discrete survival models for estimating customer churn
 - Be able to replicate the logic
 - Geometric
 - Finite mixture model

Questions?



B9651 – Marketing Analytics

Session 7: Text Analysis

Professor Hortense Fong

Last Time

- Choice Modeling
 - Model the choice process of customers
 - Use maximum likelihood estimation to estimate choice models
 - Identification: Only differences in utility matter
 - Need to set one alternative specific constant to zero
 - Need to set the coefficients of the individual characteristics to zero for one alternative
 - Can be used to predict brands bought on different purchase occasions and to compute own and cross-elasticities
 - Limitations of multinomial logit models: beware of IIA

Today: Text Analysis

1. Simple Summaries
2. Sentiment Analysis
3. Topic Modeling
4. Large Language Models

Today's Goals

Understand:

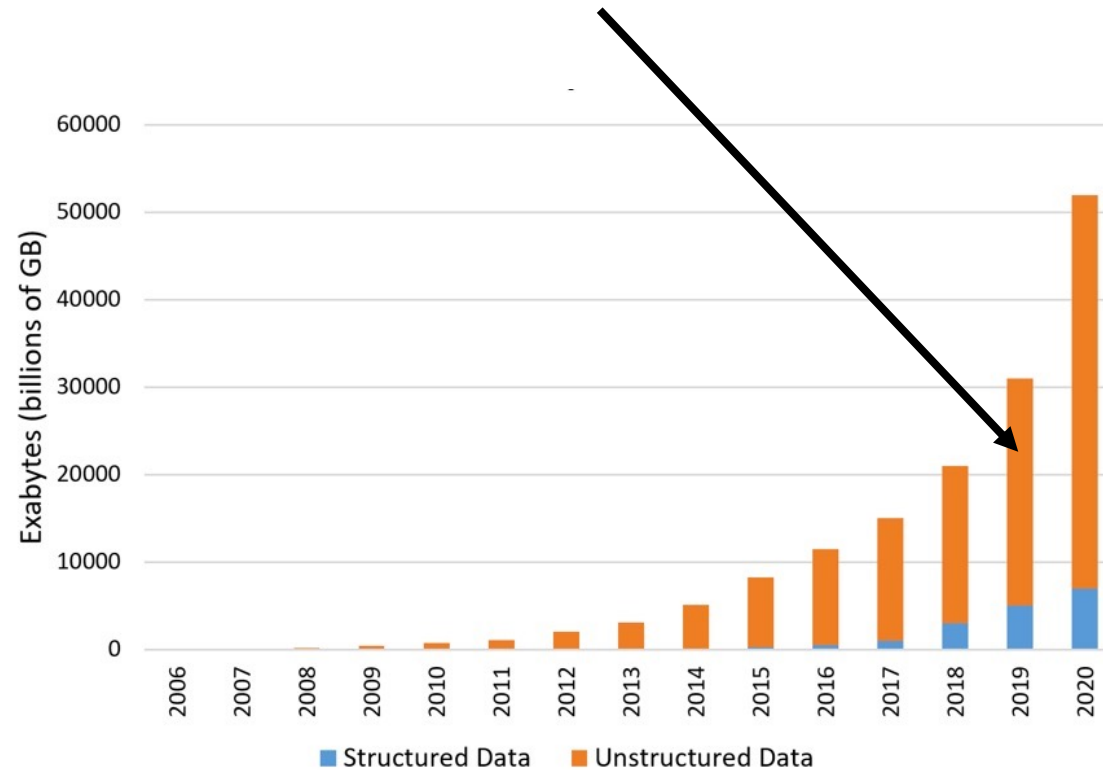
- How to extract insights from unstructured text data

Be able to:

- Implement sentiment analysis and LDA in Python

We have a lot of text data!

Big chunk of that: **text!**



Marketing implications:

- Customer reviews
- Social media
- Feedback
- Product descriptions

Can provide insights that structured data can't provide

Customer Reviews: What Do People Like/Dislike?



Star Wars: The Force Awakens (Theatrical)

MPAA Rating: PG-13 (Parents
Strongly Cautioned)

★★★★★ ~ 10,407



Chuckles

★★★★★ I waited for a good price.

October 6, 2017

Format: DVD | **Verified Purchase**

This movie continues a world changing series of films that literally helped usher in updated 'space adventure' movies.

Really well written and paced. Seems to fix in the timeline accurately and has all the elements you seek from a Star Wars movie. Wonderful vistas, intense and exciting when it should be and fun for anyone over 10 years old.

Like all used items, I wait for a while to see the price averages and when I see a good deal I jump on it. Once you have it, and if you're a sci-fi fan, it'll be on your shelf forever and will be viewed over and over with something new that you missed each time. Buy w confidence.



Amazon Customer

★★★☆☆ Ford does a great impression of himself as a cardboard cut out

August 19, 2016

Format: Blu-ray | **Verified Purchase**

Blah, the movie is nothing more than a remake of A New Hope but with nad acting and clichés of the Sci Fi genre. Ford does a great impression of himself as a cardboard cut out. Abrams is a crap ass director and should be banned from touching any movie franchise.

Customer Reviews: Customer Service



Star Wars: The Force Awakens
(Theatrical)

MPAA Rating: PG-13 (Parents
Strongly Cautioned)

★★★★★ ~ 10,407



John T.

★★★★★ **Disney Movies Anywhere makes this a great purchase**

June 2, 2016

Format: Blu-ray | **Verified Purchase**

This not a review of the film, but a review of the media enclosed.

Excellent job with this package. One purchase and I can play on any device new or old. DVD + BluRay is a good start. But now Disney allows you to own a copy of their films on every digital service you have. Amazon Video, iTunes, VUDU, etc. This is great. Because I am not a fan of playing the game of which service is going to survive. So this creates a feeling of cloud security which I am happy about.

I am also pleasantly surprised that the Extras normally found on the disc alone are present in most of the streaming services that offer this film. I hope more production companies start doing this with their films.



JustMe

★★★☆☆ **Not packaged correctly, DVD warped.**

February 9, 2019

Format: DVD | **Verified Purchase**

Absolutely love the movie, but it's warped, always have to work with it then put it back in my player. It wasn't packaged good at all, no effort was made to protect it from the very cold weather - it was warped when I got it. I paid a lot for it, but not worth the effort of mailing it back.

Social Media Monitoring



↑ Posted by u/moh804 1 day ago

16 ↓

I opened up TS4 today and got a free stuff pack. Does anyone know why?

I've been playing almost everyday. I even played a few hours yesterday. And just now when I opened up the game it said that a pack was downloading. And when I checked, all it says in the actual game screen is "Our gift to you". Does anyone know why they have me a free stuff pack? I'm not complaining, just curious. Thank you!

24 Comments Share Save 100% Upvoted

Reactions to marketing

↑ r/SimsMobile · Posted by u/rampaige666 23 hours ago

26 ↓


So much stuff I don't want or need 🤔 whyyyyyyyyy

Been playing sims mobile for over a year and I get excited for new items being released because I've been collecting the same old shit for a year.

All I wanted was the pierced braids.

I now have.... 7 benches 6 steel chairs 4 summer chill towers 7 patio chairs 3 steamy hot tubs (never got the fancy one) 5 summer umbrella tables 4 summer drink trays 3 grills (plus the mini)

Reviewing product features

 **OddsocksTango** @OddsocksTango · Apr 10

There's a lot of stuff I think should be in [#TheSims](#) base game that I'd gladly take before new expansion packs. Swatch choices that match, stairs that can be curved or built on the diagonal, more base decor like curtains. But more importantly, I'd love to see more diversity in-

2 1 8

Suggesting new features

Goal: Automatic, Interpretable Insights

Why did you come to our store on this occasion?

I just bought

Why did you come to our store on this occasion?

Why did you come to our store on this occasion?

door dining

I'm doing a
and needed

Why did you come to our store on this occasion?

My bathroom sink broke and I

Why did you come to our store on this occasion?

me right
a very
ld me

I was bored waiting for my wife to finish shopping next door.

Given lots of text, we want to...

- **Summarize** what people are talking about:
 - Topics
 - Attitudes / emotions
- Link text to **outcomes**

...without reading everything!

Challenges of Working with Text

- Language is complicated!



Amazon Customer

Sarcasm: great = bad?



Ford does a great impression of himself as a cardboard cut out

August 19, 2016

Format: Blu-ray | **Verified Purchase**

Typos: nad = bad?

Blah, the movie is nothing more than a remake of A New Hope but with nad acting and clichés of the Sci Fi genre. Ford does a great impression of himself as a cardboard cut out. Abrams is a crap ass director and should be banned from touching any movie franchise.

- “Big data” problem: thousands of words, (maybe) few data points
- Self-selection with text online: who is doing the posting?

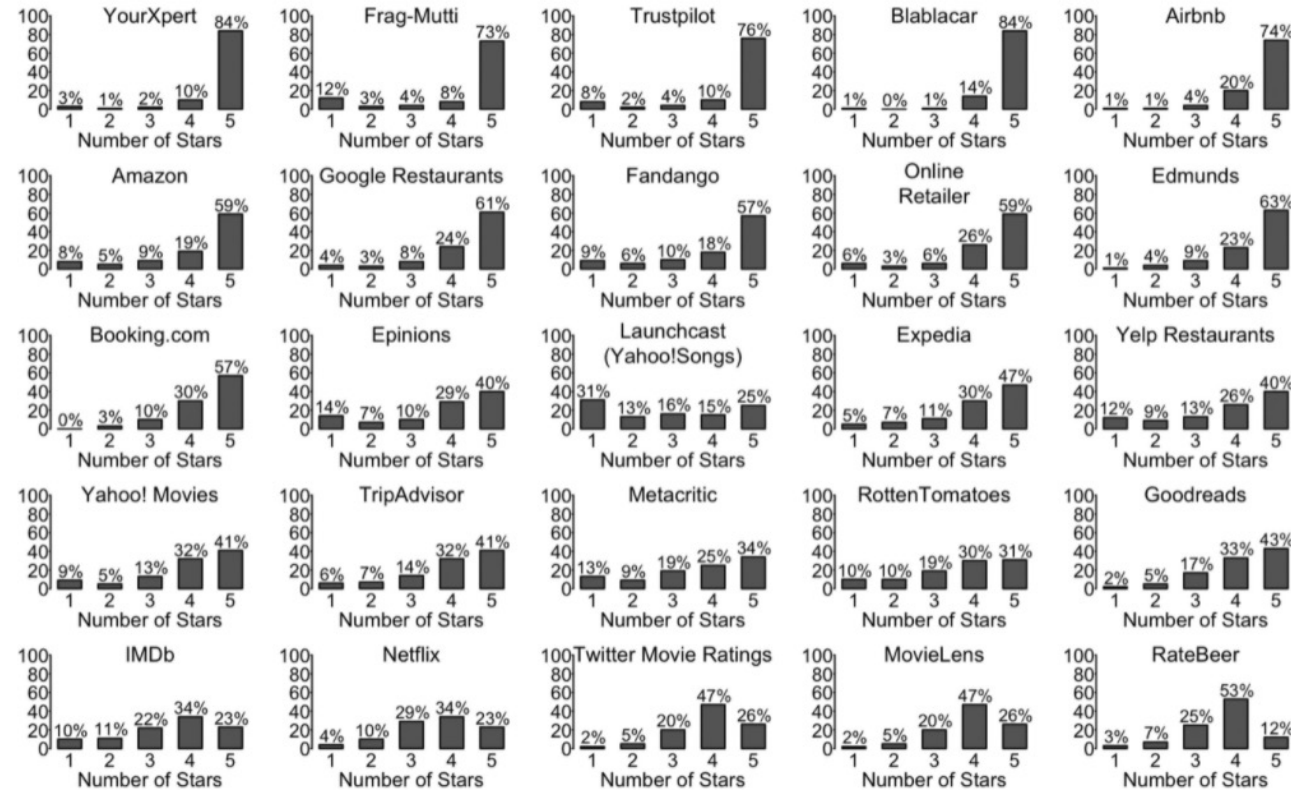
Which common reviews platform exhibits the least "five-star bias"?



Nobody has responded yet.

Hang tight! Responses are coming in.

Ratings Distributions across Platforms



Not everyone who consumes a product rates it, which can lead to selection bias on ratings!

Working with Text as Data

Step 1: Processing Text

The quick fox jumped over the lazy dog.

1. **Tokenization** – convert sentences, paragraphs, documents into units of analysis called *tokens*, typically words or n-grams (groups of n words)
2. **Cleaning** – remove punctuation, convert cases
3. **Stop Word Removal** – remove frequent words that do not carry much information (e.g., the, a, he, she, it, ...)

['The', 'quick', 'fox', 'jumped', 'over', 'the', 'lazy', 'dog.']

['quick', 'fox', 'jumped', 'over', 'lazy', 'dog']

4. **Stemming** – convert words to their base form (or root)

['quick', 'fox', 'jump', 'over', 'lazy', 'dog']

5. **Vocabulary Definition** - remove rare words and/or too common words

Not all these steps are needed! Depends on the problem.

Step 2: Bag-of-Words

- **Bag-of-words** assumption: order of the words doesn't matter

['quick', 'fox', 'jump', 'over', 'lazy', 'dog']



['over', 'quick', 'dog', 'fox', 'lazy', 'jump']

- Allows us to represent the **corpus**, or the collection of all documents, as a **term-document frequency matrix**:

	quick	fox	jump	over	lazy	dog	goat	brand	...
Doc 1	1	1	1	1	1	1	0	0	
Doc 2	1	0	0	0	0	0	2	1	
...	<i>cells = how many times token [column] appears in document [row]</i>								

Shortcoming of Bag-of-Words

“We had an amazing experience, despite the long wait and terrible weather... great food, worth ordering the pasta... No bad dishes on the menu.”

“Despite the great weather and no wait, we had a terrible experience... They have a long menu of pastas, but amazingly bad food, and no dishes worth ordering.”

Bag-of-words tokens: [amazing, experience, despite, long, wait, terrible, weather, great, food, worth, order, pasta, bad, dish, menu]

Possible solution: n-grams - look at groups of n words

- Examples: amazing experience vs. terrible experience, long wait vs. long menu, great food vs. bad food
- Two-word groups = **bi-grams** (most common)

Getting Insights from Text

Four Basic Approaches

1. Simple summaries

- Just look at word counts!
- Examples: 100 most common words, word clouds

2. Sentiment analysis

- Looking for positive or negative words, or other types of *sentiment*
- Example: amazing (positive) vs. terrible experience (negative)

3. Topic modeling

- Extracting *what* each document is about
- Example: “This review is about service quality and food.”

4. Large language models

- All of the above plus more

Text Analysis

Simple Summaries

Word Clouds

- Can limit to top X words
- Size of word in word cloud corresponds to frequency count
- Word cloud of Obama's State of the Union Speech 2016



Source: <https://www.usatoday.com/picture-gallery/news/politics/2014/01/28/what-did-obama-say-his-state-of-the-union-speeches-in-word-clouds/4966671/>

Text Analysis

Sentiment Analysis

Lexicon-based Sentiment Analysis (SA)

- Lexicon = classification of words (word types)
“happy, great, joy” = positive; “sad, terrible, bad” = negative
- For each document, compute the percentage of words falling into a certain category

Customer: I just received the TV, but the screen arrived shattered. Is there anything you can do to help?

...

Customer: Thank you for your help, I look forward to receiving the replacement soon.

Customer: I waited so long for this stupid TV, but the screen is completely shattered! I want a replacement now!!!

...

Customer: Several days?? This is ridiculous. I spent so much money on this!

SA output: **50% Positive**

100% Negative

Choosing the Lexicon

- One popular choice: **LIWC** = Linguistic Inquiry and Word Count (pronounced “Luke”). Examples:

Category	Subcategory	Example Words
Psychological	Positive emotion	Love, nice, sweet
	Negative emotion	Hurt, ugly, nasty
	– Anger	Hate, kill, annoyed
Grammar	Comparisons	Greater, best, after
	Interrogatives	How, when, what
Social	Family	Daughter, dad, aunt
Cognitive	Discrepancy	Should, would
Time	Future focus	May, will, soon

- Custom lexicon: what words matter for you?

Using the Lexicon

Customer: I just received the TV, but the screen arrived shattered. Is there anything you can do to help?

...

Customer: Thank you for your help, I look forward to receiving the replacement soon.



First person pronouns: $2/31 = 0.065$

Second person pronouns: $3/31 = 0.097$

...

Future focus: $2/31 = 0.065$

Affiliation: $2/31 = 0.065$

Money: $0/31 = 0$

Anger: $0/31 = 0$

...

Customer: I waited so long for this stupid TV, but the screen is completely shattered! I want a replacement now!!!

...

Customer: Several days?? This is ridiculous. I spent so much money on this!



First person pronouns: $3/31 = 0.097$

Second person pronouns: $0/31 = 0$

...

Future focus: $0/31 = 0$

Affiliation: $0/31 = 0$

Money: $1/31 = 0.032$

Anger: $2/31 = 0.065$

...

Logistic regression: $Y = \text{Churn}, X = \text{LIWC Proportions}$

Amazon Kindle Reviews

- Suppose you're a Brand Manager for Amazon Kindle. How do you determine whether customers like your product based on review text?

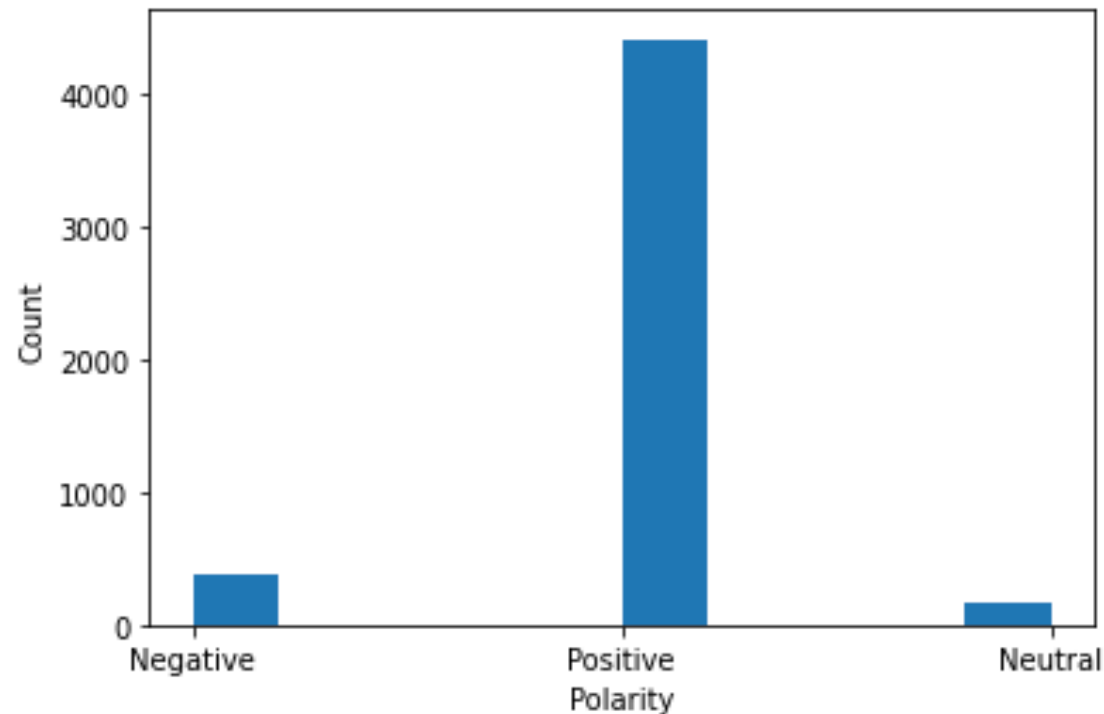
Content	Cleaned Reviews	Final Reviews	Score	Polarity
I really wanted to like the Kindle fire. I'm a...	i really wanted to like the kindle fire i m a ...	realli want like kindl fire long time ebook r...	-0.075067	Negative
love my first kindle so i purchased the fire. ...	love my first kindle so i purchased the fire n...	love first kindl purchas fire happi need conn...	0.285185	Positive

Process text

Apply LIWC or other dictionary/method

Amazon Kindle Reviews

- Suppose you're a Brand Manager for Amazon Kindle. How do you determine whether customers like your product based on review text?



Amazon Kindle Reviews

- Does the polarity correlate with the star rating?

OLS Regression Results						
=====						
Dep. Variable:	Rating	R-squared:	0.107			
Model:	OLS	Adj. R-squared:	0.106			
Method:	Least Squares	F-statistic:	119.8			
Date:	Thu, 28 Sep 2023	Prob (F-statistic):	2.03e-26			
Time:	14:36:16	Log-Likelihood:	-1635.2			
No. Observations:	1000	AIC:	3274.			
Df Residuals:	998	BIC:	3284.			
Df Model:	1					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

Score	2.1385	0.195	10.947	0.000	1.755	2.522
const	3.5122	0.058	60.990	0.000	3.399	3.625
=====						
Omnibus:	141.988	Durbin-Watson:	1.924			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	203.361			
Skew:	-1.082	Prob(JB):	6.93e-45			
Kurtosis:	3.445	Cond. No.	5.21			
=====						

- We can study the sentiment for different product attributes (e.g., screen, service)

Let's Go to Python

Sentiment Analysis

Text Analysis

Topic Modeling

Topic Modeling (Reminder)

- Automatic summarization of documents through **topics**
 - Statistical definition: topic = set of commonly co-occurring words
 - Example: in tablet reviews, “Apple, iPad, iTunes, Mac” = Apple topic
- Intuition: factor analysis for documents! Unsupervised learning
many words → few interpretable topics
- Uses:
 - Information retrieval and automatic labeling
 - Discovering patterns
 - Predicting outcomes from topics
- Most common model: **Latent Dirichlet Allocation (LDA)**
- In Python: sklearn, nltk, gensim

Jargon

- **Corpus** = set of all documents (e.g., set of all reviews)
- **Document** = set of sentences (e.g., sentences in a review)
- Words or Tokens (e.g., “noodles”, “soup”, “service”)

Review 1: I like the noodles here.

Review 2: The noodles are very chewy.

Review 3: The soup is the best part.

•
•
•

Review 100: The service is awful however.

Latent Dirichlet Allocation (LDA)

One use: as input to regression!
"Which topics are predictive of my outcome?"

Output 1: Which words belong to which topics (i.e., what are the topics)?

Note: You have to set the number of topics in advance!

Topic 2: "kindl" "fire" "amazon" "read" "book"	Topic 4 "screen" "good" "touch" "nice" "like"	Topic 7 "great" "product" "love" "purchas" "bought"	Topic 8 "ipad" "like" "much" "appl" "use"	Topic 10 "problem" "work" "day" "back" "tri"
---	--	--	--	---

Output 2: Which topics best describe each document (i.e., what percentage of the words in a given document are from topic 1, topic 2, ...)?



I love my fire and highly recommend it to anyone who wants to watch videos (netflix, hulu, amazon), read ebooks (purchased or from the local library), surf the net and play games. I work in the tech field and I LOVE apple entertainment products (I own many apple products and at work I work with several). I am very thrilled with my fire (I LOVE IT TOO!) because it works great as an entertainment product (and more affordable than my apple products). I also think the fire is a great product because of Amazons cloud and support

Topic	Proportion
1	0.09
2	0.15
3	0.10
4	0.05
5	0.05
6	0.07
7	0.17
8	0.12
9	0.07
10	0.12

Perplexity

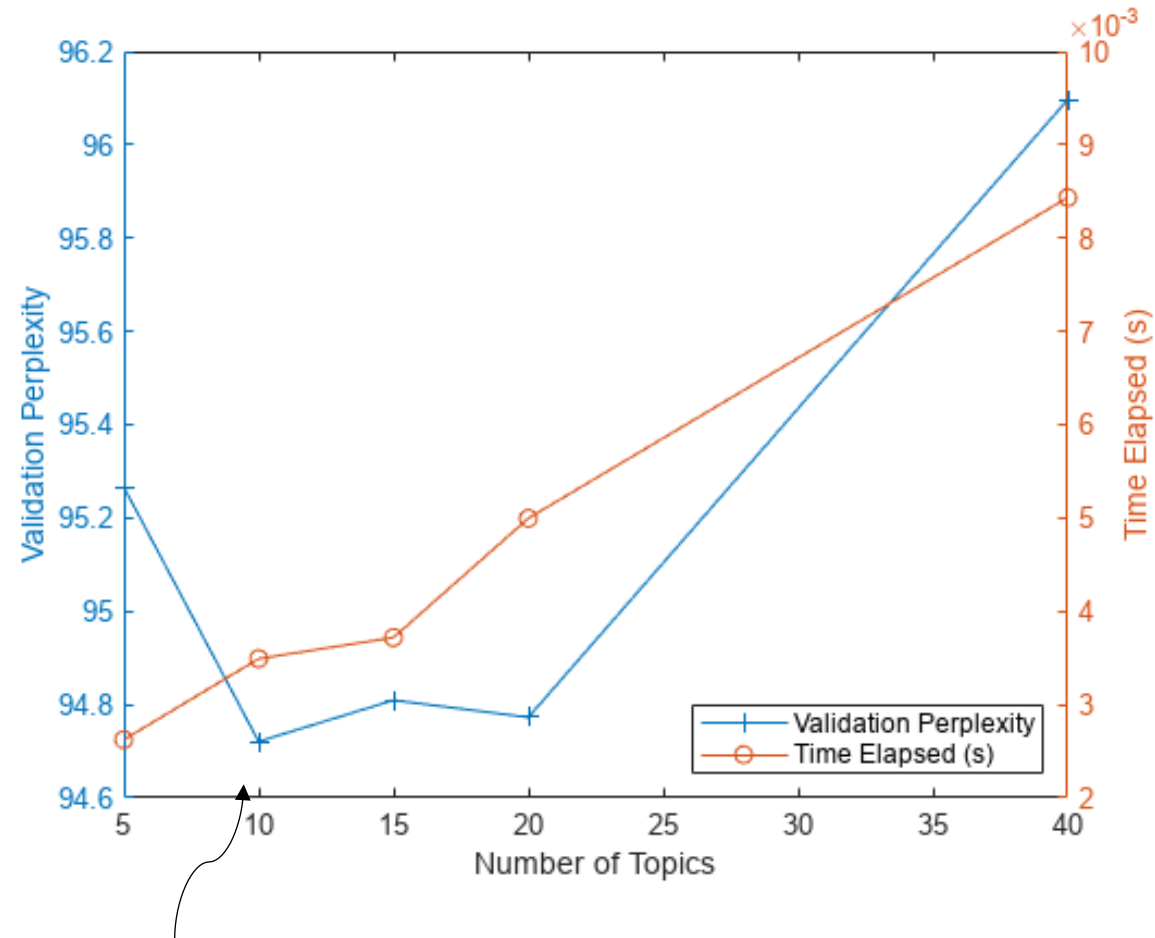
- A language model is an algorithm for assigning a probability to sequences of words
 - My cat likes to _____.
 - $P(\text{sleep} \mid \text{My cat likes to}) > P(\text{bark} \mid \text{My cat likes to})$
- How do we know if a language model is any good? **Perplexity!**

$$\text{perplexity}(X) = \exp \left\{ -\frac{1}{t} \sum_i^t \log p_{\theta}(x_i | x_{<i}) \right\}$$

- Perplexity = measure of predictive performance for a language model
 - Calculated on held-out test data
 - We split our data into two parts, one for training the model and one for testing the model
 - Often U-shaped on test data
 - Lower perplexity = better
 - Choose bottom of the U

Perplexity

Validation perplexity =
perplexity on test data



Lowest perplexity = good number of topics

Source: <https://www.mathworks.com/help/textanalytics/ug/choose-number-of-topics-for-LDA-model.html>

Let's Go to Python

Topic Modeling

Extract Topics from Corpus

```
lda_model.print_topics(num_words=5)
```

```
[(0,  
 '0.039*"tablet" + 0.037*"app" + 0.027*"android" + 0.021*"work" + 0.015*"market"'),  
(1,  
 '0.004*"america" + 0.004*"screenshot" + 0.003*"snag" + 0.003*"applib" + 0.003*"latin"'),  
(2,  
 '0.017*"kindl" + 0.017*"use" + 0.016*"app" + 0.015*"fire" + 0.014*"screen"'),  
(3, '0.005*"tf" + 0.003*"mi" + 0.003*"en" + 0.003*"repair" + 0.002*"rica"'),  
(4,  
 '0.063*"kindl" + 0.041*"fire" + 0.021*"book" + 0.020*"love" + 0.019*"read"')]
```



Weights on each of top five words per topic
Topic “4”: reading on tablet

Assign Topics to Each Document

```
count = 0
for i in range(5):
    print("doc : ",count,lda_model[doc_term_matrix][i])
    count += 1
```

```
doc : 0 [(2, 0.5271308), (3, 0.017094666), (4, 0.45255882)]
```

```
doc : 1 [(4, 0.9807361)]
```

```
doc : 2 [(0, 0.01702472), (1, 0.01666761), (2, 0.017092453), (3, 0.016667554), (4, 0.9325476)]
```

```
doc : 3 [(2, 0.43560666), (4, 0.54356545)]
```

```
doc : 4 [(0, 0.07524313), (2, 0.8612023), (4, 0.061822545)]
```

What is dominant topic for each document?

For document “1”, it is topic “4”

'love my first kindle so i purchased the fire. not happy. you need to be connected to wi-fi all the time it seems. most people assume you get wi-fi, 3-g etc. everywhere. not true!!. great toy but too dependent on wi-fi. On the orig. kindle i open my book and it's there this has to download first-if i'm at home or my office i can not get the book until i go to a hot spot.mine has trouble with the touch screen being too quick or no response.this is for those in the same type of area i'm in. plenty of black spots'

Comparing Methods

Method	Main Purpose	Pros	Cons
Topic Modeling (Latent Dirichlet Allocation)	Automatic summarization and topic identification	<ul style="list-style-type: none">• Easily interpretable• No restrictions on topics• Dimensionality reduction	<ul style="list-style-type: none">• Need to find a good number of topics• Topics can be less nuanced (e.g., may miss sentiment)
Sentiment Analysis (Lexicon-based)	Identifying sentiment, emotion, and other pre-defined concepts	<ul style="list-style-type: none">• Rooted in theory: LIWC = psychology, or custom lexicon• Easily interpretable	<ul style="list-style-type: none">• Only as good as the lexicon• Restricts the patterns you can find

Text Analysis

Large Language Models

What are Large Language Models (LLMs)?

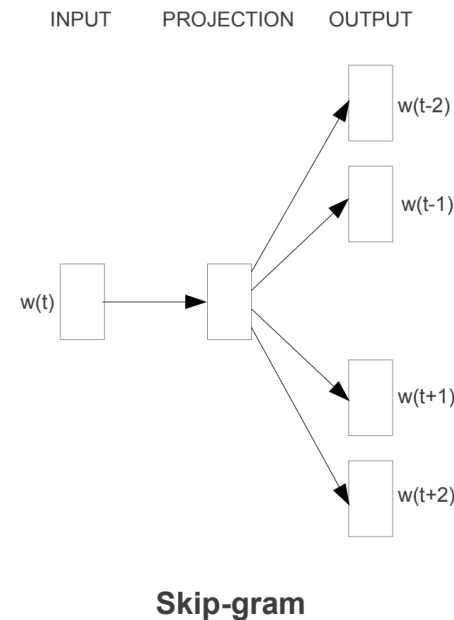
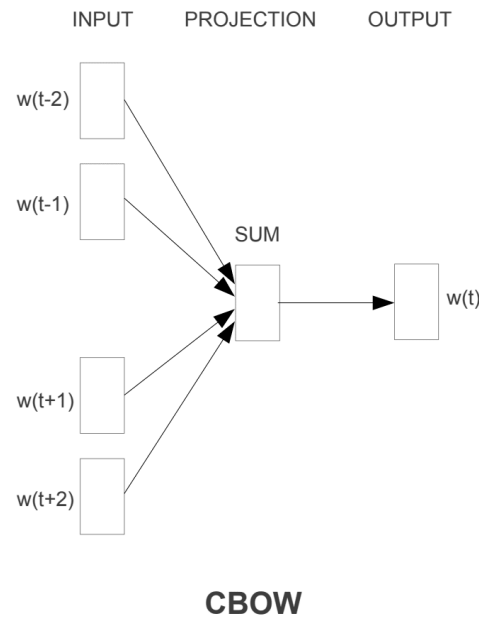
- A category of foundation models trained on immense amounts of data making them capable of understanding and generating natural language and other types of content to perform a wide range of tasks
 - Foundation model: large deep learning model trained on generalized and unlabeled data and used as a starting point for other models
- Training data: books, Wikipedia, news articles, etc. (basically, the Internet)
- Model components:
 - Word embeddings
 - Neural networks (to be discussed later in the semester)

Source: <https://www.ibm.com/topics/large-language-models>

Word Embeddings

- Word embeddings: representation of a word as a vector of numbers
 - Words with similar meaning should have similar representation
 - Word2Vec = one strategy to learn embeddings

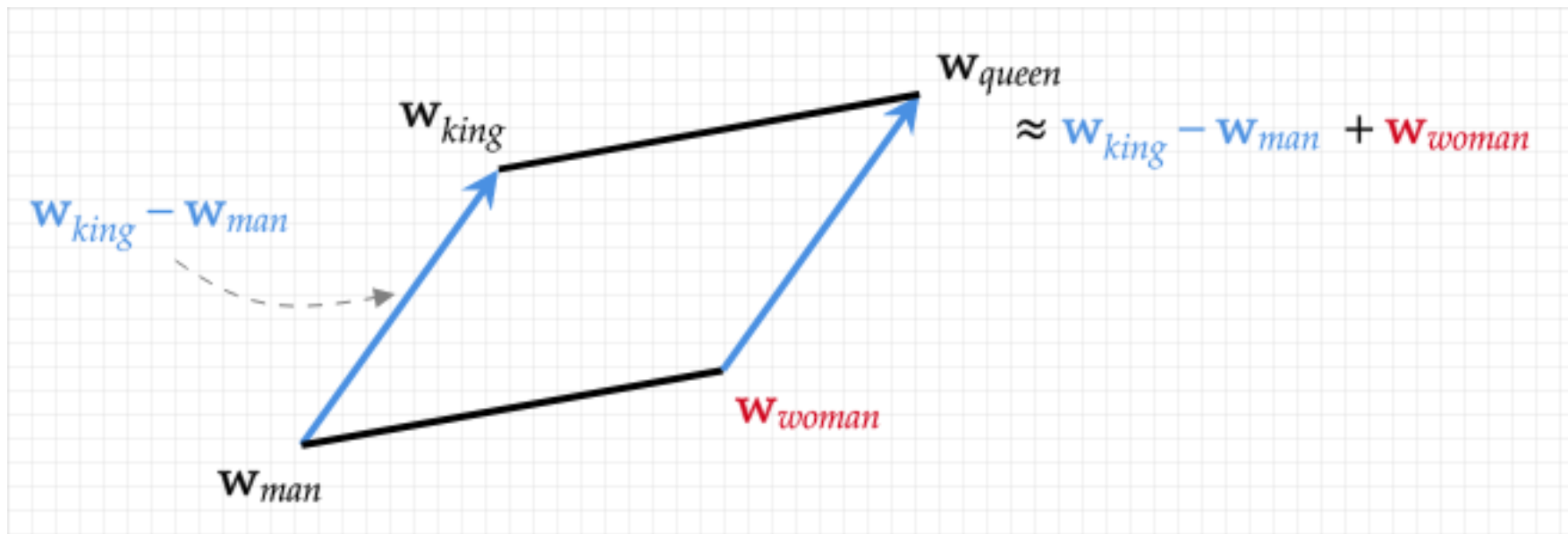
Continuous Bag-of-Words: try to predict current word based on context



Skip-gram: try to predict surrounding words based on current word

Word Embeddings

- Word embeddings: representation of a word as a vector of numbers
 - Capture semantic relationships, can perform word algebra



Source: <https://www.ed.ac.uk/informatics/news-events/stories/2019/king-man-woman-queen-the-hidden-algebraic-struct>

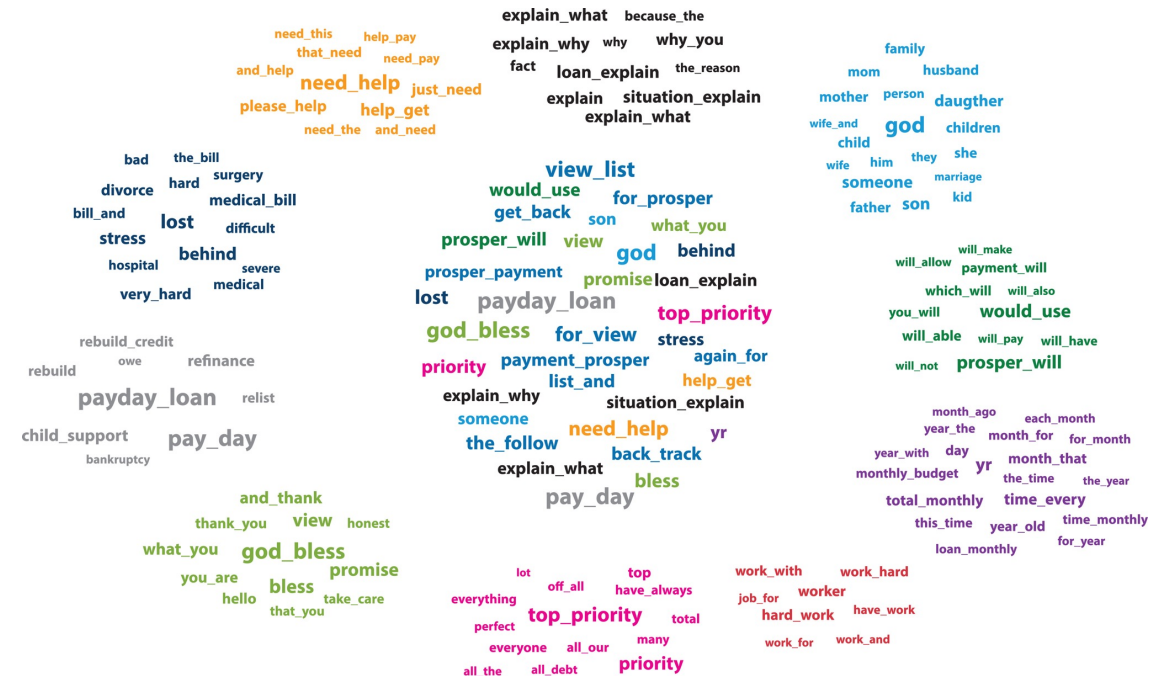
LLM Uses in Marketing

- Popular LLMs:
 - GPT by OpenAI
 - Gemini by Google
 - Claude by Anthropic
- Some uses in marketing
 - Text summarization – summarizing customer reviews, complaints, etc.
 - Sentiment analysis – extracting more nuanced sentiment (e.g., granular emotions)
 - Text generation – creating product descriptions, social media posts, emails

Text Analysis in Research

When Words Sweat: Identifying Signals for Loan Default in the Text of Loan Applications (2019)

- Use LDA to understand topics
- Use LIWC to identify writing styles most correlated with defaulting (e.g., extroverts) or repaying the loan



Result: Loan requests by defaulting borrowers are more likely to include words/themes related to borrower's family, financial and general hardship, mentions of God, mentions of the near future.

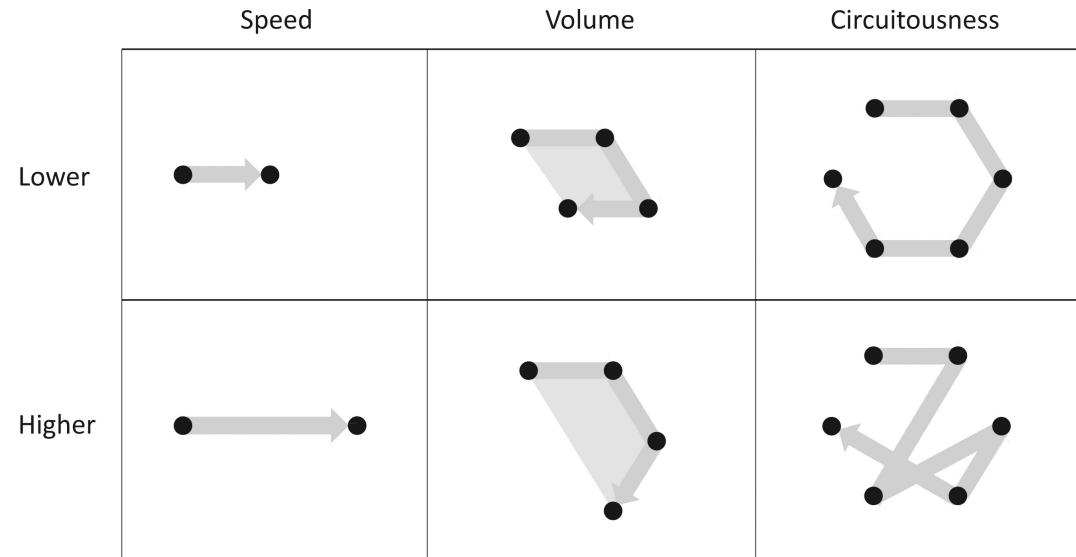
How Quantifying the Shape of Stories Predicts Their Success (2021)

Variables

- Speed
- Volume: covering lots of ground
- Circuitousness: going in circles

Strategy:

1. Chunk text
2. Look at word embeddings of each chunk and quantify above variables



Result: While movies and TV shows that move faster are liked more, TV shows that cover more ground are liked less. Academic papers that move faster are cited less, and papers that cover more ground or are more circuitous are cited more.

Predicting Story Engagement Using LLMs (work in progress)

- Decisions to continue to engage with content (e.g., book, TV series) depend on beliefs about the content to come
- Can we quantify consumers' expectations about the content to come?
- Use an LLM to simulate story continuations and quantify the imagined stories
- We find that the simulated story continuations help predict engagement decisions

Next Class

- Product – Conjoint Analysis
- Guest Speakers
 - Maria Garrido – CMO of Deezer (Tuesday, Oct 29 9:10-10:10AM, virtual)
 - Jeong-Yoon Lee - Senior Manager Applied Science at Uber (Tuesday, Oct 29 4:20-5:20PM, in-person)
 - Kaifu Zhang – Vice President and the Head of AI Initiative at Alibaba Group (Wednesday, Oct 30 9:00-10:00 AM, virtual)

Maria Garrido – CMO Deezer



With 25 years leading successful businesses, Maria is a multilingual, multinational executive with expertise in transformational leadership and purpose-driven initiatives. Renowned for driving cultural change with empathy and optimism, her background spans operational and strategic roles in Consumer Goods, Media, Communications, and Entertainment industries globally.

Maria is currently the Chief Marketing Officer of Deezer, a global leading independent music streaming platform. She oversees global brand strategy, events, market research, digital acquisition, PR & communications, and the internal creative studio. Previously, she served as Senior Vice President of Marketing at Vivendi and as Chief Insights Officer for Havas Group, driving innovation in startup ecosystems, data science and marketing technology.

Maria has also held various board and advisory roles, including Carmila, Live One, ESCP Business School, International Women's Forum and Quilt.ai, an artificial intelligence company studying the human experience.

A sought-after thought leader and prolific speaker, Maria has contributed to Bloomberg News, Cartagena Inspira, Mumbrella Australia, South Tech Summit, World Retail Congress, The Economist, The Guardian, Forbes, and many more. She has also served as president of the Entertainment Jury at Cannes Lions and Eurobest.

Passionate about empowering women, Maria supports female-founded startups, advocates for female survivors of domestic violence and is the founder of Terranam Wellness, an 18th-century manor in Spain dedicated to luxury retreats. Maria holds a Master's in International Business and a Bachelor's in International Relations. She is based in Paris with her three sons.

Jeong-Yoon Lee – Senior Manager Applied Science, Uber



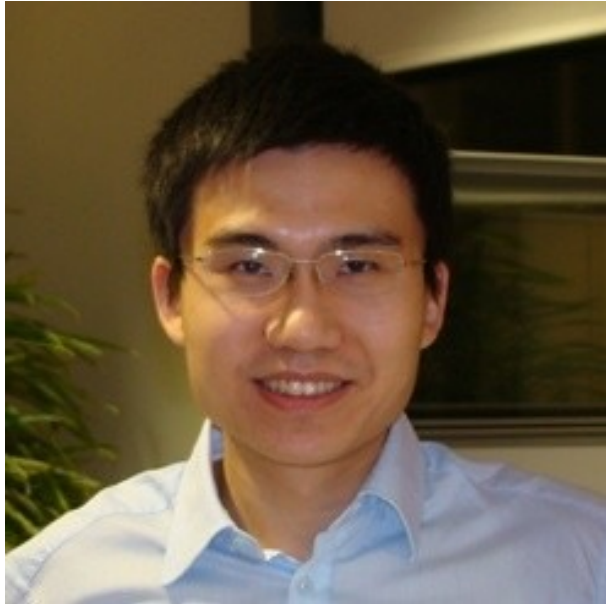
Jeong-Yoon Lee is Sr. Manager in Applied Science at Uber, where he leads research and development in personalization, targeting, optimization, and causal inference for marketing.

Jeong is also leading the CausalML project, an open-source Python package that provides a suite of causal machine-learning algorithms based on recent research. He co-organized the causal machine learning tutorial and workshops at KDD in 2021, 2023, and 2024.

As an avid participant in machine learning competitions, Jeong won the KDD Cup in 2012 and 2015 and was ranked in the top 10 at Kaggle in 2015. He also served as the KDD Cup co-chair at the 2018 ACM SIGKDD conference.

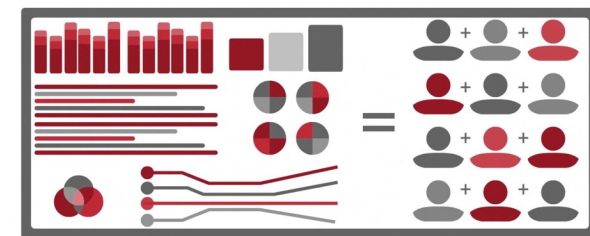
Prior to joining Uber, Jeong was Sr. Research Scientist at Netflix, Sr. Applied ML Scientist at Microsoft, and Chief Data Scientist at Conversion Logic. Jeong received his Ph.D. in Computer Science from the University of Southern California and B.S. in Electrical Engineering from Seoul National University.

Kaifu Zhang – Vice President and the Head of AI Initiative, Alibaba Group



Kaifu Zhang is currently Vice President and the Head of AI Initiative at Alibaba Group. He was Head of Operations of Taobao Marketplace from 2018 to 2021.

Before coming to the industry, he was Assistant Professor and the Xerox Junior Chair at Carnegie Mellon University. He obtained his PhD from INSEAD and BEng from Tsinghua University.



CONJOINTANALYSIS

B9651 – Marketing Analytics

Session 8: Product

Professor Hortense Fong

Last Time

- Text Analytics
 - Simple summaries
 - Sentiment analysis
 - Topic modeling
 - LLMs
- Today: 4P Analytics - Product

Today: Product

Part 1: Conjoint Analysis

1. What is conjoint analysis?
2. How to perform conjoint?
3. Applications
4. Other Approaches

Part 2: Diffusion of Innovation

1. How do we forecast the diffusion of an innovation?
2. Bass Model

Course Roadmap

STP Analytics (Identify Value)	Customer Analytics (Deliver Value)	4P Analytics (Capture Value)
Module 1	Module 2	Module 3
What datasets can we use? How can we segment and target our customers? How should we position our products/services?	How much are our customers worth? Are our customers leaving? How do our customers make choices?	How do we build a new product? How should we price our products? How do we distribute them? How do we quantify the impact of our promotions?

Conjoint Analysis

Big Picture

- The goal of conjoint: understanding trade-offs
 - How do people value product attributes?
 - What are people willing to pay for?
- Possibly the most widely used quantitative marketing research technique
 - Nearly every Fortune 500 company
 - Marketing research firms, consultants
- Use cases:
 - New product development, pricing new products, repositioning existing products, pricing product features, segmentation, new feature development...



Pricing Features – Apple vs. Samsung



- 2011-12: Apple sues Samsung for patent violations
- Five software features: quick links, universal search, background data syncing, slide-to-unlock, autocorrect
- Claim: these features are worth **\$33-40 per phone**, or approximately **\$1.5 billion** total for the 37 million infringing phones
- **How do you determine the value of slide-to-unlock?**

You hire Marketing Professor John Hauser, from MIT Sloan to run a *conjoint analysis*!



New Product Development – Designing a Hotel

- Goal: Develop a new hotel chain for travelers who are not happy with current offerings
 - Marriott was running out of sites to put their typical hotels
- What type of hotel facilities and services should be offered?
- Collaboration: Wharton and Marriott
 - Two Marketing Professors: Paul Green and Jerry Wind



Data Collection

Key hotel features:

- Building size
- Landscaping / Pool
- Food
- Room Size
- Room Quality
- Service standards
- Leisure
- Security

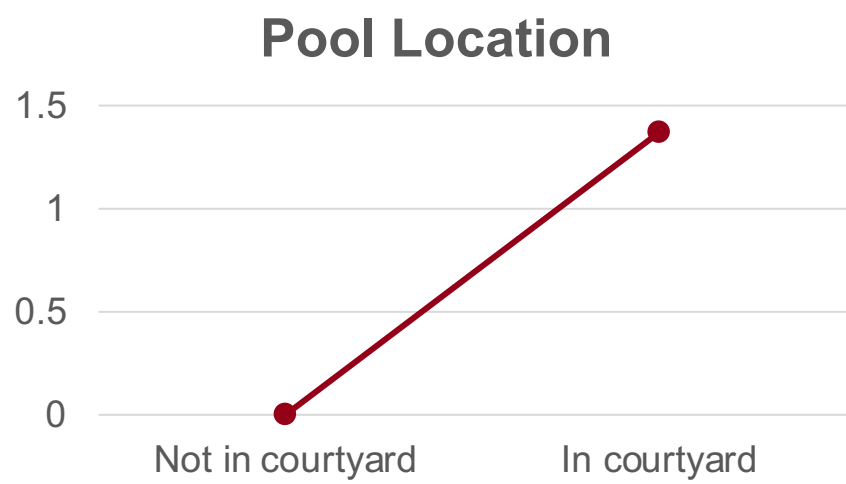
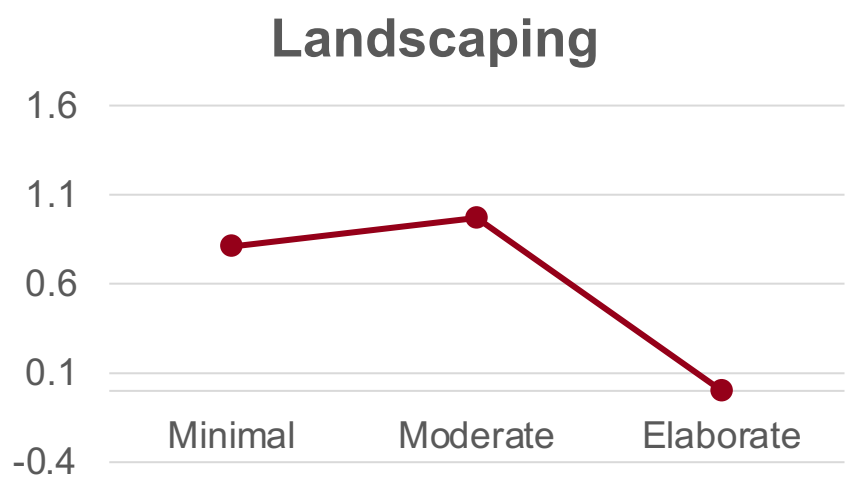
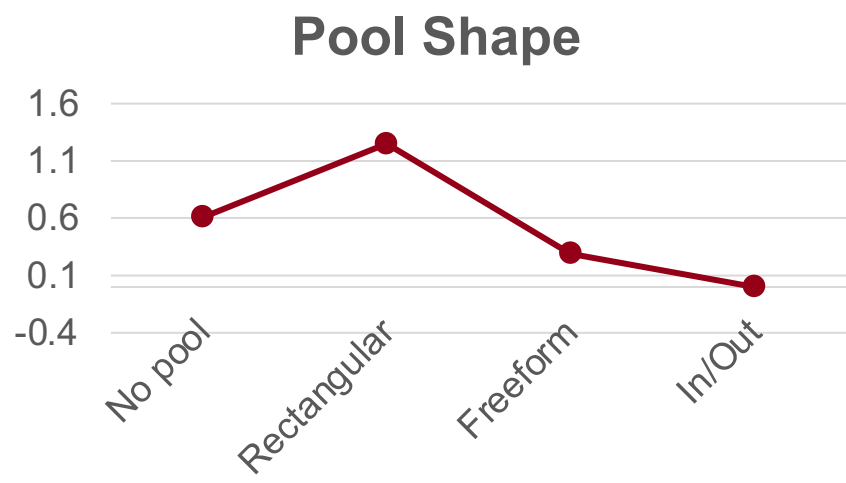
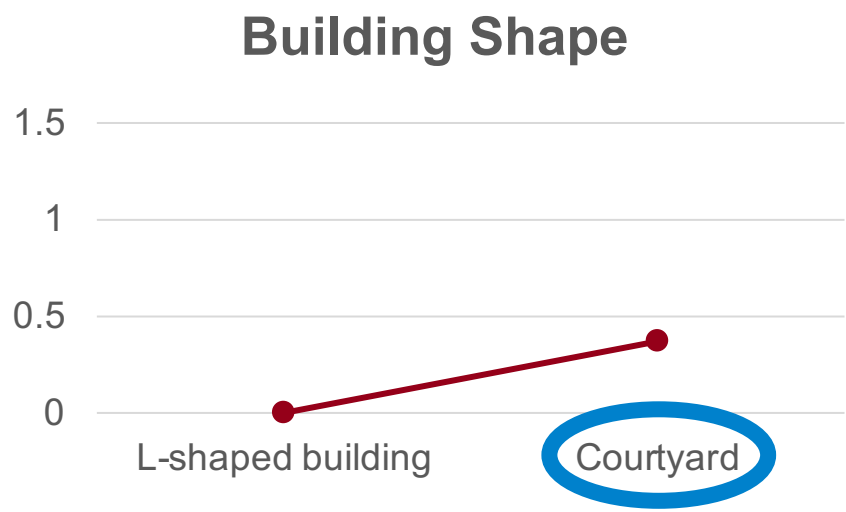
Hotel Option 1

...

Hotel Option 5

ROOM PRICE PER NIGHT IS \$44.85				
BUILDING SIZE, BAR/LOUNGE Large (600 rooms) 12-story hotel with: • Quiet bar/lounge • Enclosed central corridors and elevators • All rooms have very large windows				
LANDSCAPING/COURT Building forms a spacious outdoor courtyard • View from rooms of moderately landscaped courtyard with: — many trees and shrubs — the swimming pool plus a fountain — terraced areas for sunning, sitting, eating				
FOOD Small moderately priced lounge and restaurant for hotel guests/friends • Limited breakfast with juices, fruit, Danish, cereal, bacon and eggs • Lunch—soup and sandwiches only • Evening meal—salad, soup, sandwiches, six hot entrees including steak				
HOTEL/MOTEL ROOM QUALITY Quality of room furnishings, carpet, etc. is similar to: • Hyatt Regencies • Westin "Plaza" Hotels				
ROOM SIZE & FUNCTION Room 1 foot longer than typical hotel/motel room • Space for comfortable sofa-bed and 2 chairs • Large desk • Coffee table • Coffee maker and small refrigerator				
SERVICE STANDARDS Full service including: • Rapid check in/check out systems • Reliable message service • Valet (laundry pick up/deliver) • Bellman • Someone (concierge) arranges reservations, tickets, and generally at no cost • Cleanliness, upkeep, management similar to: — Hyatts — Marriotts				
LEISURE • Combination indoor-outdoor pool • Enclosed whirlpool (Jacuzzi) • Well-equipped playroom/playground for kids				
SECURITY • Night guard on duty 7 p.m. to 7 a.m. • Fire/water sprinklers throughout hotel				
"X" the ONE box below which best describes how likely you are to stay in this hotel/motel at this price:				
Would stay there almost all the time <input type="checkbox"/>	Would stay there almost all the time <input type="checkbox"/>	Would stay there almost all the time <input type="checkbox"/>	Would stay there almost all the time <input type="checkbox"/>	Would stay there almost all the time <input type="checkbox"/> Would stay there on a regular basis <input type="checkbox"/> Would stay there now and then <input type="checkbox"/> Would rarely stay there <input type="checkbox"/> Would not stay there <input type="checkbox"/>

Conjoint Findings



The Result: Courtyard Marriott



Courtyard San Antonio



New Feature - Let's do it together!

NEQTIQ

Search the store

Sign in
Register

0
Cart

CATEGORIES

Computers & Multimedia >

Electronic Components >

Games >

Home & Garden >

Office Supplies >

Other

Tools & DIY >


POPULAR BRANDS

HP

2-Power

Microconnect

HOME / OTHER / LENOVO 4XH0L55005 PROJECTOR MODULE BLACK 4XH0L55005



LENOVO

**Lenovo 4XH0L55005
Projector module Black
4XH0L55005**

★★★★★ (No reviews yet)

[Write a Review](#)

SKU: 4XH0L55005

UPC: 5706998646453

\$303.29

CURRENT STOCK: 1

QUANTITY:

▼

1

▲

Which of the following Tablet PCs would you prefer?



Apple with 12 hour battery life, no built-in projector, 10 inch screen, for \$500

0

Toshiba with 10 hour battery life, built-in projector, 7 inch screen, for \$400

0

Conjoint Analysis

Concepts and Jargon

Conjoint: Basics

Ask yourself: Does this always make sense?



Assumption: products are bundles of independent **attributes**

“Products = the sum of their parts”

- Attributes are composed of distinct discrete **levels**
- Each unique combination of levels for all attributes is a **profile**
- The value derived from a product is called **utility**
- Each attribute level contributes part of the total utility, the **part-worth**
- People pick the product that **maximizes utility**

Product = Sum of Attributes ↔ Utility = Sum of Part-worths

Jargon – Example

amazon.com
Prime

Brand

- ☐ HP
- ☐ Acer
- ☐ Lenovo
- ☐ ASUS
- ☐ Dell
- ☐ Microsoft
- ☐ Apple
- [See more](#)

Price

- Under \$500
- \$500 to \$600
- \$600 to \$700
- \$700 to \$800
- \$800 to \$1000
- \$1000 & Above

Deals

- ☐ Today's Deals

Computer Operating System

- Windows
- Macintosh OS
- Linux
- Chrome OS

Computer Activity Type

- ☐ Gaming
- ☐ Business
- ☐ Personal

Laptop Display Size

- ☐ 17 Inches & Above
- ☐ 16 to 16.9 Inches
- ☐ 15 to 15.9 Inches
- ☐ 14 to 14.9 Inches
- ☐ 13 to 13.9 Inches
- ☐ 12 to 12.9 Inches
- ☐ 11 to 11.9 Inches
- ☐ 11 Inches & Under

← Attribute

} 8 levels

Laptop Display

- ☐ Backlit
- ☐ LCD
- ☐ LED
- ☐ Multi-Touch
- ☐ Touchscreen Enabled

Computer Processor Type

- ☐ AMD A-Series
- ☐ AMD A10
- ☐ AMD A4
- ☐ AMD A6
- ☐ AMD A8
- [See more](#)

Computer RAM Capacity

- ☐ 64 GB & Above
- ☐ 32 GB

One **profile**: HP, \$1000 & Above +
Business + 13 to 13.9 Inches + LCD +
AMD A10 + 32GB

Steps to Conjoint Analysis

1. Study design:

- Attribute selection: what are my attributes and levels?
- Experimental design + surveys: what are my products/profiles?

2. Data collection: administer survey to target market

3. Analysis: regression

- Dependent variable = product ratings (**ratings-based conjoint**) or product choices (**choice-based conjoint - not covered in this class**)
- Independent variables = product attributes + levels
- Coefficients = part-worths

4. Insights: attributes → utility → \$

Conjoint Analysis

Study Design

Attribute Selection

- Attributes in conjoint should be...
 - unambiguous
 - useful for determining choice or preference
 - actionable
- The total number of attributes and levels should be kept low
 - Use qualitative research to decide on attributes / levels
 - 6 is the average number of attributes
 - Number of levels should reflect the sensitivity of the decision-making
- Conjoint is usually the end of the road, not the beginning!

What happens if you are missing an important attribute?

Be very skeptical of the results!

Tablet Computers – Attributes & Levels

- Brand/OS
 - a) Apple (with iOS), b) Samsung (with Android), c) Toshiba (with Android)
- Built-in projector
 - a) Yes, b) No
- Battery life:
 - a) 8 hours, b) 10 hours, b) 12 hours
- Screen size:
 - a) 7 inches, b) 10 inches
- Price
 - a) \$300, b) \$400, c) \$500

$$3 \times 2 \times 3 \times 2 \times 3 =$$

108 possible profiles

Tablet Computers – Profiles – Fractional Factorial Designs

- Problem: Showing each of the $3 \times 2 \times 3 \times 2 \times 3 = 108$ possible profiles (full factorial design) is (basically) infeasible

How do we ask fewer questions, but get the same information?

- Solution: theory of experimental designs—a **fractional factorial design**
 - The minimal number of questions to get the information we need
 - What information do we get? Ability to estimate the “main effects”
- Design determined by the number of attributes / levels
- Tool: Experimental design builder: “Conjoint_Experimental_Designs.xslm” on canvas

Interpreting the Design

Experimental Design Builder							
Parameters		Suggested Design					
Attribute	Levels	Profile	Attribute 1	Attribute 2	Attribute 3	Attribute 4	Attribute 5
1	3	(1)	1	1	1	1	2
2	2	(2)	3	1	2	2	1
3	3	(3)	3	2	2	1	2
4	2	(4)	2	2	3	1	1
5	3	(5)	1	2	1	2	3
6		(6)	3	1	3	2	3
		(7)	1	1	2	2	2
		(8)	2	2	3	2	2
		(9)	2	1	2	1	3
		(10)	2	2	2	2	3
		(11)	1	1	3	1	3
		(12)	1	2	2	1	1
		(13)	2	1	1	2	1
		(14)	2	1	1	1	2
		(15)	3	2	1	1	3

For each respondent in our conjoint study, ask them to rate...

First profile

- Level 1 of Attribute 1
- Level 1 of Attribute 2
- Level 1 of Attribute 3
- Level 1 of Attribute 4
- Level 2 of Attribute 5

Interpreting the Design

Experimental Design Builder							
Parameters		Suggested Design					
Attribute	Levels	Profile	Attribute 1	Attribute 2	Attribute 3	Attribute 4	Attribute 5
1	3	(1)	1	1	1	1	2
2	2	(2)	3	1	2	2	1
3	3	(3)	3	2	2	1	2
4	2	(4)	2	2	3	1	1
5	3	(5)	1	2	1	2	3
6		(6)	3	1	3	2	3
		(7)	1	1	2	2	2
		(8)	2	2	3	2	2
		(9)	2	1	2	1	3
		(10)	2	2	2	2	3
		(11)	1	1	3	1	3
		(12)	1	2	2	1	1
		(13)	2	1	1	2	1
		(14)	2	1	1	1	2
		(15)	3	2	1	1	3

For each respondent in our conjoint study, ask them to rate...

First profile

- Apple
- Built-in projector
- 8h Battery
- 7 inches
- \$400

Interpreting the Design

Experimental Design Builder

Parameters		Suggested Design					
Attribute	Levels	Profile	Attribute 1	Attribute 2	Attribute 3	Attribute 4	Attribute 5
1	3	(1)	1	1	1	1	2
2	2	(2)	3	1	2	2	1
3	3	(3)	3	2	2	1	2
4	2	(4)	2	2	3	1	1
5	3	(5)	1	2	1	2	3
6		(6)	3	1	3	2	3
		(7)	1	1	2	2	2
		(8)	2	2	3	2	2
		(9)	2	1	2	1	3
		(10)	2	2	2	2	3
		(11)	1	1	3	1	3
		(12)	1	2	2	1	1
		(13)	2	1	1	2	1
		(14)	2	1	1	1	2
		(15)	3	2	1	1	3

For each respondent in our conjoint study, ask them to rate...

→ Second profile

- Level 3 of Attribute 1
- Level 1 of Attribute 2
- Level 2 of Attribute 3
- Level 2 of Attribute 4
- Level 1 of Attribute 5

Interpreting the Design

Experimental Design Builder

Parameters		Suggested Design				
Attribute	Levels	Profile	Attribute 1	Attribute 2	Attribute 3	Attribute 4
1	3	(1)	1	1	1	1
2	2	(2)	3	1	2	2
3	3	(3)	3	2	2	1
4	2	(4)	2	2	3	1
5	3	(5)	1	2	1	2
6		(6)	3	1	3	2
		(7)	1	1	2	2
		(8)	2	2	3	2
		(9)	2	1	2	1
		(10)	2	2	2	2
		(11)	1	1	3	1
		(12)	1	2	2	1
		(13)	2	1	1	2
		(14)	2	1	1	1
		(15)	3	2	1	1

For each respondent in our conjoint study, ask them to rate...

→ Second profile

- Toshiba
- Built-in projector
- 10h Battery
- 10 inches
- \$300

Interpreting the Design

Experimental Design Builder						
Parameters		Suggested Design				
Attribute	Levels	Profile	Attribute 1	Attribute 2	Attribute 3	Attribute 4
1	3	(1)	1	1	1	2
2	2	(2)	3	1	2	1
3	3	(3)	3	2	2	1
4	2	(4)	2	2	3	1
5	3	(5)	1	2	1	2
6		(6)	3	1	3	2
		(7)	1	1	2	2
		(8)	2	2	3	2
		(9)	2	1	2	1
		(10)	2	2	2	2
		(11)	1	1	3	1
		(12)	1	2	2	1
		(13)	2	1	1	2
		(14)	2	1	1	1
		(15)	3	2	1	1

For each respondent in our conjoint study, ask them to rate...

Second profile

- Toshiba
- Built-in projector
- 10h Battery
- 10 inches
- \$300

15 questions total

Important: these profiles do not have to be real.

Study Design – Takeaway

- Conjoint is about trade-offs: how do people value the components of products?
- Key ideas:
 - Product = sum of *attributes* (tablet = brand + battery life + ...)
 - Attributes have fixed *levels* (brand = Apple, Samsung, Toshiba)
 - The value of a level to a consumer is called its *part-worth*
 - *Utility* of a *profile* = sum of part-worths
- How do we estimate consumers' part-worths?
 - **Step 1:** Ask them to rate many potential profiles.
 - Which profiles? Fractional factorial design
 - **Step 2:** Analyze the data!

Step 2 – Collect Data

Please rate the following tablet PCs by choosing a number on the seven-point preference scale. Use 1 to represent strong dislike and 7 to represent strong liking.

					1	2	3	4	5	6	7
Brand/OS	Built-in projector	Battery life	Screen size	Price							
Apple	Yes	8 hours	7"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	Yes	10 hours	10"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	No	10 hours	7"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	No	12 hours	7"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	No	8 hours	10"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	Yes	12 hours	10"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	Yes	10 hours	10"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	No	12 hours	10"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	Yes	10 hours	7"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	No	10 hours	10"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	Yes	12 hours	7"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	No	10 hours	7"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	Yes	8 hours	10"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	Yes	8 hours	7"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	No	8 hours	7"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Conjoint Analysis

Analyzing Ratings-based Conjoint

Data

Please rate the following tablet PCs by choosing a number on the seven-point preference scale. Use 1 to represent strong dislike and 7 to represent strong liking.

					1	2	3	4	5	6	7
Brand/OS	Built-in projector	Battery life	Screen size	Price	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	Yes	8 hours	7"	\$400							
Toshiba	Yes	10 hours	10"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Toshiba	No	10 hours	7"	\$400	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	No	12 hours	7"	\$300	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	No	8 hours	10"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	Yes	12 hours	10"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	Yes	10 hours	10"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Samsung	No	12 hours	10"	\$400	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	Yes	10 hours	7"	\$500	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	No	10 hours	10"	\$500	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	Yes	12 hours	7"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	No	10 hours	7"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	Yes	8 hours	10"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Samsung	Yes	8 hours	7"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	No	8 hours	7"	\$500	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

These are ratings for only ONE person

How did that person rate EACH profile =
Dependent Variable

Data – Indicator Variables for Levels

Please rate the following tablet PCs by choosing a number on the seven-point preference scale. Use 1 to represent strong dislike and 7 to represent strong liking.

Brand/OS	Built-in projector	Battery life	Screen size	Price	1	2	3	4	5	6	7
Apple	Yes	8 hours	7"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	Yes	10 hours	10"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Toshiba	No	10 hours	7"	\$400	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	No	12 hours	7"	\$300	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	No	8 hours	10"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	Yes	12 hours	10"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	Yes	10 hours	10"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Samsung	No	12 hours	10"	\$400	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	Yes	10 hours	7"	\$500	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	No	10 hours	10"	\$500	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	Yes	12 hours	7"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	No	10 hours	7"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	Yes	8 hours	10"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Samsung	Yes	8 hours	7"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	No	8 hours	7"	\$500	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Independent Variables

DV

Apple	Samsung	Toshiba	Projector Yes	Projector No	8 hours	10 hours	12 hours	7 inch	10 inch	\$300	\$400	\$500	Rating
1	0	0	1	0	1	0	0	1	0	0	1	0	5
0	0	1	1	0	0	1	0	0	1	1	0	0	6
0	0	1	0	1	0	1	0	1	0	0	1	0	2
0	1	0	0	1	0	0	1	1	0	1	0	0	3
1	0	0	0	1	1	0	0	0	1	0	0	1	4
0	0	1	1	0	0	0	1	0	1	0	0	1	4
1	0	0	1	0	0	1	0	0	1	0	1	0	6
0	1	0	0	1	0	0	1	0	1	0	1	0	2
0	1	0	1	0	0	1	0	1	0	0	0	1	3
0	1	0	0	1	0	1	0	0	1	0	0	1	1
1	0	0	1	0	0	0	1	1	0	0	0	1	5
1	0	0	0	1	0	1	0	1	0	1	0	0	5
0	1	0	1	0	1	0	0	0	1	1	0	0	6
0	1	0	1	0	1	0	0	1	0	0	1	0	4
0	0	1	0	1	1	0	0	1	0	0	0	1	1

These are STILL ratings for only ONE person

Stage 5: Estimate Part-worths – Dummy Coding

Please rate the following tablet PCs by choosing a number on the seven-point preference scale. Use 1 to represent strong dislike and 7 to represent strong liking.

Brand/OS	Built-in projector	Battery life	Screen size	Price	1	2	3	4	5	6	7
Apple	Yes	8 hours	7"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	Yes	10 hours	10"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Toshiba	No	10 hours	7"	\$400	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	No	12 hours	7"	\$300	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	No	8 hours	10"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	Yes	12 hours	10"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	Yes	10 hours	10"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Samsung	No	12 hours	10"	\$400	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	Yes	10 hours	7"	\$500	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	No	10 hours	10"	\$500	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	Yes	12 hours	7"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	No	10 hours	7"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	Yes	8 hours	10"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Samsung	Yes	8 hours	7"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	No	8 hours	7"	\$500	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

The removed variables are called **baseline** levels

Apple	Samsung	Toshiba	Projector Yes	Projector No	8 hours	10 hours	12 hours	7 inch	10 inch	\$300	\$400	\$500	Rating
	0	0	1		1	0		1		0	1		5
	0	1	1		0	1		0		1	0		6
	0	1	0		0	1		1		0	1		2
	1	0	0		0	0		1		1	0		3
	0	0	0		1	0		0		0	0		4
	0	1	1		0	0		0		0	0		4
	0	0	1		0	1		0		0	1		6
	1	0	0		0	0		0		0	1		2
	1	0	1		0	1		1		0	0		3
	1	0	0		0	1		0		0	0		1
	0	0	1		0	0		1		0	0		5
	0	0	0		0	1		1		1	0		5
	1	0	1		1	0		0		1	0		6
	1	0	1		1	0		1		0	1		4
	0	1	0		1	0		1		0	0		1

Three Approaches for Computing Part-worths

One regression **for all** respondents (pooled regression)



Hierarchical Bayes (HB) regression
"The best of both worlds"

Gold standard!



One regression **for each** respondent (individual-level regression)

Three Approaches for Computing Part-worths

One regression **for all** respondents (pooled regression)



Hierarchical Bayes (HB) regression
"The best of both worlds"

Gold standard!

... but beyond this class ☹



One regression **for each** respondent (individual-level regression)

Estimate Part-worths – Regression

- For ratings data, regression can be used to compute the part-worths for the attribute levels
- Attribute levels are specified in terms of dummy variables and the rating score is used as the dependent variable
- Create a “baseline” profile
 - e.g., “Apple,” “No Projector,” “12 hours battery life,” “10 inch screen,” “\$500”
 - Part-worths of the corresponding levels set to 0
 - Part-worths of other levels = deviations from this baseline profile
 - Levels in an attribute captured by dummy variables
 - Utility of baseline profile captured by the intercept

Multiple Regression: Results for One Customer

X

Y

Samsung	Toshiba	Projector	Yes	8 hours	10 hours	7 inch	300	400	Rating
0	0	1	1	1	0	1	0	1	5
0	1	1	1	0	1	0	1	0	3
0	1	0	0	0	1	1	0	1	2
1	0	0	0	0	0	1	1	0	3
0	0	0	0	1	0	0	0	0	5
0	1	1	0	0	0	0	0	0	3
0	0	1	0	0	1	0	0	1	6
1	0	0	0	0	0	0	0	1	4
1	0	1	0	0	1	1	0	0	3
1	0	0	0	0	1	0	0	0	4
0	0	1	0	0	0	1	0	0	6
0	0	0	0	0	1	1	1	0	6
1	0	1	1	1	0	0	1	0	3
1	0	1	1	1	0	1	0	1	2
0	1	0	0	1	0	1	0	0	1

Multiple Regression: Results for One Customer

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.99077057
R Square	0.98162632
Adjusted R Square	0.95712808
Standard Error	0.32707133
Observations	15

ANOVA

	df	SS	MS	F	Significance F
Regression	8	34.291479	4.28643493	40.069256	0.000119
Residual	6	0.6418539	0.10697566		
Total	14	34.933333			

PARTWORTHS		
Intercept		6.45
Brand	Apple	0.00
	Samsung	-2.60
	Toshiba	-3.58
Projector	Yes	0.09
	No	0.00
Battery life	8 hours	-1.17
	10 hours	-0.14
	12 hours	0.00
Screen	7 inch	-0.66
	10 inch	0.00
Price	\$300	0.14
	\$400	-0.03
	\$500	0.00

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	6.45	0.266865	24.1516281	3.311E-07	5.79222954	7.0982199	5.79222954	7.0982199
Samsung	-2.60	0.202898	-12.806014	1.393E-05	-3.09478812	-2.10184109	-3.09478812	-2.10184109
Toshiba	-3.58	0.2233426	-16.0231503	3.754E-06	-4.12515128	-3.03215209	-4.12515128	-3.03215209
Projector Yes	0.09	0.1720424	0.54696459	0.6041343	-0.32687156	0.51507381	-0.32687156	0.51507381
8 hours	-1.17	0.2233426	-5.24462612	0.0019298	-1.71784791	-0.62484872	-1.71784791	-0.62484872
10 hours	-0.14	0.2158158	-0.67030723	0.5276043	-0.67274527	0.38341943	-0.67274527	0.38341943
7 inch	-0.66	0.1720424	-3.81242486	0.0088385	-1.07687156	-0.23492619	-1.07687156	-0.23492619
300	0.14	0.2158158	0.67030723	0.5276043	-0.38341943	0.67274527	-0.38341943	0.67274527
400	-0.03	0.202898	-0.13152122	0.8996623	-0.52315891	0.46978812	-0.52315891	0.46978812

Multiple Regression: Results for One Customer

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.99077057
R Square	0.98162632
Adjusted R Square	0.95712808
Standard Error	0.32707133
Observations	15

ANOVA

	Regression	Residual	Total
	8	6	14
	34.291475	0.6418539	34.933333
	4.28643493	0.10697566	
	43.069256		
	0.000119		

Coefficients = Part-worths!

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	6.45	0.266865	24.1516281	3.311E-07	5.79222954	7.0982199	5.79222954	7.0982199
Samsung	-2.60	0.202898	-12.806014	1.393E-05	-3.09478812	-2.10184109	-3.09478812	-2.10184109
Toshiba	-3.58	0.2233426	-16.0231503	3.754E-06	-4.12515128	-3.03215209	-4.12515128	-3.03215209
Projector Yes	0.09	0.34655455	0.2568453	0.6041343	-0.32687156	0.51507381	-0.32687156	0.51507381
8 hours	-1.17	0.2233426	-5.24466612	0.0000000	-1.61784531	-0.72178109	-1.61784531	-0.72178109
10 hours	-0.14	0.2158155	-0.6487913	0.5193238	-0.5687156	0.38912303	-0.5687156	0.38912303
7 inch	-0.66	0.1720424	-3.81232486	0.0008285	-1.0687156	-0.23192619	-1.0687156	-0.23192619
300	0.14	0.2158155	0.6487913	0.5193238	-0.5687156	0.38912303	-0.5687156	0.38912303
400	-0.03	0.202898	-0.13152122	0.8996623	-0.52315891	0.46978812	-0.52315891	0.46978812

PARTWORTHS		
Intercept		6.45
Brand	Apple	0.00
	Samsung	-2.60
	Toshiba	-3.58
Projector	Yes	0.09
	No	0.00
Battery life	8 hours	-1.17
	10 hours	-0.14
	12 hours	0.00
Screen	7 inch	-0.66
	10 inch	0.00
Price	\$300	0.14
	\$400	-0.03
	\$500	0.00

Baseline
part-worth
=
zero

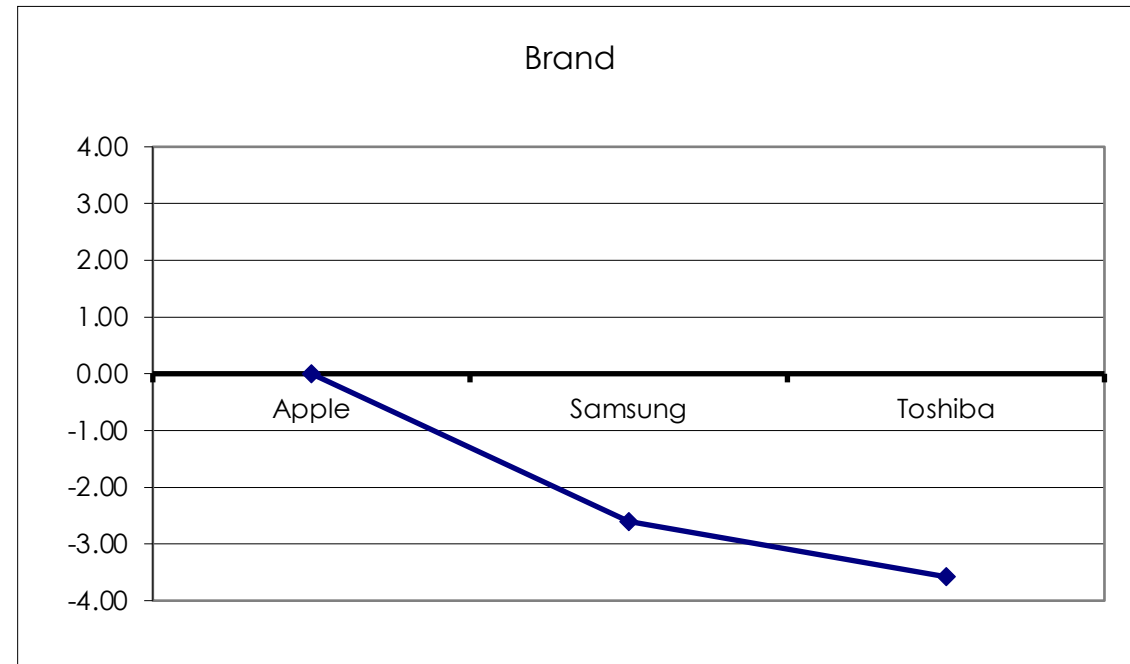
Interpretation: “gain in **utils** for this customer when the product has that feature, relative to the baseline”

Conjoint Analysis

Individual-level Insights

Part-worth Plots

Intercept		6.45
Brand	Apple	0.00
	Samsung	-2.60
	Toshiba	-3.58
Projector	Yes	0.09
	No	0.00
Battery life	8 hours	-1.17
	10 hours	-0.14
	12 hours	0.00
Screen	7 inch	-0.66
	10 inch	0.00
Price	\$300	0.14
	\$400	-0.03
	\$500	0.00



Ideal profile – Utility = $6.45 + 0 + 0.09 + 0 + 0 + 0.14 = 6.68$

Price Part-worths

Price	Part-worth
\$300	0.14
\$400	-0.03
\$500	0.00



Feature Importance

Which feature is most important to this customer?



Which attribute could change this customer's total utility by the largest amount?

PARTWORTHS		
Intercept		6.45
Brand	Apple	0.00
	Samsung	-2.60
	Toshiba	-3.58
Projector	Yes	0.09
	No	0.00
Battery life	8 hours	-1.17
	10 hours	-0.14
	12 hours	0.00
Screen	7 inch	-0.66
	10 inch	0.00
Price	\$300	0.14
	\$400	-0.03
	\$500	0.00

Importance = Range
Attribute's max level – min level

$$\text{Range} = 0.00 - (-3.58) = 3.58$$

$$\text{Range} = 0.09 - 0.00 = 0.09$$

$$\text{Range} = 0.00 - (-1.17) = 1.17$$

$$\text{Range} = 0.00 - (-0.66) = 0.66$$

$$\text{Range} = 0.14 - (-0.03) = 0.17$$

Brand is the most important!

Brand > Battery Life > Screen > Price > Projector

Percentage Importance

PARTWORTHS		
Intercept		6.45
Brand	Apple	0.00
	Samsung	-2.60
	Toshiba	-3.58
Projector	Yes	0.09
	No	0.00
Battery life	8 hours	-1.17
	10 hours	-0.14
	12 hours	0.00
Screen	7 inch	-0.66
	10 inch	0.00
Price	\$300	0.14
	\$400	-0.03
	\$500	0.00

Range = 3.58

Range = 0.09

Range = 1.17

Range = 0.66

Range = 0.17

Total = 5.67

Total Utility Variation = Sum of all Ranges

% Importance = Range / Sum of all Ranges

Intuition: What percentage of the total utility variation is this attribute responsible for?

Brand = $3.58 / 5.67 = 63.1\%$

Projector = $0.09 / 5.67 = 1.6\%$

Battery life = $1.17 / 5.67 = 20.6\%$

Screen = $0.66 / 5.67 = 11.6\%$

Price = $0.17 / 5.67 = 3.0\%$

Valuing Features

The Apple v. Samsung question: how much is a feature worth?

Next week (when we talk about pricing)!

Predicting Choices

Note: I'm ignoring the regression intercept here. If you include it, utility is the same as expected rating.

PARTWORTHS		
Intercept		6.45
Brand	Apple	0.00
	Samsung	-2.60
	Toshiba	-3.58
Projector	Yes	0.09
	No	0.00
Battery life	8 hours	-1.17
	10 hours	-0.14
	12 hours	0.00
Screen	7 inch	-0.66
	10 inch	0.00
Price	\$300	0.14
	\$400	-0.03
	\$500	0.00

Remember:

- Utility of a profile = Sum of part-worths
- Consumers maximize utility

Which product would our focal consumer choose?

A. Apple, No Projector, 12h, 10 inch, \$500

$$\text{Utility of A} = 0 + 0 + 0 + 0 + 0 = 0$$

B. Toshiba, Projector, 10h, 7 inch, \$400

$$\text{Utility of B} = -3.58 + 0.09 - 0.14 - 0.66 - 0.03 = -4.32$$

Note: You can also use logistic probabilities – next week!

Your turn!

1. Download “Conjoint_Regression_Analysis tablet PCs.xlsx” from canvas
2. Download “conjoint analysis data.xlsx” from canvas and find your ratings (using your uni)
3. Copy and paste your ratings (columns B to P) to “Conjoint_Regression_Analysis tablet PCs.xlsx” (cells I2:I16 in “Data” tab—paste special “transpose”)
4. Re-run the regression with your ratings, copy and paste results to “Analysis” tab
 - Regression Tutorial <https://www.ablebits.com/office-addins-blog/2018/08/01/linear-regression-analysis-excel/>
5. Review results and choice prediction (rows 102-103). Did the model correctly predict your choice from the beginning of class?

Did the model correctly predict your choice?

 0

Yes

0

No

0

Conjoint Analysis

Insights – All Respondents

Big picture: What do we want to learn?

Data: ratings from representative set of customers



- How important is each attribute, on average, for this market?
- What segments exist within this market, in terms of attribute importance?
- Which new product configuration would draw the most market share from existing options? (i.e., market penetration strategy)
- How can I determine the profit-maximizing price for my new product?

Big picture: What do we want to learn?

Data: ratings from representative set of customers

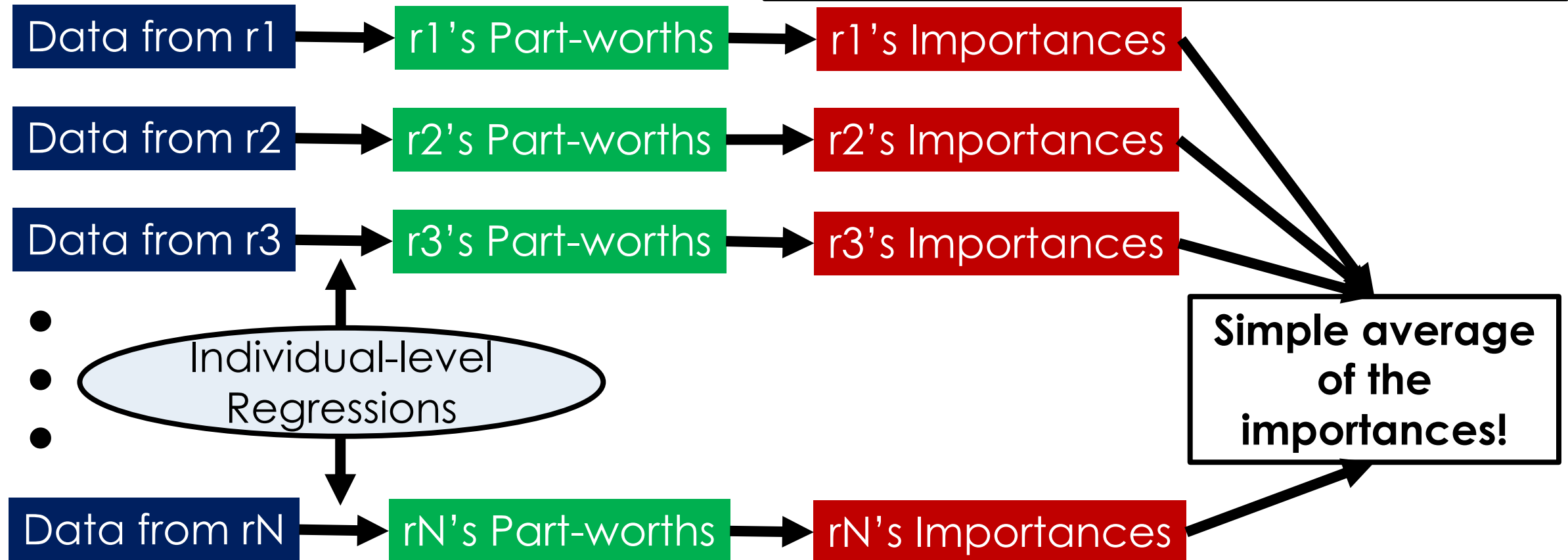


- How important is each attribute, on average, for this market?
- ~~What segments exist within this market, in terms of attribute importance?~~
← Earlier in the course!
- Which new product configuration would draw the most market share from existing options? (i.e. market penetration strategy)
- How can I determine the profit-maximizing price for my new product?

Aggregating Insights: Importance

$r1 = \text{respondent 1}$

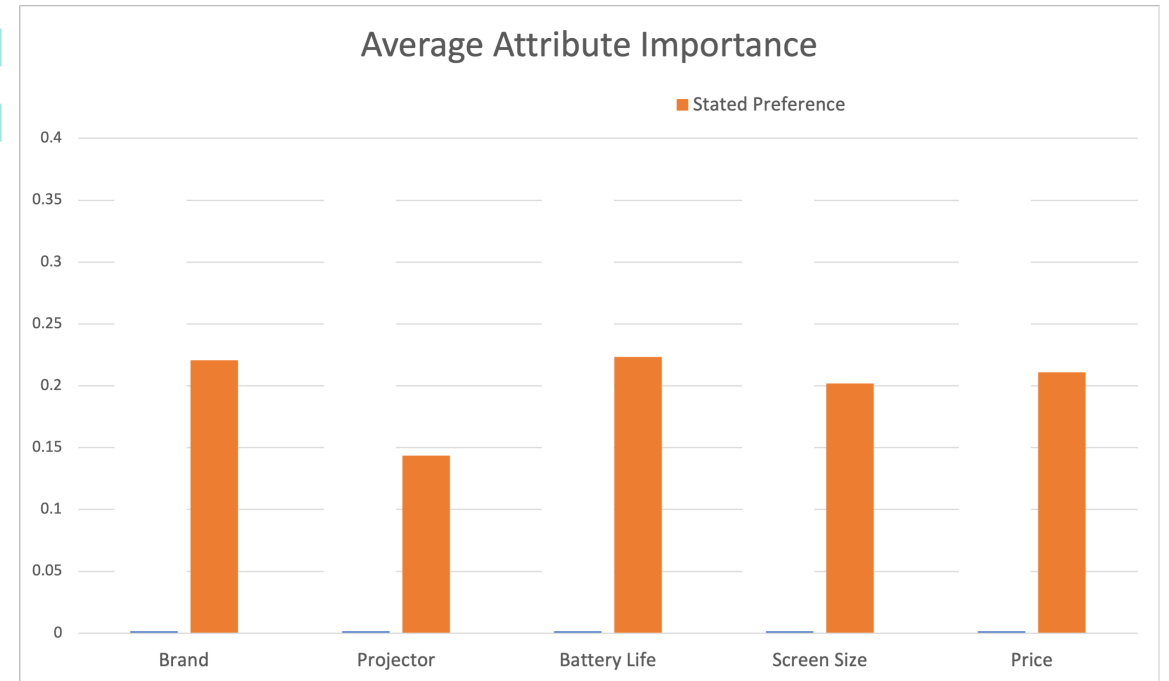
Question: How important is each attribute, on average, for this market?



Stated Importance

Rate how much importance you would give to the following features when considering tablet PCs. Use 1 to represent not at all important and 7 to represent extremely important.

	1	2	3	4	5	6	7
Brand/OS	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Built-in projector	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Battery life	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Screen size	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Price	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

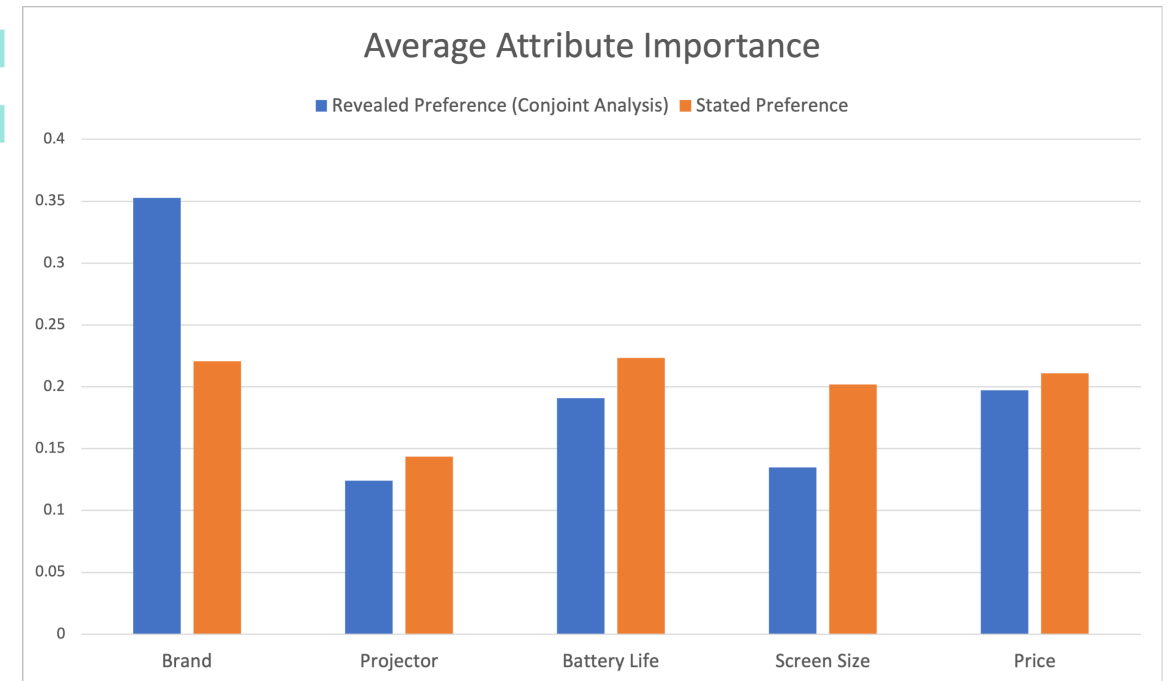


“Revealed Preferences” versus “Stated Preferences”

Rate how much importance you would give to the following features when considering tablet PCs. Use 1 to represent not at all important and 7 to represent extremely important.

	1	2	3	4	5	6	7
Brand/OS	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Built-in projector	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Battery life	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Screen size	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Price	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Brand/OS	Built-in projector	Battery life	Screen size	Price	1	2	3	4	5	6	7
Apple	Yes	8 hours	7"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	Yes	10 hours	10"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	No	10 hours	7"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	No	12 hours	7"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	No	8 hours	10"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	Yes	12 hours	10"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	Yes	10 hours	10"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	No	12 hours	10"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	Yes	10 hours	7"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	No	10 hours	10"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	Yes	12 hours	7"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apple	No	10 hours	7"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	Yes	8 hours	10"	\$300	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samsung	Yes	8 hours	7"	\$400	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toshiba	No	8 hours	7"	\$500	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

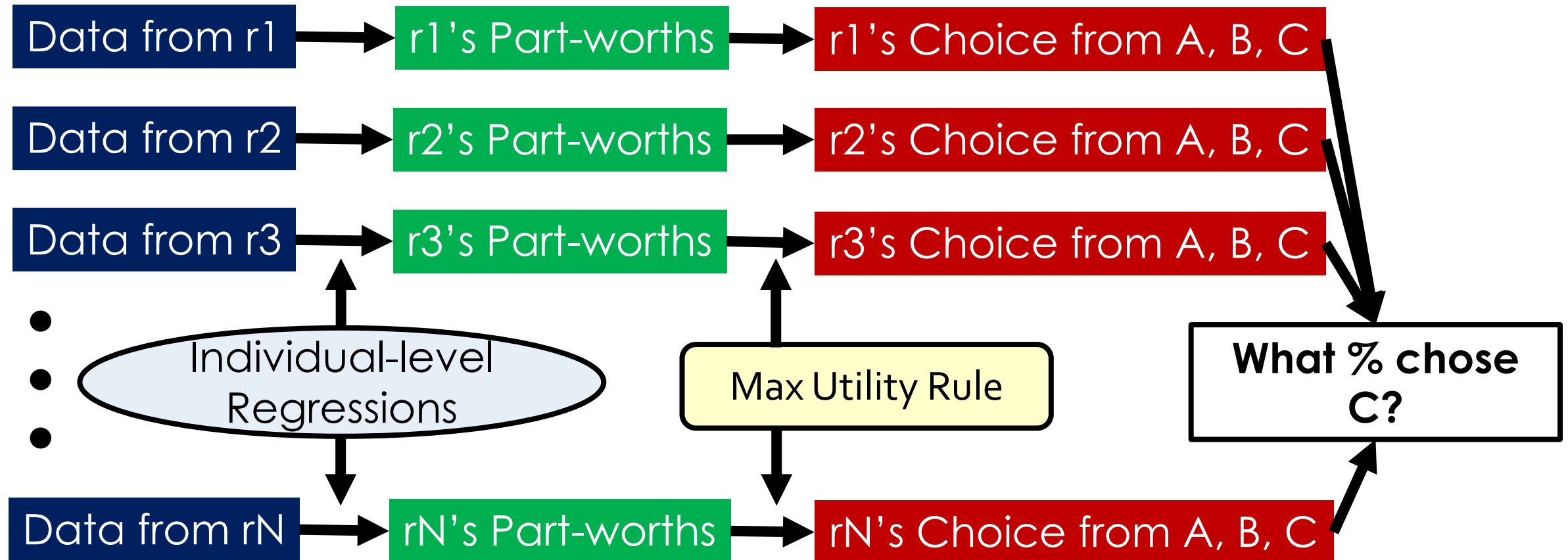


People don't always know what they truly care about

Aggregating Insights: Market Share

$r1$ = respondent 1

Question: Suppose products A and B are in the market already. How much market share will we gain if we introduce C?



Market-level analysis: market share predictions

- Define competitive set
- Predict the choice of each consumer in the sample
- Market share for each product is estimated by the proportion of consumers in the sample who would choose that product
- See “Conjoint_Market_Predictor.xlsm” on canvas

Big picture: What do we want to learn?

Data: ratings from representative set of customers



- How important is each attribute, on average, for this market?
 - Simple average
- ~~What segments exist within this market, in terms of attribute importance?~~
← Earlier in the course!
- Which new product configuration would draw the most market share from existing options? (i.e., market penetration strategy)
 - Simulation by computing market shares
- How can I determine the profit-maximizing price for my new product?
 - Simulation by changing prices

Limitations of Conjoint

- You can never include all the attributes

Apple vs. Samsung, Prof. David Reibstein's critique:

“...the research was flawed because it did not include certain important attributes such as brand name, operating system, battery life etc. He memorably said that not including these fundamental features is like doing research ‘on cupholders in cars while ignoring the manufacturer name.’ ”

Source: <https://www.trchome.com/blog/entry/how-apple-samsung-and-conjoint-came-together>

Summary: Ratings-based Conjoint

- Conjoint is about trade-offs: how do people value the components of products?
- Study design: people **rate** a set of carefully chosen profiles
- Individual-level dummy variable regressions → part-worths
- Individual-level part-worths can be aggregated to understand:
 - What do people value?
 - What segments exist in preferences?
 - What is the ideal “disruptive” product?
 - How should I price my product for maximal penetration?

More resources



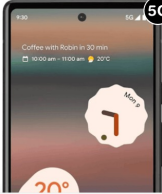



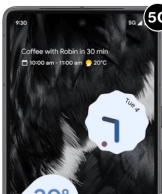



- Sawtooth's collection of technical papers:
<https://www.sawtoothsoftware.com/support/technical-papers>

Some good review ones:

- Understanding Conjoint in 15 Minutes:
<https://www.sawtoothsoftware.com/support/technical-papers/general-conjoint-analysis/understanding-conjoint-analysis-in-15-minutes-1996>
- Understanding the Value of Conjoint Analysis:
<https://www.sawtoothsoftware.com/support/technical-papers/general-conjoint-analysis/understanding-the-value-of-conjoint-analysis-2009>
- CBC Technical Paper:
<https://www.sawtoothsoftware.com/support/technical-papers/cbc-related-papers/cbc-technical-paper-2013>
- Warren Kuhfeld's technical note (using SAS):
<http://support.sas.com/techsup/technote/mr2010.pdf>

Other Conjoint Approaches

Choice-Based Conjoint


	<p>Apple iPhone SE (2022) 64GB</p> <p>2022 version of the iPhone SE features 5G connectivity, a powerful A15 Bionic chip, 4.7"...</p>	<p>4.7 inch display</p> <p>Rear Camera: 12MP</p> <p>Front Camera: 7MP</p> <p>Apple iOS 15.3</p> <p></p>
	<p>Google Pixel 6a 128GB</p> <p>Google's new 2022 model powered by Google's first processor, Tensor, and their Titan M2™...</p>	<p>6.1 inch display</p> <p>Rear Cameras: 12MP, 12MP</p> <p>Front Camera: 8MP</p> <p>Android 12</p> <p></p>
	<p>Google Pixel 6 128GB</p> <p>Google's new 2021 model powered by Google's first processor, Tensor, and their Titan M2™...</p>	<p>6.4 inch display</p> <p>Rear Cameras: 12MP, 50MP</p> <p>Front Camera: 8MP</p> <p>Android 12</p> <p></p>
	<p>Google Pixel 7 128GB</p> <p>Google's 2022 model powered by Google's newest processor, Tensor G2, featuring...</p>	<p>6.3 inch display</p> <p>Rear Cameras: 50MP, 12MP</p> <p>Front Camera: 10.8MP</p> <p>Android 13</p> <p></p>
	<p>Apple iPhone 13 mini 128GB</p> <p>Apple's 2021 model featuring new Ultra Wide and Wide cameras, new sensor-shift image...</p>	<p>5.4 inch display</p> <p>Rear Cameras: 12MP, 12MP</p> <p>Front Camera: 12MP</p> <p>Apple iOS 15</p> <p></p>

Novel Techniques

- Online computations
 - While the consumer is taking the questionnaire
 - Questionnaire becomes a “dialogue” with the respondent
 - Adaptive conjoint analysis
- Social interactions / gamification
- Physiological measures (e.g., eye movements)
- Incentive alignment

Eye Tracking

	A	B	C	D
Drink	Apple juice	Apple juice	Coke	Chocomel
Mini Candy	Toblerone	Mars	Toblerone	Snickers
Instant Soup	Tomato	Goulash	Chicken	Chicken
Sandwich Spread	Chocolate paste	Honey	Honey	Peanut butter
Healthy Snack	Mueslibar	Fruitbar	Mueslibar	Mueslibar
Price	2.00 euro	2.50 euro	2.00 euro	3.00 euro
	C	C	C	C



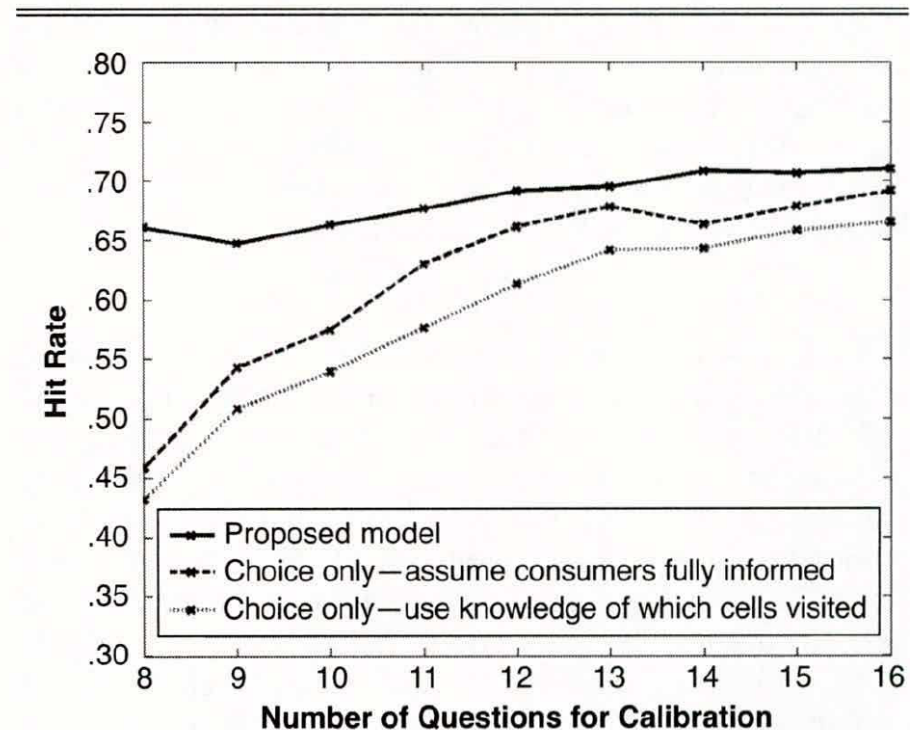
05:57

submit

Source: Yang, L., Toubia, O., & De Jong, M. G. (2015). A bounded rationality model of information search and choice in preference measurement. *Journal of Marketing Research*, 52(2), 166-183.

Learn Preferences Faster / Better

PROPOSED MODEL VERSUS CHOICE-ONLY BENCHMARKS:
AVERAGE HOLDOUT HIT RATE VERSUS NUMBER OF
QUESTIONS USED FOR CALIBRATION



Source: Yang, L., Toubia, O., & De Jong, M. G. (2015). A bounded rationality model of information search and choice in preference measurement. *Journal of Marketing Research*, 52(2), 166-183.

“Conjoint Analysis” to Assess Risk Profile

RiskSense Survey

☐ Please consider the two gambles below. Which of these two gambles would you rather play?

Option A	Option B
10% Chance to Win \$100 90% Chance to Lose \$10	90% Chance to Win \$1 10% Chance to Lose \$10
<input type="button" value="Choose A"/>	<input type="button" value="Choose B"/>

[More Information on Prizes](#)

PART 3 OF 3

Questions Completed

Source: Toubia, O., Johnson, E., Evgeniou, T., & Delquié, P. (2013). Dynamic experiments for estimating preferences: An adaptive method of eliciting time and risk parameters. *Management Science*, 59(3), 613-640.

... and Time Preferences

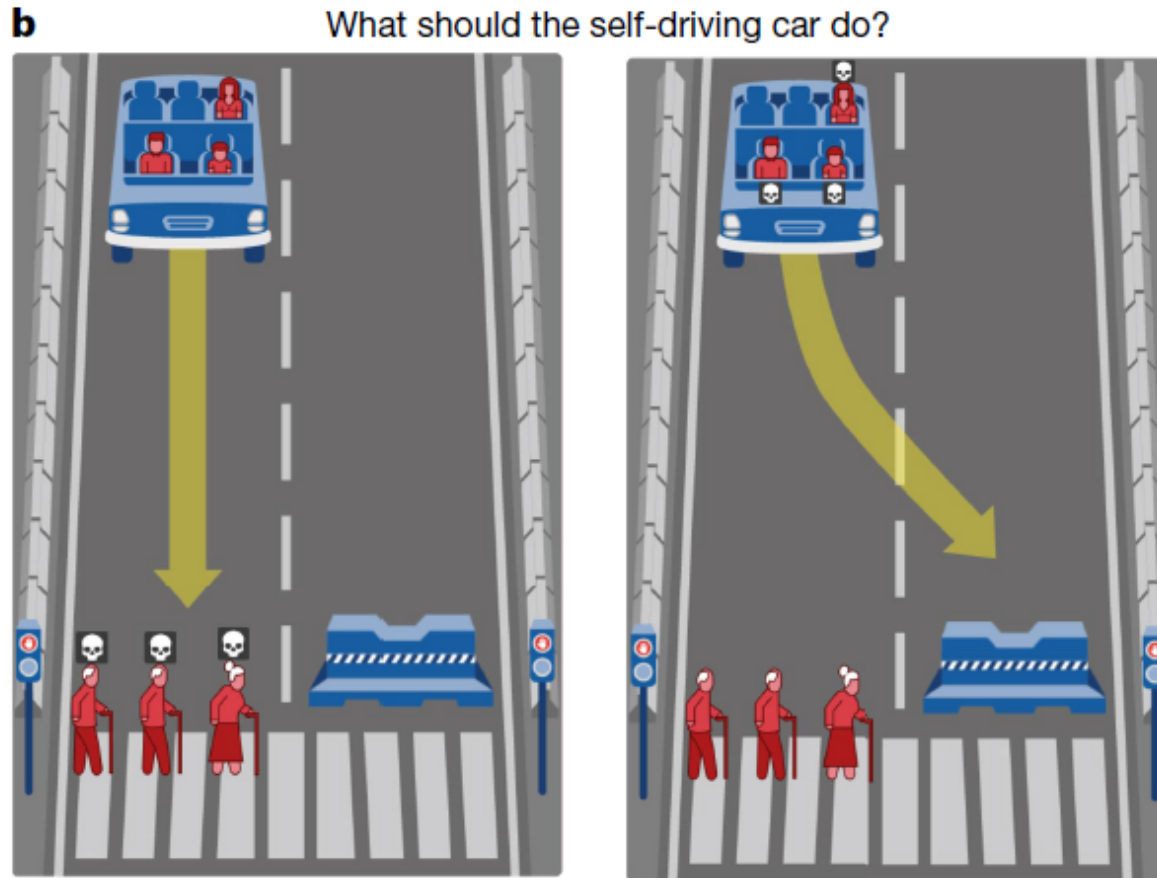
RiskSense Survey

Please consider the two options below. Which of these two options do you find more attractive?

Option A	Option B
Receive \$250 today	Receive \$300 in 3 months
<input type="button" value="Choose A"/>	<input type="button" value="Choose B"/>

[More Information on Prizes](#)
SECTION #4 OUT OF 5
Questions Completed
2/20

And Moral Preferences



Source: Awad, Edmond, et al. "The moral machine experiment." *Nature* 563.7729 (2018): 59-64.

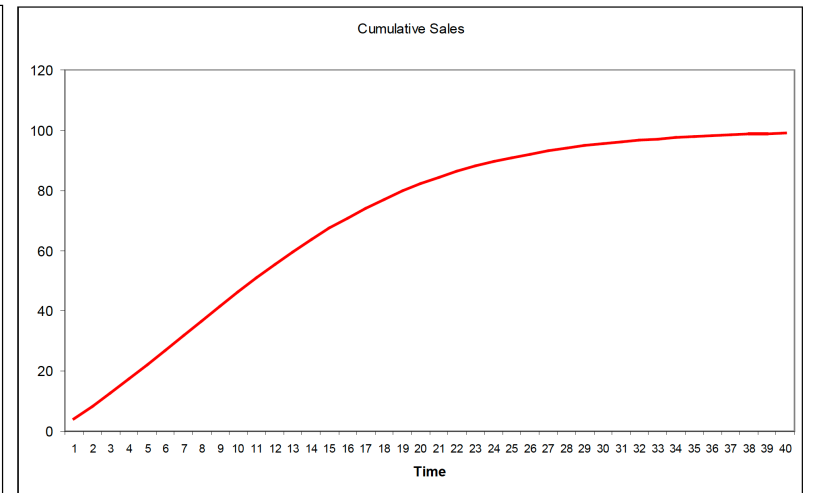
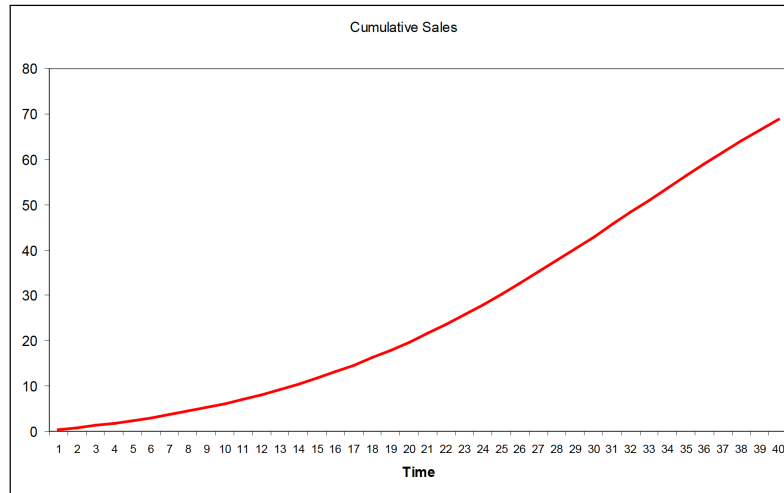
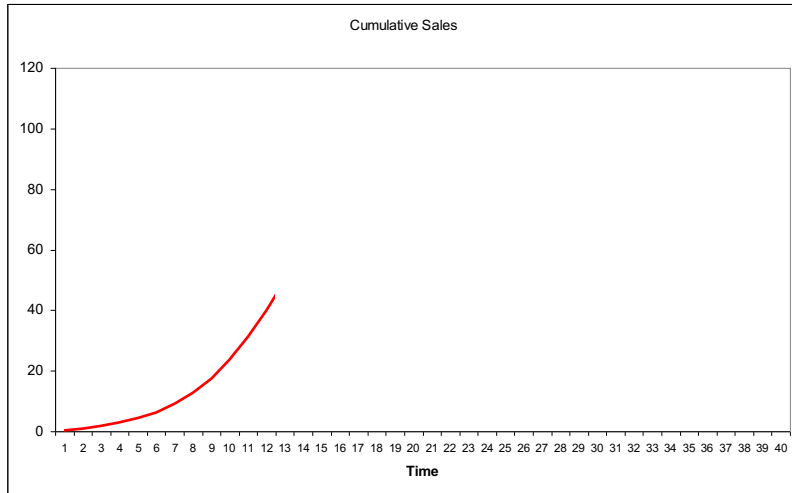
Takeaways

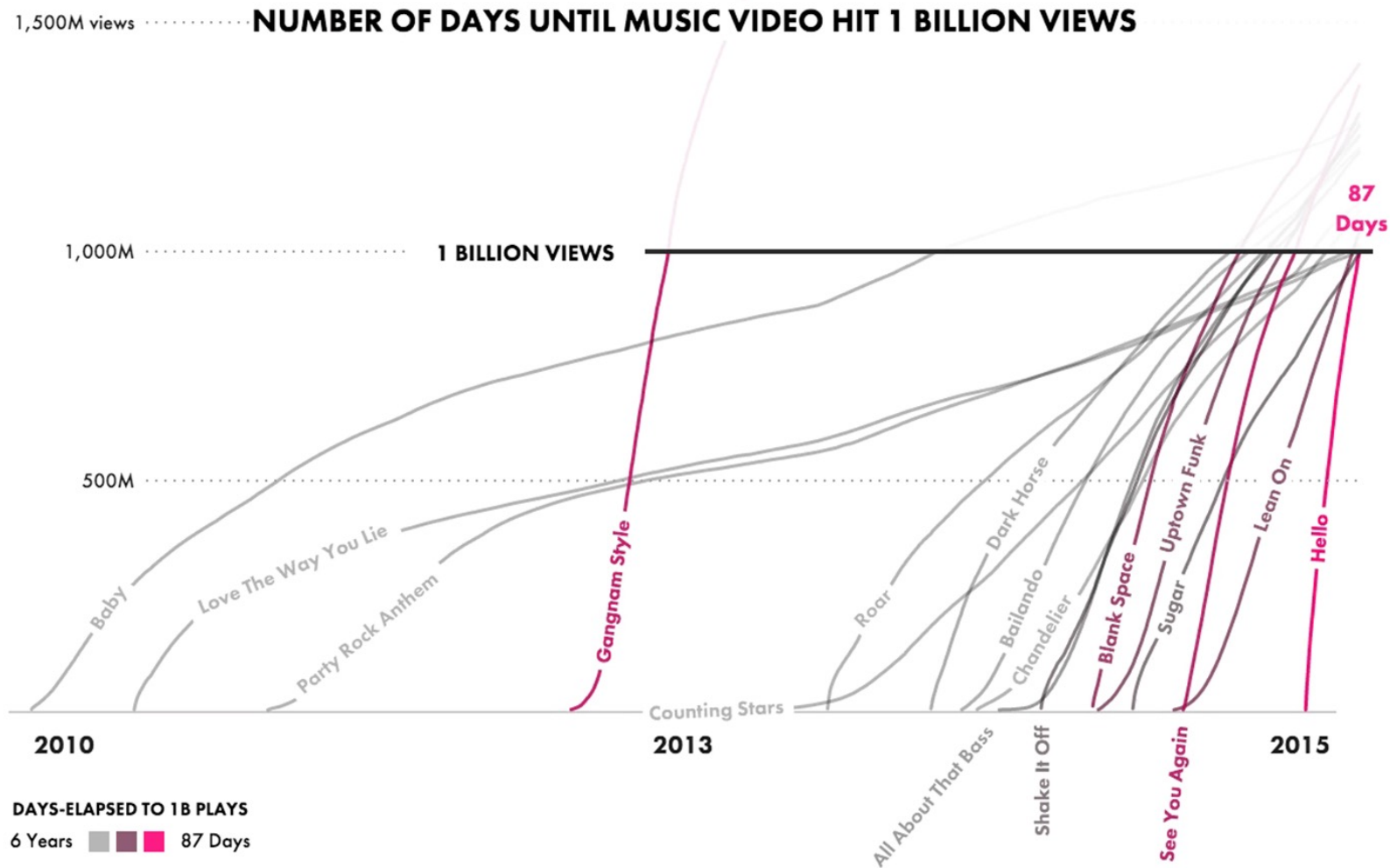
- Conjoint Analysis:
 - Tool to quantify customers' preferences for products and services
 - Allows us to assess preferences without exploring all the alternatives
- Applications
 - Assess preferences and predict choices
 - Predict market shares
 - New product design (see group assignment 2)
- Commercial Conjoint Analysis Software
 - Established market leader:
 - Sawtooth Software: <http://www.sawtoothsoftware.com/>
 - More recent entrants:
 - Epic Conjoint: <https://www.epicconjoint.com/>
 - Conjoint.ly: <https://conjoint.online/>

Diffusion of Innovation

Forecasting Diffusion of Innovation

- Our own innovation
- An innovation that complements our innovation (e.g., a platform on which we rely)





<http://youtube-trends.blogspot.co.uk/2016/01/adeles-hello-is-fastest-video-in.html>

Example: “Control and Connectivity” Market

*This Thermometer Tells Your Temperature,
Then Tells Firms Where to Advertise*



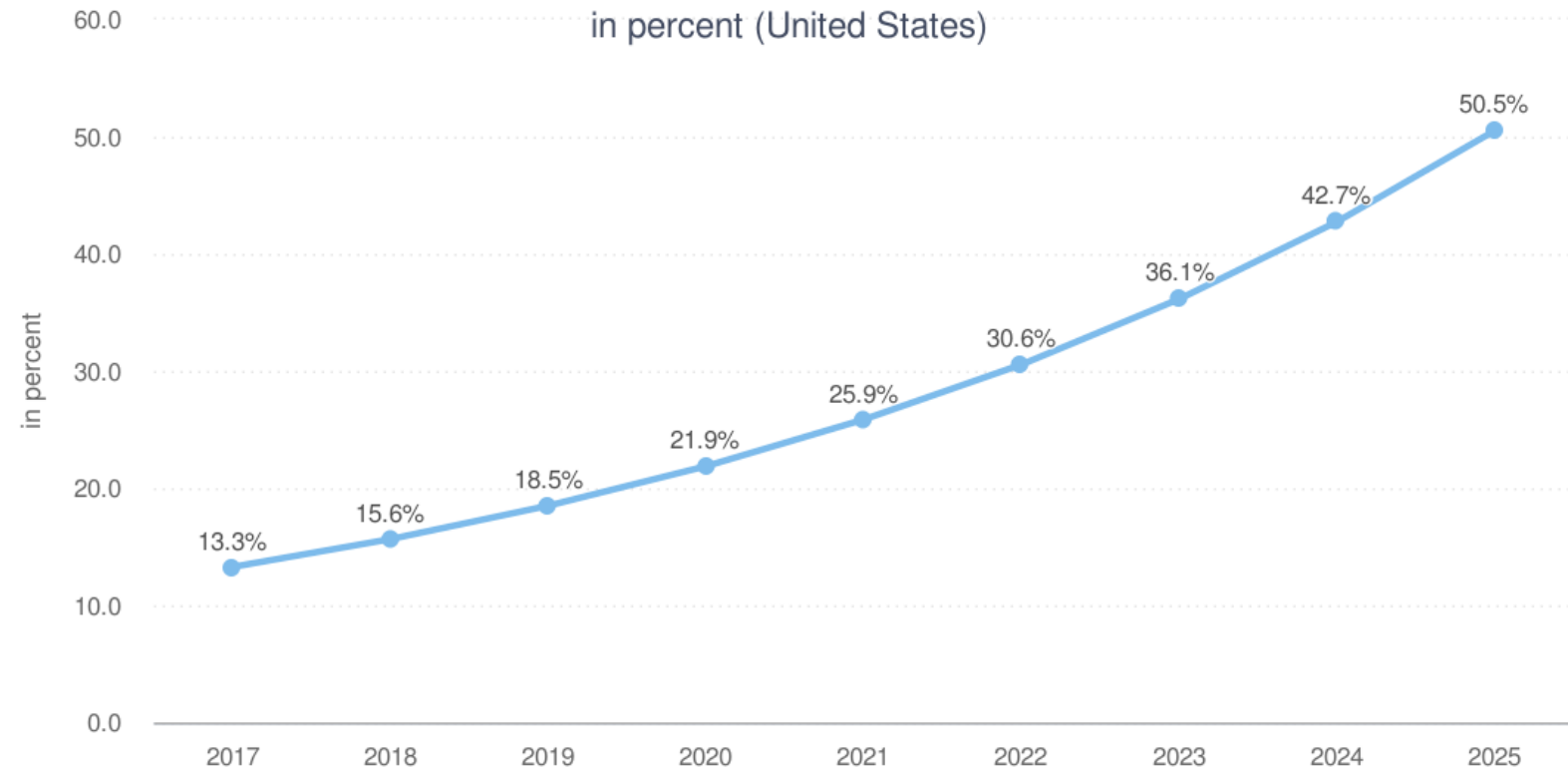
Kinsa says its smart thermometers are in more than 500,000 American households.
Tony Cenicola/The New York Times

“This flu season, Clorox paid to license information from Kinsa, a tech start-up that sells internet-connected thermometers... The data showed Clorox which ZIP codes around the country had increases in fevers. The company then directed more ads to those areas, assuming that households there may be in the market for products like its disinfecting wipes.”

- Source: New York Times, 10/23/2018

Forecast

Penetration Rate in the Control and Connectivity market



Source: Statista (Forecast adjusted for expected impact of COVID-19), October 2020

statista

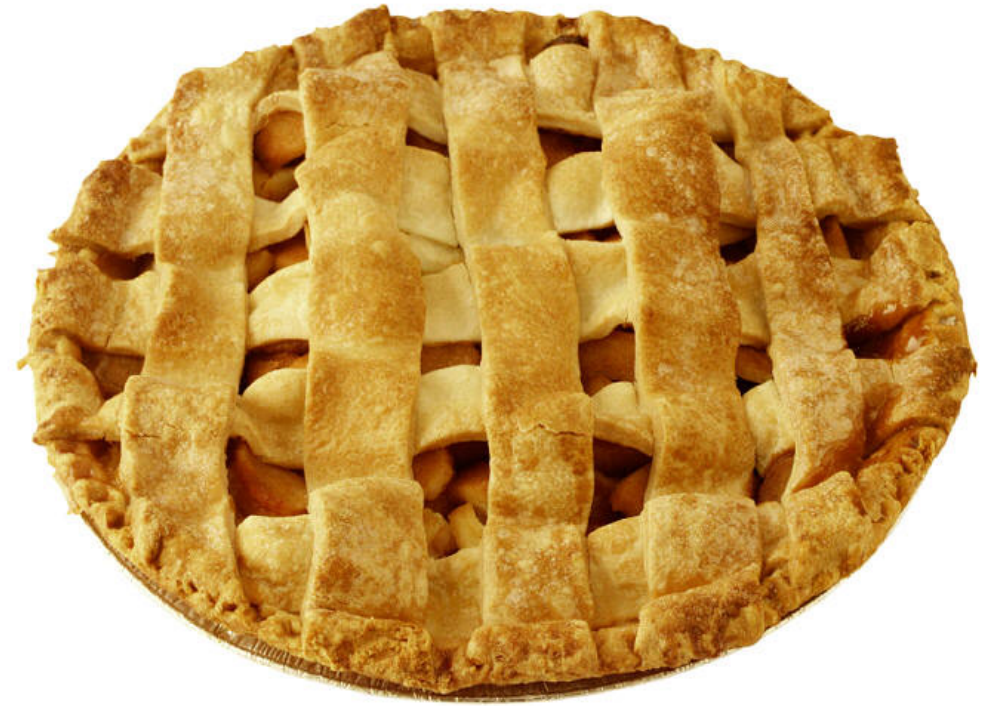
<https://www.statista.com/outlook/280/109/control-and-connectivity/united-states#market-revenue>

Understanding the Pie!

Conjoint Analysis



Bass Model

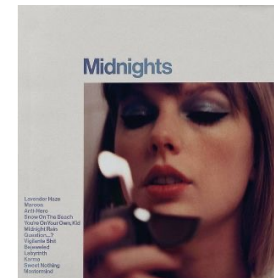


The Bass Model

- Predicts adoption curve
 - Does not study repeated sales
 - Contagion model
- We want to predict: Number of adopters in time t
- We have: Number of potential adopters (people who haven't adopted yet)
- Number of adopters in period t =
Adoption rate in period t x Number of potential adopters in period t

Possible Applications of Bass Model

- The “innovation” for which we predict diffusion could be
 - Our innovation
 - Another innovation on which our innovation relies
 - E.g., predict diffusion of wearable devices in general to assess potential of constant glucose monitor
- With the availability of daily / weekly data (e.g., apps, YouTube videos), it is easier now to apply Bass model to specific products (as opposed to new technologies)



Model Formulation

- Notation
 - t is the time period
 - $S(t)$ is the number of adopters in period t
 - p is the coefficient of innovation
 - q is the coefficient of imitation
 - $Y(t)$ is the cumulative number of adopters until t
 - m is the total market potential

$$\underbrace{S(t)}_{\text{Number of adopters}} = \underbrace{\left[p + q \frac{Y(t)}{m} \right]}_{\text{Adoption rate}} * \underbrace{[m - Y(t)]}_{\text{Number of potential adopters}}$$

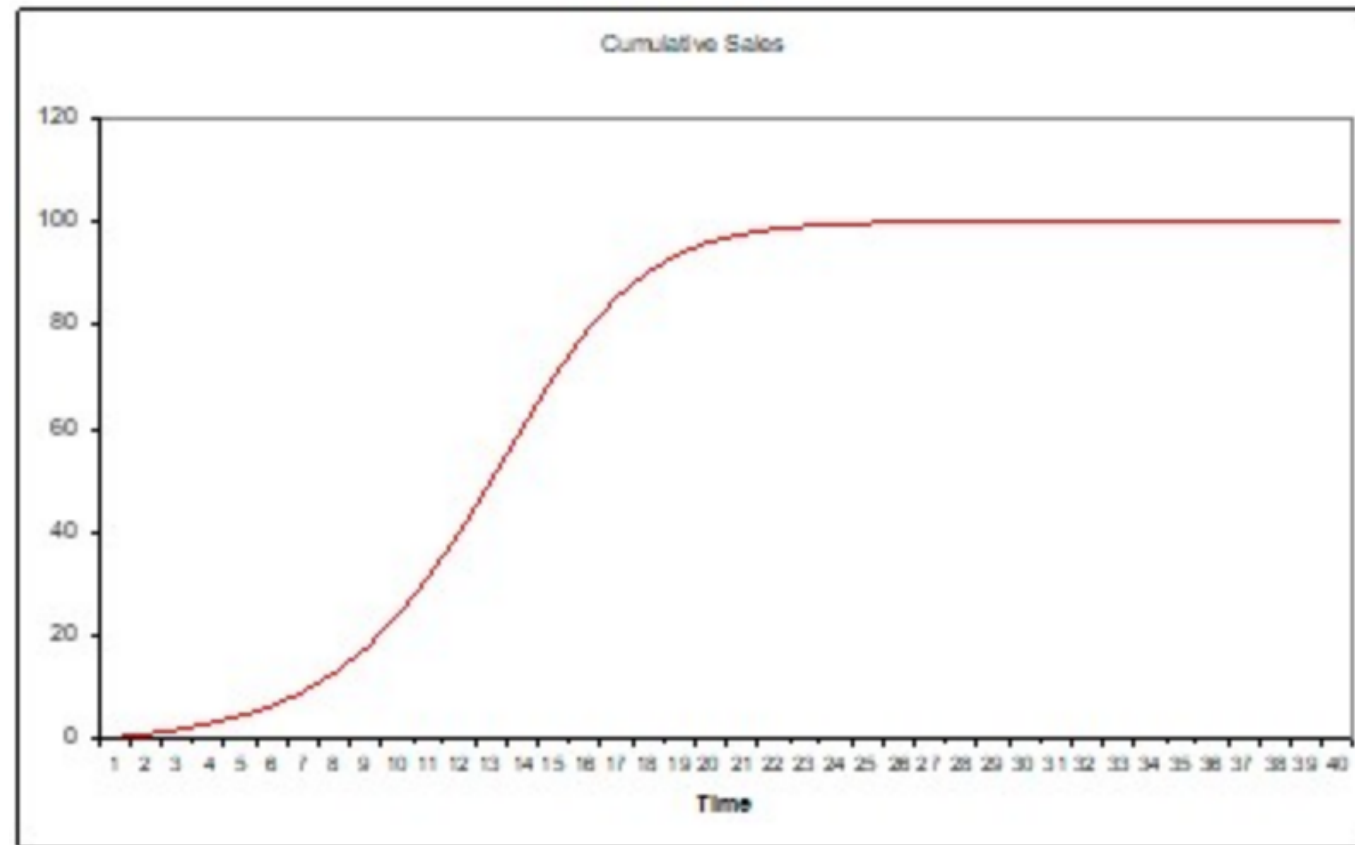
Adoption Rate

$$S(t) = \underbrace{\left[p + q \frac{Y(t)}{m} \right]}_{\text{Adoption rate}} * \underbrace{[m - Y(t)]}_{\text{Number of potential adopters}}$$

Number of adopters

- Adoption rate has two components
 - “individual component” p (probability of adopting without social influence)
 - “social” or “contagion” component $qY(t) / m$
 - increases with the number of previous adopters $Y(t)$

S-curve



$S(t) = [p + qY(t) / m][m - Y(t)]$ The parameter p influences the shape of the curve...



Mostly in the beginning

0

Mostly in the middle

0

Mostly near the end

0

Only in the first year

0

Only in the sixth year

0

$S(t) = [p + qY(t) / m] \cdot [m - Y(t)]$ The parameter q influences the shape of the curve...



Mostly in the beginning

0

Mostly in the middle

0

Mostly near the end

0

Only in the first year

0

Only in the sixth year

0

$S(t) = [p + qY(t) / m].[m - Y(t)]$ The parameter m ...

✓ 0

Only influences the end of the curve

0

"Stretches" the curve vertically

0

"Stretches" the curve horizontally

0

Rotates the curve

0

Next Class (Tuesday Classes)

- No class next Tuesday, Nov 5 (Election Day)
- Price and Place (Nov 12)
- Due:
 - Conjoint Analysis Concept Check
 - Readings
 - Montclair Video
 - Glaubinger Tree Farm
 - Group Assignment 2 (New Product Design) due Monday, Nov 18 at 8PM
 - Groups will be formed at the end of the week

Next Class (Wednesday Class)

- Bass Model + Generative AI + Explainable AI
- Due:
 - Readings
 - What every CEO should know about generative AI
 - Group Assignment 2 (New Product Design) due Monday, Nov 18 at 8PM
 - Groups will be formed at the end of the week

B9651 – Marketing Analytics

Week 9: Bass Model + Generative AI + Explainable AI

Professor Hortense Fong

Today

Part 1: Diffusion of Innovation

1. How do we forecast the diffusion of an innovation?
2. Bass Model

Part 2: Generative AI

1. Different types of GenAI models
2. How to use GenAI for market research

Part 3: Explainable AI

1. Types of explainability

Course Roadmap

STP Analytics (Identify Value)	Customer Analytics (Deliver Value)	4P Analytics (Capture Value)
Module 1	Module 2	Module 3
What datasets can we use? How can we segment and target our customers? How should we position our products/services?	How much are our customers worth? Are our customers leaving? How do our customers make choices?	How do we build a new product? How should we price our products? How do we distribute them? How do we quantify the impact of our promotions?

Diffusion of Innovation

Model Formulation

- Notation
 - t is the time period
 - $S(t)$ is the number of adopters in period t
 - p is the coefficient of innovation
 - q is the coefficient of imitation
 - $Y(t)$ is the cumulative number of adopters until t
 - m is the total market potential

$$\underbrace{S(t)}_{\text{Number of adopters}} = \underbrace{\left[p + q \frac{Y(t)}{m} \right]}_{\text{Adoption rate}} * \underbrace{[m - Y(t)]}_{\text{Number of potential adopters}}$$

How Do We Obtain the Parameters of the Bass Model?

- **m**: Demographic data (for m only)
- **p, q**: Historical analysis of analogous innovations (not necessarily in the same industry)
 - Managerial judgment
 - Average (yearly) values: $p=0.03$, $q=0.38$
- Fit early diffusion data

Some Parameter Estimates

Applications of the Bass Model

<i>product</i>	<i>years</i>	<i>p</i>	<i>q</i>
B&W TV	1946-1957	0.024	0.4
Blender	1958-1971	0.001	0.55
Calculator	1974-1986	0.0294	0.12
Can Opener	1958-1971	0.0135	0.3
Cass. Tape Deck	1974-1990	0.0179	0.18
Color TV	1954-1970	0.0001	0.66
Curling Iron	1974-1982	0.0399	0.3
Deep Fryer	1950-1961	0.0337	0.74
Digital Watch	1974-1982	0.0204	0.41
Electric Blanket	1946-1966	0.0029	0.26
Elec. Clothes Dryer	1947-1960	0.008	0.4
Electric Toothbrush	1963-1971	0.0826	0.13
Fire Extinguisher	1975-1982	0.0694	0.16
Food Processor	1978-1984	0.0271	0.51
Hair Setter	1968-1977	0.1305	0.35
Heating Pad	1922-1931	0.0274	0.26
Knife Sharpener	1957-1969	0.0655	0.5
Lawn Mower	1946-1961	0.0056	0.3
Microwave Oven	1970-1989	0.0101	0.37
Refrigerator	1922-1931	0.0047	0.42
Room A/C	1946-1960	0.0031	0.43
Slow Cooker	1973-1982	0.0654	0.32
Styling Dryer	1974-1984	0.0891	0.18
Trash Compactor	1972-1982	0.0866	0.19
Turntable	1974-1990	0.0577	0.21
VCR	1976-1990	0.0012	0.53
Waffle Iron	1922-1931	0.0128	0.43

WARNING: these are yearly values, which should only be used with yearly data

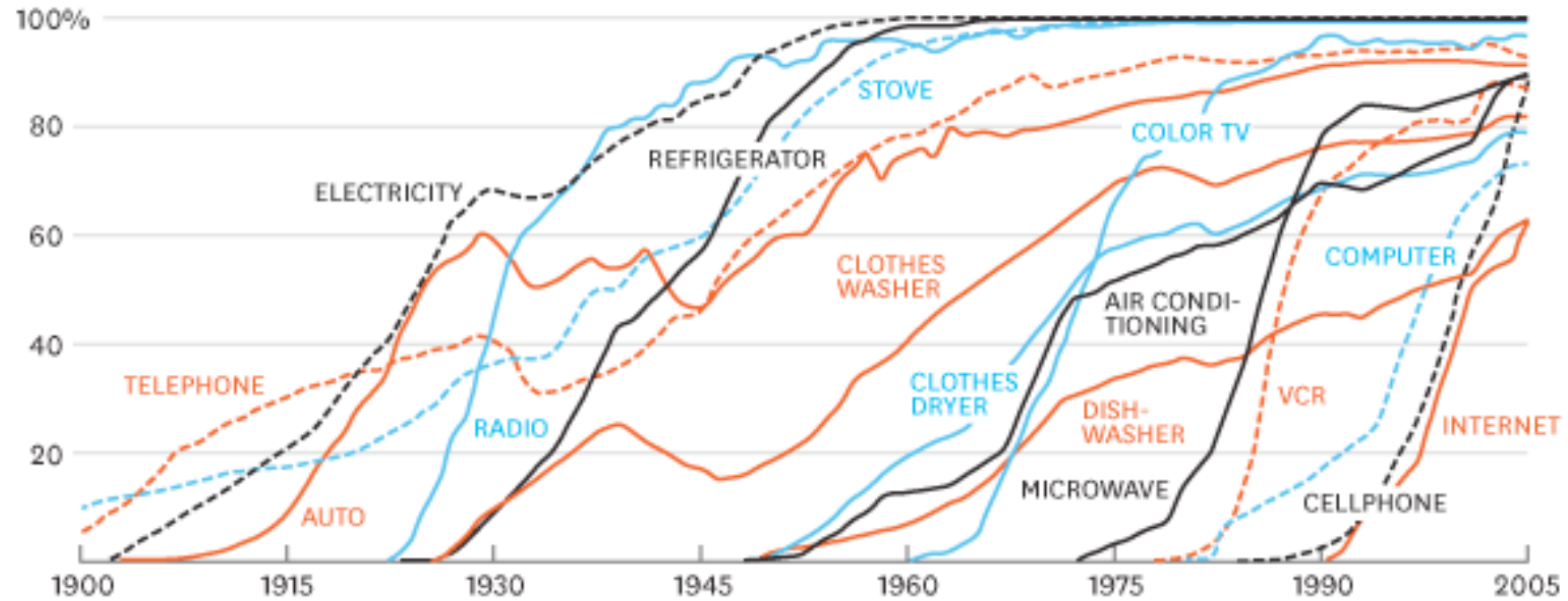
What Do We Mean by “Analogous Innovation”?

- Factors to take into account when evaluating analogies
 - Environmental situation
 - Market structure
 - Buyer behavior
 - Marketing mix strategy
 - Characteristics of innovation itself
- Some general findings
 - p higher in Europe than in US. Potential reason?
 - q higher for industrial / medical products compared to other durable products

Diffusion Speed Has Generally Increased Over Time

CONSUMPTION SPREADS FASTER TODAY

PERCENT OF U.S. HOUSEHOLDS

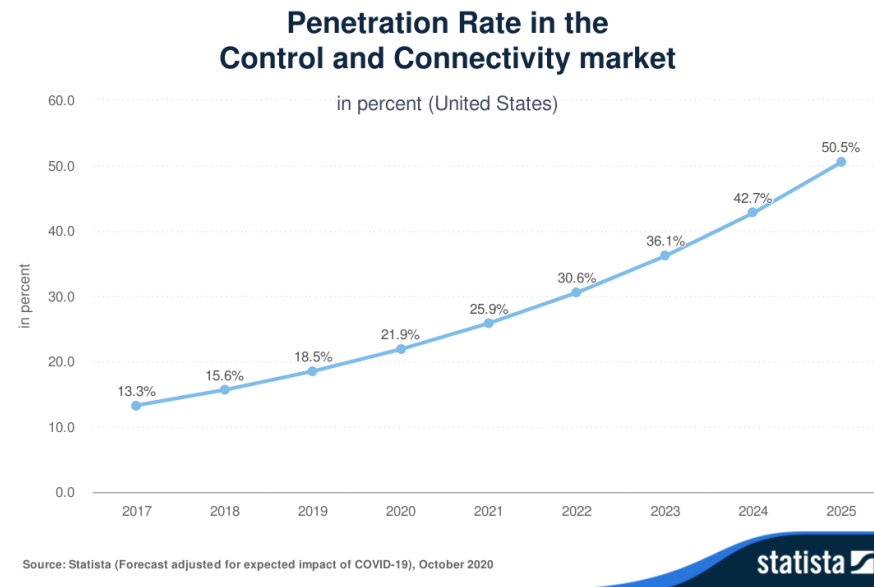
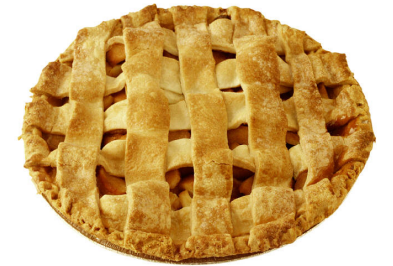


SOURCE MICHAEL FELTON, THE NEW YORK TIMES

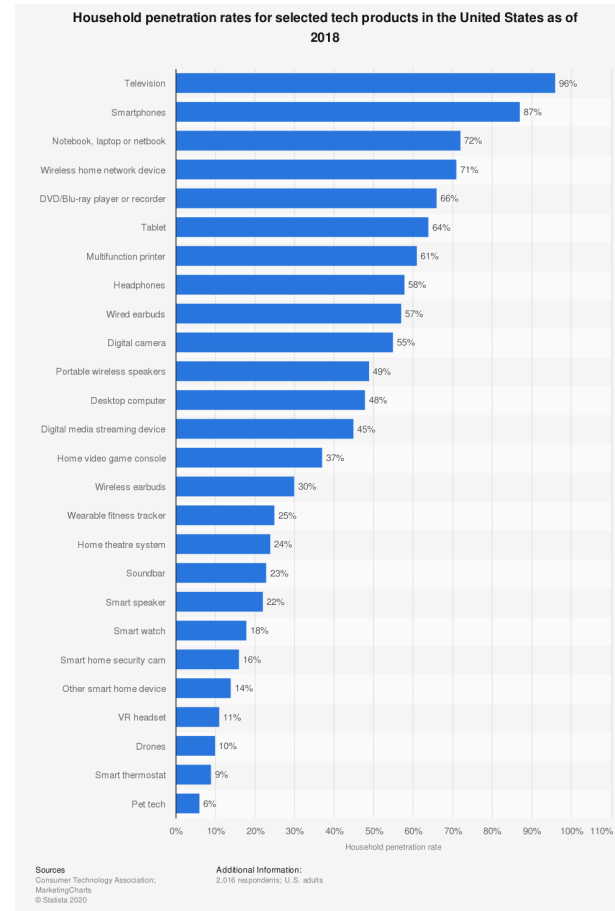
HBR.ORG

Your turn!

- Growth of the Control and Connectivity market in the US?
- Is the forecast reasonable? Aggressive? Conservative?



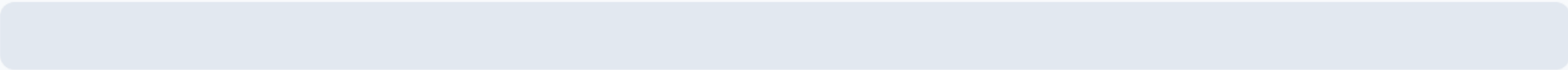
Household Penetration of Tech Products in the US as of 2018 (on canvas)



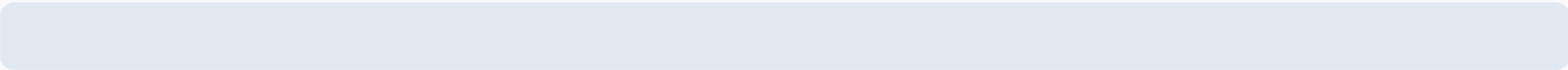
Resources

- On Canvas:
 - Bass model spreadsheet
 - Bass model analogies – determine **p, q**
 - Household Penetration of Tech Products – determine **m**
- Forecast on cumulative penetration in % of US households:
 - 2017: 13.3%
 - 2018: 15.6%
 - 2019: 18.5%
 - 2020: 21.9%
 - 2021: 25.9%
 - 2022: 30.6%
 - 2023: 36.1%
 - 2024: 42.7%
 - 2025: 50.5%
- Let's use year 2004 as $t=1$
- <http://www.entrepreneurial-insights.com/smart-home-intelligent-home-automation/>

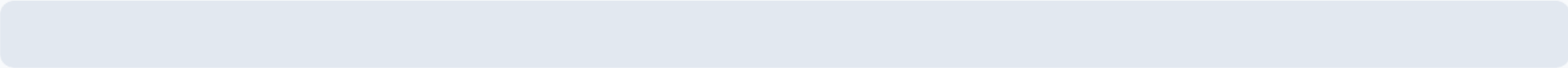
too conservative



reasonable



too aggressive



How to Use the Bass Model?

- Under which conditions will diffusion reach a given target / estimate?
 - What values of p , q , m do we need?
 - Are they reasonable based on historical data?
- What are best-case and worst-case scenarios?
- Beware of confirmation bias
 - Possible to justify a large range of forecasts by playing with the parameters

Takeaways

- Forecasting the diffusion of a new product is key to companies
- Bass model:
 - Parsimonious model to predict how big the pie is
 - Relies on understanding the extent of social influence vs potential to adopt independently
- New techniques
 - Rely on individual and social network data
 - Allow for more precise predictions

Generative AI

What is Generative AI?



ChatGPT

Generative AI refers to a category of artificial intelligence that can generate new content, based on learning from a vast amount of existing data. This content can include text, images, music, speech, or other forms of media. The key aspect of generative AI is its ability to create something new, rather than just analyzing or processing existing data.

- Text generation: essays, stories, code, ad copy, etc. (e.g., GPT)
- Image creation: text → images (e.g., DALL-E)
- Music composition: text → music (e.g., Stable Audio)
- Synthetic data: synthetic data sets that mimic real-world data, useful for training AI systems without using sensitive real-world data (e.g., gretel.ai)

Text Generation with ChatGPT

- Prompt: Please write me ad copy for a class about using generative AI for Marketing Analytics.

Unleash the Power of AI in Your Marketing: Join Our Generative AI for Marketing Analytics Class!

Transform Your Marketing Strategies with Cutting-Edge AI Tools!


Are you ready to revolutionize your marketing approach? Our comprehensive class on Generative AI for Marketing Analytics is your gateway to mastering the future of digital marketing.

What You'll Learn:

- **Harness AI for Enhanced Insights:** Dive into how generative AI can uncover deep insights from consumer data, helping you make smarter, data-driven decisions.
- **Creative Campaigns with AI:** Discover how to use AI for generating innovative marketing content, from captivating ad copies to eye-catching visuals.
- **Predictive Analytics:** Explore predictive models that forecast market trends, customer behavior, and campaign outcomes, giving you a competitive edge.
- **Personalization at Scale:** Learn to personalize marketing at an unprecedented scale, creating uniquely tailored experiences for each customer.
- **Ethical AI Use:** Understand the ethics of AI in marketing, ensuring responsible and effective use of this powerful technology.

Why Choose This Class?

- **Expert Instructors:** Learn from professionals who are pioneering AI in marketing.
- **Hands-On Experience:** Gain practical experience with the latest AI tools and platforms.
- **Networking Opportunities:** Connect with like-minded professionals and potential collaborators.
- **Future-Ready Skills:** Equip yourself with skills that are increasingly essential in the modern marketing landscape.

 **Special Bonus:** Access to an exclusive community forum for ongoing support and knowledge sharing.

Who Should Enroll?

- Marketing Professionals seeking to upskill.
- Business Owners looking to leverage AI for growth.
- Students and Enthusiasts eager to explore AI in marketing.

 **Enroll Now** and be at the forefront of the AI revolution in marketing!

[Sign Up Today and Transform Your Marketing Approach with AI!]

Image Creation with DALL-E

- Prompt: Please create an ad for a Marketing Analytics with Generative AI class.



Music Composition with Stable Audio

- Prompt: Soundtrack to most exciting marketing class you have ever taken.



How have you used GenAI?



Nobody has responded yet.

Hang tight! Responses are coming in.

The Effect of GenAI on Creativity

ScienceAdvances

[Current Issue](#) [First release papers](#) [Archive](#) [About](#) ▼

Generative AI enhances individual creativity but reduces the collective diversity of novel content

[ANIL R. DOSHI](#)  AND [OLIVER P. HAUSER](#)  [Authors Info & Affiliations](#)

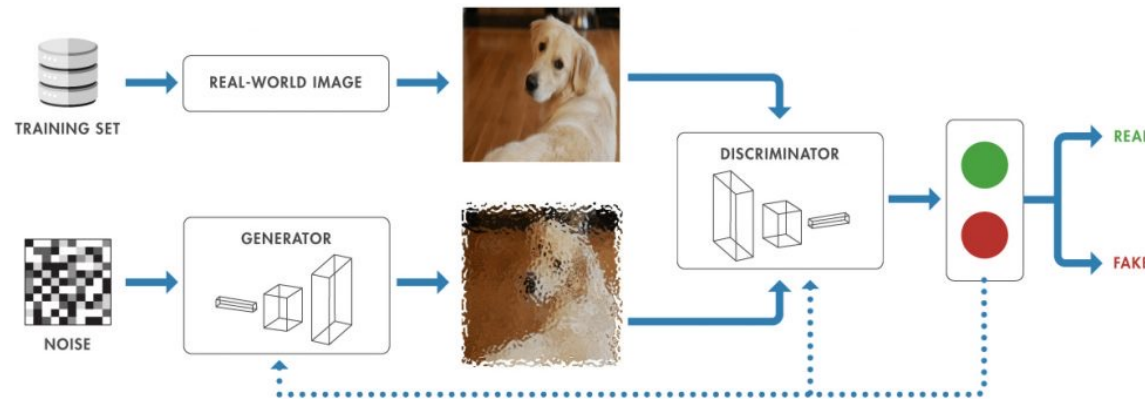
SCIENCE ADVANCES • 12 Jul 2024 • Vol 10, Issue 28 • DOI: 10.1126/sciadv.adn5290

- Some writers obtained story ideas from an LLM, others did not
- Those with access to an LLM wrote stories that were more creative, better written, and more enjoyable
- On net, the stories were less diverse

Generative Models

Generative Adversarial Networks (GANs)

- A neural network architecture used to generate new data that resembles the training data
- Two neural networks compete against each other in a zero-sum game framework
 - Generator: generates new data instances
 - Discriminator: evaluates the data it receives and labels the data as real (from training data) or fake (from generator)

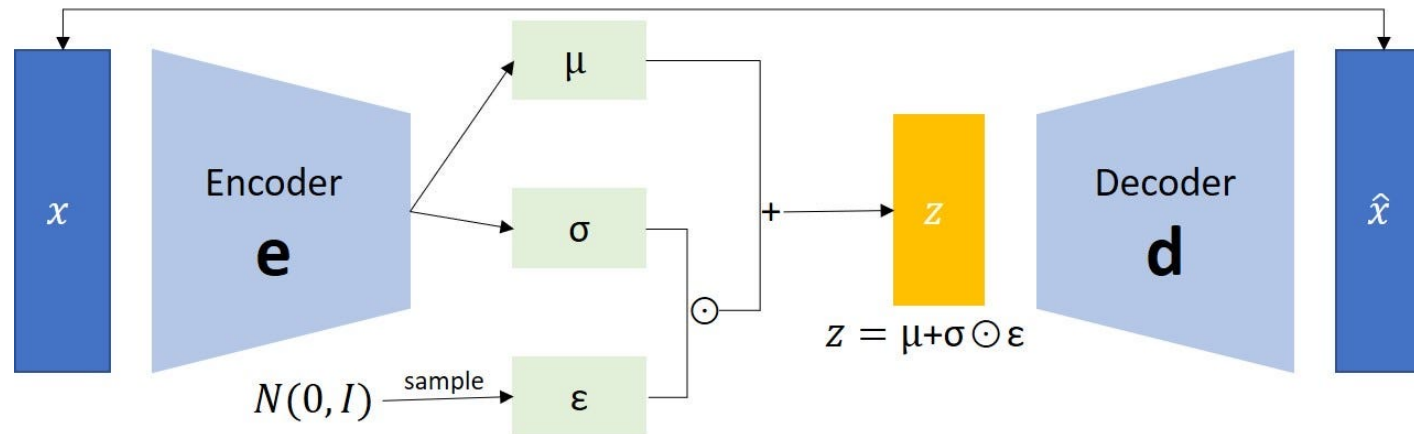


Generative Adversarial Networks (GANs)

- A neural network architecture used to generate new data that resembles the training data
- Two neural networks compete against each other in a zero-sum game framework
 - Generator: generates new data instances
 - Discriminator: evaluates the data it receives and labels the data as real (from training data) or fake (from generator)
- Applications: image generation, style transfer, data augmentation

Variational Autoencoders (VAEs)

- A neural network architecture that learns deep latent (hidden) representations of data
 - Encoder: takes input data and compresses it into a smaller, dense representation (i.e., latent space)
 - Learns the parameters (mean and variance) of a probability distribution representing each data point in the latent space
 - Decoder: takes this compressed data and reconstructs the original input from this representation



Variational Autoencoders (VAEs)

- A neural network architecture that learns deep latent (hidden) representations of data
 - Encoder: takes input data and compresses it into a smaller, dense representation (i.e., latent space)
 - Learns the parameters (mean and variance) of a probability distribution representing each data point in the latent space
 - Decoder: takes this compressed data and reconstructs the original input from this representation
- Applications: image generation, music generation, data compression, dimensionality reduction

Transformers

- A neural network architecture that is efficient in handling sequences of data (e.g., text) and processing different parts of the sequence in parallel
 - Attention mechanism: weighs the importance of different parts of the input data differently
 - Good at handling long-range dependencies in the data
- Applications: NLP, time-series analysis

Applications in Marketing

Perceptual Maps

Language Models for Automated Market Research: A New Way to Generate Perceptual Maps [Li, Castelo, Katona, Savary]

- Explores the potential of generative language models to substitute for human participants in Perceptual Analysis



- How well does market research based on large language models (LLMs) replicate market research based on human respondents?

Data for Perceptual Analysis – Human vs. AI

- Data based on surveys of brand similarity or attributes
 - “How similar are the brands BMW and Audi?”
 - 1 = not similar at all, 9 = very similar
- **Surveys** require large sample sizes, are expensive, and time-consuming
- **Language models** trained to replicate human responses
 - “How similar are the brands BMW and Audi?”

Fill in the blank:

“The car brand BMW is similar to the car brand” ← Measure co-occurrences of brands to develop **brand similarity score**

“The most comfortable car brand is...” ← Frequency of brand captures **brand attribute score**

Using Text for Perceptual Maps

- Count frequency of words after a prompt → frequency matrix

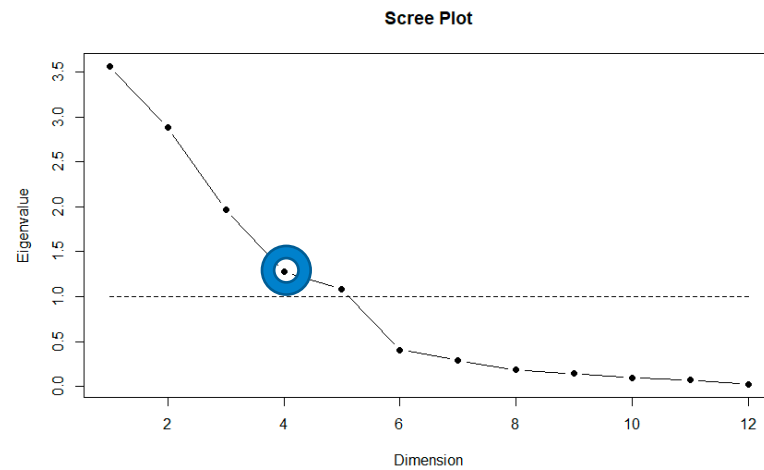
The most eco-friendly car brand is...

	Eco-friendly	Sporty	Reliable
Audi	5	6	4
BMW	6	4	6
Toyota	7	2	9

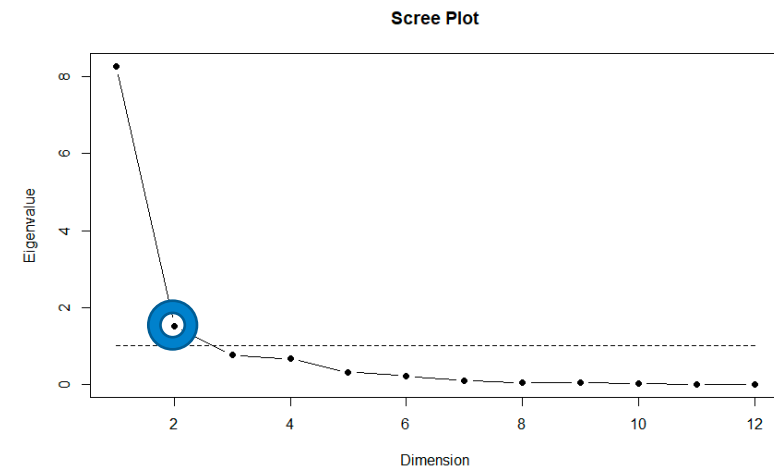
- PCA

Attribute-based Analysis

- 21,000 AI responses, 1,050 human responses

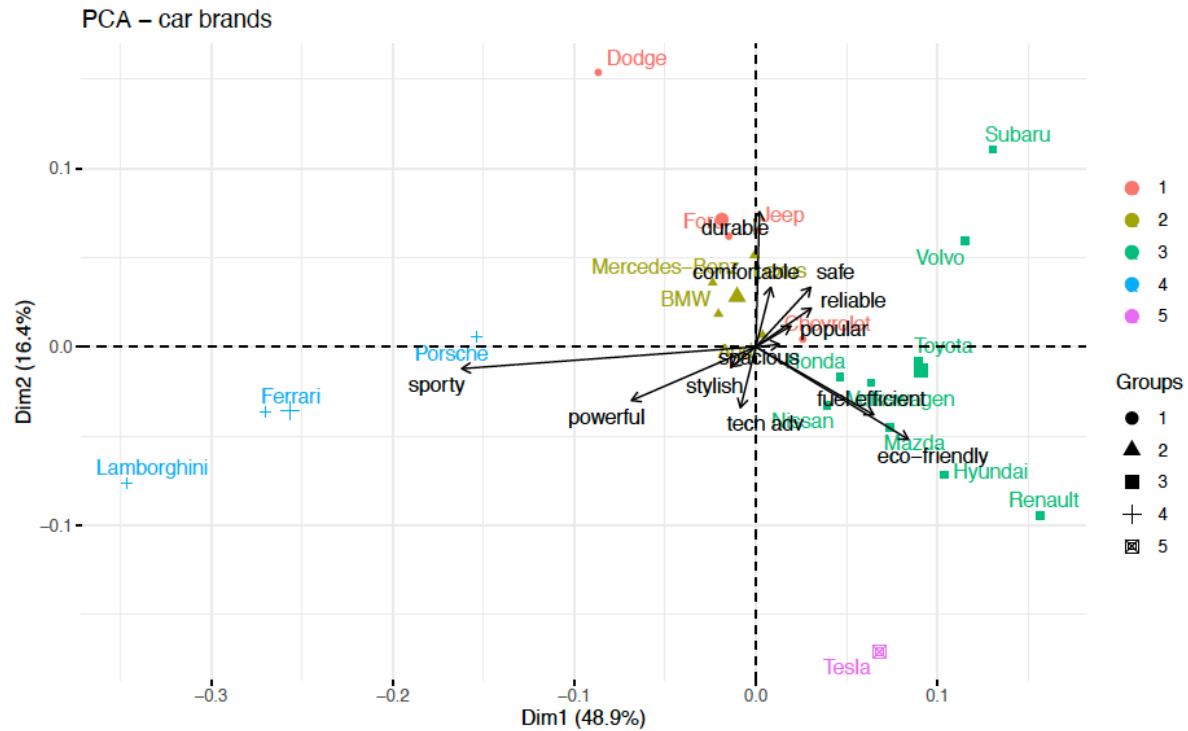


(a) AI Open-ended

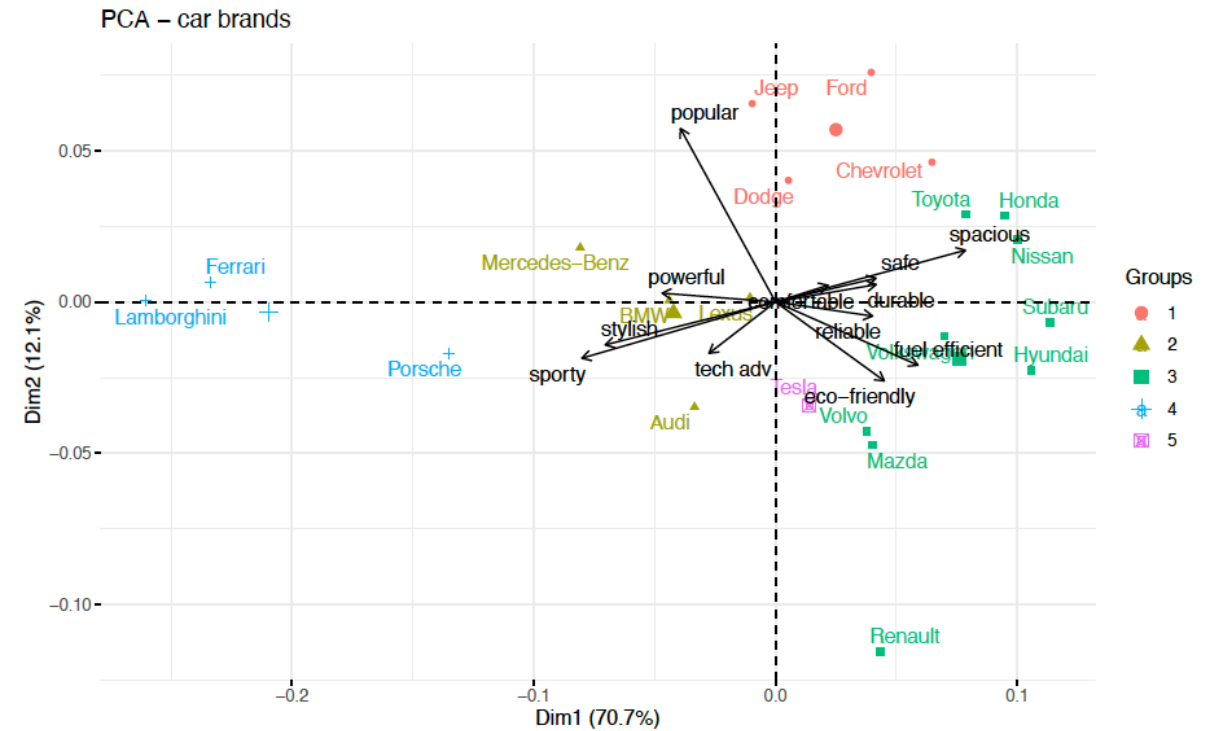


(b) Human

Perceptual Maps



(a) AI Open-ended



(b) Human

Language Models for Automated Market Research: A New Way to Generate Perceptual Maps

- Shows that using data from AI generates outputs that are close to those generated from human surveys
- Complement or substitute?
 - Complement: pilot test survey with AI before conducting survey with humans
 - Substitute: only use data from AI
- AI-generated data benefits: scale, speed, volume
 - Can answer more nuanced questions based on demographic variables or contextual variation that would be prohibitively expensive or infeasible with human respondents
- Limitations: training data, models currently do not capture time dimension

Applications in Marketing

WTP/Conjoint Analysis

Using GPT for Market Research [Brand, Israeli, Ngwe]

- Measure the extent to which GPT exhibits fundamental properties of consumer demand
- Goal: use GPT to generate a distribution of responses (not just a single one)
 - Set temperature on GPT to its max value for text completion (1.0) to maximize variation across responses
- Procedure:
 - Prompt GPT to fill in responses of a survey question
 - Describe any relevant attributes of the customer (e.g., income)

Responses Adhere to Economic Theory

- Preferences implied by GPT's responses appear to adhere to economic theory
 - Downward sloping demand curves
 - Flatter demand curve for higher-income customers
 - State dependence
 - Diminishing marginal utility

Responses Produce Realistic WTP

- Distribution of WTP for products generates reasonable values for multiple attributes
 - Conjoint with full set of options for each brand (i.e., don't need fractional factorial design because GPT not limited in same way humans are)

	GPT	Humans
Fluoride	\$3.40	\$3.27
No Aluminum	\$0.99	\$1.97

GPT = way to learn about consumer preferences in a fast, low-cost manner

Prompt Engineering

Example of Good vs. Bad Prompt

- Prompt engineering – providing clear instructions to a generative model to get what you want
- Bad Prompt: “I want to cook something.”
- Good Prompt: “Acting as an expert home cook, for someone who enjoys spicy Indian food and has only 25 minutes to prepare dinner, could you recommend a recipe including a list of ingredients and step-by-step instructions?”

Best Practices

- Role/identity – define the persona you want the AI to emulate
 - E.g., Python coding instructor
- Context – provide background
 - E.g., You are teaching a group of students 45-60 years old who have limited coding experience
- Prompt – provide a **specific** ask
- Output format – request a specific format
 - E.g., Use bullet points
- Examples – provide examples (few-shot learning) to help the AI learn what you are looking for

In-class Activity in Groups

Guess that Brand!

Groups of 4

Use ChatGPT to write a positioning statement for a brand

Positioning Statement Reminder

- Positioning statement
 - Who is the product for?
 - What does the product have to offer?
 - How is the product different?
- Example: **Volvo** is a car brand with a wide range of customers. Their cars drive like any other vehicle and, generally last a long time – but what Volvo is known for is “keeping your family safe.” Their core target audience is parents, they build products for families, and their unique point of difference is safety. Even though they do other things well, they want to own and claim the market position of safety.

In-class Activity in Groups

Guess that Brand!

Groups of 4

Use ChatGPT to write a positioning statement for a brand

Come up with a group name

Paste statement (without brand name) to Google slide

Can other teams correctly guess your brand?

https://docs.google.com/presentation/d/14Uv2AABW2EI_Ipdias3GZy9uLakIRksXcAe6WRVX1E/edit?usp=sharing

GenAI Challenges & Ethical Implications

The Dark Side of GenAI

- Deepfakes: realistic-looking but synthetic media



The Dark Side of GenAI

- Intellectual Property and Copyright: use of copyrighted material to train generative AI models without permission

Franzen, Grisham and Other Prominent Authors Sue OpenAI

The suit, filed with the Authors Guild, accuses the A.I. company of infringing on authors' copyrights, claiming it used their books to train its ChatGPT chatbot.

 Share full article  



The Dark Side of GenAI

- Bias and Fairness: inherit and amplify biases present in their training data

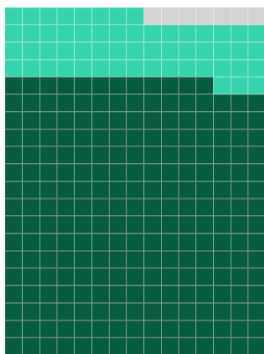
**HUMANS ARE BIASED.
GENERATIVE AI
IS EVEN WORSE**

Stable Diffusion's text-to-image model amplifies stereotypes about race and gender – here's why that matters

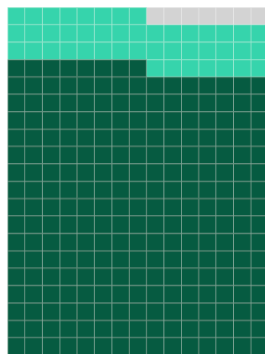
Perceived Gender: ■ Man ■ Woman ■ Ambiguous

High-paying occupations

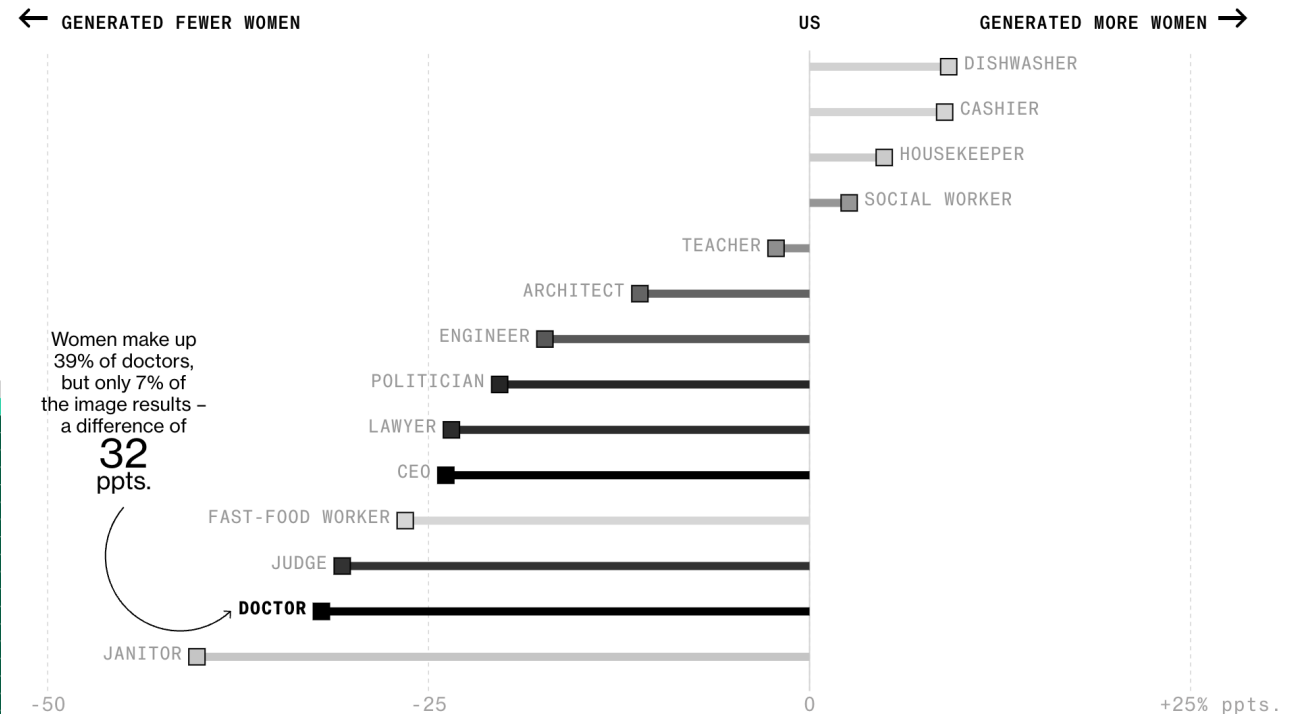
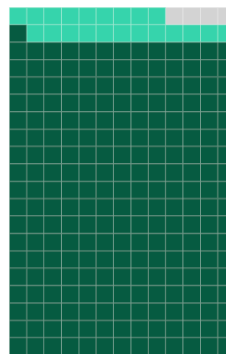
ARCHITECT



LAWYER



POLITICIAN



The Dark Side of GenAI

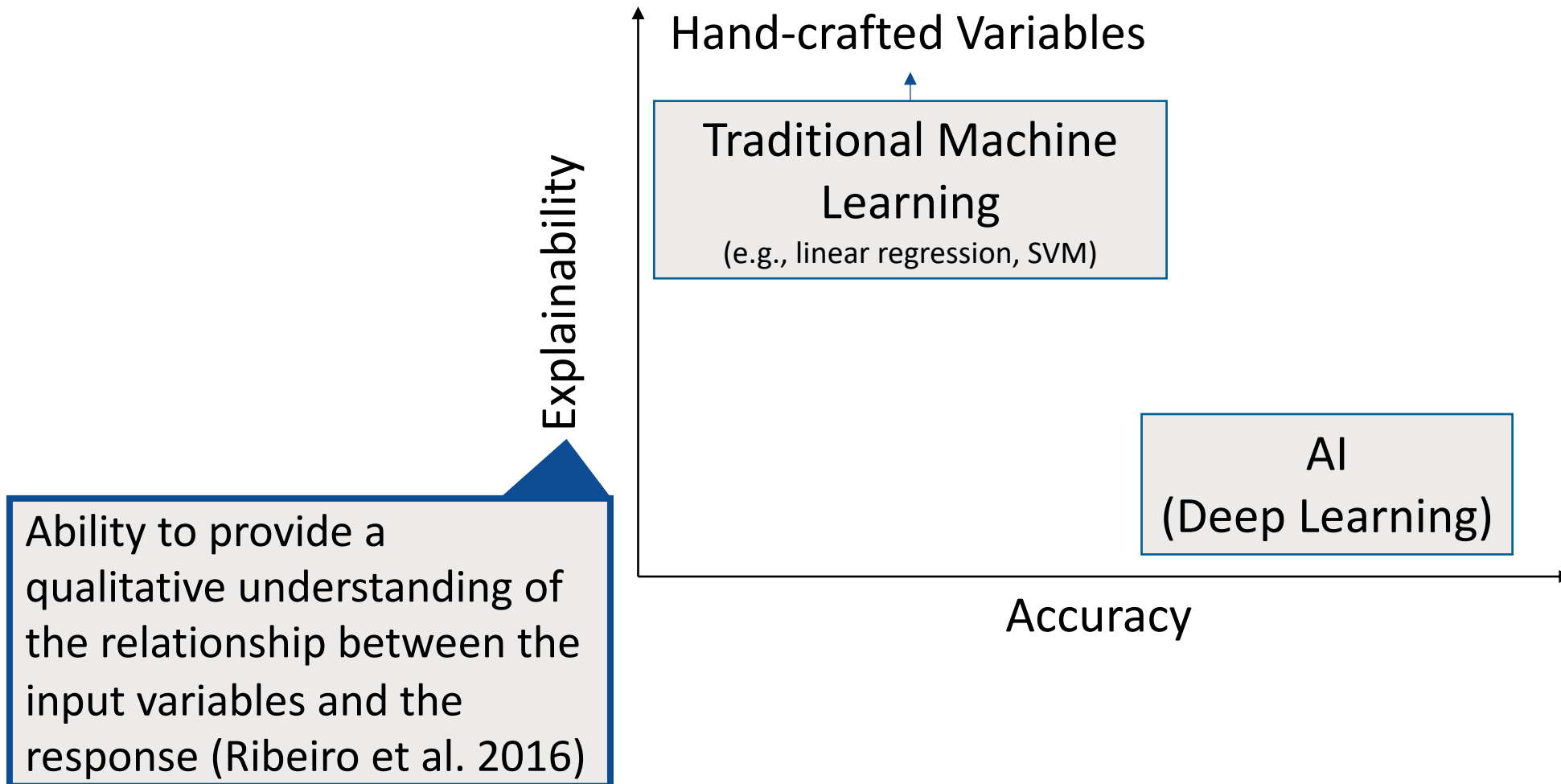
- Deepfakes: realistic-looking but synthetic media
- Intellectual Property and Copyright: use of copyrighted material to train generative AI models without permission
- Bias and Fairness: inherit and amplify biases present in their training data
- Privacy concerns: training data could capture private information
- ...

Explainable AI

Types of Models

- Transparent:
 - Decision trees
 - Linear regression
 - K-nearest neighbors
 - Etc.
- Opaque/Black-box:
 - Random forests
 - Deep learning
 - Etc.

Improved performance comes at the cost of explainability



Why do we care about explainability?

Explainability is important for:

- Managers to have **trust** in predictions → deploy model at scale
- **Generalizability**/robustness of model in other settings
- “If your system doesn’t work and you don’t know why it’s quite hard to improve it.” – Uber AI researcher
- Ethical and fairness concerns

F Forbes

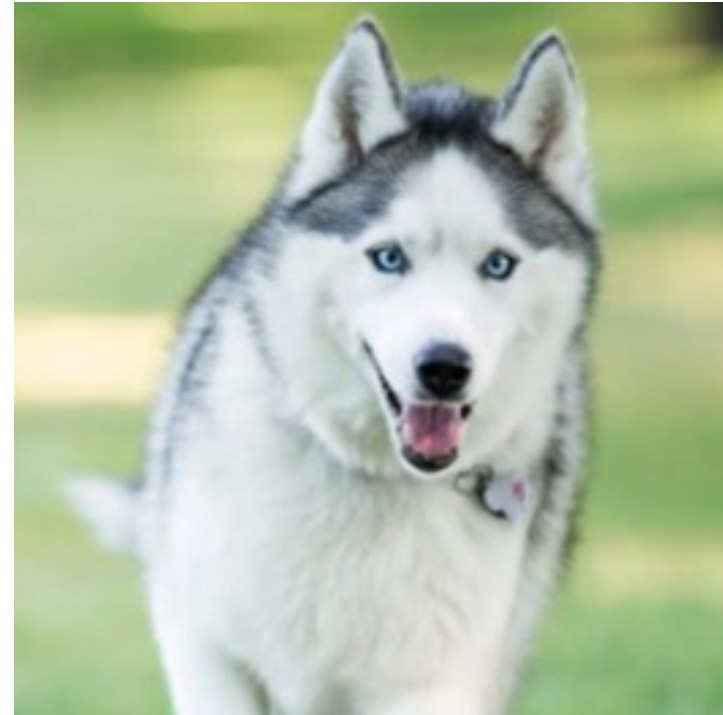
Nobody Can Explain For Sure Why ChatGPT Is So Good At What It Does, Troubling AI Ethics And AI Law

Wondered how it is that ChatGPT and other generative AI are so good at what they do? AI researchers and AI makers are also unsure and unable...



Why do we care about explainability?

Can we classify wolves vs. huskies (breed of dog)?



Types of Model Explainability

Text Explanation – LIME

- LIME = Local Interpretable Model-Agnostic Explanations
- Explanation by simplification – approximates an opaque model locally by building a linear model or decision tree around the predictions of an opaque model



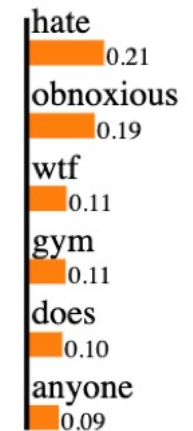
Input: does anyone else hate
obnoxious gym goers? #loud #wtf
Label: anger

Prediction probabilities

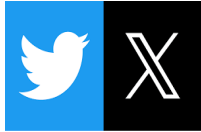


NOT anger

anger



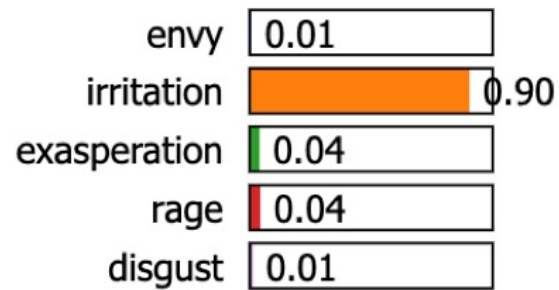
Text Explanation – LIME



Input: does anyone else hate obnoxious gym goers? #loud #wtf

Label: Irritation

Prediction probabilities

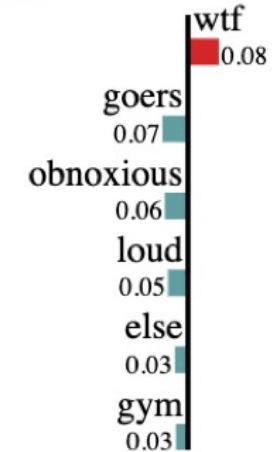
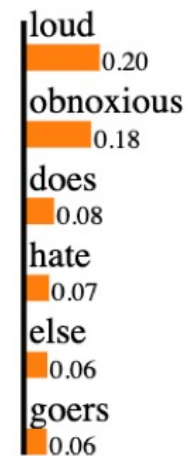


NOT irritation

irritation

NOT rage

rage

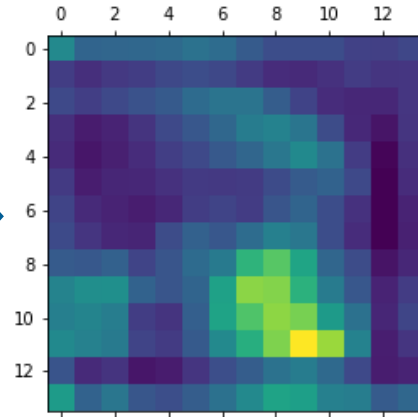


Visual Explanation – Grad-CAM

Input



Grad-CAM Heatmap



Sharp teeth important
in classification!

Output

Great White
Shark

Whale Shark

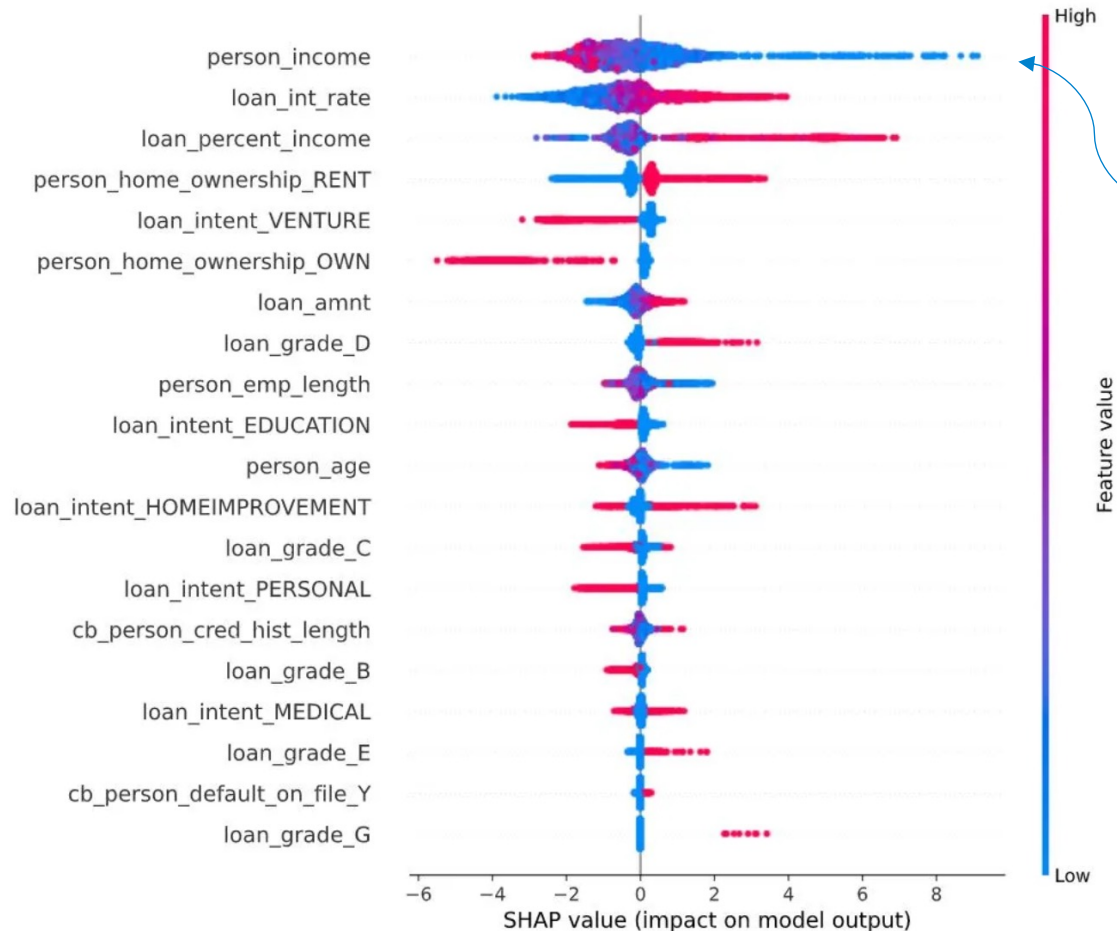
Feature Relevance Explanation – SHAP

- Quantifies the influence of each input variable (i.e., ranking of importance scores)
- SHAP = Shapley Additive exPlanations
 - Build a linear model around the instance to be explained
 - Shapley value of a feature = average expected marginal contribution to the model's decision

Feature Relevance Explanation – SHAP

Prediction: likelihood of loan default

Ordered by feature importance



Low income is positively associated with default

Counterfactual Explanation

- Assuming a data point was classified as negative (positive), how can we generate a new data point, as similar as possible to the original one, that the model would classify as positive (negative)?
- By comparing the new data point with the original we learn why the model changed its decision

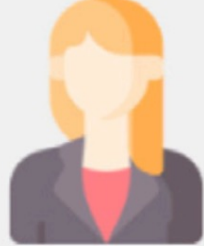
Explainability Affects Many Stakeholders

Data Scientist



- Understand the model
- De-bug it
- Improve its performance

Business Owner



- Understand the model
- Evaluate fit for purpose
- Agree to use

Model Risk



- Challenge the model
- Ensure its robustness
- Approve it

Regulator



- Check its impact on consumers
- Verify reliability

Consumer



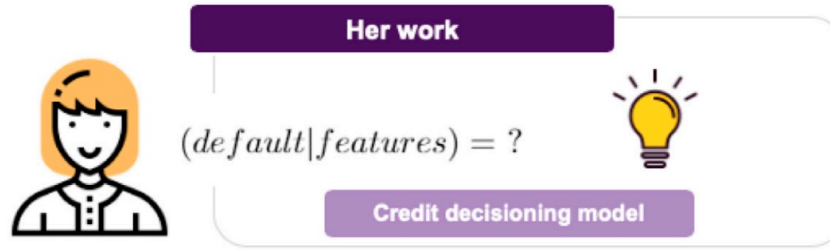
- “What is the Impact on me?”
- “What actions can I take?”

Jane, The Data Scientist

Principles and Practice of Explainable Machine Learning

Belle and Papantonis (2021)

Jane's Job



- Builds ML models for loan approvals (i.e., would like to predict the likelihood of default)
- Goal: maximize performance while maintaining explainability

Logistic Regression

Transparent, feature's coefficients tell you feature's impact on predicting loan default

Random Forest

Better performance but difficult to interpret the model's decisions

Which model would you choose?

0

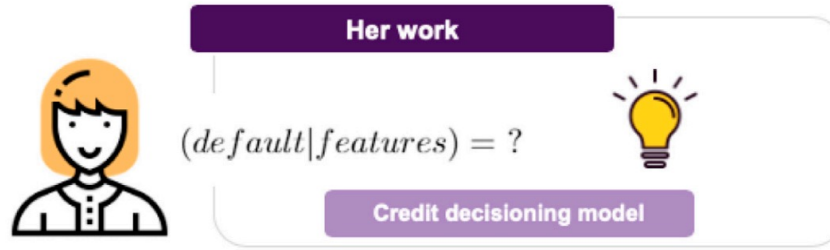
Logistic Regression

0

Random Forest

0

Jane's Job



- Builds ML models for loan approvals (i.e., would like to predict the likelihood of default)
- Goal: maximize performance while maintaining explainability

Logistic Regression

Transparent, feature's coefficients tell you feature's impact on predicting loan default

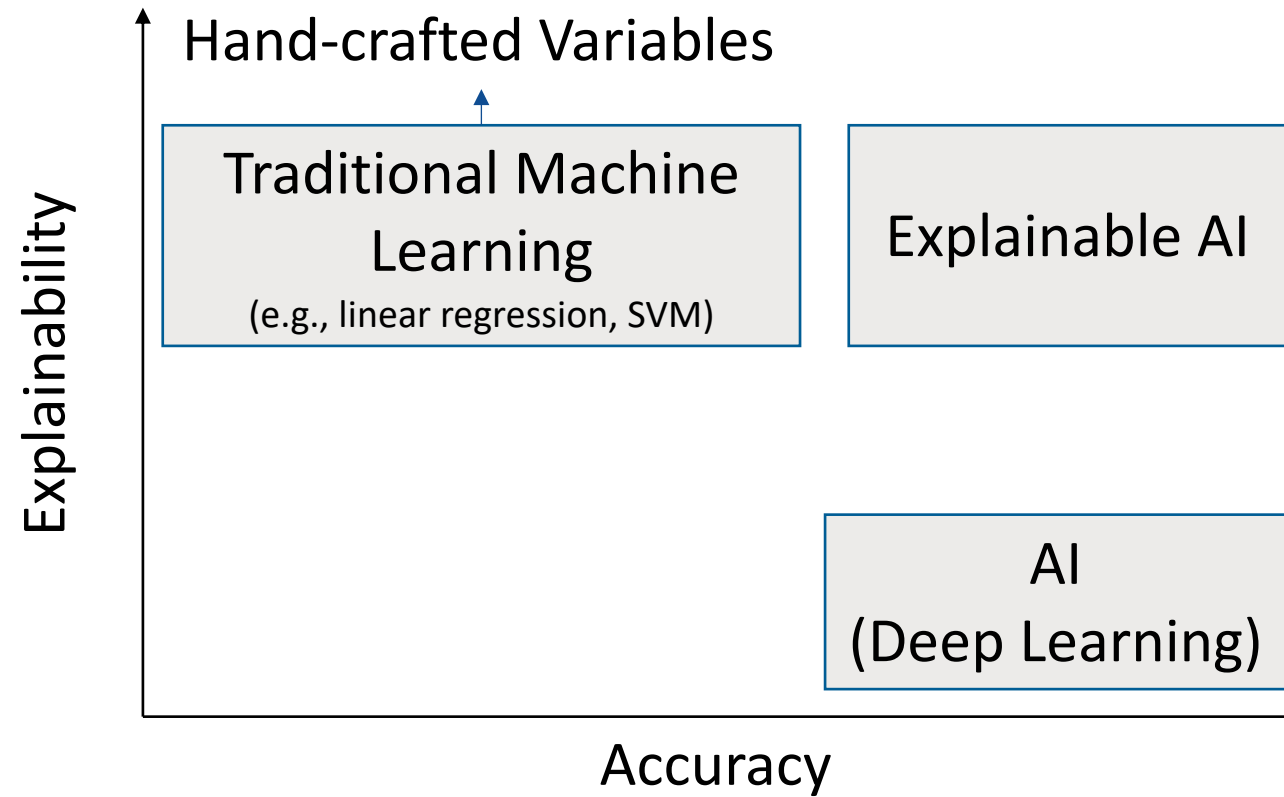
Random Forest

Better performance but difficult to interpret the model's decisions



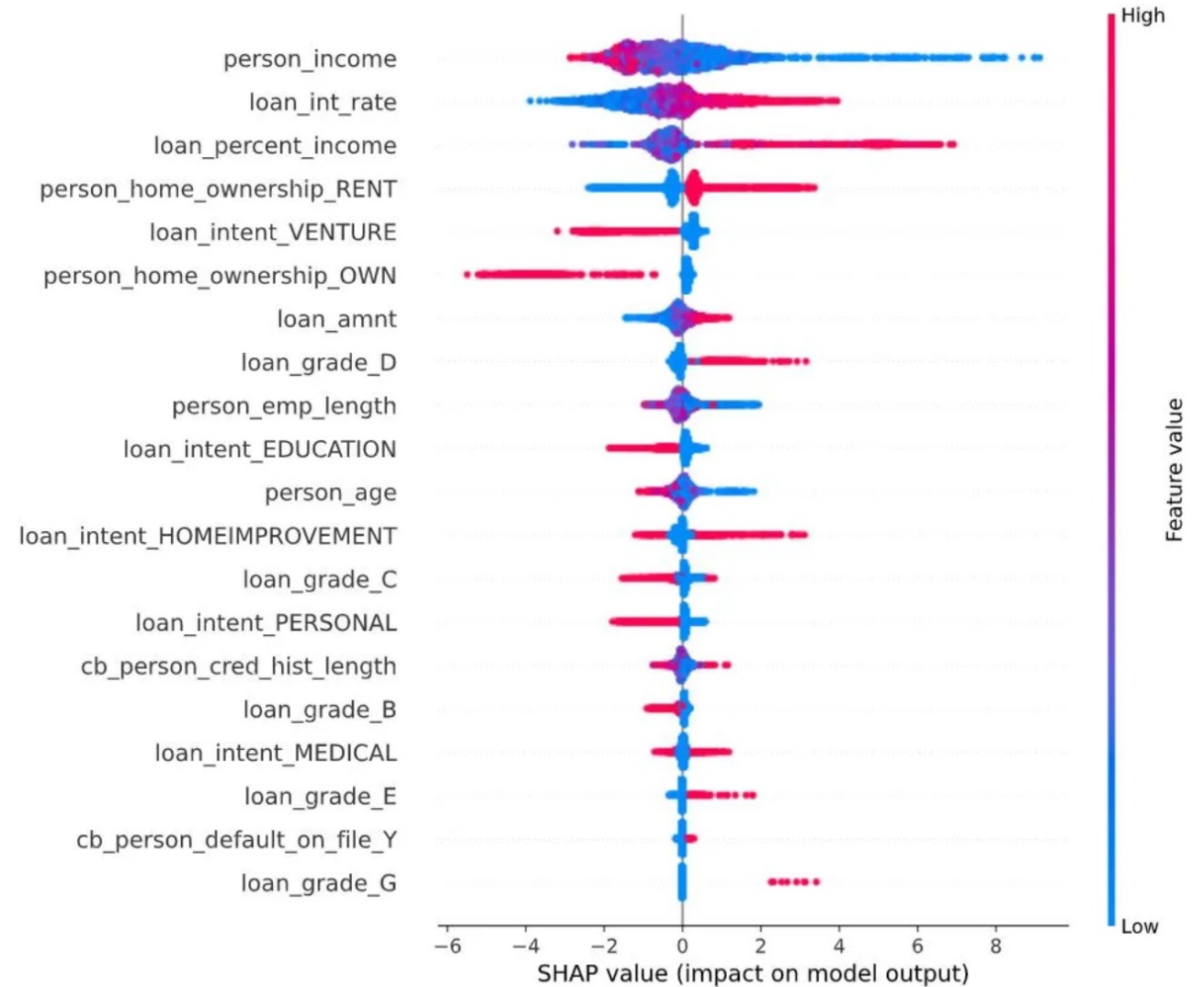
Better performance so she picks this one!

Goal



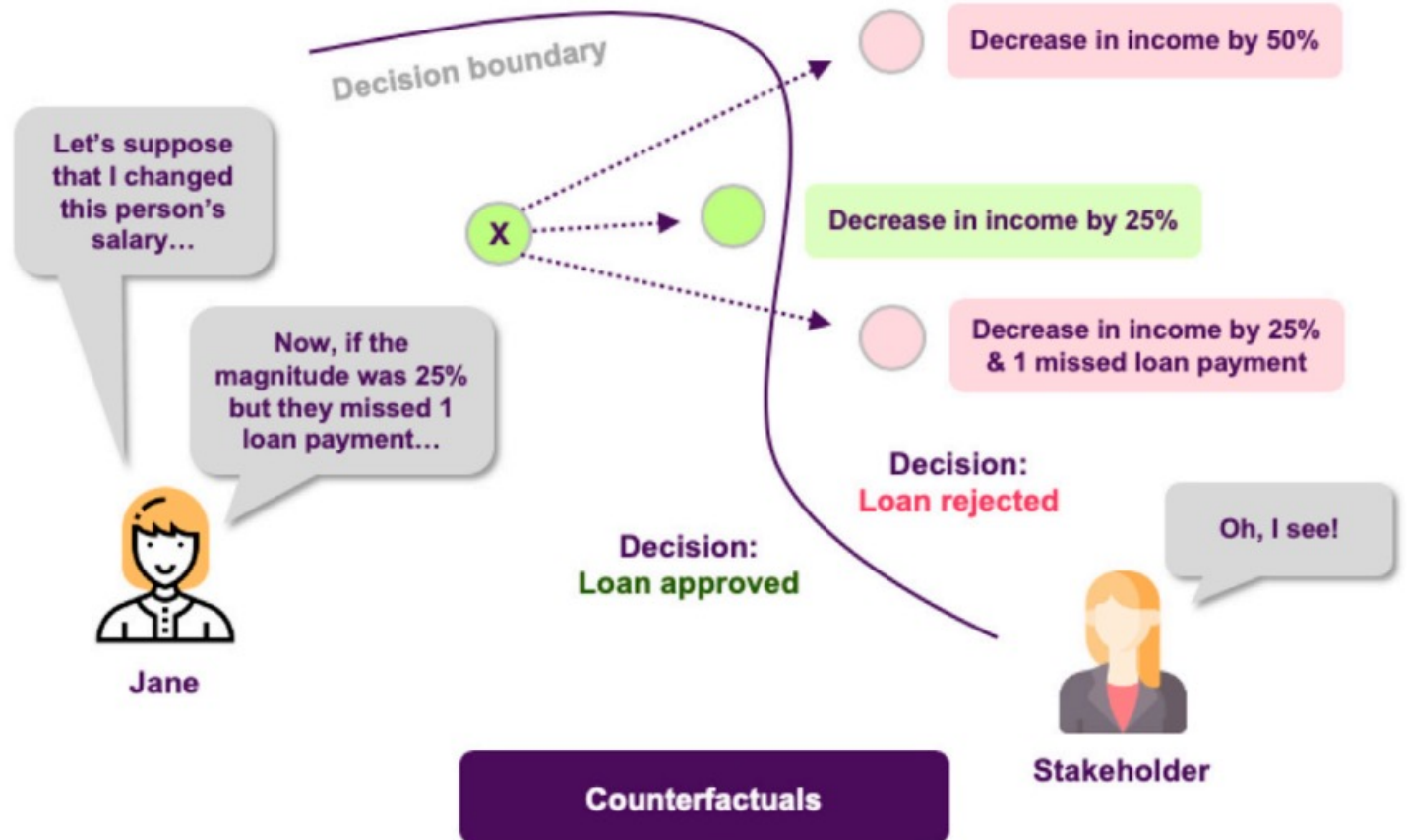
Tools for Explainability

- SHAP values to explain decision made by the model



Tools for Explainability

- Would different applications have been approved?
- Counterfactual analysis
- Learn that the missed loan payment resulted in loan rejection



In-Class Group Activity

As algorithms make more decisions, in what cases do you think it is important for customers to understand *why* algorithms made specific decisions?



B9651 – Marketing Analytics

Session 10: Price + Place

Professor Hortense Fong

Last Time

- Conjoint
 - Tool to quantify customers' preferences for products and services
 - Steps
 - Applications
 - Consumer preference + importance of attributes
 - Market share – my slice of the pie
- Diffusion of Innovation
 - Forecasting product adoption
 - Bass model
 - Predict how big the pie is
- Today: 4P Analytics – Price + Place

Today: Price & Place

Part 1: Pricing

1. Estimating the Demand Curve
 1. Direct Measures
 2. Conjoint Analysis
2. Estimating a Demand System: Montclair Video

Part 2: Place

1. Glaubinger Case
2. Omnichannel Behavior

Today's Goals

Understand:

- What is the demand curve
- What approaches can we use to estimate the demand curve
- How firms can optimize their distribution channels
- How to recognize the different effects associated with an omnichannel distribution

Be able to:

- Estimate a demand curve with direct measures and conjoint analysis
- Analyze different distribution channels

Course Roadmap

STP Analytics (Identify Value)	Customer Analytics (Deliver Value)	4P Analytics (Capture Value)
Module 1	Module 2	Module 3
What datasets can we use? How can we segment and target our customers? How should we position our products/services?	How much are our customers worth? Are our customers leaving? How do our customers make choices?	How do we build a new product? How should we price our products? How do we distribute them? How do we quantify the impact of our promotions?

Pricing

Direct Measure

Demand Curve Estimation

- Demand curve links price to demand (quantity)
- Goal: Optimize profit which is a function of demand curve
- We need to estimate the demand function
 - How demand changes with price
- How?
 - Direct measures – stated preference
 - Conjoint analysis – revealed preference



Demand Curve Estimation Using Direct Measures

How much would you be willing to pay for this product?

- This product
 - Increases blood circulation
 - Relaxes stressed muscles
 - Relieves aches and pains
- Ask how much a person is willing to pay
- What are the drawbacks?
 - This is not a realistic experience
 - How can we incentivize the respondent to be truthful?



Experimental Procedures

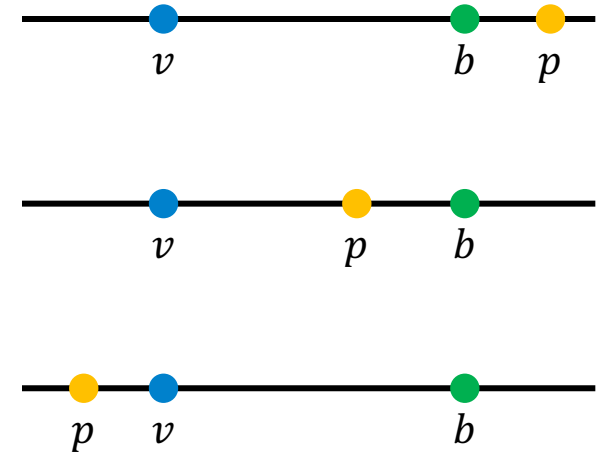
- Becker DeGroot Marschak (BDM) procedure
 - Incentive Aligned
 - Subject formulates a bid
 - Bid is compared to a price determined by a random number generator
 - If the subject's bid is greater than or equal to the price, he or she pays the (random) price and receives the item
 - If the subject's bid is lower than the price, he or she pays nothing and receives nothing
- What is the optimal response for the consumer?

BDM Optimal Response

v = true valuation, b = bid formulated by subject, p = random price

Case 1: $b > v$ (overbidding)

- Case 1a: $p > b$
 - Subject doesn't get the item (payoff = 0)
 - Same as if subject bid v since $p > b > v$
- Case 1b: $b \geq p > v$
 - Subject gets the item (payoff = $v - p < 0$)
 - Worse off than bidding v (payoff 0)
- Case 1c: $b > v \geq p$
 - Subject gets the item (payoff = $v - p \geq 0$)
 - Same as bidding v since what you pay is p



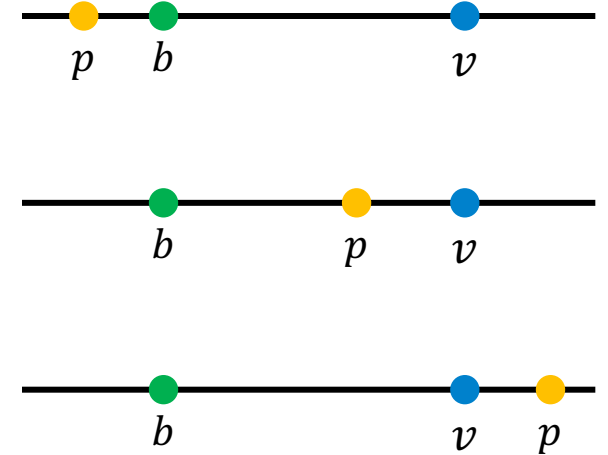
Note that what you pay is p and NOT b

BDM Optimal Response

v = true valuation, b = bid formulated by subject, p = random price

Case 2: $b < v$ (underbidding)

- Case 2a: $b \geq p$
 - Subject gets the item (payoff = $v - p \geq 0$)
 - Same as bidding v
- Case 2b: $v \geq p > b$
 - Subject doesn't get the item (payoff = $0 \leq v - p$)
 - Worse off than bidding v
- Case 2c: $p > v > b$
 - Subject doesn't get the item (payoff = 0)
 - Same as bidding v



The optimal bid is always to bid your true valuation v !

Your Turn: How Should This Product Be Priced?

- Assume marginal cost = \$0.50
- Data available on canvas (“Head Massager WTP”)
- Construct demand curve and find optimal price
 - Using data without incentives
 - Using data with incentives
- Note: we do not observe fixed costs so we need to find the price that maximizes **contribution** rather than profit
 - Profit = Demand x (Price – Marginal Cost) – Fixed Cost
 - **Contribution** = Demand x (Price – Marginal Cost)

How do the "optimal prices" compare? Incentive-aligned (BDM) vs. non-incentive aligned



Price suggested by incentive-aligned data > Price suggested by non-incentive aligned data

0

Price suggested by incentive-aligned data = Price suggested by non-incentive aligned data

0

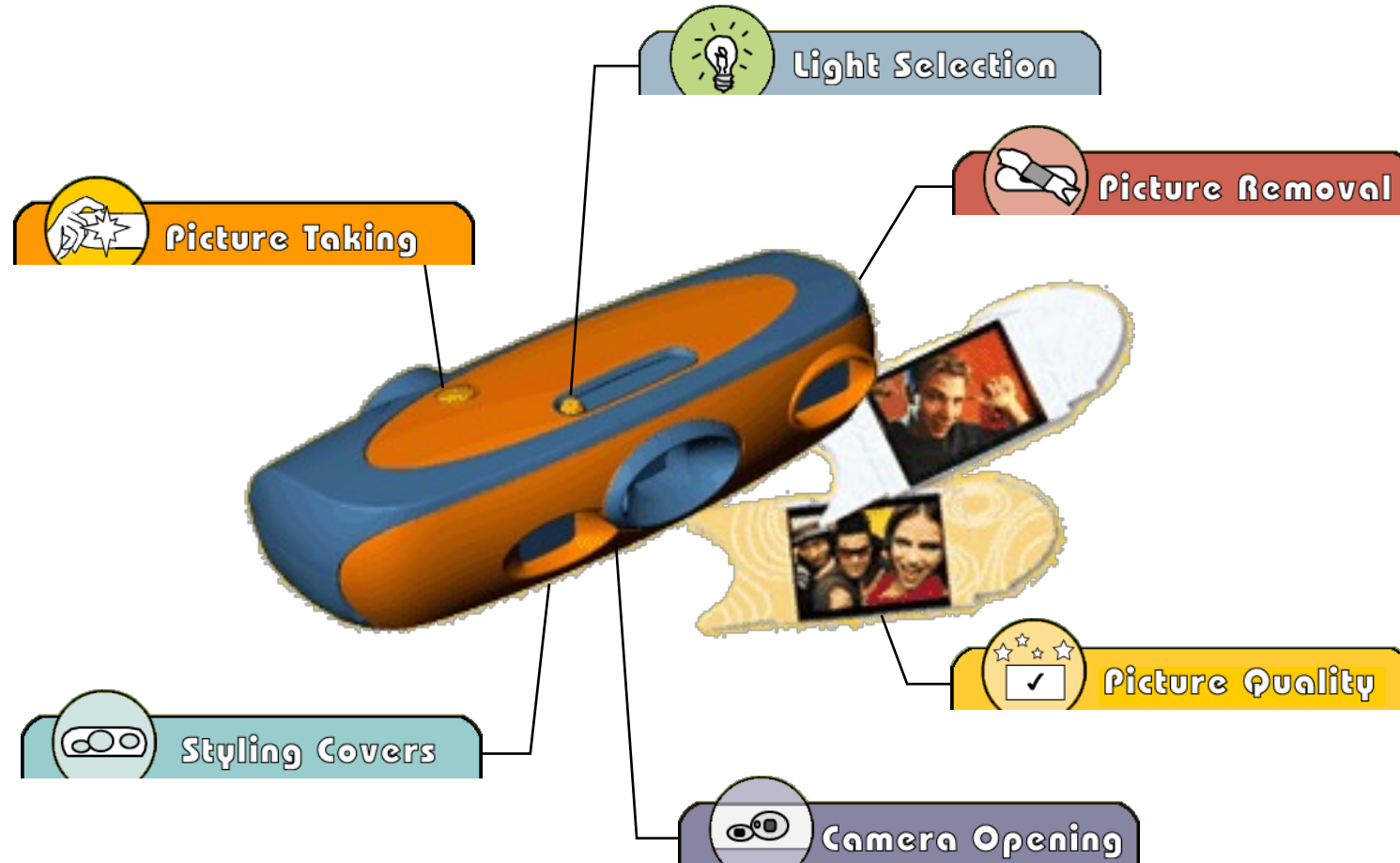
Price suggested by incentive-aligned data < Price suggested by non-incentive aligned data

0

Pricing

Conjoint Analysis

Demand Curve Estimation Using Conjoint Analysis







Conjoint Survey and Results



8 Questions

Ready for something a little different? Here's what to do:

1. We are going to show you 8 different pairs of cameras with some of the features and options we just showed you.
2. For each of the 8 pairs, compare the 2 different cameras.
3. Click on the white circle below to tell us how much you like one camera compared to the other.
4. Touch the yellow dots to see what each white circle means
5. Try the example question below...

Features	Camera A	Camera B
 Price	\$34.99	\$24.99
 Picture Removal	Manual	Automatic
 Picture Taking	2 Step	1 Step
 Styling Covers	Changeable	Permanent

Click on the feature icons for a reminder.

Need the scale? Touch the yellow dot

I like A completely more than B

I like B completely more than A

Next

Back

Help

Binary Attributes	Utility (Part-worth)
Price (\$35 vs. \$25)	22.65
Removable Covers	21.15
Picture Quality	30.51
Auto vs. 2-step	-0.84

How much more is this consumer willing to pay for removable covers

0

$21.15 \times 10 / 22.65 = \9.34

0

\$21.15

0

$22.65 / 21.15 = \$1.07$

0

$22.645 / 10 \times 21.151 = \47.90

0

I have no idea

0

Inferring Willingness to Pay Using Conjoint Analysis






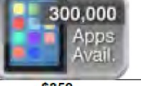
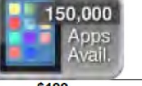
\$10 = 22.65 utils →

Exchange rate: 1 util = 10/22.65 dollars

Variables	Utility	WTP
Price (\$35 vs. \$25)	22.65	\$10.00
Removable Covers	21.15	\$9.34
Picture Quality	30.51	\$13.47
Auto vs. 2-step	-0.84	-\$0.37

Valuing Features – Apple vs. Samsung

TABLE 2. A SAMPLE CHOICE TASK.

With this camera:	12 MP Rear Camera, HD Video Recording, Autofocus, 2 MP Front Camera, Zoom 	8 MP Rear Camera, HD Video Recording, Autofocus, 2 MP Front Camera 	8 MP Rear Camera, HD Video Recording, Autofocus 	3 MP Rear Camera, Standard Video Recording, Autofocus 
With this size and weight:	10 inches, 2 lbs. 	9 inches, 1.75 lbs. 	8.5 inches, 1.5 lbs. 	7 inches, 1 lb. 
With this touchscreen:	Full Multi-Touch, Auto-Switch (1 to 2 Fingers), Rubberband, Tap to Re-center after Zoom 	Very Limited Multi-Touch, Auto-Switch (1 to 2 Fingers), Rubberband, Tap to Re-center after Zoom 	Full Multi-Touch, Rubberband, Tap to Re-center after Zoom 	Full Multi-Touch 
With this storage/memory:	64 GB (16,000 songs or 23,000 photos) 	32 GB (8,000 songs or 12,000 photos) 	16 GB (4,000 songs or 6,000 photos) 	8 GB (2,000 songs or 3,000 photos) 
With this connectivity:	WiFi, Bluetooth, MicroUSB, HDMI 	WiFi, Bluetooth, MicroUSB 	WiFi, Bluetooth 	WiFi 
With this number of apps available:	600,000 	450,000 	300,000 	150,000 
With this price:	\$659 	\$499 	\$359 	\$199 



Valuing Features – Apple vs. Samsung

Apple vs. Samsung, Prof. David Reibstein's critique:

“...the research was flawed because it did not include certain important attributes such as brand name, operating system, battery life etc. He memorably said that not including these fundamental features is like doing research ‘on cupholders in cars while ignoring the manufacturer name.’ ”

Apple vs. Samsung, Prof. Tulin Erdem's critique:

“You are elevating artificially the importance, the value of these things. They are not even in the radar screen of consumers. These are very granular ... and they wouldn't drive demand. ”

Apple was awarded \$539 million

Source: <https://www.trchome.com/blog/entry/how-apple-samsung-and-conjoint-came-together>

Break

Pricing

Montclair Video

Montclair Video

- Montclair video has decided to launch 4 plans:
 - 1-, 2-, 3- and 4- at a time, all with Blu-Ray available
- How much should they charge for each plan (monthly fee)?
- Head Massager case:
 - Build demand curve → build profit curve → find optimal price

Montclair Video

- Montclair: similar approach but some complications
 - Competitors
 - Our demand depends on our price and on our competitors' prices
 - Customers also have the choice of not buying anything
 - Multiple products
 - Our demand for each product will be influenced by the prices of the other products
 - Preference data for different features of the product, not for a fixed set of products
 - Instead of one X axis (price) and one Y axis (demand) we have many X axes (features of our and our competitors' products) and many Y axes (demand for each product in the line)
- Our approach
 - Create a demand system rather than a demand curve (will capture competition and feature-based products)
 - Optimize prices jointly (will capture product line optimization)

Your Turn: Montclair Video

- Data available on canvas (“Montclair Video inclass”)
 - Conjoint analysis results
- Compute profit (contribution) and find the price that maximizes profit
 - We need to jointly optimize profit

Multinomial Logit Reminder

$$Prob(\text{choose option } A) = \frac{\exp(U_A)}{\exp(U_A) + \exp(U_B) + \exp(U_C) + \exp(U_D)}$$

Option	Utility	Exp(Utility)	Choice Probability
A	1.00	2.72	0.03
B	2.00	7.39	0.09
C	3.00	20.09	0.24
D	4.00	54.60	0.64

Your Turn: Montclair Video – Market Shares

- Since we used a choice-based conjoint, market share can be computed in two ways
- Max Utility Rule: this is the same approach as we saw before
 - Assign each respondent to the product that maximizes utility
- Expected Market Share: **only** possible with **choice-based conjoint**
 - For each option o , expected market share is the sum of probabilities to choose that option across all respondents
 - Let X_i^o be a random variable that equals 1 if respondent i chooses o and 0 otherwise
 - $\mathbb{E}[MS_o] = \mathbb{E}[X_1^o + \dots + X_n^o] = \mathbb{E}[X_1^o] + \dots + \mathbb{E}[X_n^o] = \sum_{i=1}^n P(\text{respondent } i \text{ chooses } o)$
 - $P(\text{respondent } i \text{ chooses } o)$ corresponds to the MNL probability

Suppose Montclair were to charge the same prices as Netflix. How many consumers would choose each Montclair plan (1-, 2-, 3-, 4- at a time)?



25,18,5,7

0

9,8,12,6

0

13,4,3,1

0

What set of prices (1-, 2-, 3-, 4- at a time) would maximize Montclair's total contribution?

0

\$9.12, \$14.13, \$19.43, \$28.27

0

\$9.13, \$13.77, \$19.50, \$31.85

0

\$8.67, \$15.12, \$20.04, \$29.95

0

None of the above

0

Takeaways

- How do we measure customers' willingness to pay?
 - Incentive-aligned methods are not always easy to use
 - Conjoint analysis is a viable option (e.g., preferred in litigation cases)
- Conjoint analysis
 - Can also be used to simulate market shares
 - Market share simulations help with profit maximization
- Need to consider “incentive compatibility” when pricing product line
- Today's spreadsheets can be easily adjusted to other scenarios

Break

Place

Case: Glaubinger Tree Farm

- Assume that Glaubinger's marginal cost is identical under both channels (using retailer and DTC)
 - Glaubinger needs to understand the price he would charge and the corresponding profits under both approaches
1. What price should Glaubinger charge to consumers in the direct-to-consumer channel? Assume a linear demand function $Q=a+b*P$.
 2. What wholesale price should Glaubinger charge to Green Acres Nursery in the traditional channel in 2009?

Price in DTC Channel

- Notation:

- Wholesale price W
- Marginal cost C
- Retail price P
- Quantity sold $Q = A + BP$



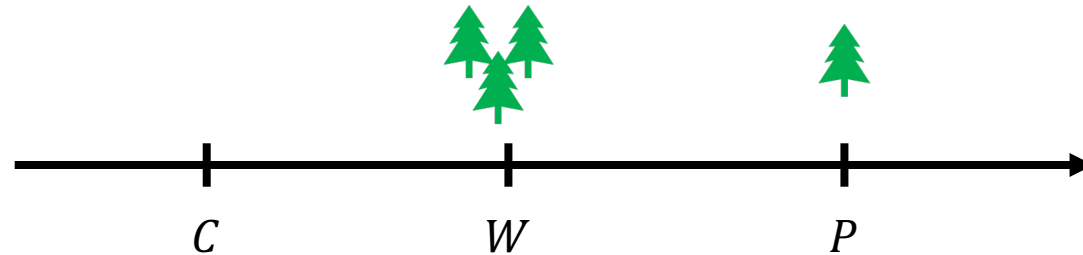
- Process: Wholesaler sets a retail price and sells to customers
- Optimal price in direct-to-consumer channel:
 - Wholesaler maximizes profit: $(P - C) * (A + BP)$
 - Why? In DTC, $W = P$
 - Derivative wrt P : $A + 2BP - BC = 0$ iff $P = (BC - A)/(2B)$

Price in Traditional Channel

- Notation:

- Wholesale price W
- Marginal cost C
- Retail price P
- Quantity sold $Q = A + BP$

Process: Wholesaler sets a price W and sells to the retailer. The retailer then sets a price P and sells to customers.



- Optimal price in Traditional channel: **we work backward!**

- Retailer reacts to W and maximizes $(P(W) - W) * (A + BP(W))$
 - Derivative: $BP(W) - BW + A + BP(W) = 0$ iff $P(W) = (BW - A)/(2B)$
- Wholesaler sets W to maximize $(W - C) * (A + BP(W)) = (W - C) * \left(A + B \frac{BW - A}{2B}\right)$
 - Derivative: $\frac{A + BW}{2} + (W - C) * \frac{B}{2} = 0$ iff $W = \frac{BC - A}{2B}$

Your Turn

Plot the data

Calculate A, B, and profit in both business models

Marginal cost = \$10

Case: Glaubinger Tree Farm

- Glaubinger is in fact uncertain about what its marginal cost would be under the direct-to-consumer channel (due to additional distribution costs)
- What is (approximately) the maximum direct-to-consumer marginal cost to Glaubinger under which it would still be profitable to switch to the direct-to-consumer channel?

Uncertain Marginal Cost

- Notation:

- Wholesale price W
- Uncertain marginal cost \tilde{C}
- Retail price P
- Quantity sold $Q = A + BP$

- Optimal profit in direct-to-consumer channel:

$$\left(\frac{B\tilde{C} - A}{2B} - \tilde{C}\right) * \left(A + B \frac{B\tilde{C} - A}{2B}\right) = -\frac{(A + B\tilde{C})^2}{4B}$$

DTC Profit: $(P - \tilde{C}) * (A + BP)$

$$P = (B\tilde{C} - A)/(2B)$$

- Optimal profit in traditional channel:

$$\left(\frac{BC - A}{2B} - C\right) * \left(A + \frac{1}{2}\left(B \frac{BC - A}{2B} - A\right)\right) = -\frac{(A + BC)^2}{8B}$$

Wholesaler Profit: $(W - C) * \left(A + \frac{BW - A}{2}\right)$

$$W = \frac{BC - A}{2B}$$

- At worst, profits must be equal: set profits equal and solve for \tilde{C}

$$\tilde{C} = \frac{1}{B} * \left(\frac{A + BC}{\sqrt{2}} - A\right)$$

Place

Omnichannel

Omnichannel Behavior: Impact of New Touchpoint

MANAGEMENT SCIENCE

Vol. 56, No. 1, January 2010, pp. 4–24
ISSN 0025-1909 | EISSN 1526-5501 | 10 | 5601 | 0004

informs

DOI 10.1287/mnsc.1090.1066
© 2010 INFORMS

Cost Structure, Customer Profitability, and Retention Implications of Self-Service Distribution Channels: Evidence from Customer Behavior in an Online Banking Channel

Dennis Campbell, Frances Frei

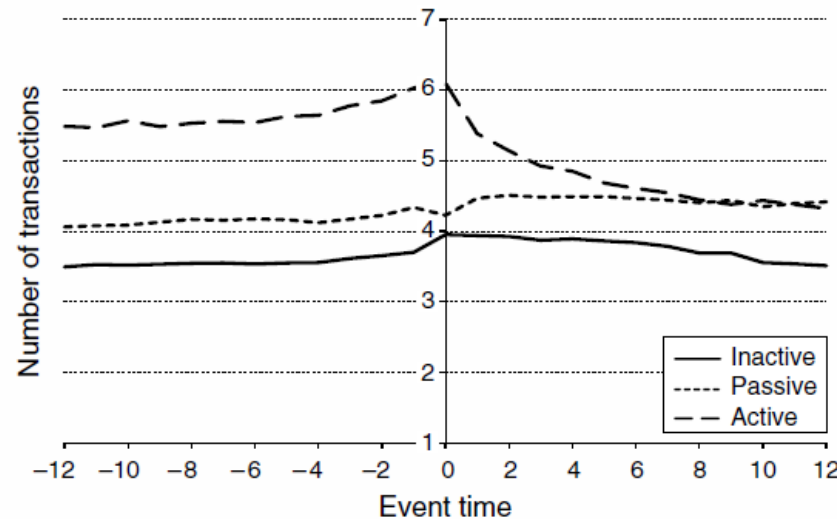
Harvard Business School, Harvard University, Boston, Massachusetts 02163
{dcampbell@hbs.edu, ffrei@hbs.edu}

- Sample of 200,000 retail banking customers at large retail bank
- 30-month period (12/04-05/07)
- Online Banking introduced in the middle of the period
- Inactive customers: did not adopt online banking
- Passive customers: adopted online banking but infrequent usage
- Active customers: adopted online banking and frequent usage

Substitution Effect

- Online banking adopters relatively more likely to **reduce** their use of self-service touchpoints:
 - ATM and automated call centers (VRU)

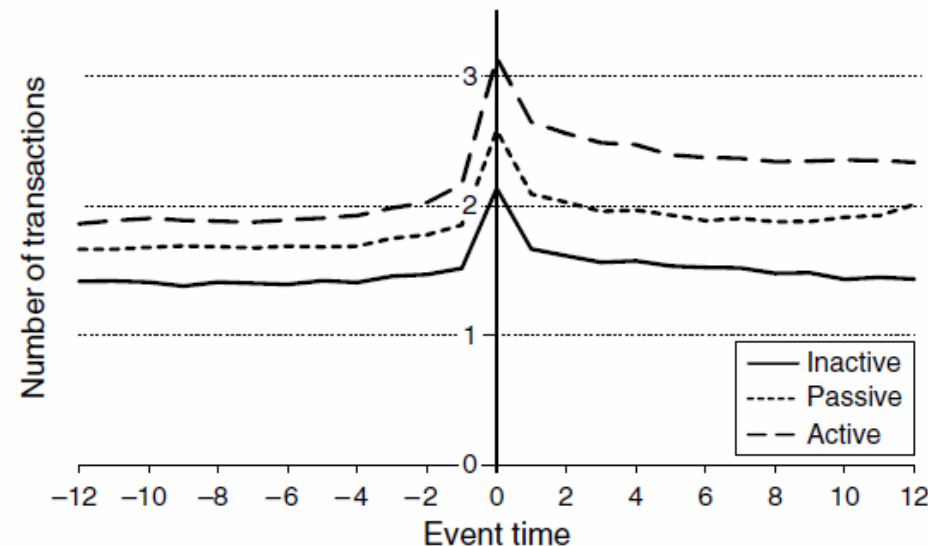
Figure 1(a) Transactions in Offline Self-Service Channels:
ATM and VRU



Augmentation Effect

- Online banking adopters relatively more likely to **increase** their use of assisted-service touchpoints:
 - branch and call centers

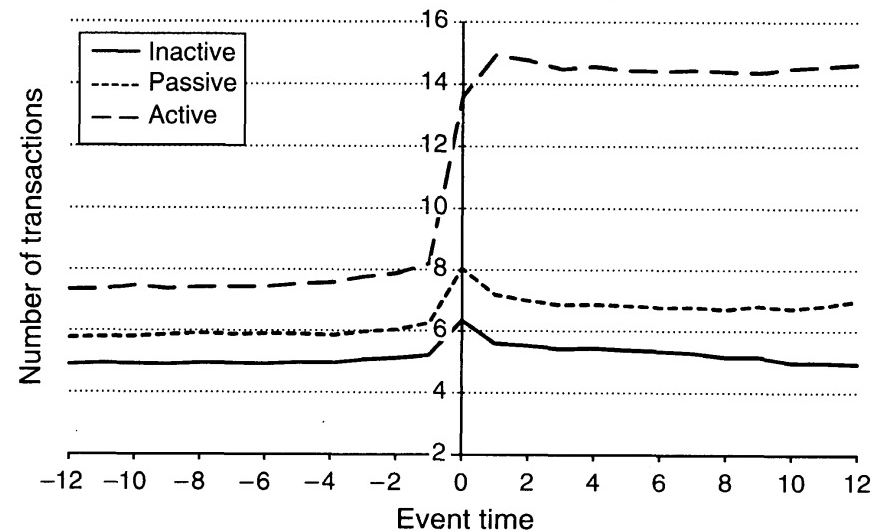
Figure 1(b) Transactions in Offline Assisted-Service Channels:
Branch and Call Center



Volume Effect

- Online banking adopters relatively more likely to **increase** their total volume of transactions

Figure 1(c) Total Transaction Volume: Online and Offline Channels



Net Impact on Cost to Serve

- Active adopters of online banking became relatively **more costly** to serve after the introduction of online banking

Table 5 Estimate of Change in Normalized Cost to Serve for Active Adopters

Channel	Normalized cost (\$)	Preadoption mean transactions	Preadoption normalized cost (\$)	Change (%)	Postadoption normalized cost (\$)
Branch	1.00	1.75	1.75	9	1.90
Call center	0.94	0.31	0.29	4	0.30
ATM	0.31	3.41	1.06	-10	0.95
VRU	0.18	2.9	0.52	-59	0.21
Online	0.09	0	0.00	N/A	0.90
Total			3.62		4.27

Notes. This table shows the estimated change in normalized cost to serve for the sample of active adopters. The unit cost of performing a transaction in each channel is normalized by the unit cost of a branch transaction to preserve confidentiality in National Bank's cost estimates. Preadoption mean transactions by channel are taken from Table 2. The estimates of percentage change in transactions attributable to adoption of online banking for each channel are computed using the coefficient estimates from Table 3. The postadoption normalized cost estimate in the final column assumes the average level of monthly postadoption transactions for active adopters of 10 per month.

Impact on Retention

- Retention of online banking adopters was **higher** than retention of customers who did not adopt online banking

Table 7 Descriptive Statistics by Online Status for a Random Sample of Customers as of January 2004

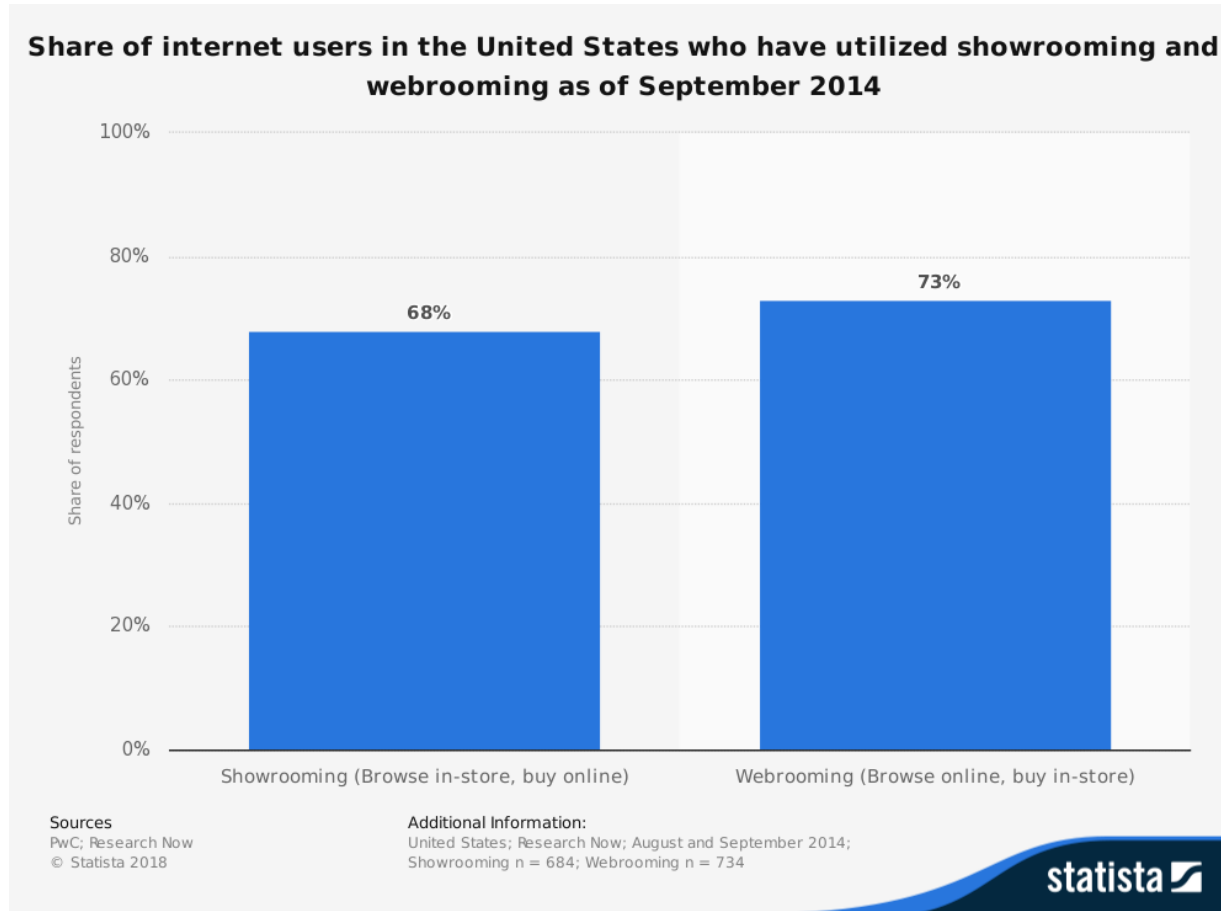
	Offline customers (<i>N</i> = 69,871)		Online customers (<i>N</i> = 30,129)	
	Mean	Std. dev.	Mean	Std. dev.
Retention—1 year	0.86	0.35	0.90	0.30
Retention—2 years	0.78	0.42	0.84	0.37
Retention—3 years	0.71	0.45	0.79	0.41
Tenure	10.33	12.25	8.29	7.54
Age	45.34	20.77	38.90	13.79
Number of deposit products	1.02	0.97	1.79	1.23
Number of loan products	0.32	0.58	0.53	0.75
Number of investment products	0.11	1.94	0.04	0.30
Deposit account balances	8,781	53,778	10,530	58,718
Loan account balances	3,240	17,911	4,127	19,487
Investment account balances	1,935	29,204	483	6,587
Annual customer profitability	196.8	1,056	217.9	1,037

Notes. “Offline” customers = customers who have not adopted online banking as of December 2003. “Online” customers = customers who have adopted the online channel as of December 2003. All differences in means between these groups are significant at least at the 10% level using *t*-tests.

Omnichannel Behavior: Impact of New Touchpoint

- *Substitution effect*: use new channel instead of old
- *Augmentation effect*: use older channels more because of new channel
- *Volume effect*: increase in total transaction volume
- Even if the new touchpoint has a lower marginal cost compared to existing touchpoints, cost to serve customers might increase
- Impact on CLV depends on impact on retention and revenue

Omnichannel in Retail: Showrooming, Webrooming



Omnichannel: Online to Offline

WARBY PARKER
eyewear



HOME TRY-ON

5 days
5 pairs
100% free

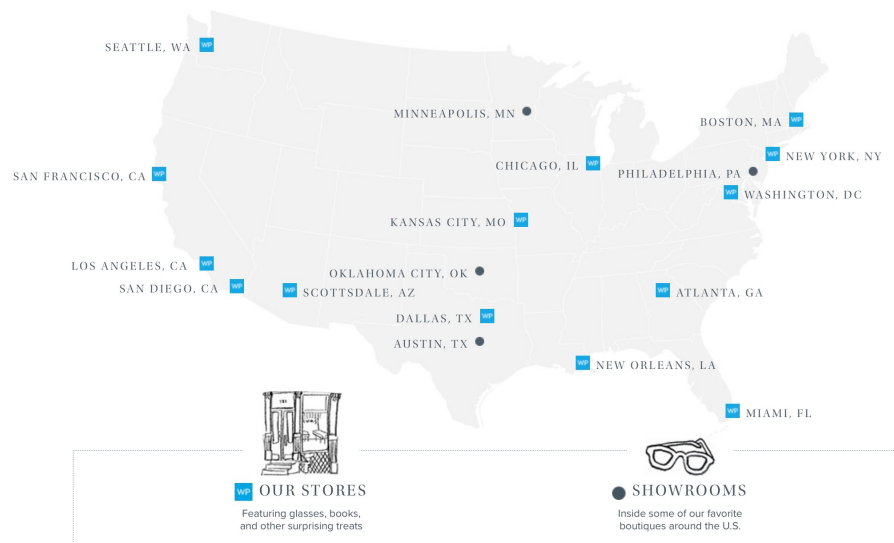
GET STARTED

Sources: Bell, Gallino, Moreno (2014), "How to win in an omnichannel world," *MIT Sloan Management Review*; Bell, Gallino, Moreno (2017), "Offline showrooms in omnichannel retail: Demand and operational benefits," *Management Science*.

Warby Parker's “Display Only” Showrooms

Retail Locations

Our retail spaces combine the snappy ease of online ordering with the fun and serendipity of real-life shopping (with a photo booth or two). We can't wait to meet you!



Warby Parker's "Display Only" Showrooms: Impact

- Warby Parker opened 20 new display only stores across US between February 2010 and March 2013
 - Customer tries product in store, order fulfilled online
- Where new stores were opened:
 - Overall sales in area increased by 9% ("volume effect")
 - Website sales increased by 3.5% ("augmentation effect")
 - Sales through "Home Try-On" program went down by 5.5% ("substitution effect")
 - But conversion rate of "Home Try-On" increased by 0.8%
 - The "pickier" customers use "display only" stores
 - Returns decreased by 1.2% - better matching

Omnichannel Initiative by Tesco



<https://www.youtube.com/watch?v=fGaVFRzTTP4>

Bonobos

MIT Sloan Management Review

The Store Is Dead— Long Live the Store


Legacy offline stores and next-generation online retailers are each finding their way to a new kind of shopping experience: the showroom.


BY DAVID R. BELL, SANTIAGO GALLINO, AND ANTONIO MORENO

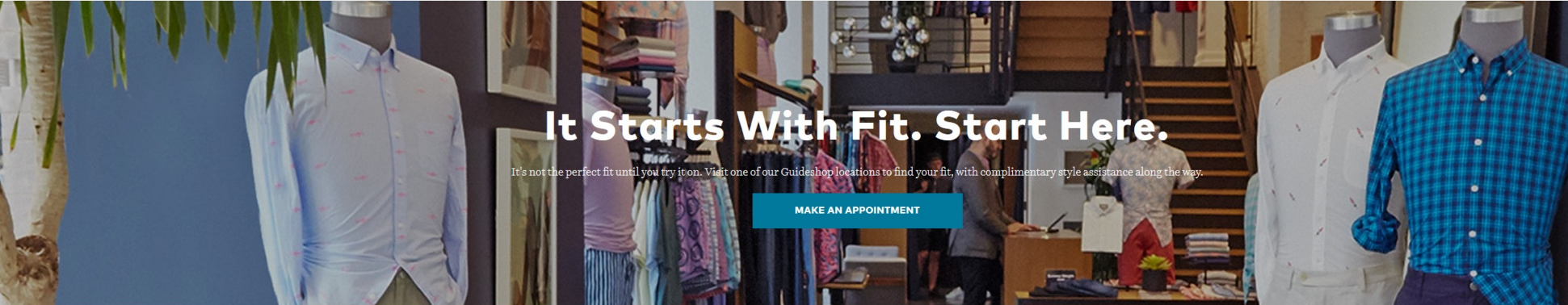
BONOBOS



Bonobos' "Guide Shops"

[Shop](#)[Final Sale](#)[Extended Sizes](#)[Gift Guide](#)[Visit Us](#)






It Starts With Fit. Start Here.

It's not the perfect fit until you try it on. Visit one of our Guideshop locations to find your fit, with complimentary style assistance along the way.


[MAKE AN APPOINTMENT](#)

How does it work?




1-on-1 Attention

A knowledgeable Guide will walk you through the entire Bonobos assortment.




Find The Right Fit

Fit Matters. A Guide will help you find your perfect fit from all of our available options.



Walk Out Hands-Free

A Guide will place your order and it will be shipped directly to your home or office for free.



Returns & Exchanges

Need to make a return or exchange? A guide will process your returns in-store and help with any size or fit questions, free of charge.

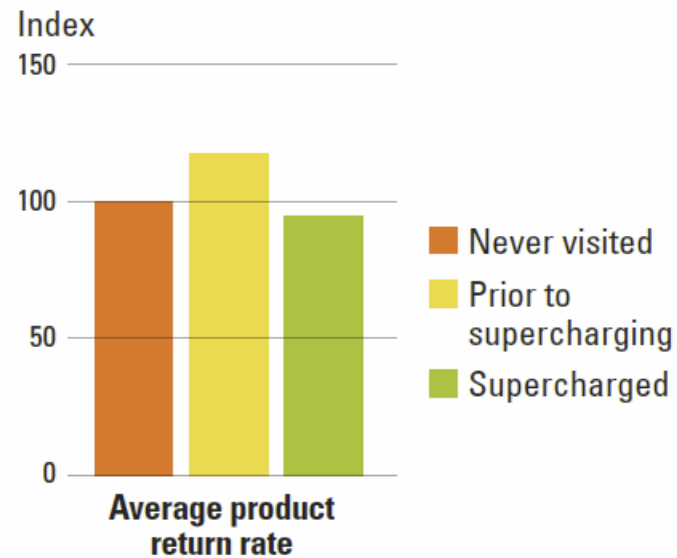
In the area? Walk-ins are always welcome. Weekends tend to be busy, so we suggest booking ahead if you can.

Source: <https://bonobos.com/guideshop>

Showroom Visits and Returns

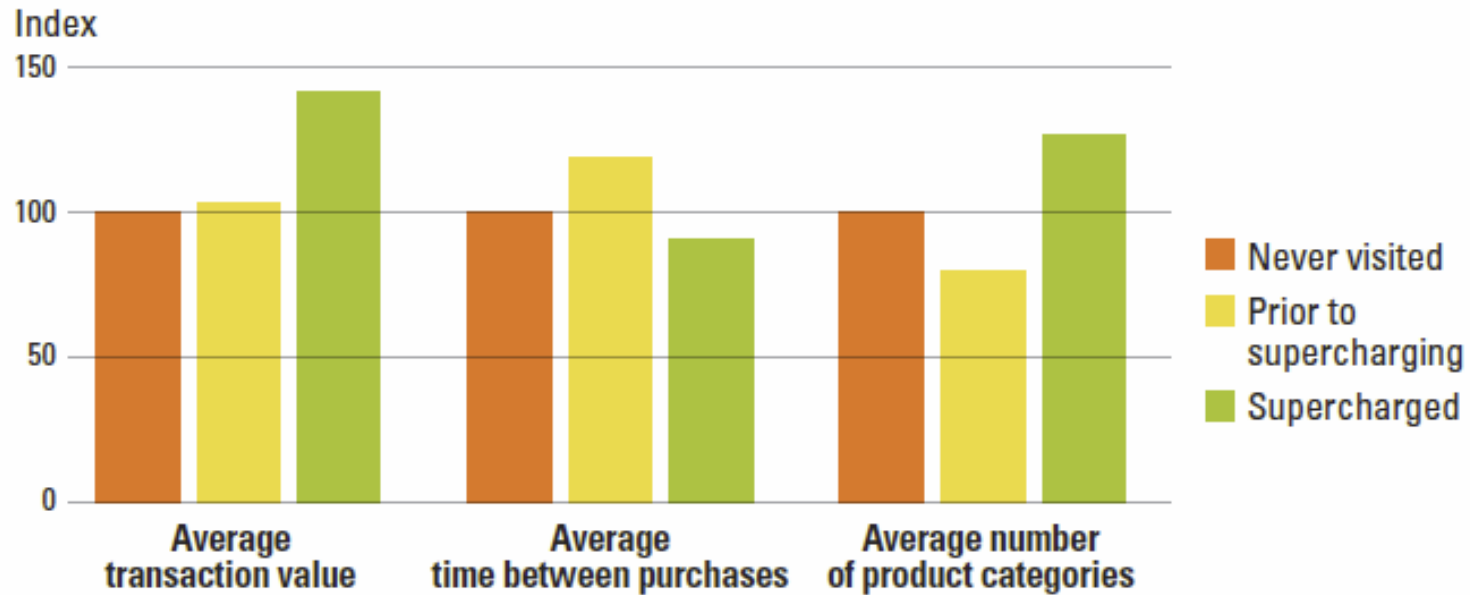
EFFECT OF SHOWROOM VISITS ON RETURNS

Customers who experience the brand offline and are thereby “supercharged” are less likely to return the items they purchase.



Shopping Patterns

EFFECT OF SHOWROOM VISITS ON SHOPPING PATTERNS



Customers who experience the brand offline and are thereby “supercharged” spend more per transaction, shop more frequently, and buy in more categories.

Takeaways

- Double Marginalization
 - When two players at different steps of the supply chain each optimize their own profits and apply a mark-up to their prices
- Omnichannel
 - *Substitution effect*: use new channel instead of old
 - *Augmentation effect*: use older channels more because of new channel
 - *Volume effect*: increase in total transaction volume

Next Class

- Promotion
- Due:
 - Pricing Concept Check due before next class
 - Second Group Assignment due Nov 18 at 8PM
 - Readings
 - Rocket Fuel: Measuring the Effectiveness of Online Advertising
 - Chirpin' Tavern's Coupon Promotion
 - Begin working on Second Individual Assignment (4P Analytics)
 - Due Nov 23 at 8PM



B9651 – Marketing Analytics

Session 11: Advertising & Promotion

Professor Hortense Fong

Logistics

- Final Thurs, Dec 12 @ 2:30-4:30PM
 - One-page cheat sheet (front and back)
 - Week 6 onwards
 - Calculator
- Second individual assignment (4P Analytics) due Sat, Nov 23 at 8PM
- Third individual assignment (Arteq) due Mon, Dec 2 at 8PM

Last Time

- Price
 - Incentive-Aligned Direct Measure - BDM
 - Conjoint
- Place
 - Distribution Channels
 - Omnichannel Distribution
- Today: 4P Analytics – Advertising & Promotion

Today: Advertising & Promotion

Part 1: Advertising

1. Search vs Display
2. Measuring Ad Effectiveness
 1. Attribution Problem
 2. Rocket Fuel
3. Novel Approaches

Part 2: Promotion

1. Evaluating Price Promotions

Today's Goals

Understand:

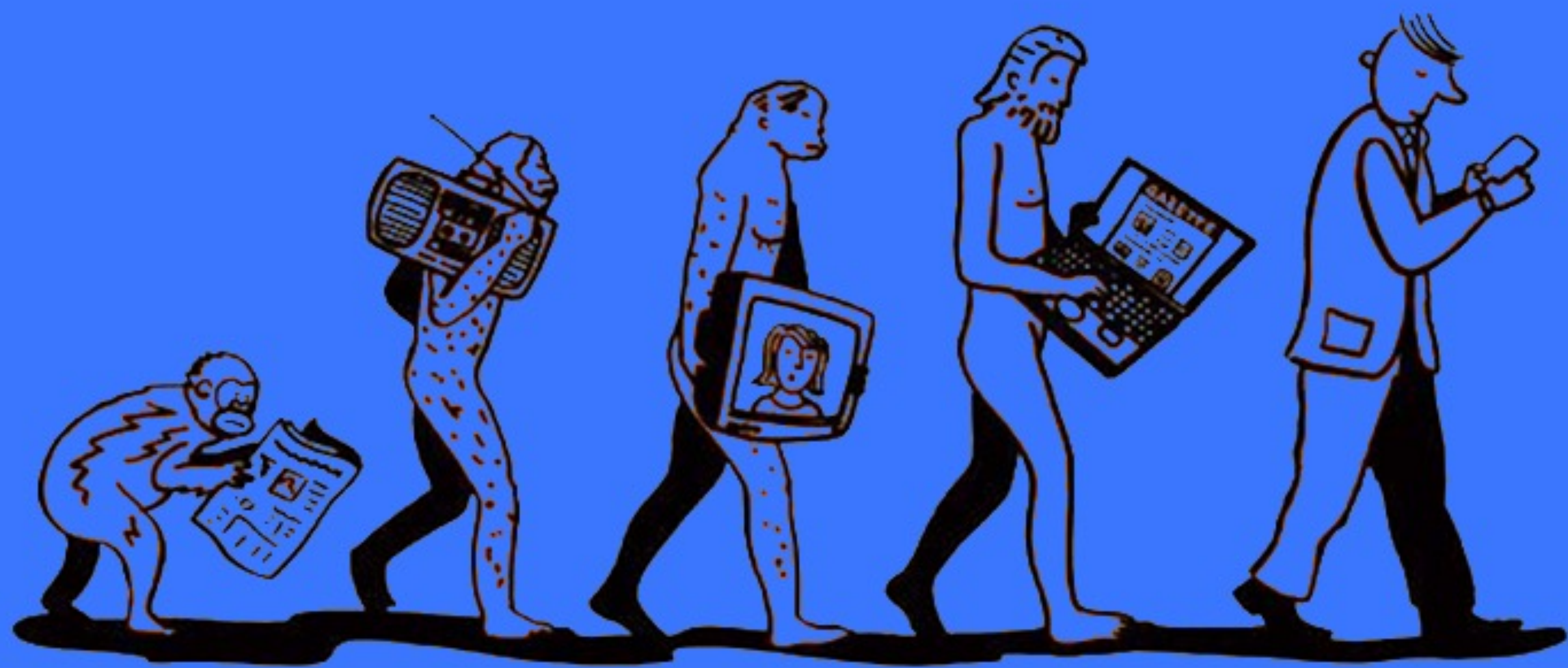
- The difference between display and search ads
- The attribution problem and its link with the purchase funnel
- The different factors that affect the evaluation of a promotion

Be able to:

- Run an A/B test and evaluate the effectiveness of marketing efforts
- Assess the effectiveness of a promotion through linear regression

Course Roadmap

STP Analytics (Identify Value)	Customer Analytics (Deliver Value)	4P Analytics (Capture Value)
Module 1	Module 2	Module 3
What datasets can we use? How can we segment and target our customers? How should we position our products/services?	How much are our customers worth? Are our customers leaving? How do our customers make choices?	How do we build a new product? How should we price our products? How do we distribute them? How do we quantify the impact of our promotions?



Traditional Advertising

TV Advertising: Media Exposure Measurement

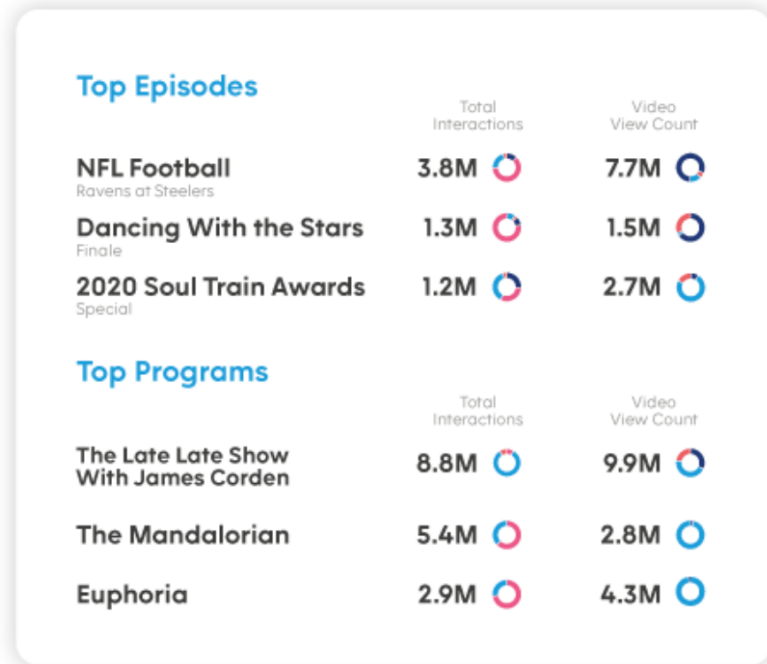
Ad Context Measures:

- Rating = Number of TVs tuned to a particular program / Total households with TV
- Share = Number of TVs tuned to a particular program / Total TVs that are on
- <https://www.nielsen.com/top-ten/>

Ad Measures:

- Reach = Percent of target audience exposed at least once to the campaign
- Frequency = Average number of times an individual who is reached in the target audience is exposed to the message
- Gross Rating Points = Reach x Frequency

Nielsen Social



Social Content Ratings® TV

The leading social solution for networks, streaming providers, content producers, advertisers and agencies looking to understand the impact of their social media strategies, talent and the buzz generated by fans talking about their programs.

Learn more >

Source: <http://www.nielsensocial.com/socialcontentratings/weekly>

Online Advertising

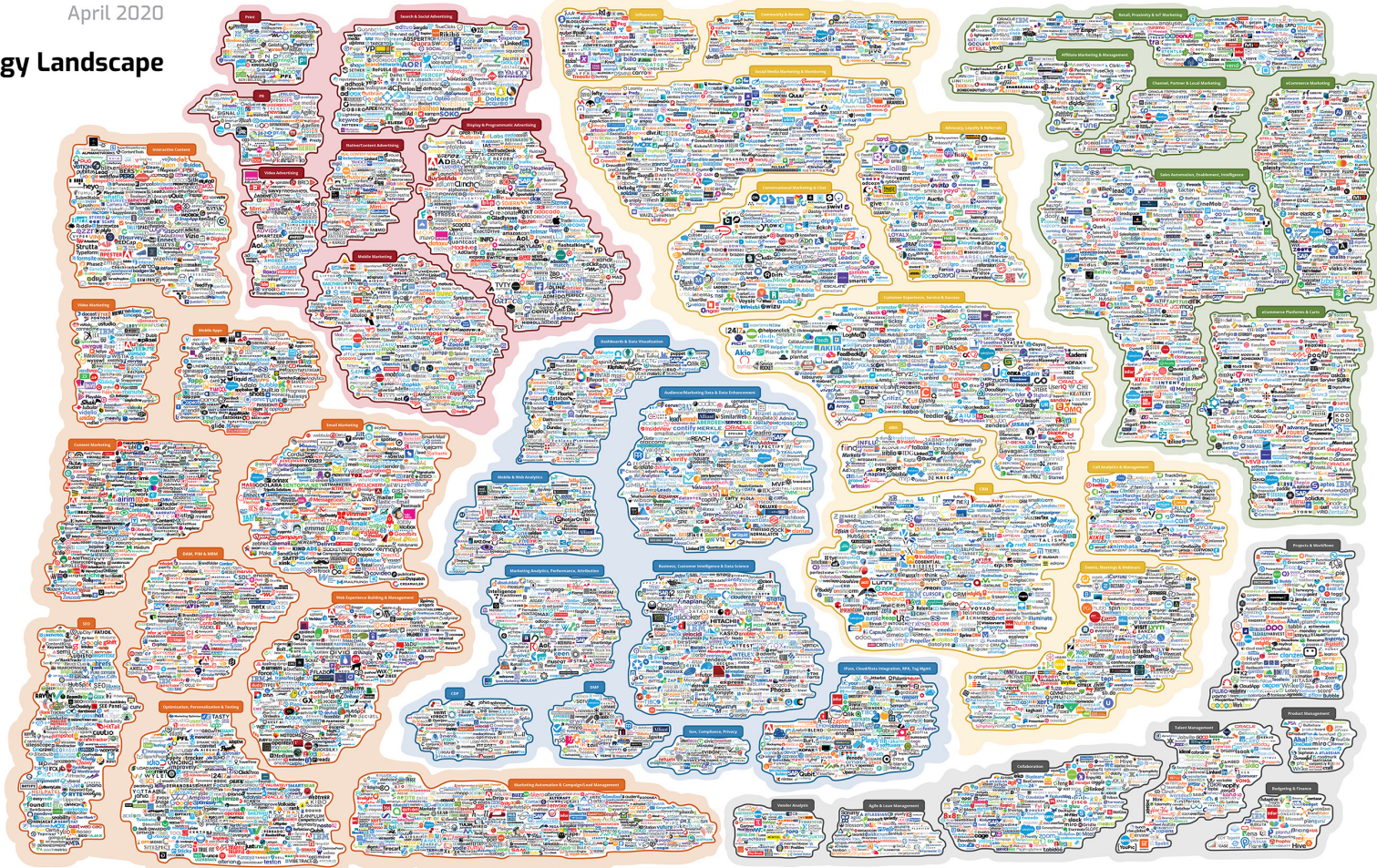
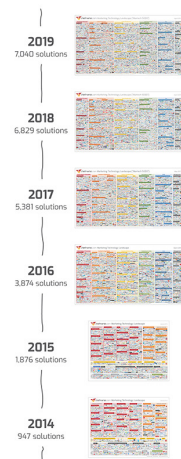


April 2020

Marketing Technology Landscape The Martech 5000

Total Solutions	8,000
Advertising & Promotion	922
Content & Experience	1,936
Social & Relationships	1,969
Commerce & Sales	1,314
Data	1,258
Management	601

Access all the data of this landscape & more at martech5000.com

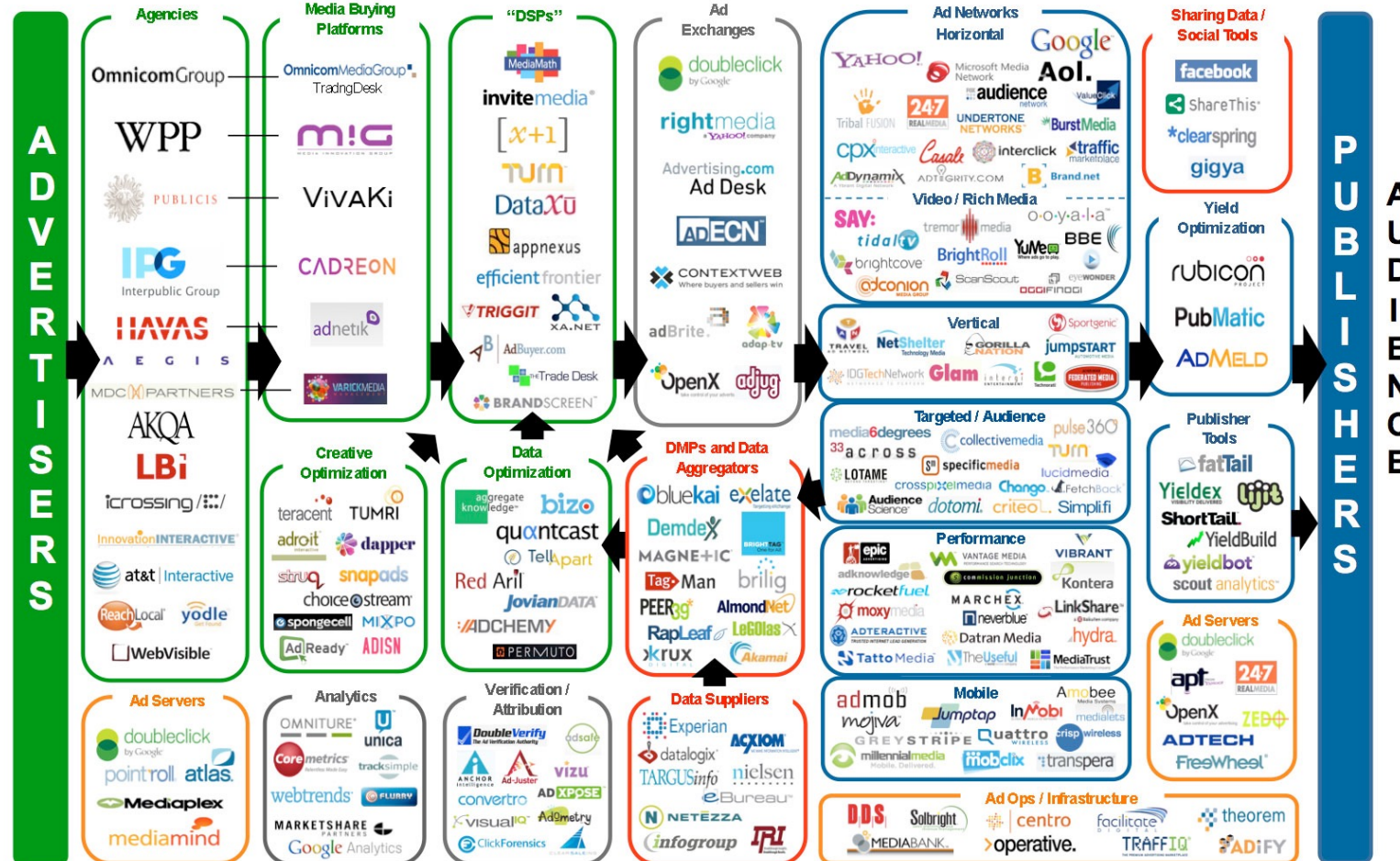


Copyright © 2020 Marketing Technology Media, LLC. See <https://chiefmartec.com/2020/04/marketing-technology-landscape-2020-martech-5000/> for details and sources.

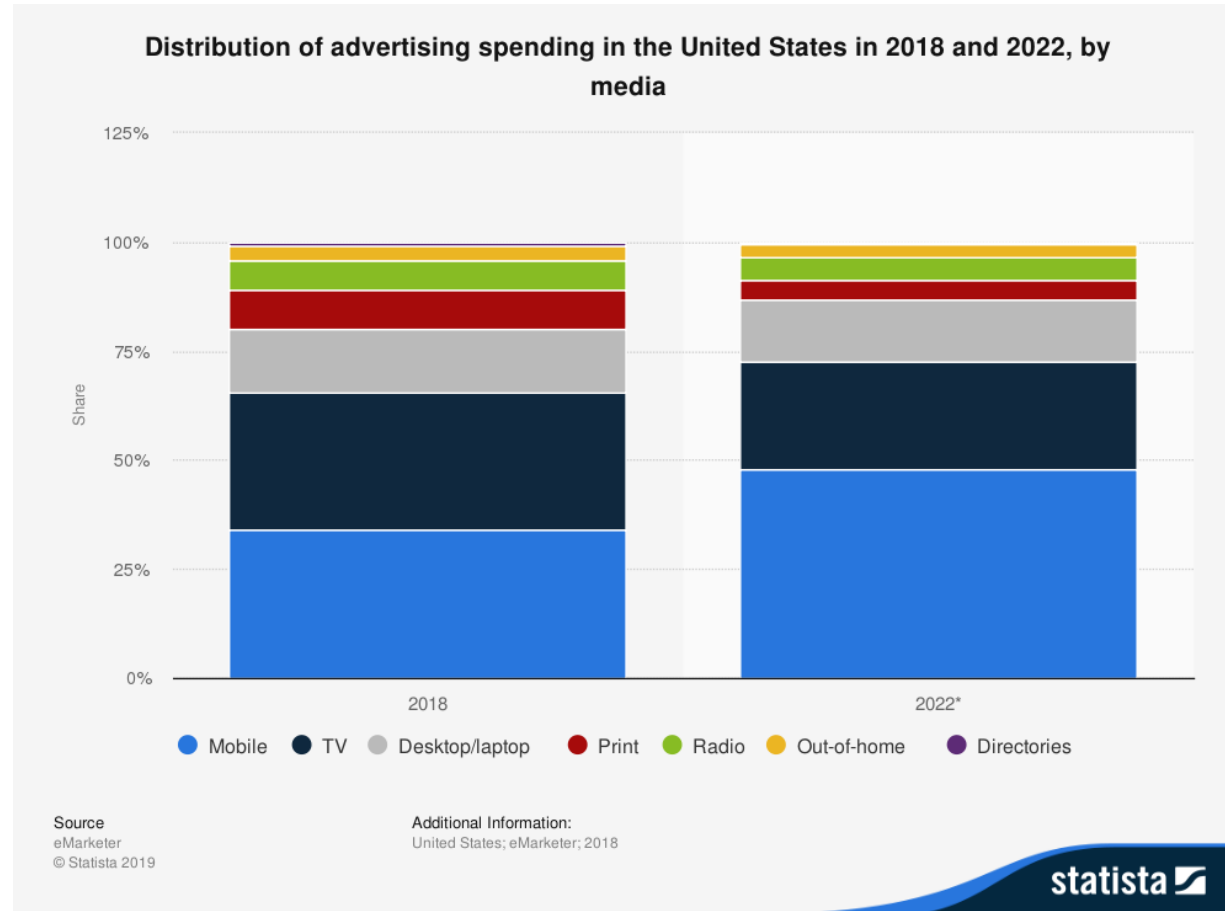
Produced by Scott Brinker (@chiefmartec) and Blue Green Brands (@bluegreenbrands).

Adtech

Display Advertising Technology Landscape



Media Mix



Accounting for Multiple Screens

United States

CLIENT L



INSIGHTS SOLUTIONS NEWS CENTER ABOUT

Search All Nielsen



NIELSEN DIGITAL AD RATINGS OVERVIEW FEATURES AND BENEFITS NEWS

NIELSEN DIGITAL AD RATINGS

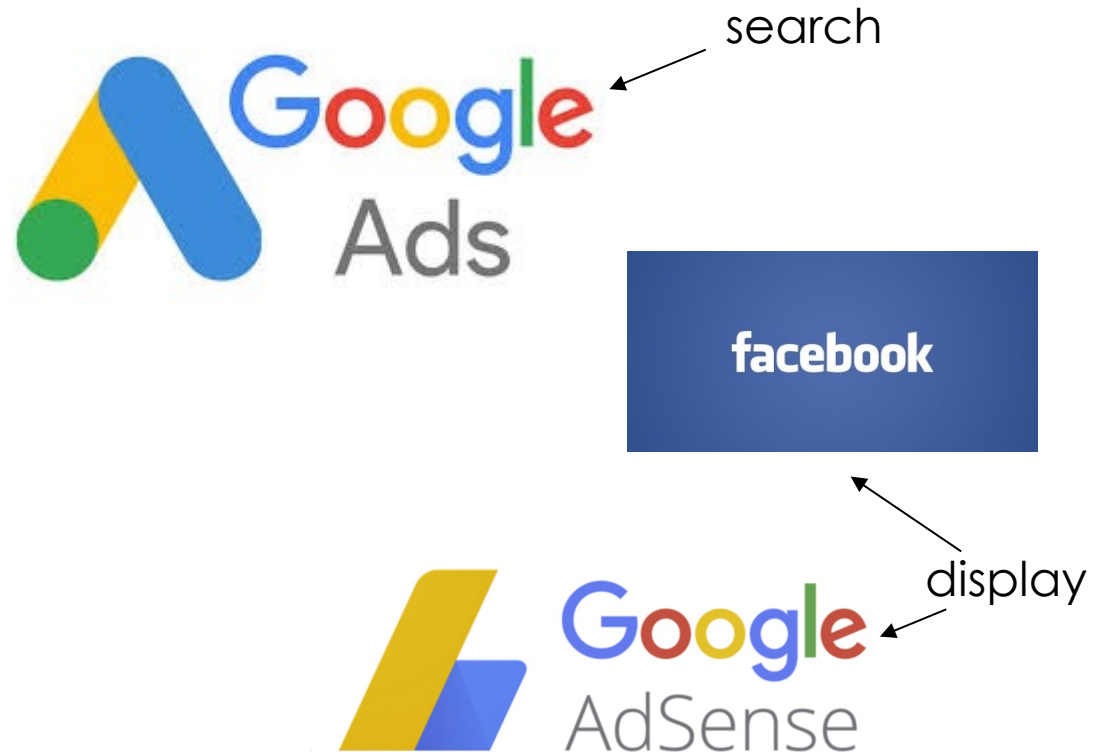
THE INDUSTRY STANDARD FOR DIGITAL AD MEASUREMENT



Source: <https://www.nielsen.com/solutions/audience-measurement/digital-ad-ratings/>

Digital Advertising: Overview

- Search vs. Display
- Target based on:
 - Search
 - Demographics
 - Content of the page
 - Context
 - Behavior of user
 - Any combination of the above
- Pay per
 - Impression (CPM)
 - Click (CPC)
 - Action (CPA)



Types of Online Advertising

Display

- Banner ads
 - Standardized ad shapes with images
 - Normally not related to content
 - Targeted based demographic, behaviour, etc
- Content linked ads
 - Related to content on page

Search

- Search linked ads
 - Related to search terms



[\\$150 Cash - Limited Time](#)
Open a BBVA Compass Checking Account Online. Apply Now.
[BBVACompass.com](#)

[Earn \\$150 Cash When You](#)
Open A BBVA Compass Bank Checking Account. Apply Online Now.
[BBVACompass.com](#)

Display Advertising: Content-based Targeting

18 Oct 2010 at 10:48 AM / NEWS

Columbia Reminds B-School Students To Brush Teeth, Remove “Tacky Cufflinks” Before Interfacing With Potential Employers

By BESS LEVIN

This is the second in a **two-part series** from CBS re: not “acting in a socially undesirable way runs a strong risk of branding you as undesirable.” This apparently includes wearing gold (instead of the more desirable silver) watches. Next up, the person sending these memos will pick one student at random, show up at his or her apartment, forcibly shampoo and condition their hair and pick out said student’s clothes prior to the next recruiting event.

To: 'ibc-members@bettyblue.gsb.columbia.edu'
Cc: ibc-board@bettyblue.gsb.columbia.edu
Subject: [IBC] Personal hygiene

Dear 1st Year Members,

It has come to our attention (through complaints from IBC board representing firms they are going to full time) that some of you may not have followed personal hygiene basics during recruiting events. We understand that it is an incredibly intense recruiting period, and is very hard to find time for yourself, but this is a friendly reminder on some dress code and personal hygiene basics:

- § Brush your teeth regularly, or have a mint/mouth refreshers before going to recruiting events (avoid chewing gums)
- § Carry anti-perspirant with you if you are worried about sweating. Don't wear too much cologne/perfume
- § Carry a sewing mini-toolkit, in case your suit hems need an emergency sewing
- § Professional haircuts
- § No backpacks with you
- § Men – no tacky cufflinks or watches (with no crazy patterns, silver is preferable to

38 tweets

retweet

193

Share

Twitter Facebook Email RSS

More subscription options...

Display Ads

<https://www.youtube.com/watch?v=iqhu2GWPCXk>



Search-based Advertising: Organic vs. Paid

The image is a screenshot of a Google search results page for the query "machine learning course". The search bar at the top shows the query and the Google logo. Below the search bar, there are tabs for "All", "Videos", "Books", "Images", "News", and "More". The results show "About 1,980,000,000 results (0.50 seconds)".

There are three main results visible:

- Ad · https://www.mygreatlearning.com/**
MIT IDSS - 12 Week Program - Machine Learning Program
Learn Data Science & **Machine Learning** skills from world-renowned MIT faculty. Learn **Machine Learning**, Deep **Learning**, NLP, Recommendation Systems, and more. Apply Now.
[Data Science & ML Program](#) · [What You'll Learn](#) · [Program Benefits](#) · [Program Faculty](#)
- Ad · https://online-exec.cvn.columbia.edu/course_machine/learning**
Columbia's 5-Month ML Program - Machine Learning Course
Acquire the applied **machine learning** knowledge needed to solve real business challenges. Enroll in a 5-month **machine learning** program developed by Columbia Engineering Exec Ed. 6-8 hours per week. Flexible Payment Option. **Course Fees:** USD 2,350.
- https://www.coursera.org > ... > Machine Learning**
Supervised Machine Learning: Regression and Classification
Build and train supervised **machine learning** models for prediction and binary classification tasks, including linear regression and logistic regression The ...
[Machine Learning](#) · [Machine Learning Specialization](#) · [Imperial College London](#)

Red arrows indicate the type of results:

- A red arrow points from the word "Paid" to the first two results (MIT IDSS and Columbia's 5-Month ML Program).
- A red arrow points from the word "Organic" to the third result (Coursera's Supervised Machine Learning).

Search Advertising

- Search Engine Marketing: bid on keywords so that your ad shows up on search engines
- Rank depends on
 - Cost per click (i.e., how much you bid)
 - Quality Score = $f(\text{expected click through rate})$ (Google)
- Search Engine Optimization: Fine tuning your website so that it shows up in organic search
 - Rank depends upon relevance / influence

Online Advertising Costs

- Cost per thousand impressions (CPM) = Ad rate web sites charge for displaying an ad 1000 times
 - Mille = Thousand in Latin
- Cost per click (CPC) = Ad rate charged when the surfer responds to a displayed ad
- Cost per lead / action (CPA) = Charged only if an ad viewer “converts” to a sale

Rocket Fuel Case

Measuring Effectiveness of Digital Advertising

- Advertise to ~590K users
- Conversion rate=2.54% of users (15K/590K)
- 14.5 M impressions (from case)
- CPM=\$9
- What is the Cost of Acquiring Customer (CAC)?

$$\text{CAC} = \frac{\text{How much did I pay?}}{\text{How many customers did I acquire?}} =$$

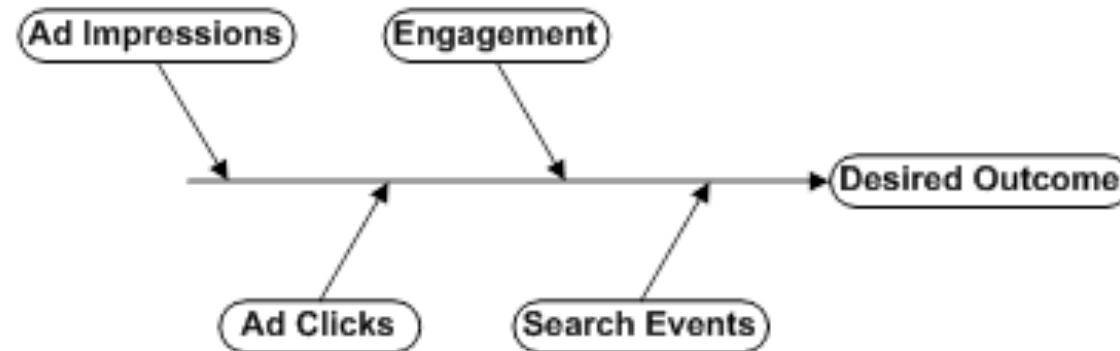
- Effectiveness (ROI): \$40/\$8.71 ~ 4.6
- Why is this WRONG?
 - We do not know the counterfactual. What is the baseline?

Attribution

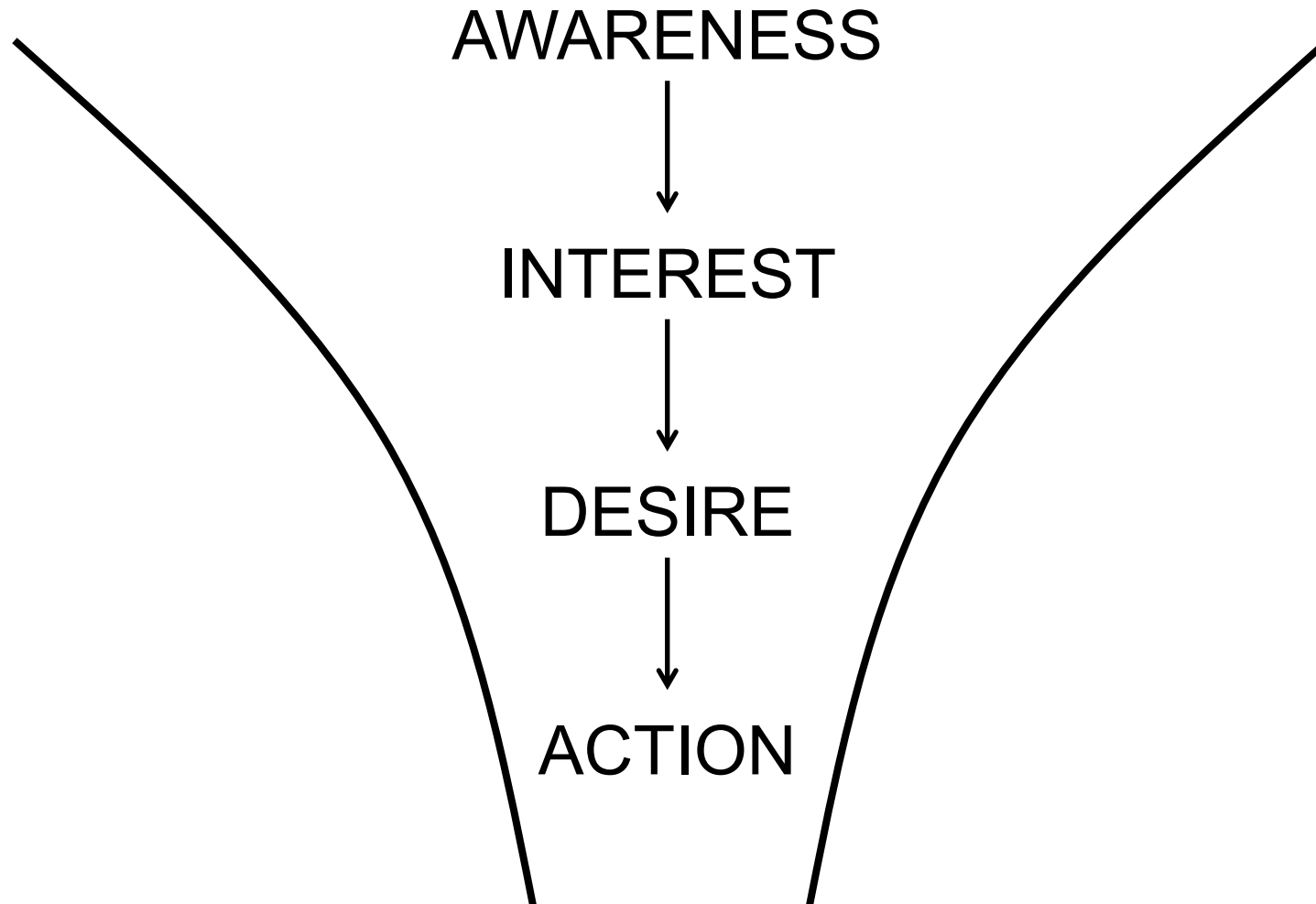
Attribution Problem

“Attribution is the process of identifying a set of user actions (“events”) across screens and touch points that contribute in some manner to a desired outcome, and then assigning value to each of these events.”

(source: <https://www.iab.com/guidelines/iab-attribution-hub/>)



Attribution and the Purchase Funnel



Attribution and Advertising

- Different communication goals require different ad vehicles
 - TV, sponsorship and display advertising for awareness, higher in the choice “funnel”
 - Search advertising for converting interested customers
- Need to map out the whole acquisition process to increase ad effectiveness
 - Identify areas with the lowest conversion rates and try to improve on them
 - Choose appropriate ad vehicles for this part of the funnel
- Attributing sales to a particular ad is a hard problem
 - Knowing that a person bought after clicking on an ad is not enough – need to find the actual incremental number of sales for each ad
 - This requires experimentation (A/B testing)

Causal Research

Types of Marketing Research

Exploratory Research

(Ambiguous Problem)

“Our sales are declining and we do not know why.”

Descriptive Research

(Aware of Problem)

“What kinds of people are buying our products?”

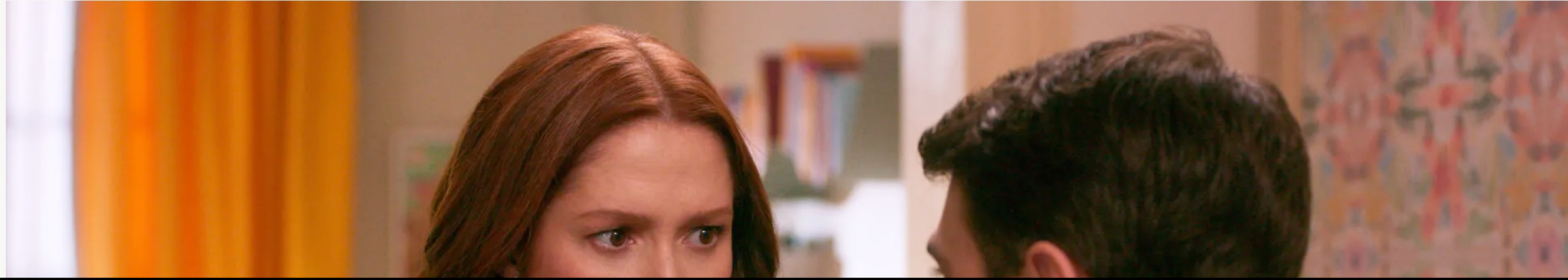
“Who buys our competitors’ products?”

Causal Research

(Problem Clearly Defined)

“Will buyers purchase more of our product in a new package?”

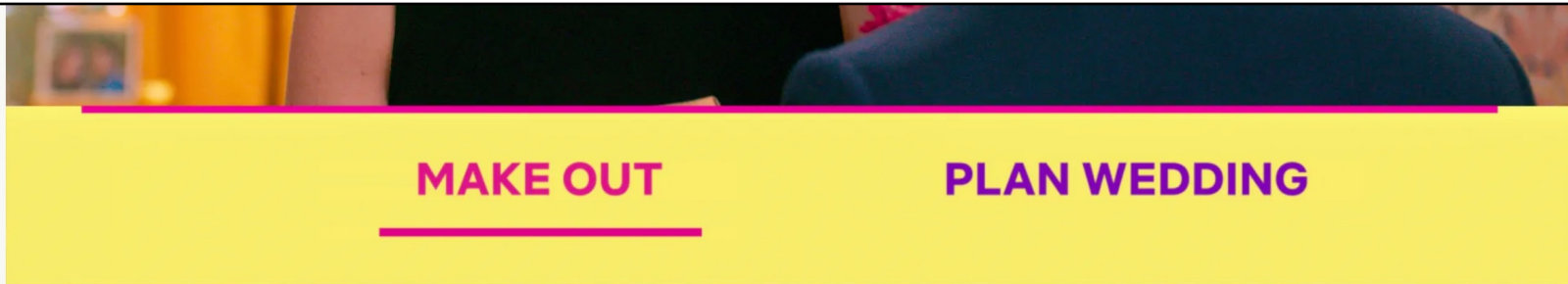
Thinking Counterfactually



Causal inference is all about understanding “alternate universes”

To know: “If we do X, what happens?”

We also need to know: “If we don’t do X, what happens?”



Correlation and Causation

- Correlation = relationship between two variables
“increased sales are correlated with our promotions”
- Causation = one variable producing an effect in another variable
“the promotion caused an increase in sales”

Correlation \neq Causation

What's needed to say X causes Y?

1. Correlation: evidence of association between X and Y
2. Temporal antecedence: X must occur before Y
3. No third factor driving both X and Y: control of other causal factors

This is usually the tough one!

Important: These are necessary conditions for saying X causes Y, not sufficient conditions

Causal Research

A/B Tests

The Rule of “-tions”

No causation without (quasi-)experimentation.

We can guarantee that an observed effect is a causal effect by running an experiment where ...

No experimentation without manipulation.

... subjects are assigned to different conditions in which causal variables are systematically and differentially manipulated.

And randomization is necessary

What is an A/B Test?

Control group (A)



Outcome: 100

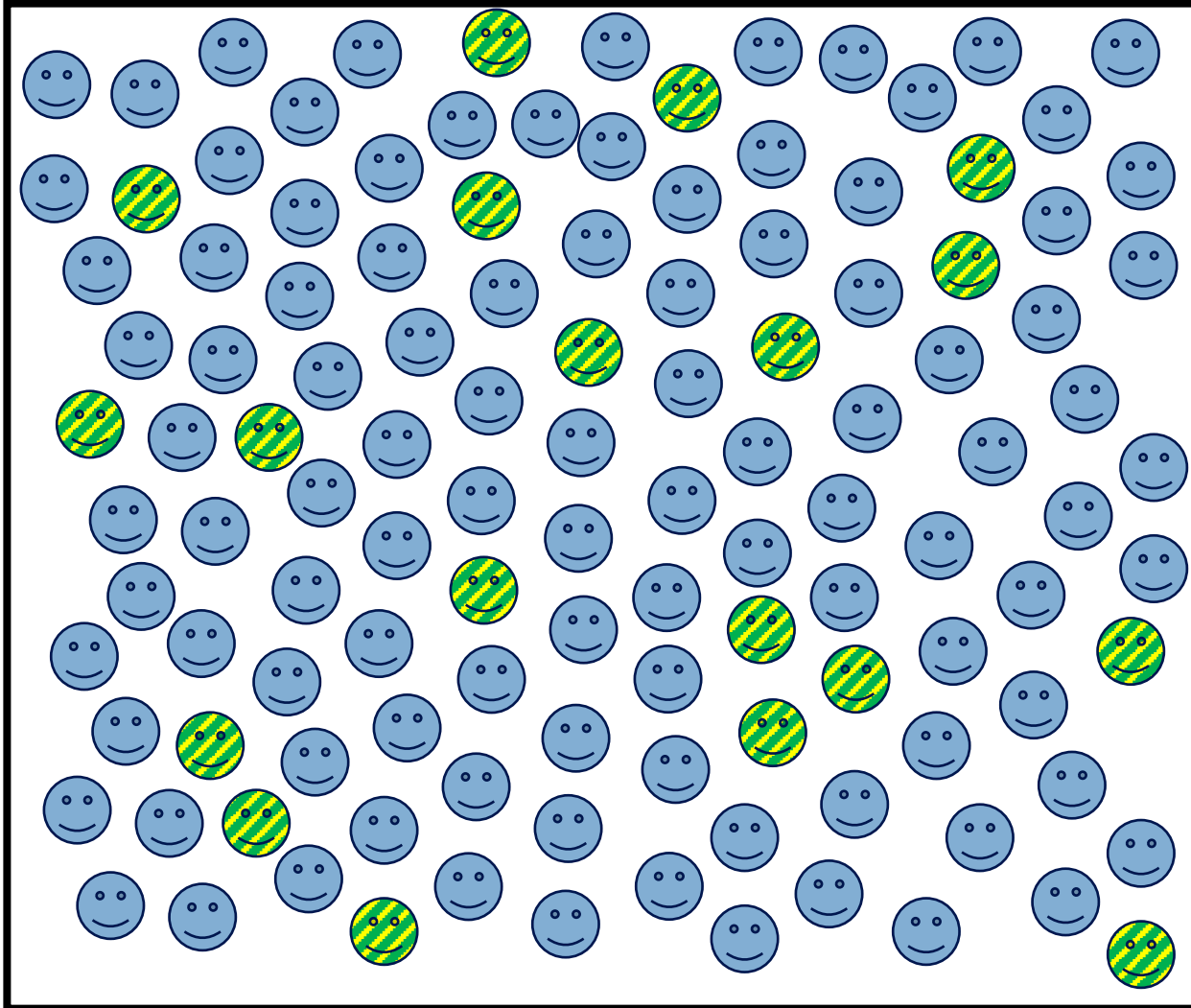
Treatment group (B)



Outcome: 150

- Important considerations:
 1. Randomization: Assignment into groups is random
 2. Manipulation: One variant at a time versus many variants at once
 3. Statistical significance of the difference in outcomes
- If 1-3 hold, we can attribute the increase in the outcome to the ad

Randomization



Idea: Randomly sample the population

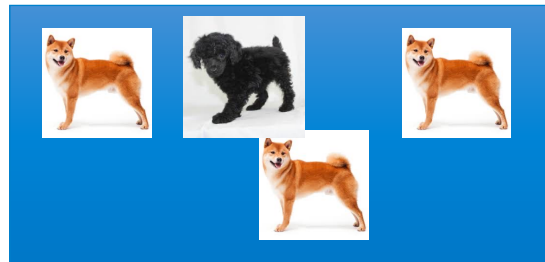
- Every person in the population has an equal chance of being in the treatment group

Randomization: Perils and Checks

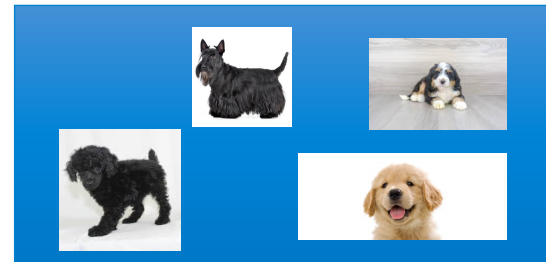
Even when properly randomized, systematic differences may still exist!

- Always check for differences between conditions (**randomization checks**)!
 - Statistical tests (t-tests, linear regression)

Control group (A)



Treatment group (B)



- Can also identify badly done experiments

Necessary Sample Size Calculation – Example

A marketing manager for a start-up wishes to discover the **proportion** of teenagers in US who own an iPhone. If the manager wants a 95% confidence interval with acceptable sampling error of 0.05, how many teenagers must be sampled?

- Objective: Find the **number** of teenagers that must be sampled to **effectively** estimate the **proportion** of iPhone owners

Necessary Sample Size Calculation – 1 Proportion

- The 95% confidence interval for the population proportion **p** is

$$\hat{p} \pm 1.96 \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

- To estimate **p** to within $\pm E$ set $1.96 \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = E$ where E is the acceptable amount of sampling error
- Necessary sample size is $n = \frac{1.96^2 \hat{p}(1 - \hat{p})}{E^2}$
 - Must specify E and the desired confidence level
 - Conservative estimate of $p(1-p)$ = largest possible value of $p(1-p)$, which occurs when $p = 0.5$
 - Can also use an educated guess for \hat{p}

Necessary Sample Size Calculation – 1 Proportion

A marketing manager for a start-up wishes to discover the proportion of teenagers in US who own an iPhone. If the manager wants a 95% confidence interval with acceptable sampling error of 0.05, how many teenagers must be sampled?

- $E = 0.05$
- $n = \frac{1.96^2 \hat{p}(1-\hat{p})}{E^2} =$
- Conservative necessary sample size = 385
- If you have a prior belief about \hat{p} , you can use a smaller sample size

Necessary Sample Size – Difference Between Groups

- **Power:** probability of correctly identifying a difference between two groups in the study sample when one genuinely exists in the populations from which the samples were drawn
 - Generally desired to be 80%-95%

- If we have 2 equally sized groups of n individuals:

$$n = \frac{p_1(1-p_1)+p_2(1-p_2)}{(p_1-p_2)^2} c_{p,power}$$

- If groups unequal, adjust above:

$$N' = \frac{N(1+k)^2}{4k}$$

- N = total sample size ($2n$), k = ratio of two groups
- Sample sizes of two groups are $\frac{N'}{1+k}$ and $\frac{kN'}{1+k}$

$c_{p,power}$

P	Power (%)			
	50	80	90	95
0.05	3.8	7.9	10.5	13.0
0.01	6.6	11.7	14.9	17.8

P-value: how likely that the observed effect in the sample is due to chance
The smaller the p-value, the larger the sample size needed

Source: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC137461/>

Break

Causal Research

Analysis

Difference in Proportions

- Confidence interval for a difference in proportions:

$$\hat{p}_1 - \hat{p}_2 \pm 1.96 \times \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$$

- Statistical significance measures the “strength of statistical evidence” in support of some claim
 - Null hypothesis: $p_1 - p_2 = 0$
- The p-value is a measure of statistical significance
 - The p-value is the smallest value of α such that the confidence interval does not include 0 (or another hypothesized value)
 - The smaller the p-value, the stronger the evidence that our estimate is different from the hypothesized value

Your Turn

Calculate the difference in conversion rates for Rocket Fuel

A/B Testing: Users Randomly Assigned to Control vs. Test

	Test	Control
Number of users who did not convert	550154	23104
Number of users who converted	14423	420
Sample size	564577	23524
Conversion rate	2.55%	1.79%

- Does display advertising have a statistically significant positive effect on conversion rate?
 - i.e., is the conversion rate in the test group significantly larger than the conversion rate in the control group?

Do you agree with the following statement: "The campaign had a statistically significant positive effect on conversion rate"?



0

Yes

0

No

0

A/B Testing: Users Randomly Assigned to Control vs. Test

	Test	Control
Number of users who did not convert	550154	23104
Number of users who converted	14423	420
Sample size	564577	23524
Conversion rate	2.55%	1.79%

- Does display advertising have a statistically significant positive effect on conversion rate?
 - i.e., is the conversion rate in the test group significantly larger than the conversion rate in the control group?
- Was the campaign profitable?
- What was the ROI of the campaign?

Quantifying the Effectiveness of Promotions

- Promotion Effectiveness = Incremental Revenue – Incremental Cost
- Incremental Revenue = Revenue with Promotion – Revenue without Promotion
- Incremental Cost = Cost with Promotion – Cost without Promotion
- ROI = Promotion Effectiveness / Incremental Cost

Campaign Effectiveness and ROI

- Difference in conversion rates between test and control: 0.77%
- Number of users in test group: 564,577
- Value of converted user (from case): \$40
- CPM (from case): \$9
- Incremental revenue from campaign: ?
- Total number of impressions: 14,597,182
- Cost of impressions: ?
- Effectiveness of campaign (incremental revenue-incremental cost): ?
- ROI of campaign (effectiveness/incremental cost): ?

What was the ROI of the campaign?



<0

0

Between 0 and 15%

0

Between 15 and 30%

0

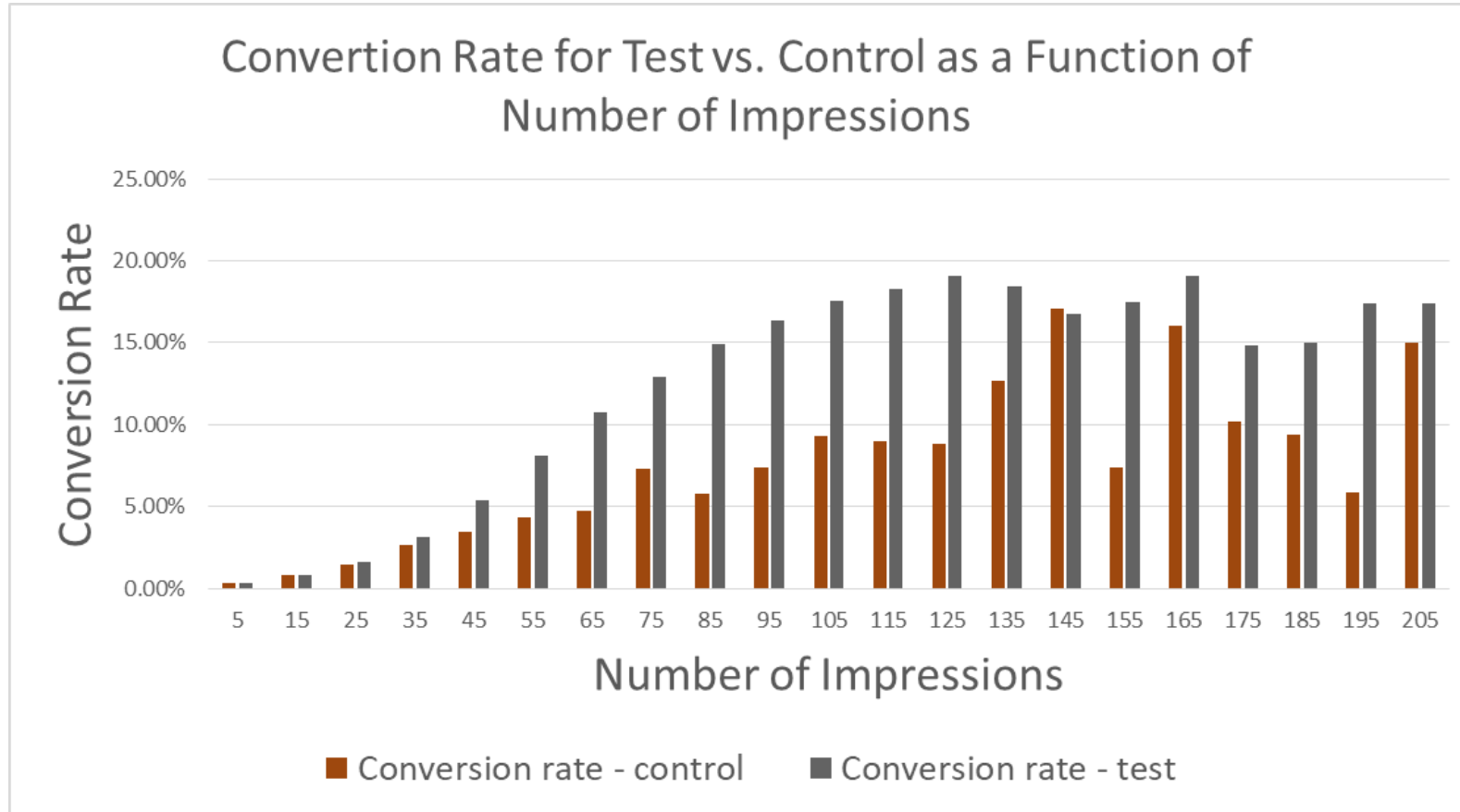
Higher than 30%

0

Attribution and Experimentation – Rocket Fuel Case

- A/B testing
 - Establishes a causal link between the display advertising campaign and the increased revenue in the test group
 - Assumptions
 - No spillover from test to control users
 - Only difference between two groups = test was exposed to display advertising
- However,
 - Effect of the display advertising campaign might still have been influenced by **other** campaigns the firm was conducting at the same time
 - E.g., there may have been a search advertising campaign going on, which increased the effectiveness of the display campaign
 - The ROI of display advertising might change if the firm changed other aspects of its marketing mix

Number of Impressions vs. Conversion



Do you agree with the following conclusion: "The number of exposures has a positive impact on conversion"?



0

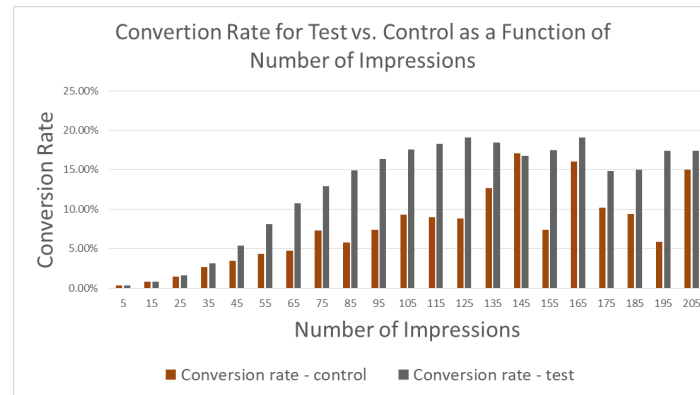
Yes

0

No

0

Number of Impressions vs. Conversion



- The number of impressions was NOT varied experimentally
- In each condition:
 - Because of targeting, customers who are more likely to convert may be exposed to more ads
 - A correlation between number of impressions and conversion rate does NOT mean that more impressions cause higher conversion rate

Causal Research

Analysis using Regression

Average Treatment Effect

- Average Treatment Effect: $I(purchase_i) = \alpha + \beta I(treatment_i) + \epsilon_i$

$I(treatment_i)$	
0 = control	α
1 = treatment	$\alpha + \beta$

Effect of treatment = β

Is β statistically significant at the 0.05-level?

Heterogeneous Treatment Effects

- Average Treatment Effect: $I(purchase_i) = \alpha + \beta I(treatment_i) + \epsilon_i$
- Heterogeneous Treatment Effects:
 - By student status (binary variable):

$$I(purchase_i) = \alpha + \beta I(treatment_i) + \gamma I(student_i) + \delta I(treatment_i) \times I(student_i) + \epsilon_i$$

$I(treatment_i)$	$I(student_i)$	
0 = control	0 = not student	α
0 = control	1 = student	$\alpha + \gamma$
1 = treatment	0 = not student	$\alpha + \beta$
1 = treatment	1 = student	$\alpha + \beta + \gamma + \delta$

Heterogeneous Treatment Effects

- Average Treatment Effect: $I(purchase_i) = \alpha + \beta I(treatment_i) + \epsilon_i$

- Heterogeneous Treatment Effects:

- By student status (binary variable):

$$I(purchase_i) = \alpha + \beta I(treatment_i) + \gamma I(student_i) + \delta I(treatment_i) \times I(student_i) + \epsilon_i$$

- By income (continuous):

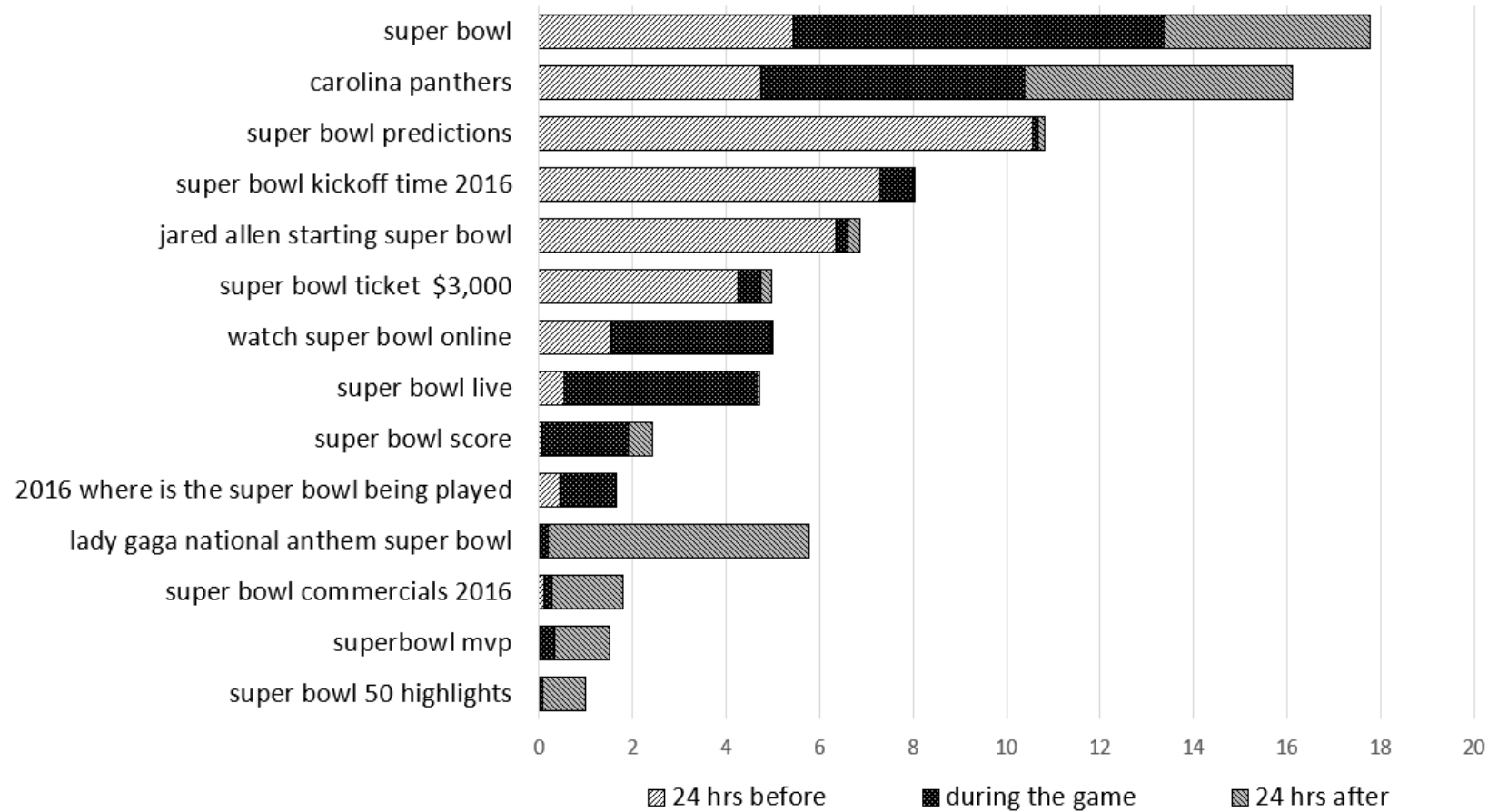
$$I(purchase_i) = \alpha + \beta I(treatment_i) + \gamma income_i + \delta I(treatment_i) \times income_i + \epsilon_i$$

Novel Approaches

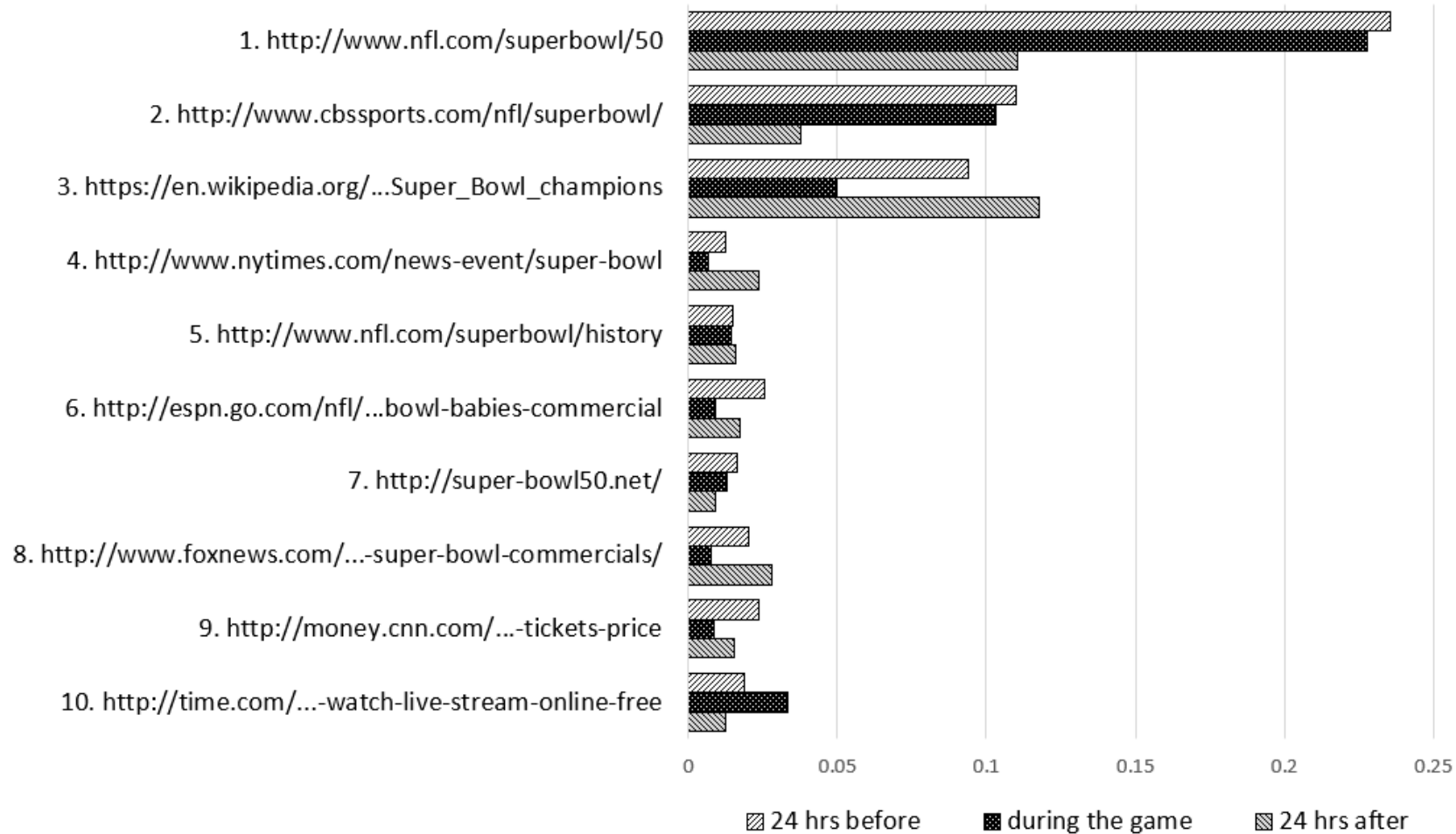
Content-Based Search

- How can we infer what people are really searching for, based on their search queries?
- Applications: Search Engine Optimization, Search Engine Marketing
 - Show the most relevant content to the right consumer at the right time
- Collaboration with Microsoft Research
 - Bing searches related to 14 popular TV shows: content of search queries, volume of search, content of top search results, click-through rates
- Develop a model that extracts relevant topics, and estimates consumers' preferences over topics based on their queries and context (e.g., before, during, after a show is aired)
- Liu, Toubia, Hill (2020), "Content-Based Model of Web Search Behavior: An Application to TV Show Search", *Management Science*

Most Popular Search Queries for Super Bowl 2016



Click-Through Rate of Top Search Results for Query “super bowl”



Topics for Super Bowl



Topic 1: Watch Online



Topic 2: Prediction



Topic 3: Falcons



Topic 4: Winners



Topic 5: Rivalry



Topic 6: Commercial



Topic 7: Patriots



Topic 8: Halftime Show

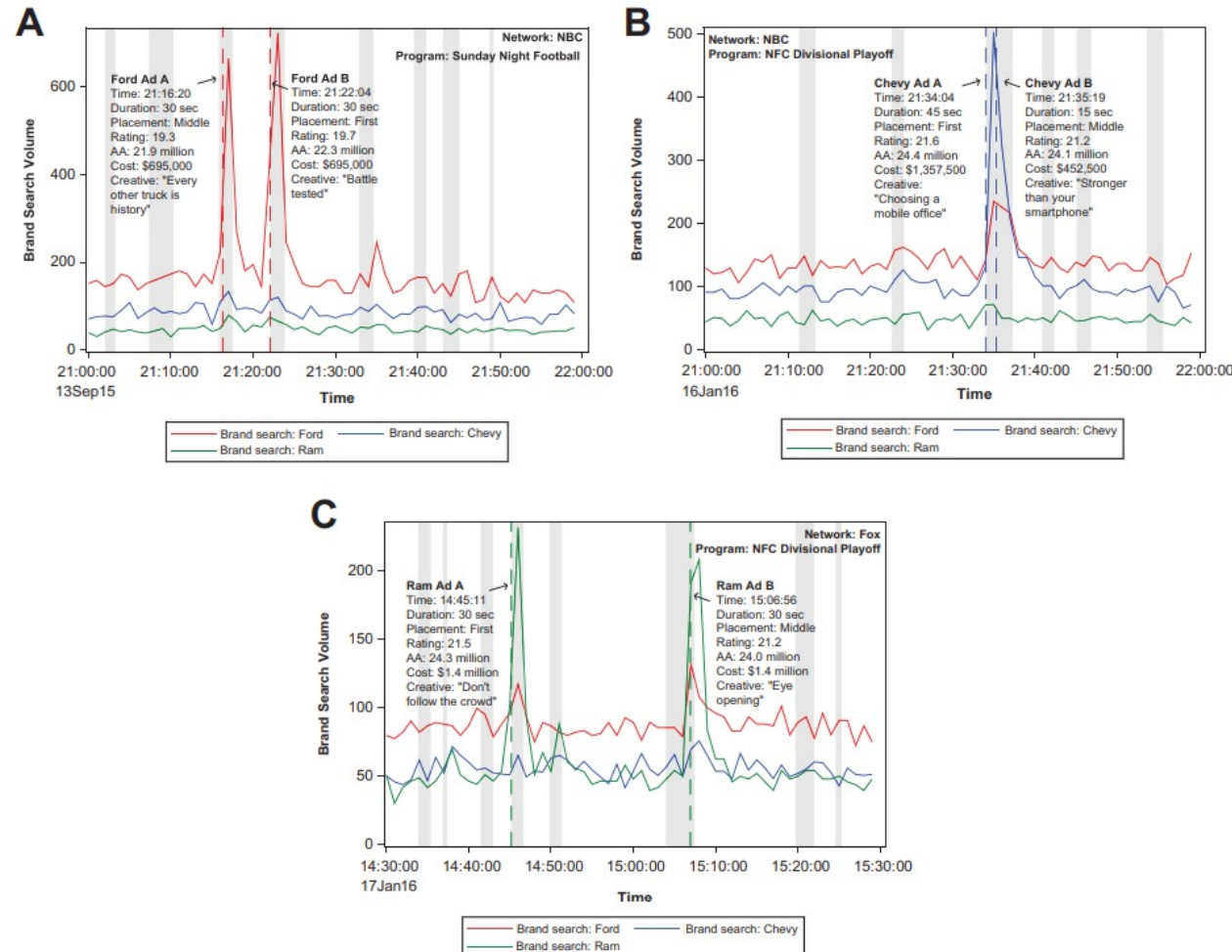
Preferences Over Topics Corresponding to Query “super bowl”

Topic	before	during	after
1. Watch Online	1.6623	6.8292	0.0247
2. Prediction	1.5412	0.6810	0.0202
3. Falcons	0.2522	1.0020	1.6623
4. Winners	6.8292	0.0247	0.2325
5. Rivalry	0.6810	0.0202	0.2522
6. Commercial	1.0020	1.6624	6.8292
7. Patriots	0.0247	0.2325	0.6810
8. Halftime Show	2.1343	1.0788	1.6448

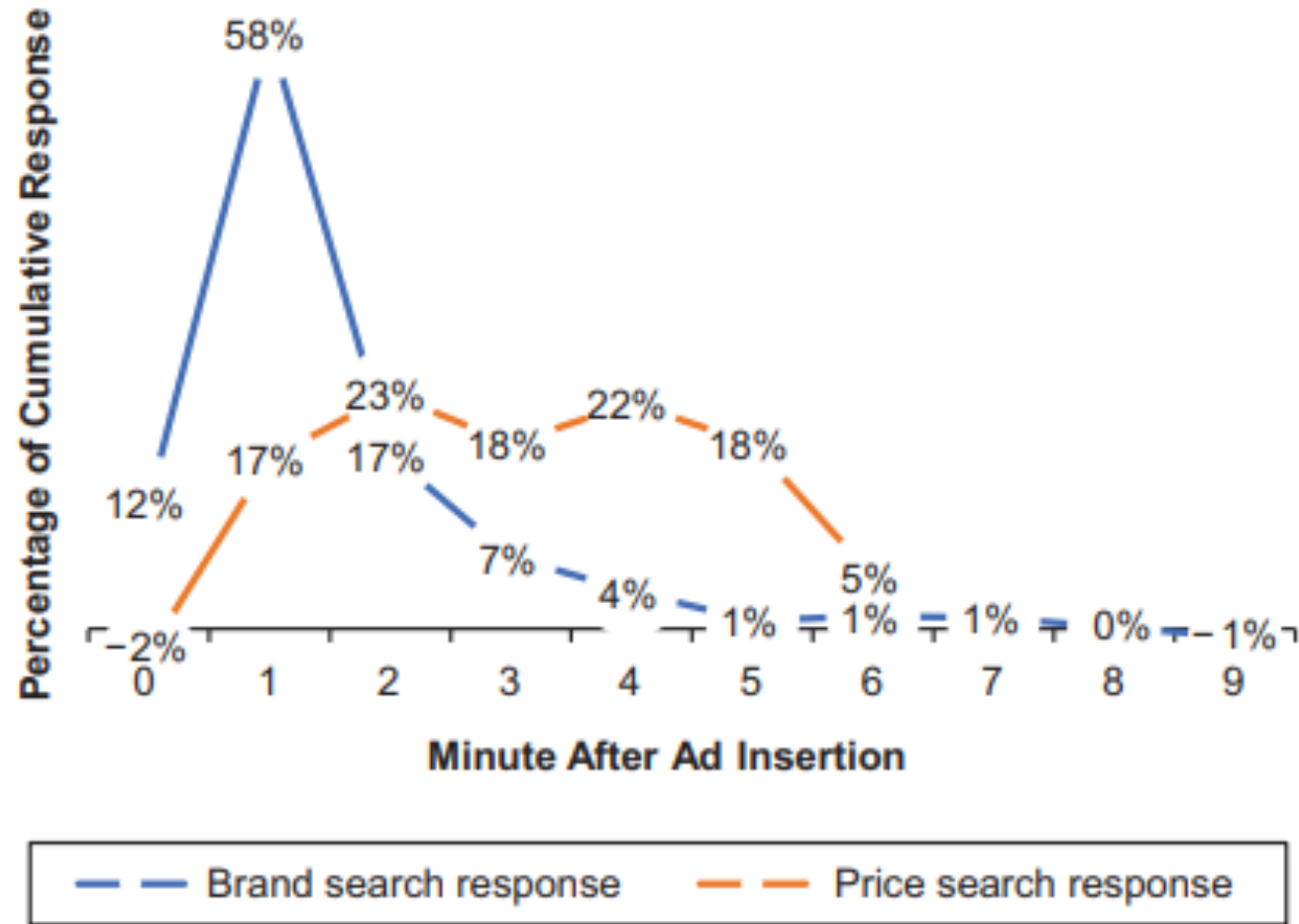
Immediate Response to TV Ads

- Traditionally, response to TV ads is not observed until much later
- Nowadays, people can react to an ad by searching on their phones
 - Allows us to measure effectiveness “on the spot”
- Du, Xu and Wilbur (2019), Immediate Responses of Online Brand Search and Price Search to TV Ads”, *Journal of Marketing*

National TV Ads and Post-ad Brand Search Spikes



Brand vs Price Search



Impact on Competitors

Minute After Ad Insertion	Brand Search Response Per One Million Impression-Minutes						Price Search Response Per One Million Impression-Minutes					
	Chevy ↓ Ford	Ram ↓ Ford	Ford ↓ Chevy	Ram ↓ Chevy	Ford ↓ Ram	Chevy ↓ Ram	Chevy ↓ Ford	Ram ↓ Ford	Ford ↓ Chevy	Ram ↓ Chevy	Ford ↓ Ram	Chevy ↓ Ram
0	.40	.81*	.60*	.07	.64*	.53	-.15	.19	.05	.42*	.21*	-.21
1	4.85*	2.31*	2.59*	.33	2.15*	1.54*	.01	.02	.14	.18	.06	-.04
2	2.18*	.54*	.69*	.10	.94*	.63*	-.41*	.15	-.05	.60*	.22*	-.07
3	.42	.87*	-.01	.09	.50*	.15	.21	-.18	-.13	-.12	.20*	-.23*
4	.83*	-.32	-.04	.37	-.01	.70*	-.56	-.15	.21	-.53*	.28*	.55*
5	-.48	.13	.42*	.13	-.04	.29						
Total	8.20*	4.35*	4.25*	1.09*	4.18*	3.84*	-.91	.03	.23	.55*	.98*	-.01
Avg. elasticity	.01	.01	.04	.003	.05	.01	-.01	.0003	.01	.01	.12	-.0004

* $p < .01$.

Takeaways

- Attribution is a key problem
 - Purchase funnel
- A/B testing is key to understanding ad effectiveness
 - Beware of the distinction between causality and correlation
 - You can only assess causality with randomization and experimentation
- Measures of effectiveness and ROI

Break

Promotion

Sales Promotions

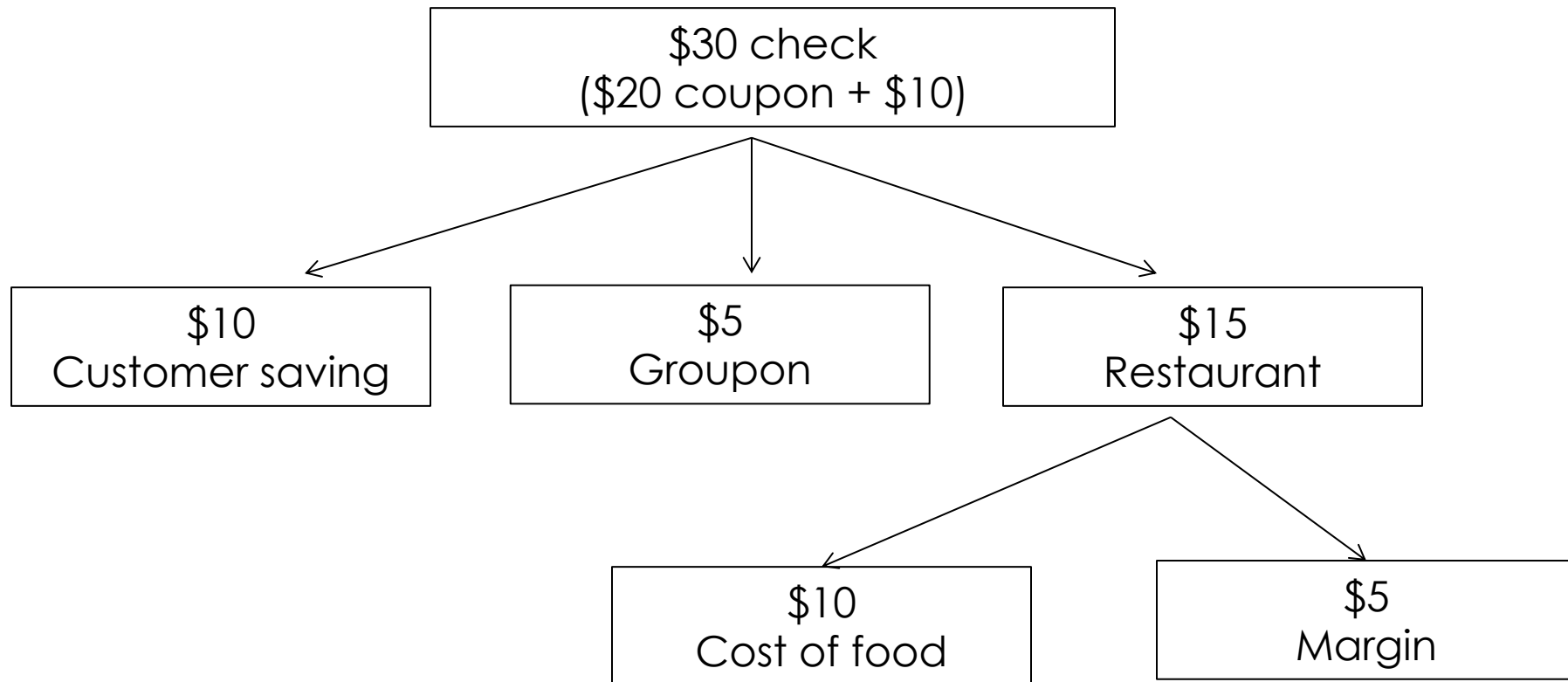
- Short term inducements designed to have a direct impact on the buying behavior of end-users and/or trade



Chirpin' Tavern Case

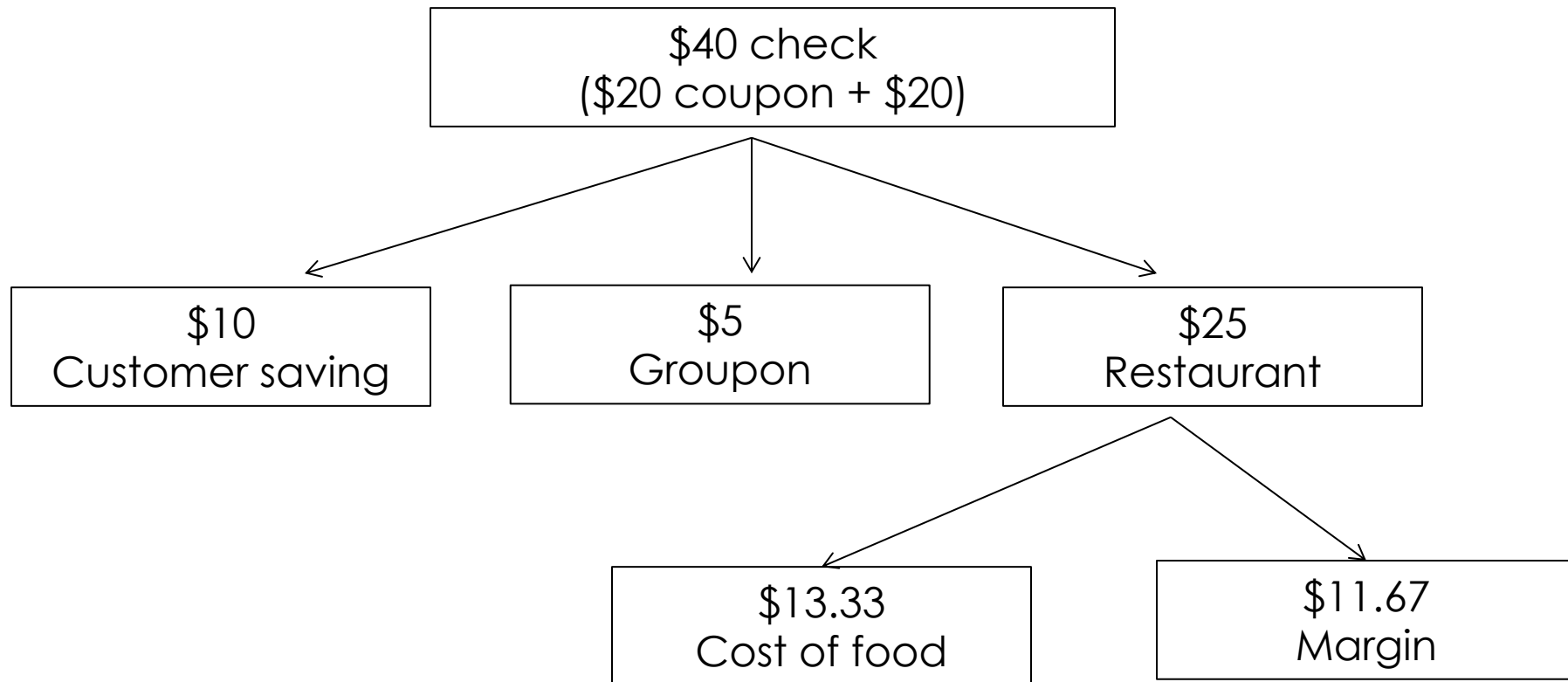
“Evaluating Price Promotions”

The Economics of a Redeemed Groupon



“Evaluating Price Promotions”

The Economics of a Redeemed Groupon



Not so simple because...

Sales Promotions: Some Considerations

- Exposure effect
 - Customer may be positively impacted by exposure to offer even if does not buy coupon (e.g., advertising)
- Opportunity cost
 - Does selling a unit with the promotion prevent from selling a unit without the promotion (e.g., restaurant sold out)?
- Is the sale incremental?
 - Are we giving a better deal to a customer who would have purchased anyway?
- What is the long-term impact of the promotion?
 - Trial vs. repeat?
 - Loyalty?
 - Change reference price / expectations?
 - Customers wait for next promotion?
 - Stockpiling (for Consumer Packaged Goods)?
 - ...

Chirpin' Tavern Data

- We will compare the sales in the months with promotion to the sales in the months without promotion
- We do not have an experiment
- Observe seasonality so need to include seasonality in the model
 - Winter dummy and summer dummy
- Observe a positive trend in restaurant sales
 - Time trend

Quantifying the Effectiveness of Sales Promotions

- Promotion Effectiveness = Incremental Revenue – Incremental Cost
- Incremental Revenue = Revenue with Promotion – Revenue without Promotion
 - “baseline” revenue
 - Revenue should be the revenue in the entire period in which the promotion might have an effect, not just the duration of the promotion
- Incremental Cost = Cost with Promotion – Cost without Promotion

Approach

1. Model (Regress) Revenue as a function of...
 - See “Chirpin’ Tavern inclass.xlsx”
2. Calculate incremental revenue from each promotion
 - Incremental sales + revenue from non-redeemed coupons
3. Calculate incremental cost of each promotion
 - Cost of food + \$15 per redeemed coupon (check value vs. what restaurant actually gets)
4. Calculate promotion effectiveness (incremental revenue-incremental cost)

What is the effectiveness of a Groupon promotion for the restaurant (long-term incremental revenue – incremental cost)?



>+\$4,000

0

Between +\$2,000 and +\$4,000

0

Between +\$0 and +\$2,000

0

Between -\$2,000 and +\$0

0

Between -\$4,000 and -\$2,000

0

<-\$4,000

0

Promotion as Advertising

- Sahni, Zou and Chintagunta (2016) “Targeted discount offers serve as advertising? Evidence from 70 field experiments,” *Management Science*
- 70 randomized experiments on large online ticket resale platform
- Offers sent to members by email
- Finding: 90% of gains from promotional offers are NOT through redemption
 - Carryover effects (over time): people spend more the week after the promotion expires
 - Spillover effects (across categories): people buy more non-promoted products (i.e., tickets not covered by the promotion)
- Promotions serve as a form of “advertising”

Takeaways

- Price promotions can have various short-term and long-term effects on sales (positive or negative)
- Measuring the effectiveness of price promotions requires counterfactual predictions:
 - What would sales be with vs. without the promotion?
- Regression analysis based on historical data allows making these predictions
 - Isolate the impact of each promotion on current and future sales
- Analysis is based on very basic data, available to any business!

Pricing Concept Check

Time permitting

Pricing Concept Check

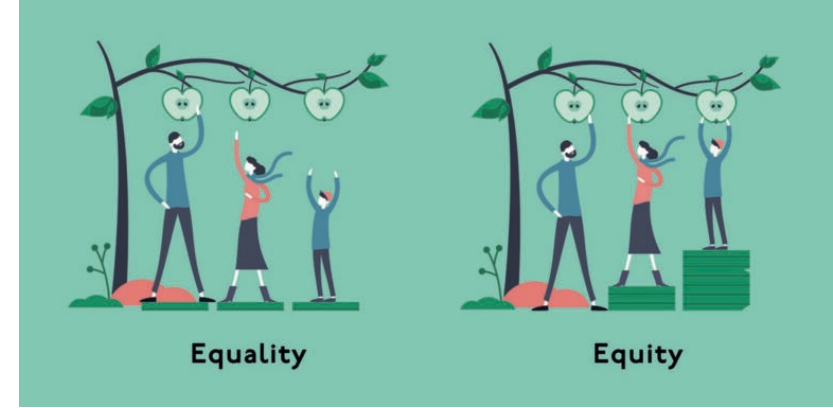
- You should compute willingness to pay (WTP) based on the conjoint partworths, assuming a price increase from a baseline of \$500 (i.e., the “exchange rate” should be based on the partworth for \$600 vs. \$500), and assuming that the baseline for the keyboard is Touch (i.e., the utility for the virtual keyboard should be based on the partworth for Virtual vs. Touch).
- Suppose Toshiba decides to sell a virtual keyboard as a separate accessory compatible with most tablet PCs. In all questions below, assume a marginal cost of \$50.
- Assume that WTP for the virtual keyboard is the same when it is sold as an accessory vs. built in. Based on the results of the conjoint analysis survey, which of the following prices for the virtual keyboard as an accessory would give rise to the highest contribution for Toshiba?

Pricing Concept Check

- For each consumer, determine USD to utils exchange rate
 - Consumer 1: $\$100 = 0.67 \text{ utils}$
 - Consumer 3: $\$100 = .33 \text{ utils}$
- Determine WTP for virtual keyboard:
 - Consumer 1: $0.17 \text{ utils} * \$100 / 0.67 \text{ utils} = \25
 - Consumer 3: $0 \text{ utils} * \$100 / 0.33 \text{ utils} = \0
- Calculate demand at \$125, \$150, \$175, \$200
- Calculate profit and select price with highest profit

Next Class

- Tuesday: Generative AI & Explainable AI
- Wednesday: No class
- Due:
 - Second individual assignment (4P Analytics) due Sat, Nov 23 at 8PM
 - Third individual assignment (Artea) due Mon, Dec 2 at 8PM



B9651 – Marketing Analytics

Session 12: Artefa & Fairness

Professor Hortense Fong

Today: Artea, Fairness, & Review

Part 1: Artea Case

Part 2: Fairness

1. Definitions
2. Calculate Measures of Fairness
3. Ways to Mitigate Bias

Part 3: Review

1. Review
2. Course Evaluations
3. Exam
4. Words of Advice

Today's Goals

Understand:

- A/B test results
- Definitions of fairness

Be able to:

- Consider fairness in decision making

Artea

Artea: Designing Targeting Strategy

- Goal: Increase Artea's sales by using data from website visitors
- Run A/B test to see if sending 20% off coupons increases sales
- Examine the data from the A/B test:
 - What are your insights from the test? Did revenues increase? Did transactions?
 - Which of the new users should receive a coupon?
 - Would Artea be able to increase transactions / revenues with this campaign? By how much?

Successful Randomization?

- Logistic regression of treatment (coupon) assignment on user features (channel, # past purchases, etc.)

	coef	std err	t	P> t	[0.025	0.975]
-----	-----	-----	-----	-----	-----	-----
const	0.4829	0.023	21.025	0.000	0.438	0.528
num_past_purch	0.0019	0.003	0.620	0.535	-0.004	0.008
spent_last_purchase	8.77e-05	0.000	0.624	0.533	-0.000	0.000
weeks_since_visit	0.0037	0.003	1.192	0.233	-0.002	0.010
browsing_minutes	-0.0002	0.001	-0.217	0.828	-0.002	0.002
shopping_cart	-0.0183	0.016	-1.163	0.245	-0.049	0.013
channel_2	0.0189	0.019	0.994	0.320	-0.018	0.056
channel_3	-0.0016	0.017	-0.096	0.923	-0.035	0.032
channel_4	0.0297	0.034	0.865	0.387	-0.038	0.097
channel_5	0.0054	0.048	0.113	0.910	-0.088	0.099

- Low coefficients and not statistically significant → Successful random assignment

A/B Tests: Causality vs. Correlation

```
## =====
##                               Dependent variable:
##                               -----
##                               trans_after
##                               (1)                (2)
## -----
## test_coupon1                0.026** (0.012)      0.026** (0.011)
## channel_acq Facebook                0.113*** (0.014)
## channel_acq Instagram              0.108*** (0.013)
## channel_acq Referral               0.130*** (0.026)
## channel_acq Other                 0.159*** (0.036)
## num_past_purch                   0.057*** (0.002)
## spent_last_purchase              -0.0004*** (0.0001)
## weeks_since_visit                -0.020*** (0.002)
## browsing_minutes                 0.003*** (0.001)
## shopping_cart                   0.173*** (0.012)
## Constant                       0.126*** (0.008)   -0.065*** (0.018)
## -----
## Observations                   5,000              5,000
## R2                             0.001              0.177
## Adjusted R2                    0.001              0.176
## Residual Std. Error    0.419 (df = 4998)    0.381 (df = 4989)
## F Statistic            4.872** (df = 1; 4998) 107.660*** (df = 10; 4989)
## =====
## Note:                          *p<0.1; **p<0.05; ***p<0.01
```

- All coefficients are statistically different than zero
- Getting a coupon has a positive association of 0.026 with # of transactions
- Having an item in the cart has a positive association of 0.173 with # of transactions
- Does this mean that if we send a coupon, we increase transactions?
- Does this mean that if we put something in shoppers' carts or force them to abandon cart, we increase transactions?

Learnings from the A/B Test

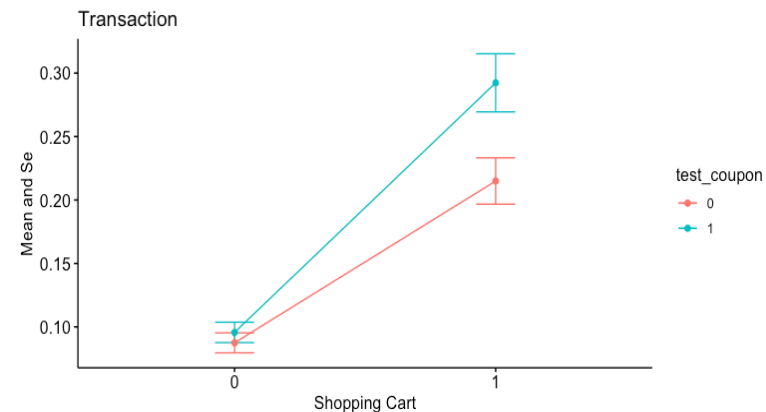
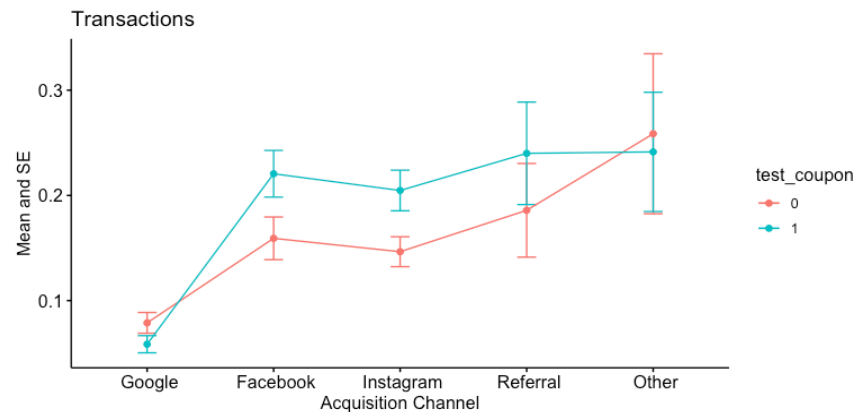
- The coupon causes a 0.026 pp increase in transaction likelihood
- The coupon does not have a significant effect on revenues
 - Negative coefficient but not statistically significant

Using the results of an A/B test we can:

- Implement a global change (one size fits all)
 - Less common for coupons, more common for design element such as the look and feel of a website, the steps to checkout, etc.
- Understand heterogeneity in response to the intervention (personalization)
 - By comparing different sub-groups of the population across treatment and control we can learn which groups are most affected by the intervention

Learnings from the A/B Test

- Some meaningful interactions:
 - Channels of acquisition (coupon increased revenue and transactions among those acquired through instagram)
 - Shopping cart status (coupon increased revenue and transactions among those with product in shopping cart)

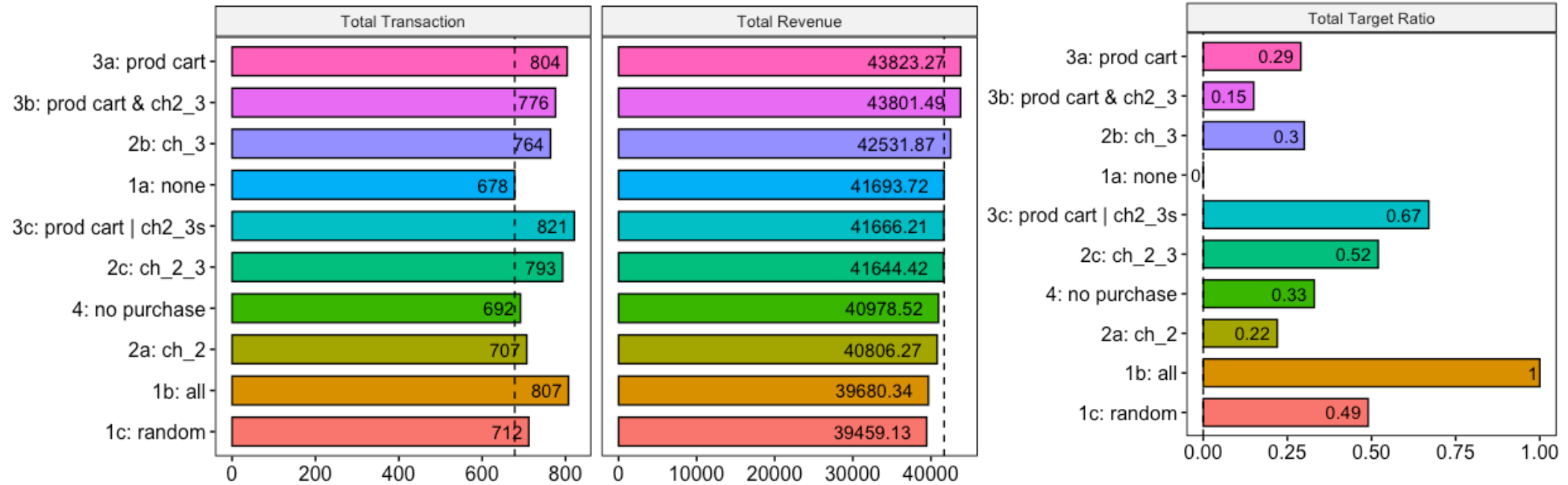


Given the insights, what is your targeting strategy recommendation?

What Are Some Potential Targeting Strategies?

- No coupon
- Give everyone a coupon
- Randomly target the coupon
- Target customers from Facebook
- Target customers from Instagram
- Target customers from Facebook and Instagram
- Target customers with products in cart
- Target customers with products in cart from Facebook and Instagram
- Target customers with products in cart or from Facebook and Instagram
- Target customers with no purchase history
- Others...

Comparison of Potential Targeting Strategies



There are tradeoffs between transaction, revenues, and % of targeted.

What should Alex do?

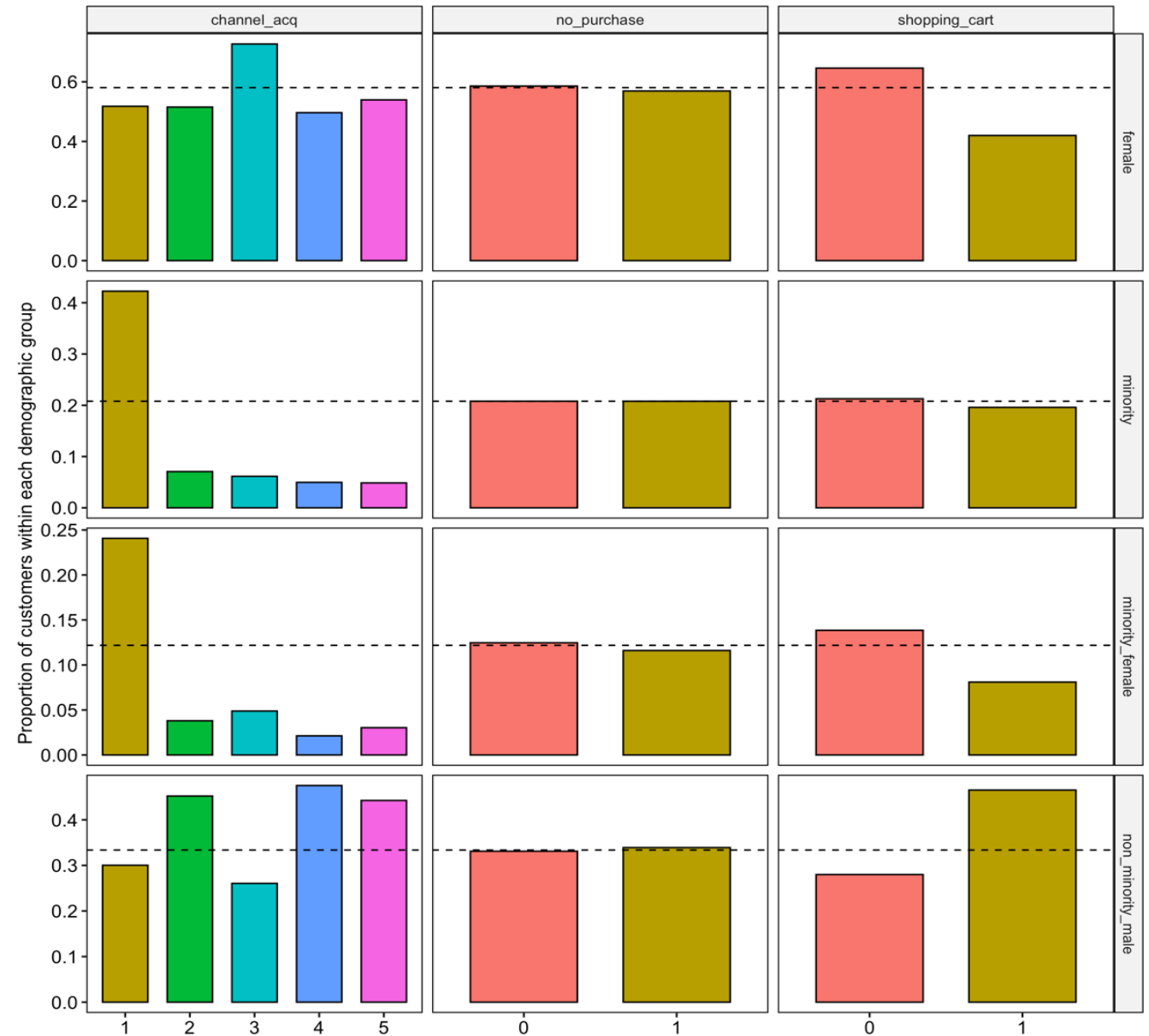
Artea (C): Potential Discrimination through Algorithmic Targeting

Using Demographics

- Upon signup, users indicated their gender, race, and ethnicity
- Should we use this data? How?

Differences in Consumer Behavior

- Are there implications of these differences on your selected targeting policy?
- Channel 1 = Google
- Channel 2 = Facebook
- Channel 3 = Instagram
- Channel 4 = Referral
- Channel 5 = Other



Would Artea be discriminating if they targeted promotions based on channel acquisition?

0

Yes

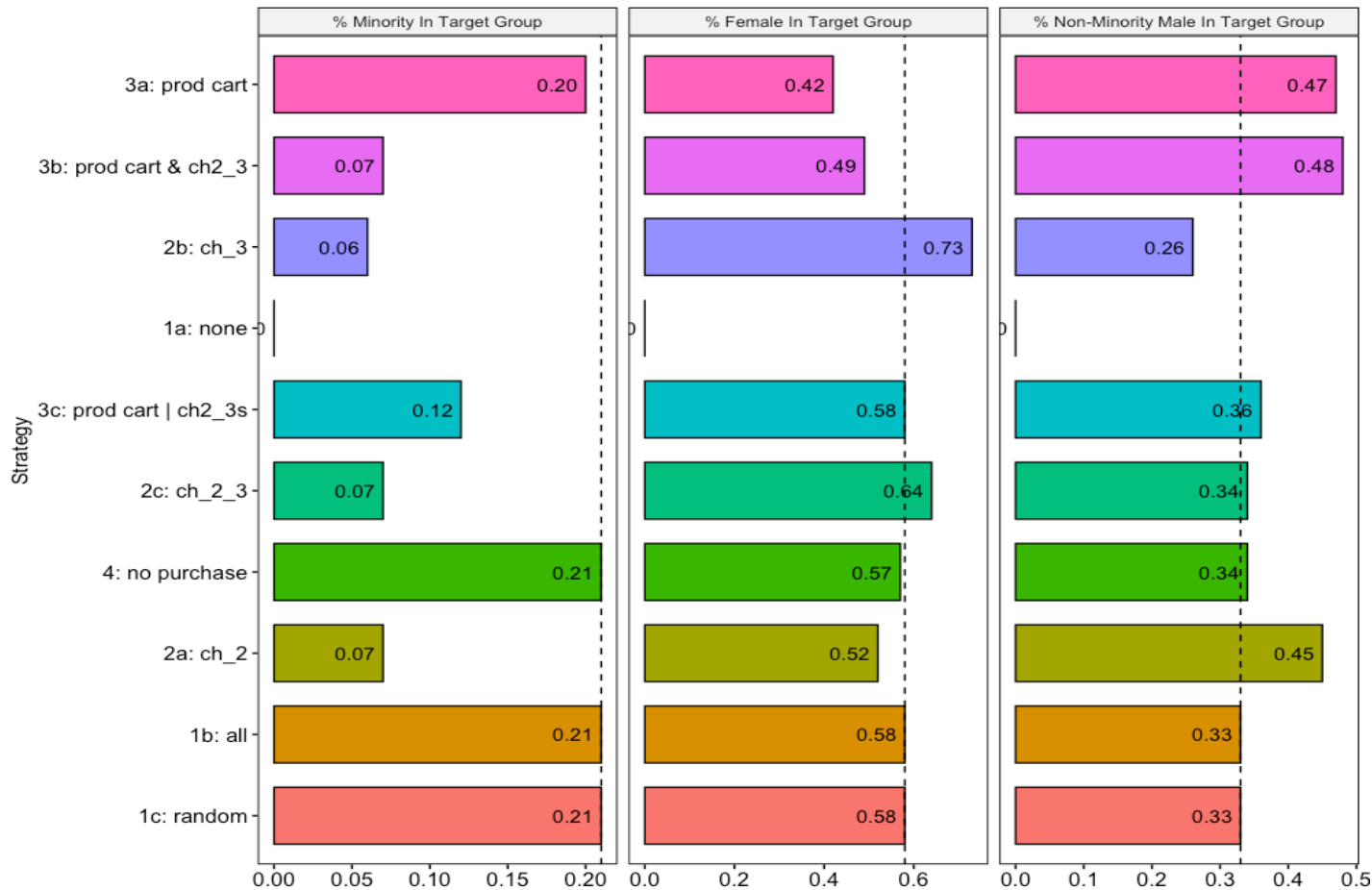
0

No

0

Impact of Demographics on Targeting Strategies

Baseline data contains 58% female and 21% minority



Listed in order of predicted total revenue, dashed line = average in the data

- How should Campbel react to these data patterns?
- Would Artea be discriminating if they targeted the promotions based on channel of acquisition?

Algorithmic Bias in Marketing

- Typically think of it as a result of:
 - Perpetuating historical disparities or human biases
 - Result of unrepresentative data
- Here, we see that it happens due to underlying correlations in the data despite all intentions to be fair
- This type of problem is likely to occur in incidents of personalization – very common in marketing applications and E-Commerce

Algorithmic Bias in Marketing

- An algorithm using the data to produce a targeting policy disproportionately targets a specific population due to underlying correlations in the data
- Is this a problem?
 - Legally no. Ethically?
- Sunlight is the best disinfectant –
 - Without demographics data we don't even know there is a problem...
- Should we target at all? Should we personalize? How?
- What do you do as a manager if an issue like this comes to your attention?
- Should you let an algorithm make the decisions for you?

All important issues to think about. No one best/right answer – but understanding the details will make you a better decision maker!

Key Takeaways

- A/B tests can be used to test a global change or to identify heterogeneity in response to the intervention and implement a more personalized approach
- Targeted strategies enable firms to efficiently use marketing resources and increase profitability (or any other outcome of interest) as a result
- Targeting and personalization practices may result in policies that discriminate against certain types of customers
- Algorithmic bias can occur even when demographic data is not collected
- It is sometimes difficult to identify that a bias is occurring. Marketers need to evaluate the implications of their policies to identify these biases

Break

Fairness

AI Can be Biased

Women less likely to be shown ads for high-paid jobs on Google, study shows

Automated testing and analysis of company's advertising system reveals male job seekers are shown far more adverts for high-paying executive jobs



Source: <https://www.theguardian.com/technology/2015/jul/08/women-less-likely-ads-high-paid-jobs-google-study>

Apple Card Investigated After Gender Discrimination Complaints

A prominent software developer said on Twitter that the credit card was “sexist” against women applying for credit.

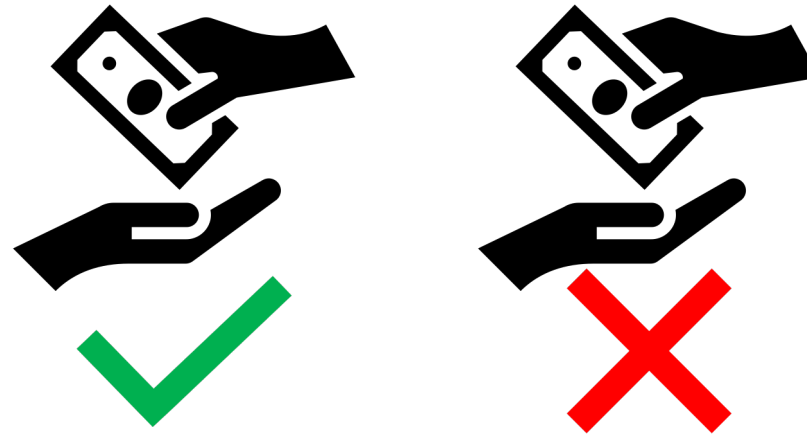
 Give this article  



Source: <https://www.nytimes.com/2019/11/10/business/Apple-credit-card-investigation.html>

Motivation

- Action: provide loan or not



- Data $X = (A, B)$
 - A = socially salient attributes (also referred to as “sensitive” or “protected” attributes)
 - Examples: race, gender
 - B = all other attributes
 - Examples: income, loan size, years of employment

Fairness Through Unawareness

- Use only B (non-sensitive attributes) to determine whether to give a loan
 - I.e., don't use sensitive attributes in prediction problem
- Implicit assumption: B is independent of A
- Benefit: a type of procedural fairness (“mainstream” fairness)

Fairness Through Unawareness

- What's the problem? Perpetuation of historical bias
 - If the data used to train the system includes **past loan approvals** that were influenced by bias (conscious or unconscious), the system learns to replicate these patterns (even if do not use sensitive attribute A)
 - One can often use B to predict $A \rightarrow$ system implicitly uses A
 - Example: Use employment length to predict loan approval
 - Men historically given more loans
 - Don't use gender but employment length highly correlated with gender (e.g., women taking time off for children)
 - Lower employment length less likely to be approved for a loan
 - The system, unaware of these correlations, indirectly discriminates by giving more loans to men
 - This will provide fewer opportunities to women and exacerbate the gender gap



Fairness Definitions

Many Potential Definitions of Fairness

- Important to clearly identify what you mean by fairness
- Is it fair to make a loan decision by not using gender?
- Is it fair if equal proportions of men and women receive loans?
- Is it fair if equal proportions of men and women who pay back their loans receive loans?
- Let's be precise!

Problem Setup

- Data $X = (A, B)$
 - A = socially salient attributes (also referred to as “sensitive” or “protected” attributes)
 - Examples: race, gender
 - B = all other attributes
 - Examples: income, loan size, years of employment
- True label Y
 - Example: indicator for loan repayment
 - Not known at time of business decision
 - Can also think about non-binary response
- Predicted value \hat{Y}
 - Learn a model f to predict Y from X
 - Example: predict whether loan will be repaid
- Based on \hat{Y} , decide on a business decision
 - Example: provide loan?

Problem Setup - Continued

- True label Y
 - Example: indicator for loan repayment
 - Not known at time of business decision
- Predicted value \hat{Y}
 - Example: predict whether loan will be repaid

Prediction \hat{Y}	True label Y	Notion $P(\hat{Y} Y)$
$\hat{Y} = 1$	$Y = 1$	True positive rate
$\hat{Y} = 0$	$Y = 1$	False negative rate
$\hat{Y} = 1$	$Y = 0$	False positive rate
$\hat{Y} = 0$	$Y = 0$	True negative rate

Statistical Parity

- A classifier satisfies statistical parity if all groups have equal probability of being assigned to the positive predicted class (\hat{Y} is independent of A)
 - Also known as disparate impact, independence
 - $P(\hat{Y} = 1|A = i) = P(\hat{Y} = 1|A = j)$ for all groups i, j
 - Example: equal proportion of men and women receive loans
- Implicit assumption: expression of belief that “relevant” traits are independent of A



Statistical Parity

- Limitation: can be gamed by carefully selecting x% from one group and carelessly selecting x% from another
 - Carefully select which men receive loan but randomly select which women receive loan
 - Achieves statistical parity **BUT** uncreditworthy women who get loan can't pay it off (high false positive rate) and creditworthy women don't get loan (high false negative rate)



Statistical Parity

- Hard to achieve exactly $P(\hat{Y} = 1 | A = i) = P(\hat{Y} = 1 | A = j)$
- Allow for some margin:
 - Ratio of rate of favorable outcome for the unprivileged group to that of the privileged group
 - Ideally 1
 - Might consider fair so long as ratio at least 0.8



Demographic Parity

- A classifier satisfies demographic parity if A is equally distributed amongst those selected
 - $P(A = i | \hat{Y} = 1) = P(A = j | \hat{Y} = 1)$ for all groups i, j
 - Example: of those who receive high credit limit, half are men and half are women
 - Note: order is reversed compared to statistical parity
 - Ensures a diverse pool is selected
- Implicit assumption: expression of belief that “relevant” traits are independent of A



Demographic Parity

- Limitation: can be gamed by carefully selecting n from one group and carelessly selecting n from another
 - Carefully select which men receive loan but randomly select which women receive loan
 - Achieves demographic parity **BUT** uncreditworthy women who get loan can't pay it off (high false positive rate) and creditworthy women don't get loan (high false negative rate)



Separation

- A classifier satisfies separation if Y depends on A only to the extent justified
 - \hat{Y} does not depend meaningfully on A given Y
 - For all groups i, j
 - $P(\hat{Y} = 1|A = i, Y = 1) = P(\hat{Y} = 1|A = j, Y = 1)$ **AND**
 - $P(\hat{Y} = 1|A = i, Y = 0) = P(\hat{Y} = 1|A = j, Y = 0)$
 - Groups experience same TP and FP rates
 - Example:
 - Of those that repay loan, equal % of men and women receive loan **AND**
 - Of those who do not repay loan, equal % of men and women receive loan

Separation

- Implicit assumption: some assumption on A is justified (allows us to violate statistical parity)
- Limitation: does not consider contextual conditions that led to this situation
 - Will not close gap between gender over time



Repay



Default



Repay



Repay



Repay



Default

Sufficiency

- Given a score, the probability of the true variable being 1 should be the same for each group
 - Y does not depend meaningfully on A given \hat{Y}
 - For all groups i, j
 - $P(Y = 1|A = i, \hat{Y} = 1) = P(Y = 1|A = j, \hat{Y} = 1)$
 - Also known as calibration
 - Example:
 - Of those that receive loan, equal % of men and women repay loan



Repay



Repay



Repay



Repay



Which Definition of Fairness Should be Used?

- It depends on what is trying to be achieved and the law
 - Consider equality of outcome vs. equality of opportunity
 - Equality of outcome: fairness in terms of results, minimize disparity, interventions may undermine merit-based rewards
 - Equality of opportunity: fairness in terms of process, differences in outcome will occur based on individual effort and ability
- There are many more definitions of fairness
- Impossibility result – one classifier cannot handle all notions of fairness

Source: <https://fairware.cs.umass.edu/papers/Verma.pdf>

Is Our Model Biased?

- A model to predict grad school admission
 - Predicted Label
 - 0 = reject
 - 1 = accept
 - True Label
 - 0 = rejected
 - 1 = accepted
- Does the data satisfy Statistical Parity? Demographic Parity? Separation? Sufficiency?

Person #	Gender	True Label	Predicted Label
1	M	1	1
2	M	1	1
3	F	0	0
4	M	0	1
5	F	1	0
6	M	1	0
7	M	1	1
8	F	1	1
9	M	0	0
10	F	0	0

Is fairness achieved for statistical parity?

 0

Yes

0

No

0

Calculating Fairness Metrics

- Statistical parity:
 - $P(\hat{Y} = 1|A = i) = P(\hat{Y} = 1|A = j)$ for all groups i, j
 - Acceptance rate of males and females should be the same
 - Acceptance rate of males: $4/6 = 66.7\%$
 - Acceptance rate of females: $1/4 = 25\%$
 - Ratio of acceptance rates: $25\%/66.7\% = 0.37 < 0.80$
 - Does not satisfy statistical parity

Person #	Gender	True Label	Predicted Label
1	M	1	1
2	M	1	1
3	F	0	0
4	M	0	1
5	F	1	0
6	M	1	0
7	M	1	1
8	F	1	1
9	M	0	0
10	F	0	0

Calculating Fairness Metrics

- Demographic parity:
 - $P(A = i | \hat{Y} = 1) = P(A = j | \hat{Y} = 1)$ for all groups i, j
 - Equal number of males and females should be accepted
 - Number of accepted males: 4
 - Number of accepted females : 1
 - Does not satisfy demographic parity

Person #	Gender	True Label	Predicted Label
1	M	1	1
2	M	1	1
3	F	0	0
4	M	0	1
5	F	1	0
6	M	1	0
7	M	1	1
8	F	1	1
9	M	0	0
10	F	0	0

Calculating Fairness Metrics

- Separation:
 - $P(\hat{Y} = 1 | A = i, Y = 1) = P(\hat{Y} = 1 | A = j, Y = 1)$ **AND**
 - $P(\hat{Y} = 1 | A = i, Y = 0) = P(\hat{Y} = 1 | A = j, Y = 0)$
 - Equal TPR and FPR
 - TPR for males: $3/4 = 75\%$
 - TPR for females : $1/2 = 50\%$
 - FPR for males: $1/2 = 50\%$
 - FPR for females: 0%
 - Does not satisfy separation

Person #	Gender	True Label	Predicted Label
1	M	1	1
2	M	1	1
3	F	0	0
4	M	0	1
5	F	1	0
6	M	1	0
7	M	1	1
8	F	1	1
9	M	0	0
10	F	0	0

Calculating Fairness Metrics

- Sufficiency:
 - $P(Y = 1|A = i, \hat{Y} = 1) = P(Y = 1|A = j, \hat{Y} = 1)$
 - Males: $3/4 = 75\%$
 - Females: $1/1 = 100\%$
 - Does not satisfy sufficiency
 - Model not as accurate at predicting appropriate qualifications of males compared to females → model biased towards males

Person #	Gender	True Label	Predicted Label
1	M	1	1
2	M	1	1
3	F	0	0
4	M	0	1
5	F	1	0
6	M	1	0
7	M	1	1
8	F	1	1
9	M	0	0
10	F	0	0

So What? Strategies to Achieve Fairness

Preprocessing Data

Inprocessing –
Constrained
Optimization

Postprocessing Outputs

- Alter the feature space to be uncorrelated with the protected attribute
- Incorporate fairness constraint into the optimization problem
- Manipulate the classifier to be uncorrelated with the protected attribute

Use relevant definition of fairness to confirm intervention successful.

AI Fairness 360

- <https://aif360.res.ibm.com/data>

Break

Review

Final Exam

- Will be held on 12/12 2:30-4:30 in Geffen
 - You will have two hours to complete the exam
- Choice Modeling onwards
- One page (front and back) cheat sheet will be allowed
- Exam will contain both multiple choice and written answers
- No need for Python or Excel
- Bring a calculator

Week 6 – Choice Models

- Definitions of choice set
- Understand the modeling approach and be able to interpret the results
 - Utility function dependent on consumers' characteristics and alternative attributes
 - Understand the random utility model and be able to write down the model
 - Beware of different portions of the systematic utility and identification issues
 - Only differences in utility matter
 - Understand MLE and how to use BIC to compare models
 - Know how to calculate choice probabilities
- Elasticities and IIA

Consider a discrete choice model over a choice set consisting of 4 alternatives. The dataset contains the choices made by consumers, a Price attribute for the brands and a single consumer demographic (Age). Assume that the systematic component of the utility contains the alternative-specific intercepts, the main effects of Price and Age, and an interaction between these two variables.

As only differences in utilities matter, how many coefficients in the model are identifiable?

0

6

0

7

0

8

0

9

0

Start the presentation to see live content. For screen share software, share the entire screen. Get help at pollev.com/app

Week 7 – Text Analysis

- Simple Summaries
- Sentiment Analysis
- Topic Modeling
- Large Language Models

Week 8 - Product

- Conjoint Analysis
 - Product design concepts (attributes, levels, profiles)
 - Design a conjoint study and collect the data
 - Stated vs. revealed preference methods
 - Analyze the data to generate insights (attribute importance, market shares, ...)
- Diffusion of innovation
 - Bass Model + assumptions and parameters + how to estimate

t is the time period

p is the coefficient of innovation

q is the coefficient of imitation

$Y(t)$ is the cumulative number of adopters until t

m is the total market potential

$$S(t) = \underbrace{\left[p + q \frac{Y(t)}{m} \right]}_{\text{Adoption rate}} * \underbrace{[m - Y(t)]}_{\text{Number of potential adopters}}$$

Number of adoptersNumber of potential adopters

What are partworths in conjoint analysis?



An estimate of the value of a conjoint analysis survey to the firm running the survey

0

The minimum number of respondents required to make sure the survey was worth running

0

A set of parameters that quantify how much each consumer values each level of each attribute

0

The set of product descriptions to be shown to consumers taking the survey

0

None of the above

0

Week 9 – Generative and Explainable AI

- Types of generative models
 - Generative Adversarial Networks
 - Variational Autoencoders
 - Transformers
- Potential use of GenAI in Market Research
- Risks associated with GenAI
- Why is explainability important?
- Tools for explainability
 - LIME for text, Grad-CAM for images, SHAP values for structured data

Week 10 – Price & Place

- Estimate the demand curve
 - Direct measures: survey, BDM procedure... be able to build demand curve and find optimal price given WTP
 - Conjoint analysis: be able to infer WTP using conjoint
- Estimate a demand system (Montclair Video case)
 - Incentive compatibility
- How to choose a channel?
 - Direct-to-consumer vs. retailer (Glaubinger case)
 - Double marginalization
- Omnichannel behavior
 - Substitution, augmentation, and volume effects

Attribute	Partworth
Price (\$500 vs. \$400)	20
Built-in projector	30

How much more is this consumer willing to pay for a built-in projector?

\$30

0

\$67

0

\$100

0

\$150

0

None of the above

0

Start the presentation to see live content. For screen share software, share the entire screen. Get help at pollev.com/app

In omnichannel marketing, when a mobile banking customer uses their smartphones to locate the nearest ATM, this is an example of:



the substitution effect

0

the augmentation effect

0

the volume effect

0

Week 11 – Advertising & Promotion

- Online advertising: display vs. search ads
- Understand and be able to compute CPM, CPC, CPA
- Attribution problem
 - A firm may reach the consumer through different touchpoints
 - It's difficult to attribute an acquisition, purchase to a single channel
- Conduct an A/B test and measure the ROI and profitability of a campaign
 - $ROI = \text{Campaign Effectiveness} / \text{Incremental Cost}$
 - Important: you can only assess causality on aspects that you experiment on
 - Having an experiment does not mean everything will be causal
- Long term impact of promotions
 - Be able to compute the effectiveness of price promotions given the output of a regression

If you pay \$100 for 10,000 impressions and get 20 clicks, what is your CPM?

0

\$1

0

\$5

0

\$10

0

\$100

0

None of the above

0

Below are the results of a regression in which the dependent variable was sales in \$. What is the total incremental sales per promotion?



-\$4,000

 0

-\$1,000

 0

\$1,500

 0

\$2,500

 0

\$5,000

 0

Start the presentation to see live content. For screen share software, share the entire screen. Get help at pollev.com/app

	Coefficient	P-value
Intercept	10,000	<0.05
Winter	-2,000	<0.05
Fall	-1,000	<0.05
Spring	500	<0.05
Promotion this month	2,500	<0.05
Promotion last month	-1000	<0.05

Week 12 – Artea & Fairness

- Artea
 - How to analyze results of an A/B test
 - Heterogeneous treatment effects
- Fairness
 - Definitions of fairness

Class Ratings

Please submit your class ratings

Last semester, students who didn't respond to the survey could only access their grades much later than students who did respond to the survey

Class Takeaway

- Marketing analytics is at the intercept of conceptual/substantive questions and quantitative modeling
- Marketing is about changing the actions of the company
 - Product design, pricing, choice of channel, advertising, etc.
- Effective marketing implementation requires two components:
 - Understanding factors to account for
 - Using the right tools to describe the problem and assess the actions' effectiveness
 - Exploratory and descriptive research + lab/field/natural experiments + ML

My Objectives for this Course

- Marketing and customer analytics have changed drastically
 - Old marketers: proficient with concepts & business acumen but less about technique
 - Old engineers/data scientists: proficient with technique but not as exposed to concepts
 - Future marketer/data analyst/data scientist: able to recognize concepts and the right tools for a business problem
 - Marketing Analytics hasn't failed yet!
- Introduce you to important marketing
 - Datasets
 - Frameworks
 - Tools
 - Concepts

Some Advice...

- Problems faced by companies are difficult
- We need to be organized and pragmatic when working with data
 - Always take the time to obtain some model-free evidence (descriptive stats)
 - There is a trade-off between prediction (full ML model) and interpretability
 - “Insights” problems: interpretability is key
 - Prediction problems: you want to use the most parsimonious model that offers the best prediction (and ideally some interpretation)
 - One strategy: start with simple models and increase complexity as you go
- Be a team player
 - A good teacher when you have some expertise
 - A good student when you don't and learn from your colleagues
- Keep on learning!

Last Words of Wisdom...

- Marketing job # 1: Facilitate a profitable exchange between the customer and the firm
- But it all starts with the customer, so be **customer-centric**

I hope you had a great learning experience!

