

# Causality & Advertising

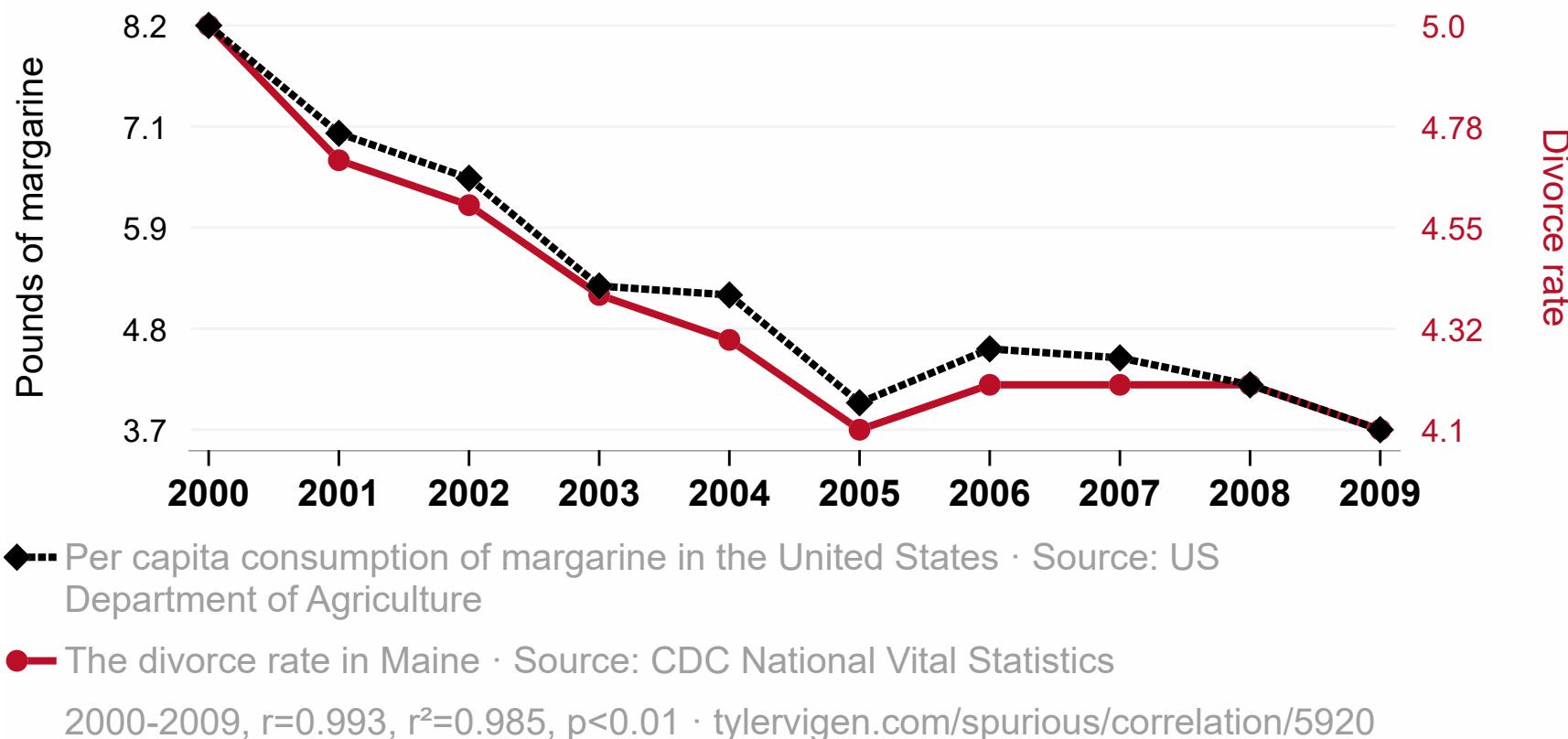
UCSD MGTA 451-Marketing

Kenneth C. Wilbur

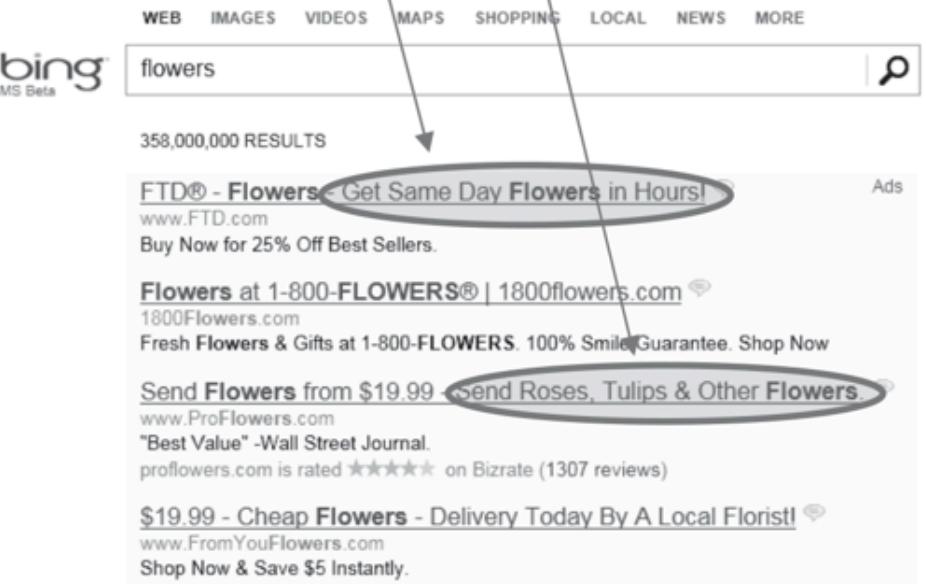
# Per capita consumption of margarine

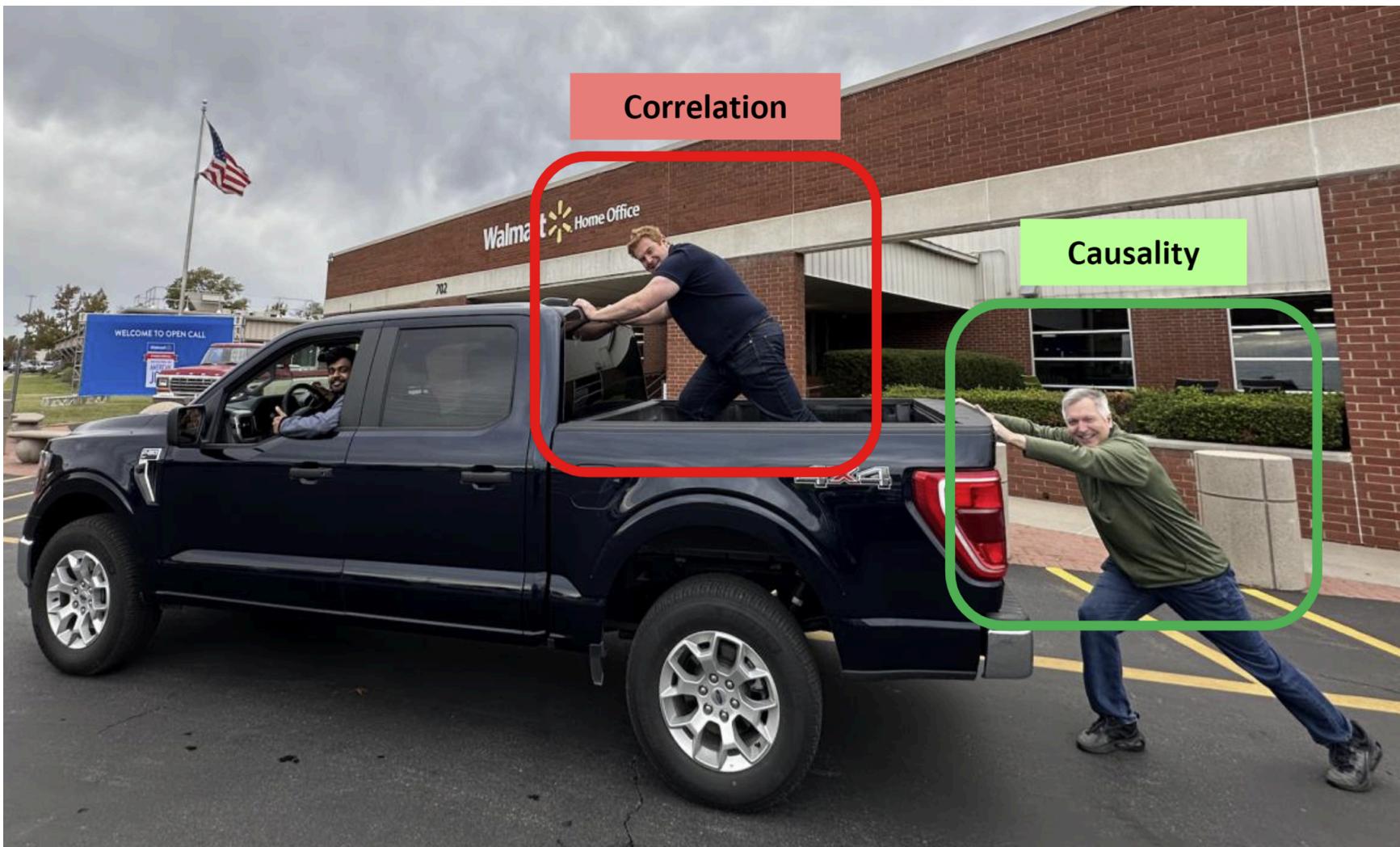
correlates with

## The divorce rate in Maine



- Suppose 10 outcomes, 1000 predictors, N=100,000 obs
  - Outcomes might include visits, sales, reviews, ...
  - Predictors might include customer attributes, session attributes, ...
- Suppose everything is noise, no true relationships
  - The distribution of the 10,000 correlation coefficients would be Normal, tightly centered around zero
  - A 2-sided test of `{corr == 0}` would reject at 95% if  $|r| > .0062$
- We should expect 500 false positives
  - What is a 'false positive' exactly?
- In general, what can we learn from a significant correlation?

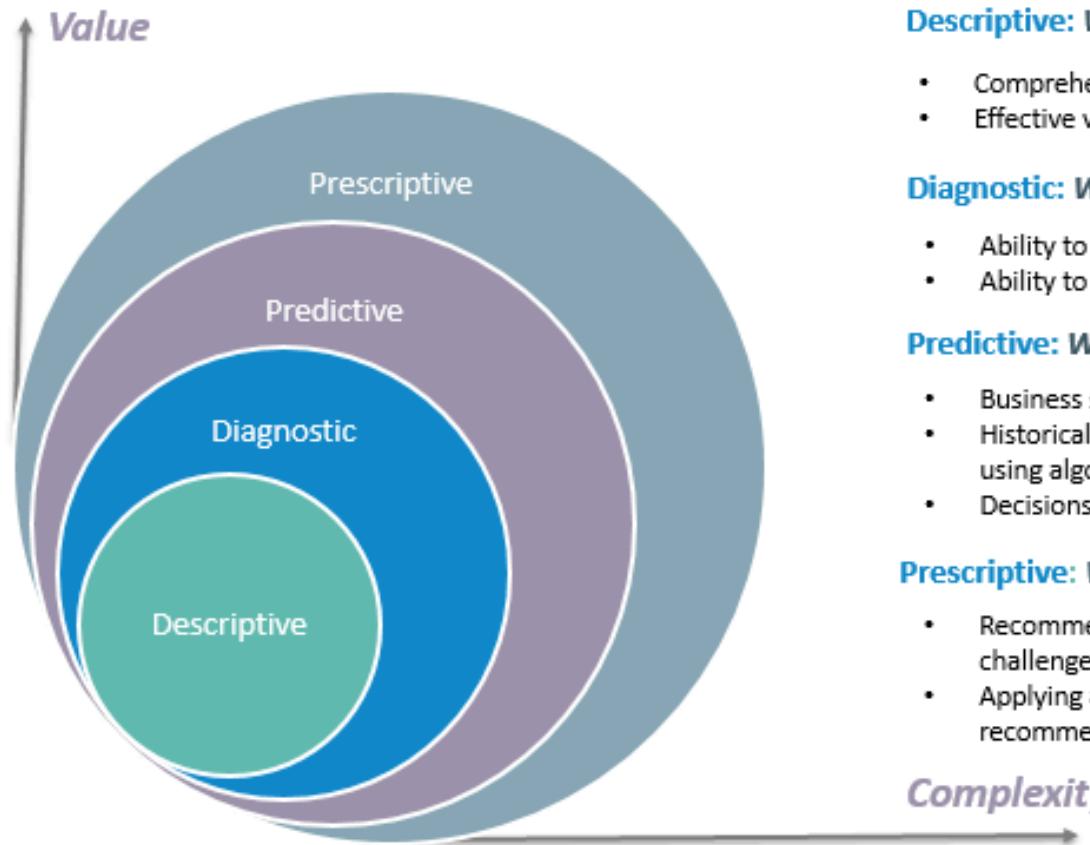




# Agenda

- Causality
- Experiments
- Quasi-experiments
- Correlations
- Ad/sales frameworks

## 4 types of Data Analytics



### What is the data telling you?

#### **Descriptive:** *What's happening in my business?*

- Comprehensive, accurate and live data
- Effective visualisation

#### **Diagnostic:** *Why is it happening?*

- Ability to drill down to the root-cause
- Ability to isolate all confounding information

#### **Predictive:** *What's likely to happen?*

- Business strategies have remained fairly consistent over time
- Historical patterns being used to predict specific outcomes using algorithms
- Decisions are automated using algorithms and technology

#### **Prescriptive:** *What do I need to do?*

- Recommended actions and strategies based on champion / challenger testing strategy outcomes
- Applying advanced analytical techniques to make specific recommendations

# Causal Inference

- Suppose we have a binary “treatment” or “policy” variable  $T_i$  that we can “assign” to person  $i$ 
  - Examples: Advertise, Serve a design, Recommend
  - “Treatment” terminology came from medical literature
- Suppose person  $i$  could have a binary potential “response” variable  $Y_i(T_i)$ 
  - Examples: Visit site, Click product, Add to Cart, Purchase, Rate, Review
  - Looks like the marketing funnel model we saw previously
- Important:  $Y_i$  may depend fully, partially, or not at all on  $T_i$ , and the dependence may be different for different people

# Why care?

- We want to maximize profits  $\pi_i(Y_i(T_i), T_i)$
- Suppose  $Y_i = 1$  contributes to revenue; then  $\frac{d\pi_i}{dY_i} > 0$
- Suppose  $T_i = 1$  is costly; then  $\frac{d\pi_i}{dT_i} = \frac{\partial\pi_i}{\partial Y_i} \frac{\partial Y_i}{\partial T_i} + \frac{\partial\pi_i}{\partial T_i}$
- We have to know  $\frac{\partial Y_i}{\partial T_i}$  to optimize  $T_i$  assignments
  - Called the "treatment effect" (TE)
- Profits may decrease if we misallocate  $T_i$

# Fundamental Problem of Causal Inference

- We can only observe either  $Y_i(T_i = 1)$  or  $Y_i(T_i = 0)$ , but not both, for each person  $i$
- This is a missing-data problem that we cannot resolve. We only have one reality
  - The case we don't observe is called the "counterfactual"

# So what can we do?

1. Experiment. Randomize  $T_i$  and estimate  $\frac{\partial Y_i}{\partial T_i}$  as  $\text{avg } Y_i(T_i = 1) - Y_i(T_i = 0)$

- Creates new data; costs time, money, attention; deceptively difficult to design and then act on

2. Use assumptions & data to estimate a “quasi-experimental” average treatment effect using archival data

- Requires expertise, time, attention; difficult to validate; not always possible

3. Use correlations: Assume past treatments were assigned randomly, use past data to estimate  $\frac{\partial Y_i}{\partial T_i}$

- Easiest approach; but T is only randomly assigned when we run an experiment, so what exactly are we doing here?

4. Fuhgeddaboutit, go with the vibes, do what we feel

# How much does causality matter?

- How hard should we work?
- Organizational returns or costs of getting it right?
- Data thickness: How likely can we get a good estimate?
- How does empirical approach fit with organizational analytics culture? Can we act on what we learn?
- Individual considerations: promotion, bonus, reputation, career; Will credit be stolen or blame be shared?
- Accountability: Will ex-post attributions verify findings? Will results threaten or complement rival teams/execs?

# Ad/sales example: Experiment

1. Randomly assign ads to customer groups on a platform; measure sales in each group,  
e.g. [Link](#)

- Often called "incrementality" in ad/sales context
- Pros: AB testing is easy to understand, easy to implement, easy to validate
- Cons: Can we trust the platform's "black box"? Will we get the data and all available insights? Could platform knowledge affect future ad costs?

2. Randomly assign messages within a campaign

3. Randomly choose times, places, segments or combinations; compare treated times to controls

4. Randomize over budgets and bids

5. Randomly choose platforms, publishers, behavioral targets, etc., to compare RoAS across options

RoAS = Return on Ad Spend. Usually,  $\text{RoAS} = \text{Sales} / \text{AdSpend}$ . Sometimes,  $= (\text{Sales} - \text{AdSpend}) / \text{AdSpend}$

# 2. Ad/sales example: Experiment

## Key issues for any experimental design:

- Always run A:A test first. Validate the infrastructure before trusting a result
- Can we agree on the opportunity cost of the experiment? "Priors"
- How will we act on the (uncertain) findings? Have to decide before we design
  - Example: Suppose we estimate RoAS at 0.5 with a 95% confidence interval of [0.45, 0.55]. Or, suppose we estimate RoAS at 0.5, but we have a 95% confidence interval of [-2.1], 3.1]. What will we do with this information?

# Before quasi-experiments: Vocab

**Model:** Mathematical relationship between variables that simplifies reality, eg  $y=xb+e$

**Identification strategy:** Set of assumptions that isolate a causal effect  $\frac{\partial Y_i}{\partial T_i}$  from other factors that may influence  $Y_i$

- A system to compare apples with apples, not apples with oranges

We say we “identify” the causal effect if we have an identification strategy that reliably distinguishes  $\frac{\partial Y_i}{\partial T_i}$  from possibly correlated unobserved factors that also influence  $Y_i$

If you estimate a model without an identification strategy, you should interpret the results as correlational

You can have an identification strategy without a model, e.g.

$$\text{avg } Y_i(T_i = 1) - Y_i(T_i = 0)$$

Usually you want both. Models help with quantifying uncertainty and estimating treatment effects by controlling for relevant observables

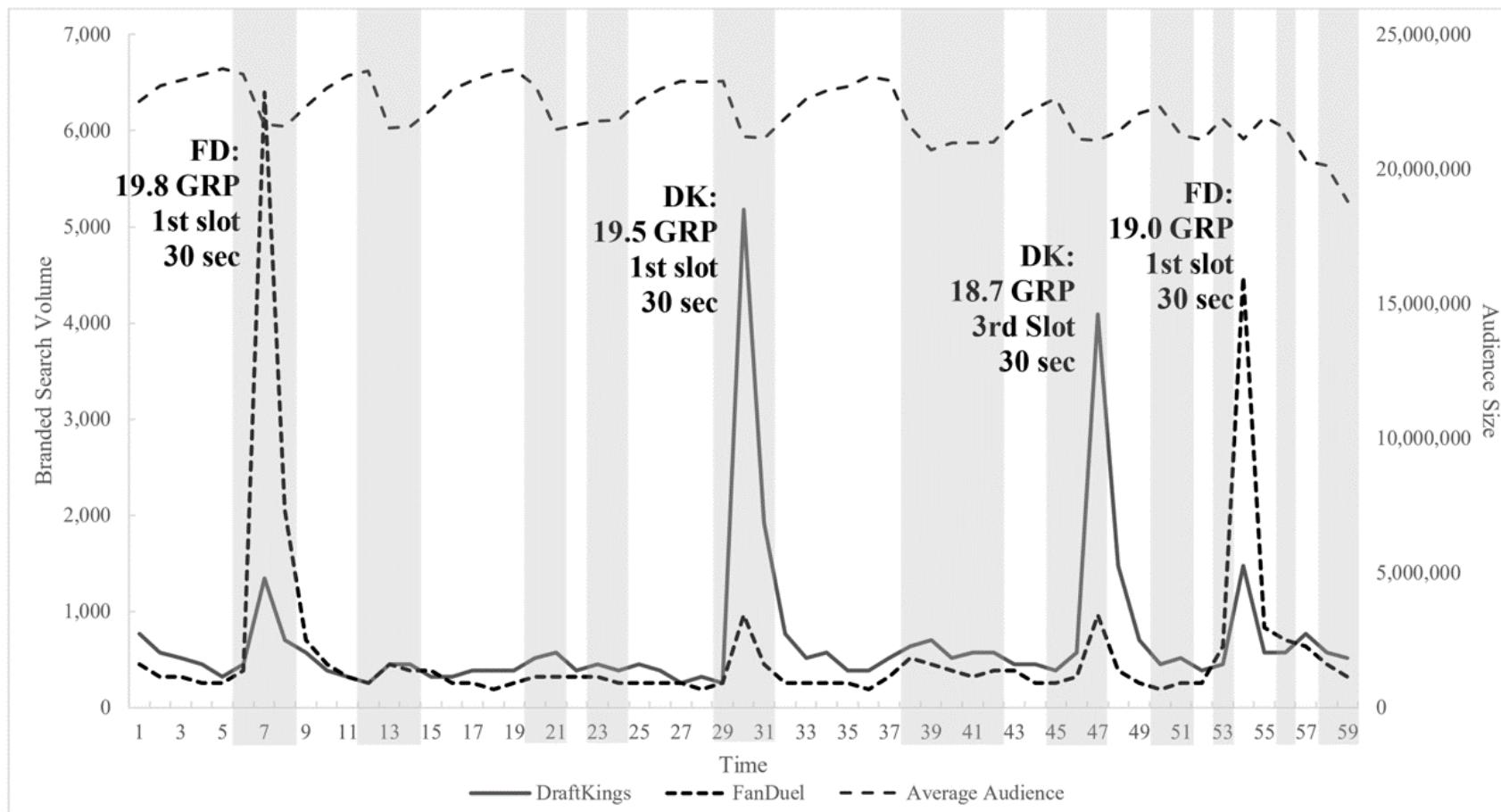
## 2. Ad/sales: Quasi-experiments

Goal: Find a “natural experiment” in which  $T_i$  is “as if” randomly assigned, to identify  $\frac{\partial T_i}{\partial Y_i}$

Possibilities:

- Firm starts, stops or pulses advertising without changing other variables, especially when staggered across times or geos
  - Competitor starts, stops or pulses advertising
  - Change in ad prices, availability or targeting for exogenous reasons that don't affect Y (e.g., election or outage)
  - Discontinuous ad copy/campaign changes without changing other variables

# DraftKings and FanDuel TV ad effects on Google Search



# Ad/sales: Quasi-experiments (2)

Or, construct a “control group”

- Customers or markets with similar demand trends where the firm never advertised
- Competitors or complementors with similar demand trends that don’t advertise

Techniques like “difference in differences,” “synthetic control,” “regression discontinuity,” “matching,” and “instrumental variables” can help

In each case, we try to predict our missing counterfactual data, then estimate the causal effect as observed outcomes minus predicted outcomes

# 3. Ad/sales example: Correlational

Just get historical data on  $Y_i$  and  $T_i$  and run a regression

Most people use OLS, but Google's CausalImpact R package is also popular

The implicit assumption is either that there are no unobservables that influence both  $Y_i$  and  $T_i$ ; or that we are OK with a correlation

"Better to be vaguely right than precisely wrong"

But are we the guy in the truck bed?

### 3. Problem 1 with $\text{corr}(\text{ad}, \text{sales})$

Advertisers try hard to maximize ad effects

We select messages, media, times, contexts, target audiences, publishers (etc.) to maximize potential impact

Sometimes, we use ML algos that max  $\text{corr}(\text{ad}, \text{sales})$

So, if you assume that past  $T_i$  was random, it's functionally equivalent to assuming that the advertising team is incompetent

- This type of selection/treatment problem is common in marketing

### 3. Problem 2 with $\text{corr}(\text{ad}, \text{sales})$

Suppose you wisely advertise only in places where you expect a big response

E.g. you advertise surfboards in coastal cities with big waves, but not in landlocked cities

You'll get a big correlation between ads and sales, partly because of differences in addressable market sizes

More ads in san diego, more surfboard sales in san diego

$\text{Corr}(\text{ad}, \text{sales})$  would overestimate the causal effect of ads on sales, and likely lead you to spend too much on ads in san diego

Many, many firms basically do this

# Why are some teams OK with $\text{corr}(\text{ad}, \text{sales})$ ?

## 1. Some worry that if ads go to zero $\rightarrow$ sales go to zero

- For small firms or new products, this may be good logic
- If this is true, then the correlation is the causal effect
- Directly suggests an inexpensive experimental design
- Claim directly implies a huge customer satisfaction problem. Happy customers usually share their experiences with others. If you really believe this is true, try a referral program

## 2. Some firms assume that correlations indicate direction of causal results

- The guy in the truck bed is not pushing backwards right?
- But what is this assumption based on?
- Direction is only part of the picture; what about effect size?
- Unfortunately, if it's hard to get a causal estimate, then this assumption is probably ill founded

# Why are some teams OK with $\text{corr}(\text{ad}, \text{sales})$ ?

## 3. CFO and CMO negotiate ad budget

- CFO asks for proof that ads work
- CMO asks ad agencies, platforms & marketing team for proof
- CMO sends proof to CFO ; We all carry on
- Things may change if ad effects team reports to CFO

## 4. Few rigorous analytics cultures or ex-post checks

- Downside of lost sales may exceed downside of foregone profits

## 5. Estimating causal effects of ads can be pretty difficult

- Many firms lack expertise, patience, execution skill
- Ad/sales tests may be statistically inconclusive, especially if small
- Tests are often designed without subsequent actions in mind

# Why are some teams OK with $\text{corr}(\text{ad}, \text{sales})$ ?

## 6. Platforms often provide correlational ad/sales estimates

- Which one do you think is larger, correlational or experimental ad effect estimates?
- Which one would most client marketers prefer?
- Platform estimates are typically "black box" without neutral auditors
- Sometimes platforms respond to marketing executive demand for good numbers
- "Nobody ever got fired for buying [famous platform brand here]"

## 7. Historically, most advertising work was done by agency professionals

- Agency compensation usually relies on spending, not incremental sales
- Principal/agent problems are common
- "Advertising attribution" is all about maximizing credit to ads
- These days, most marketers have in-house agencies, and split work

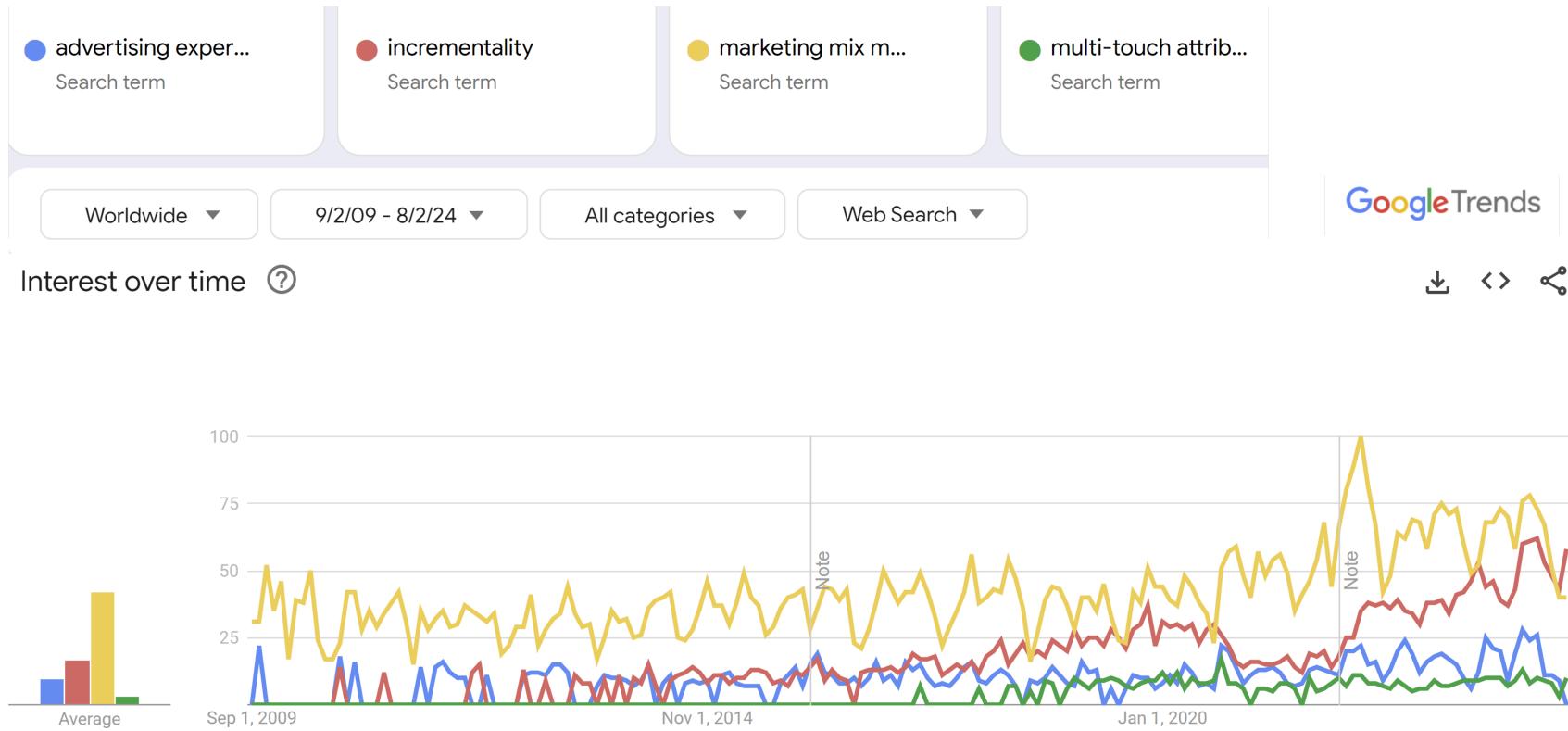
# What is incrementality?



Incrementality refers to the measure of the additional impact or value generated by a specific action, campaign, or intervention beyond what would have occurred naturally without it. In marketing and advertising, incrementality is often used to determine the effectiveness of campaigns by comparing the results of those exposed to the campaign versus a control group that was not exposed. This helps in understanding the true value and ROI of marketing efforts.

## Key Points of Incrementality

1. **Causal Inference:** Incrementality is rooted in causal inference, aiming to isolate the effect of a specific action from other factors.
2. **Control Groups:** A key methodology involves using control groups to measure what would have happened in the absence of the intervention.
3. **Lift:** Incrementality is often expressed as "lift," representing the increase in desired outcomes (sales, conversions, engagement, etc.) due to the campaign.
4. **A/B Testing:** Commonly used techniques to measure incrementality include A/B testing, where one group is exposed to the treatment, and the other is not.
5. **Attribution Models:** Incrementality is crucial for accurate attribution models, ensuring that credit is assigned correctly to the actions that drive results.



- I believe we're at the start of a generational shift

### 2023 Advertising To Sales Ratios by Industry Sector

Industry Sector	Ad to Sales Ratio %	Ad Growth %	Sales Growth %
Agriculture, Forestry, Fishing	0.40	-17.79	17.64
Mining, Extraction	0.07	13.28	-3.45
Construction	0.35	26.94	0.53
Manufacturing	2.70	10.47	0.69
Transportation, Communications, Utilities	3.60	4.26	5.31
Wholesale Trade	1.27	7.81	0.70
Retail Trade	2.42	11.69	4.09
Finance, Insurance, Real Estate	2.31	-0.05	8.57
Services	4.06	11.27	11.86
<b>All sectors combined</b>	<b>2.83</b>	<b>9.01</b>	<b>4.77</b>

**Advertising Ratios & Budgets** is the source for the above data. This detailed report covers over 2,500 companies and 315 industries with fiscal 2023 and 2022 advertising budgets and revenue, 2023 ad-to-sales ratio and ad-to-profit ratio, as well as 2023 annual growth rates in ad spending and sales. Use it to track competition, win new ad agency clients, set and justify ad budgets, sell space and time or plan new media ventures and new products. Includes industry and advertiser ad spending rankings and data on over 350 non-U.S. headquartered companies. Bought by major advertising agencies, media companies, advertisers and libraries. Published May 2024.

Advertising Sales Ratios - SAI Books

- Typical net margin: 8-10% ([see Damodaran](#))
- Ads finance information production

Here, [Sam Yadegar](#), CEO of HawkSEM, delivers expert advice on marketing budget estimates and how to balance your profit margin with your marketing goals for ultimate return on investment (ROI).

## How much revenue should I spend on marketing?

**Your marketing budget needs anywhere from 5-20% of your revenue to thrive. Generally, 5-10% is enough to sustain, but you'll need 11-20% into data-driven marketing campaigns to grow.**

Campaigns can span [digital marketing](#), traditional advertising, outreach events, tools, technology, in-house marketing team salaries, and agency services.

That said, arbitrary spending reaps minimal rewards, if any. To truly drive results, you've gotta spend wisely.

# Marketing Mix Model

- The “marketing mix” consists of quantifiable marketing efforts, such as product line, length and features; price and price promotions; advertising, PR, social media and other communication efforts; retail distribution intensity and quality; etc.
- A “marketing mix model” quantifies the relationship between marketing mix variables and outcomes
  - Idea goes back to the 1950s
  - E.g., suppose we increase price & ads at the same time
  - Or, suppose ads increased demand, and then inventory-based systems raised prices
- A “media mix model” quantifies numerous advertising efforts and relates them to sales
  - For example, suppose the brand bought ads from 000s of publishers
  - Confusingly, both abbreviated MMM and often feature similar structures
- MMM goal is to quantify past marketing mix effects, to better inform future efforts

# MMM elements

Typically, MMM uses geo/time data

- Outcome: usually sales. Could include more funnel metrics (visits, leads, ...)
- Predictors: Marketing mix factors under our control, plus competitor variables, seasonality, macroeconomic factors, + any other demand shifters

Model structure is usually some type of panel regression, vector autoregression, or bayesian model

- Often includes lags, interactions, and other nonlinear terms
- Regressions typically estimate marginal effects, not average effects
- Nonlinearities built into the model, such as Inc or Dec returns to ad spend, can drive key results

MMM often used to retrospectively evaluate advertising media and copy, advertising interactions, and inform future ad budgets

- MMM can be used for prediction but should be calibrated appropriately

# MMM Challenges

- Without experimental data, MMM results are correlational
- Data availability, accuracy and granularity are all critical
- MMM requires sufficient variation in predictors, else it cannot estimate coefficients
- “Model uncertainty” : Results can be strongly sensitive to modeling choices
- For much more, see this [MSI White Paper](#) or the [MMM Wikipedia article](#)

# Other Popular Ad/Sales Approaches

Remember, model <> identification strategy

- Lift Tests
- Multi-touch attribution (MTA)
  - Seeks to allocate "credit" for sales across advertising touchpoints
  - Related: First-touch attribution, last-touch attribution
- Cookie-based approaches vs. Google's Privacy Sandbox
- Ghost ads
- Other platform-provided experimentation tools



**Kenneth Wilbur** • You

Professor of Marketing and Analytics at University of California, San Diego...

4d • Edited •

MSBA student asked a great question. How would you answer?

Suppose you understand the importance of incrementality in advertising measurement, but everyone you work with prefers correlational measurements, and some actively discourage experiments. What should you do?

 **Robert Olinger** 4d ...  
Assistant Dean, Institutional Collaboration at Duke University - The Fuqu...

Ask the colleagues to teach you more about correlational measurements. Listen to them first, then ask what they have learned about incrementality. This is a psychological problem more than a preference, so use psychology to address it.

Like 4 · Reply 3

 **Rachel Fagen** 4d ...  
COO | Co-Founder | Partner | Advisor



Like · Reply

 **Kenneth Wilbur** Author 4d ...  
Professor of Marketing and Analytics at University of California, Sa...

**Robert Olinger** could you say more about what you mean by psychological problem?

Like · Reply

 **Robert Olinger** 4d ...  
Assistant Dean, Institutional Collaboration at Duke University - Th...

**Kenneth Wilbur:** I believe if experimentation is actively discouraged, this is due to aversion, a desire to feel comfortable, a desire to feel right, loss aversion, etc. The way you phrase the argument sounds like a lack of openness to listen--so my advice is you need to open up the colleagues--the best way to do that is by listening to them, understanding as best you can their expertise and approach--then engage their curiosity toward something new--the experimentation has to seem like it was their idea--so focus on engaging curiously with the colleagues, and when there is an openness ask questions related to the ideas you want included. Have them think about it... This is the way to shift preferences--persistent nudging.

Like 2 · Reply

 **Joel Person** 4d ...  
Research Scientist at Spotify | Causal Inference, Machine Learning and D...

You could demonstrate the value of experimentation for the business use case, for instance by showing via simulation that correlational evidence can lead to incorrect decisions (product launches, rollouts, etc) but that causal estimates from experiments get it right. You could even attach a relevant business metric (dollar value, engagement, reach, etc) ...see more

Like 4 · Reply

 **Dean Eckles** (edited) 4d ...  
scientist & statistician; faculty at MIT

One option: Consider looking for a new job. The number of firms with people who get A/B testing has expanded a lot. Fits with avoiding being the smartest person in the room.  
(Of course, there are other good options... but as a person in a junior role, this is one of the better ones.)

Like 5 · Reply

 **Brett Gordon** 4d ...  
Professor of Marketing at Kellogg School of Management | Amazon Sch...

Definitely bring in academics as outside consultants ;-)

Like 6 · Reply

 **Nirzar Bhaidkar** 4d ...  
Executive Paid Search @ GroupM | AI-Driven Marketing

Propose small scale pilot experiments to demonstrate the value of incremental measurement without significant resource investment.

Like 1 · Reply

 **Ayman Farahat** 10h ...  
Principal Scientist at Amazon



Like · Reply

 **Brad Shapiro** 3d ...  
Professor at The University of Chicago Booth School of Business

Generally agree with **Dean Eckles**. But depends on their reason for discouraging experimentation. If it is a genuine lack of understanding, I would try and be persuasive, show examples of how correlational assessments might lead you astray, etc. If it is an agency problem whereby they feel they need to mislead their management in order to keep their jobs, I'd say look for another job.

Like 1 · Reply

 **Michael Cohen** 2d ...  
Customer Centric Privacy Protecting Marketing AI

Change the way they are compensated or incentivized to be aligned with marginal economics of business aligned kpis.

Like 4 · Reply

# Takeaways

- Fundamental Problem of Causal Inference: We can't observe all data needed to optimize actions. This is a missing-data problem
- Approaches: Experiments, Quasi-experiments, Correlations, Ignore
- Experiments are the gold standard, but are costly and difficult to design, implement and act on
- It's likely that most large firms spend too much on ads



# Going deeper

- [What is Incrementality? And How Do We Measure it in 2024?](#)
- [Inferno: A Guide to Field Experiments in Online Display Advertising](#): Covers frequent problems in online advertising experiments
- [Inefficiencies in Digital Advertising Markets](#): Discusses digital RoAS estimation challenges and remedies
- [The Power of Experiments](#): Goes deep on digital test-and-learn considerations
- [New Developments in Experimental Design and Analysis \(2024\)](#) by Athey & Imbens
- [Mostly Harmless Econometrics](#): Covers quasi-experimental techniques

