On the Path to Causal Inference

Mark Farragher

TRIPTEASE

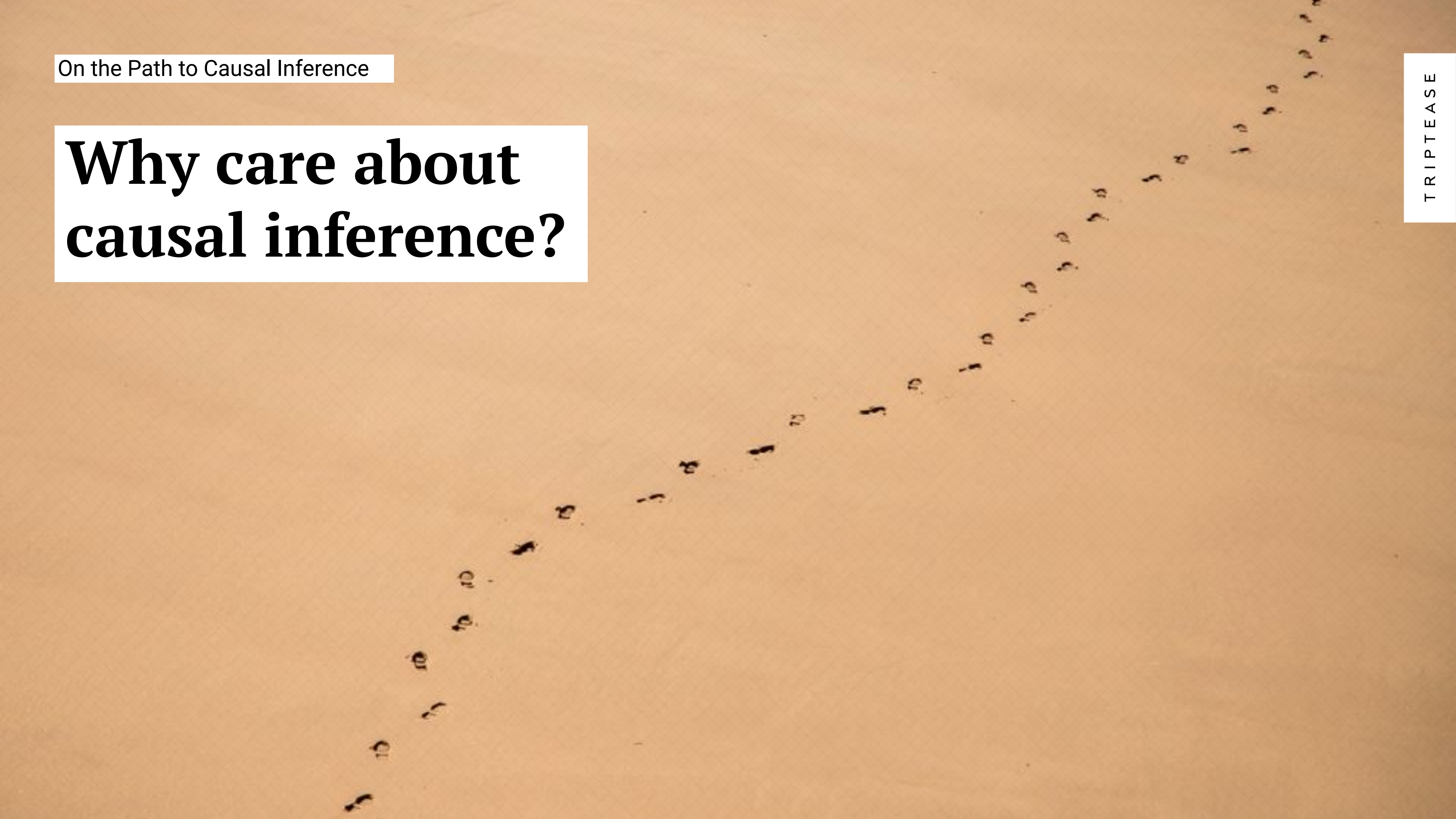
Agenda

- ➔ Why care about causal inference?
- ➔ The Ladder of Causation
- ➔ The causal diagram
- ➔ Types of biases that affect causal models

About me

- ➔ Background in quant economics
- ➔ …then moved into data science
- ➔ T R I P T E A S E (travel tech)
- ➔ Project: `appelpy`

Why care about causal inference?



TRIPTEASE

Causal questions are ubiquitous

- Social sciences
- Epidemiology
- Executives want ‘actionable insights’

TRIPTEASE

TRIPTEASE

How does passive smoking affect the mortality of non-smokers?

What is the effect of minimum wage laws on employment?

What are the causes of customer churn at a software company?

On the Path to Causal Inference

Ladder of causation

TRIPTEASE

Ladder of causation

- Counterfactual ('what if' I'd done X1 instead of X2?)
- Intervention (what happens to Y when I do X1?)
- Association (how do X1 and Y relate?)

TRIPTEASE

In the news today!

TRIPTEASE

JAMA Network Open

JAMA Network Open Enter Search Term

Original Investigation | Pediatrics

July 12, 2019

Association of Preterm Birth and Low Birth Weight With Romantic Partnership, Sexual Intercourse, and Parenthood in Adulthood

A Systematic Review and Meta-analysis

Marina Mendonça, PhD[1]; Ayten Bilgin, PhD[1,2]; Dieter Wolke, PhD[1,3]

» Author Affiliations | Article Information

JAMA Netw Open. 2019;2(7):e196961. doi:10.1001/jamanetworkopen.2019.6961

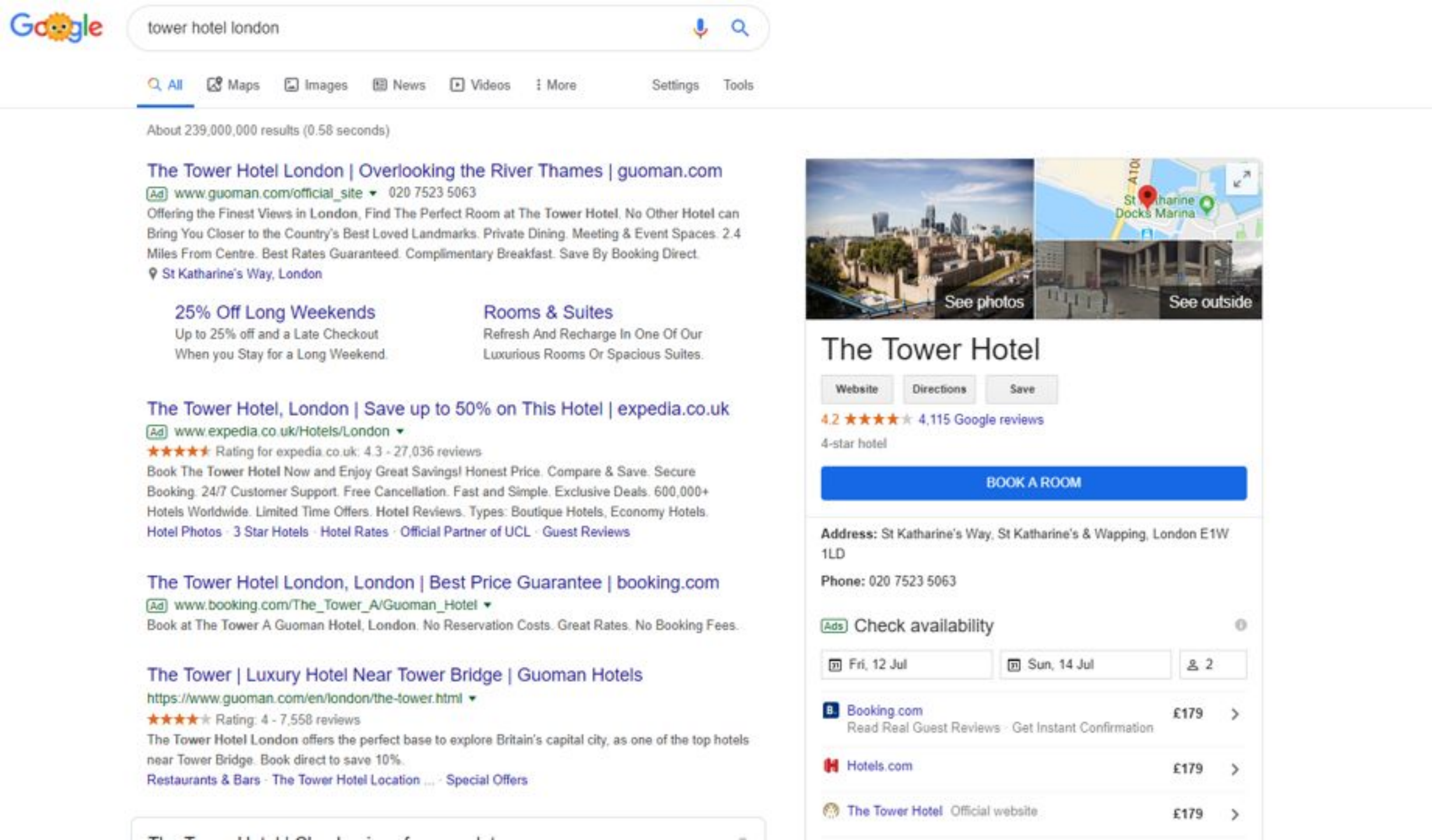
- → Paper talks about ‘association’
- → Press release uses language of causation (subconsciously?)

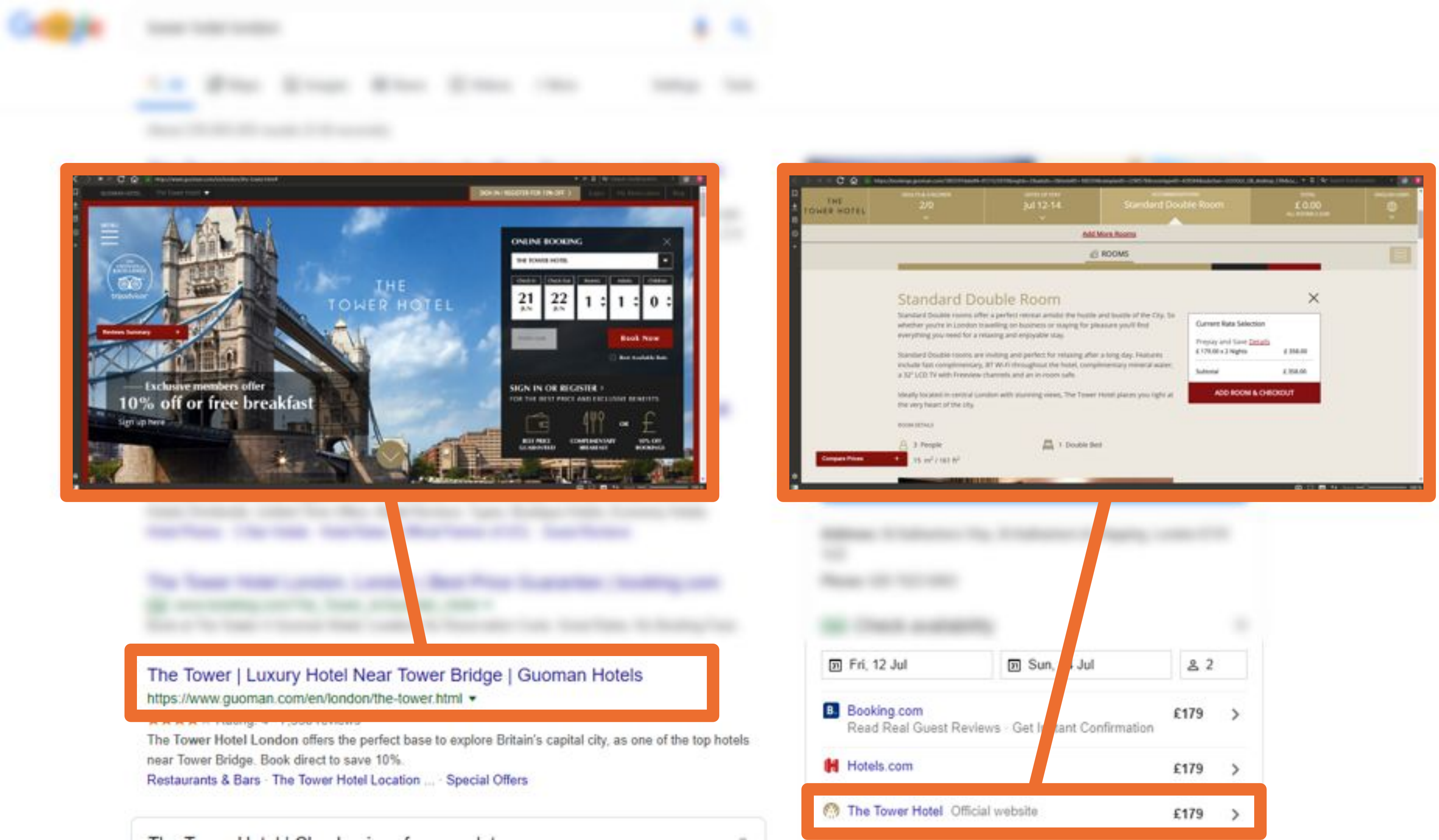
Causal diagram

On the Path to Causal Inference

A night in the Tower?

TRIPTEASE


tower hotel london
All
Maps
Images
News
Videos
More
Settings
Tools
About 239,000,000 results (0.58 seconds)
The Tower Hotel London | Overlooking the River Thames | guoman.com
Ad www.guoman.com/official_site 020 7523 5063
Offering the Finest Views in London, Find The Perfect Room at The Tower Hotel. No Other Hotel can Bring You Closer to the Country's Best Loved Landmarks. Private Dining. Meeting & Event Spaces. 2.4 Miles From Centre. Best Rates Guaranteed. Complimentary Breakfast. Save By Booking Direct.
St Katharine's Way, London
25% Off Long Weekends
Up to 25% off and a Late Checkout When you Stay for a Long Weekend.
Rooms & Suites
Refresh And Recharge In One Of Our Luxurious Rooms Or Spacious Suites.
The Tower Hotel, London | Save up to 50% on This Hotel | expedia.co.uk
Ad www.expedia.co.uk/Hotels/London
Rating for expedia.co.uk: 4.3 - 27,036 reviews
Book The Tower Hotel Now and Enjoy Great Savings! Honest Price. Compare & Save. Secure Booking. 24/7 Customer Support. Free Cancellation. Fast and Simple. Exclusive Deals. 600,000+ Hotels Worldwide. Limited Time Offers. Hotel Reviews. Types: Boutique Hotels, Economy Hotels.
Hotel Photos · 3 Star Hotels · Hotel Rates · Official Partner of UCL · Guest Reviews
The Tower Hotel London, London | Best Price Guarantee | booking.com
Ad www.booking.com/The_Tower_A/Guoman_Hotel
Book at The Tower A Guoman Hotel, London. No Reservation Costs. Great Rates. No Booking Fees.
The Tower | Luxury Hotel Near Tower Bridge | Guoman Hotels
https://www.guoman.com/en/london/the-tower.html
Rating: 4 - 7,558 reviews
The Tower Hotel London offers the perfect base to explore Britain's capital city, as one of the top hotels near Tower Bridge. Book direct to save 10%.
Restaurants & Bars · The Tower Hotel Location ... · Special Offers
See photos
See outside
The Tower Hotel
Website
Directions
Save
4.2 4,115 Google reviews
4-star hotel
BOOK A ROOM
Address: St Katharine's Way, St Katharine's & Wapping, London E1W 1LD
Phone: 020 7523 5063
Ads Check availability
Fri, 12 Jul
Sun, 14 Jul
2
Booking.com
Read Real Guest Reviews · Get Instant Confirmation
£179
Hotels.com
£179
The Tower Hotel Official website
£179



ONLINE BOOKING
21
22
Book Now
THE TOWER HOTEL
Exclusive members offer
10% off or free breakfast
SIGN IN OR REGISTER
THE TOWER HOTEL
Standard Double Room
ADD ROOM & CHECKOUT
The Tower | Luxury Hotel Near Tower Bridge | Guoman Hotels
https://www.guoman.com/en/london/the-tower.html
The Tower Hotel London offers the perfect base to explore Britain's capital city, as one of the top hotels near Tower Bridge. Book direct to save 10%.
Restaurants & Bars · The Tower Hotel Location ... · Special Offers
Fri, 12 Jul
2
Booking.com
Read Real Guest Reviews
£179
Hotels.com
£179
The Tower Hotel Official website
£179

TRIPTEASE

Hotel website funnel (example)

TRIPTEASE

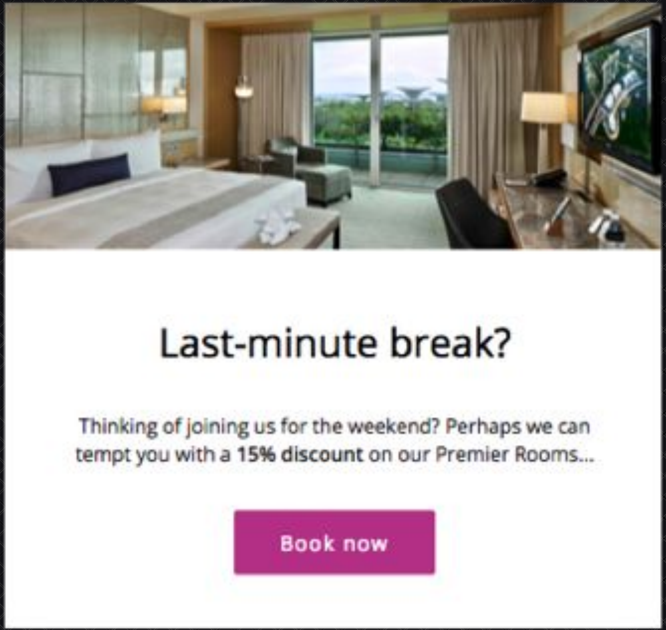
What could affect conversion rate?

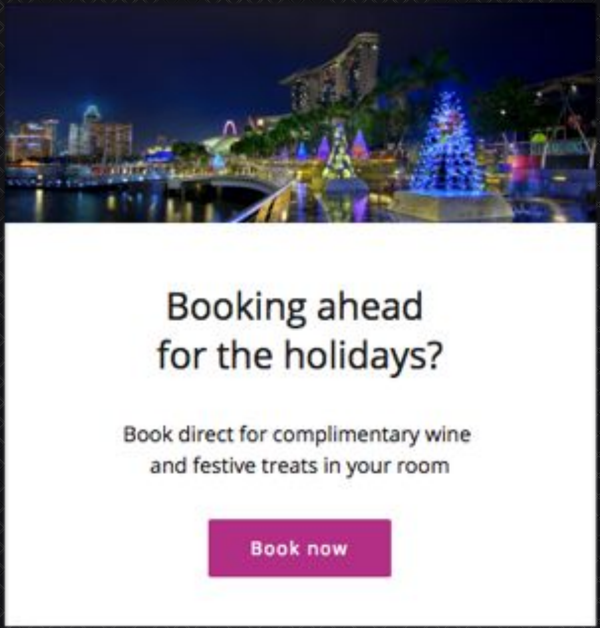
- ➔ Seasonality (hour of day, day of week, week of year)
- ➔ Traffic source (organic, paid, etc.)
- ➔ Device
- ➔ Price factors
- ➔ …and much more!

Consider two theories

- **Personalised messages** on hotel websites have a positive effect on conversion rate.
- **Searchers from Google** have a higher likelihood of booking a hotel room.

Messages

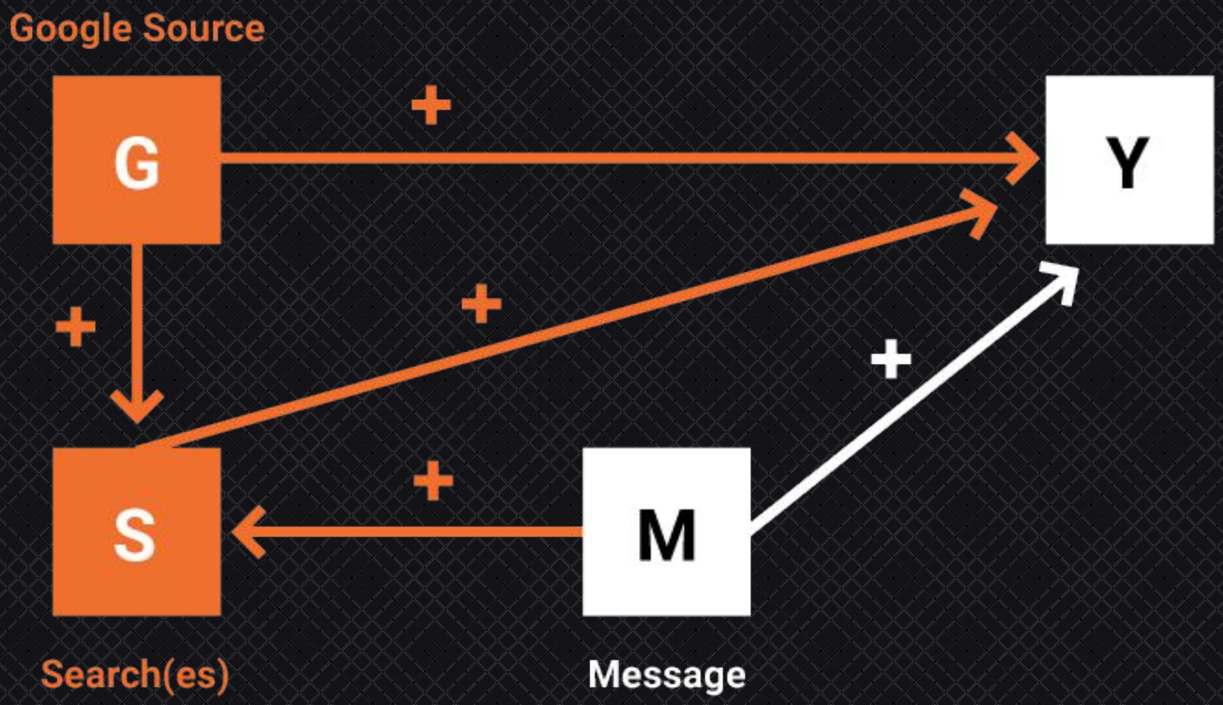

Last-minute break?
Thinking of joining us for the weekend? Perhaps we can tempt you with a **15% discount** on our Premier Rooms...
Book now


Booking ahead for the holidays?
Book direct for complimentary wine and festive treats in your room
Book now

Experimentation

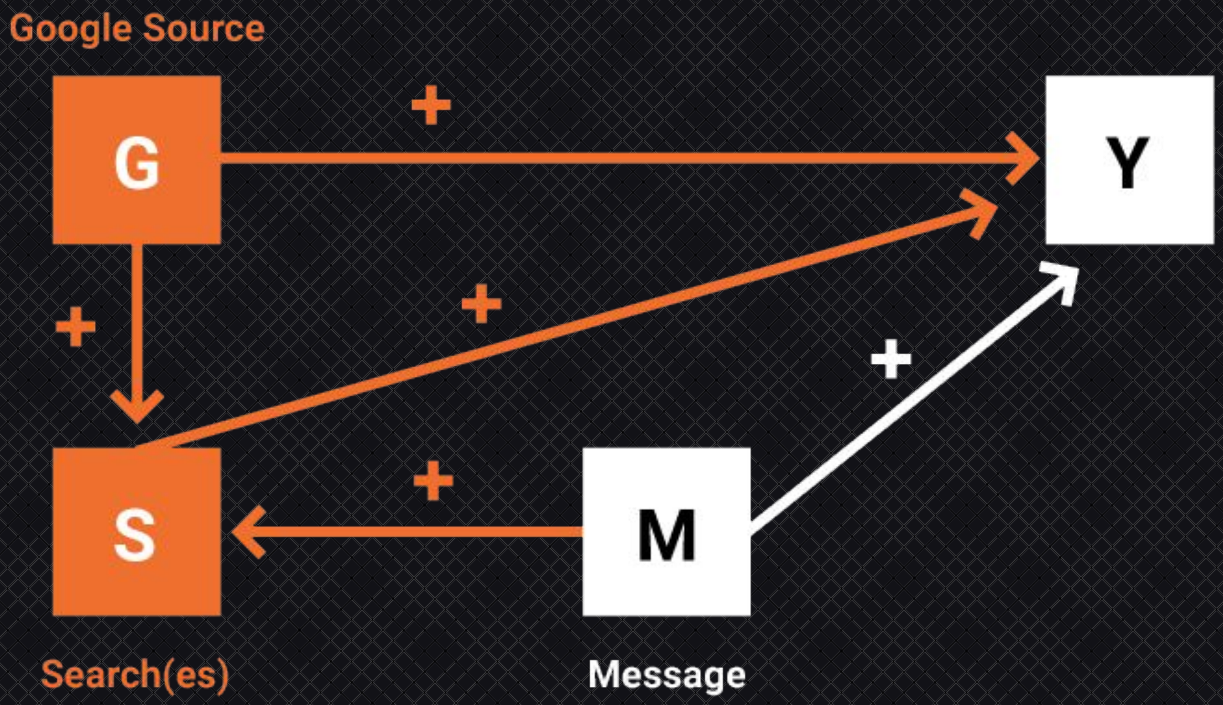
- ➔ Can the treatment be randomly assigned?
- ➔ **Personalised messages**: YES (but the scale of data can be an issue).
- ➔ **Searchers from Google**: NO.

Diagram: effect of *messages*



TRIPTEASE

Diagram: effect of *messages*



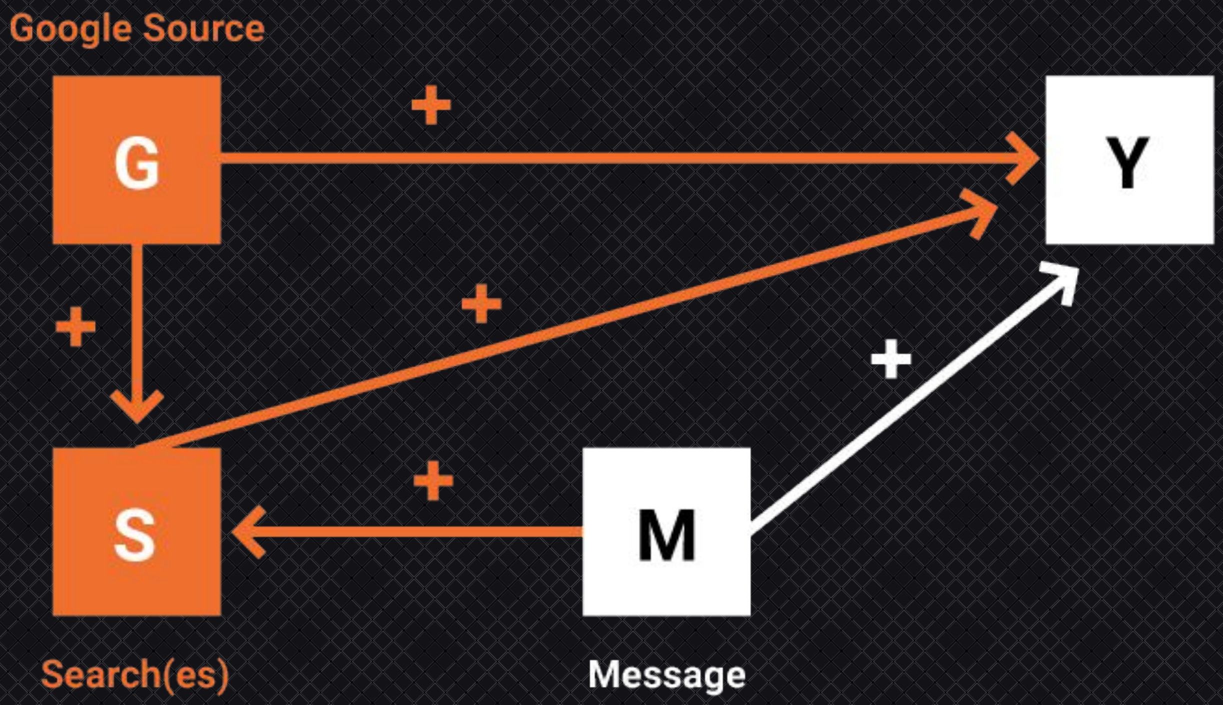
M → S → Y

No collider

Block with {S}

TRIPTEASE

Diagram: effect of *messages*



$M \longrightarrow S \longleftarrow G \longrightarrow Y$

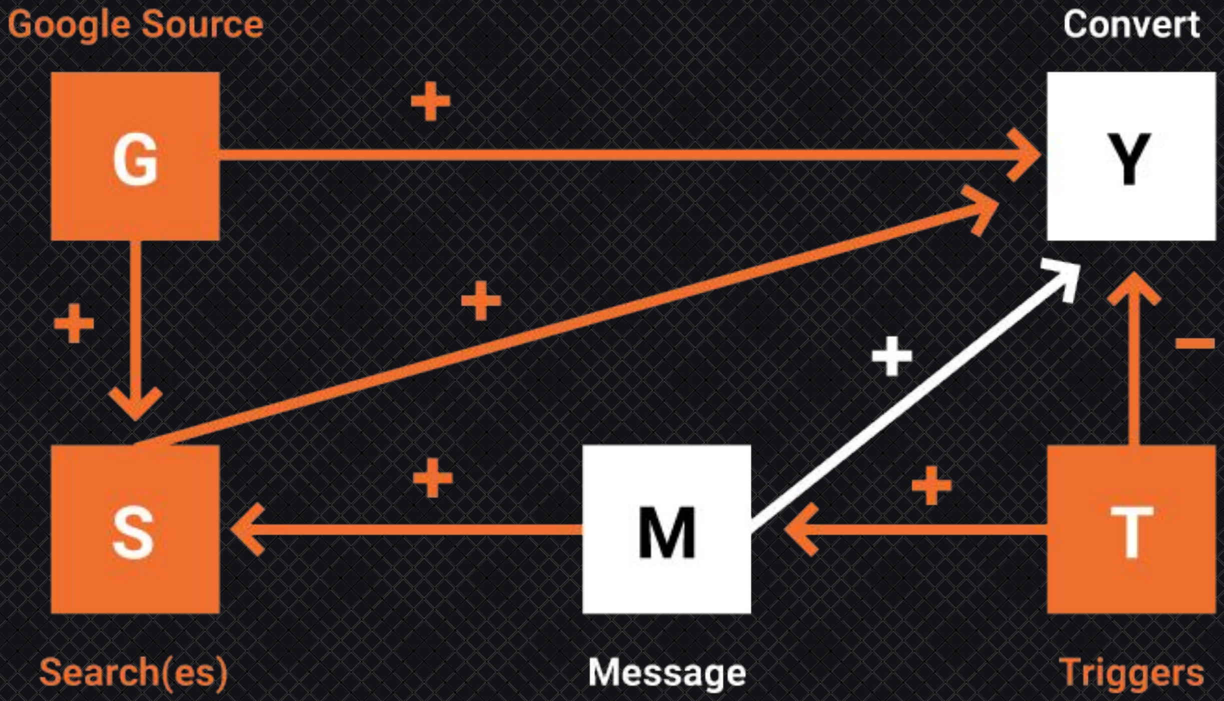
S is collider here!

$M \longrightarrow |S| \longleftarrow G \longrightarrow Y$

Block with {G}

TRIPTEASE

Diagram: enter message triggers



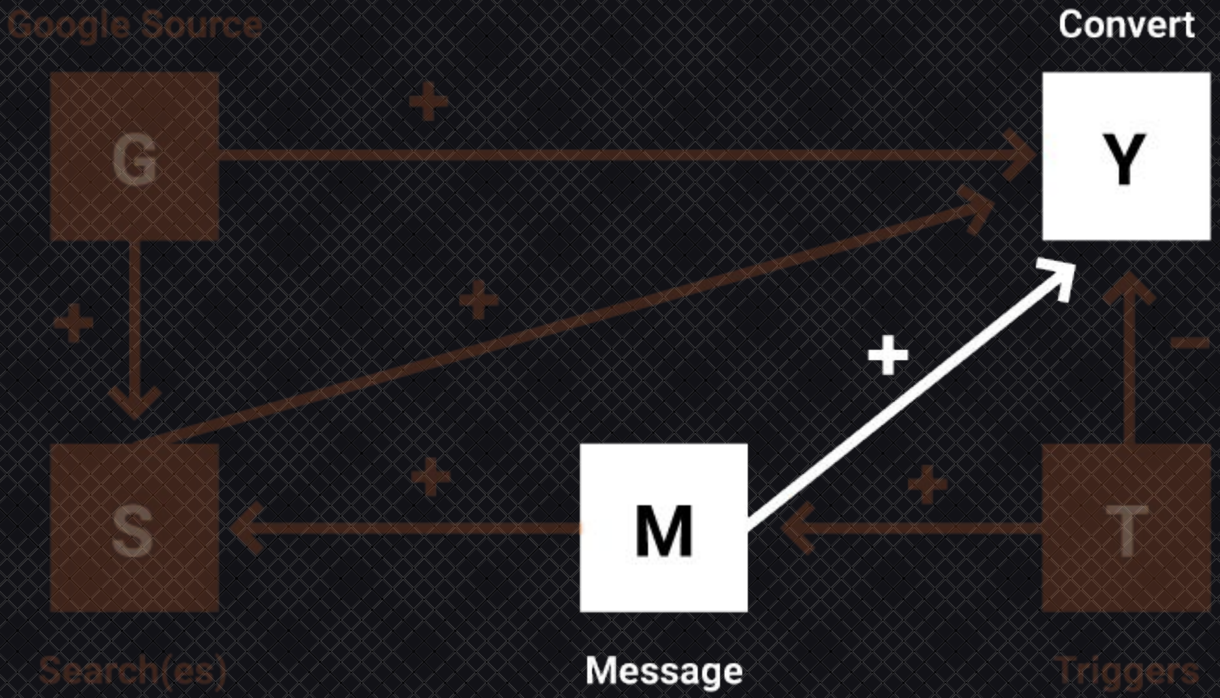
$M \leftarrow T \rightarrow Y$

No collider

Block with {T}

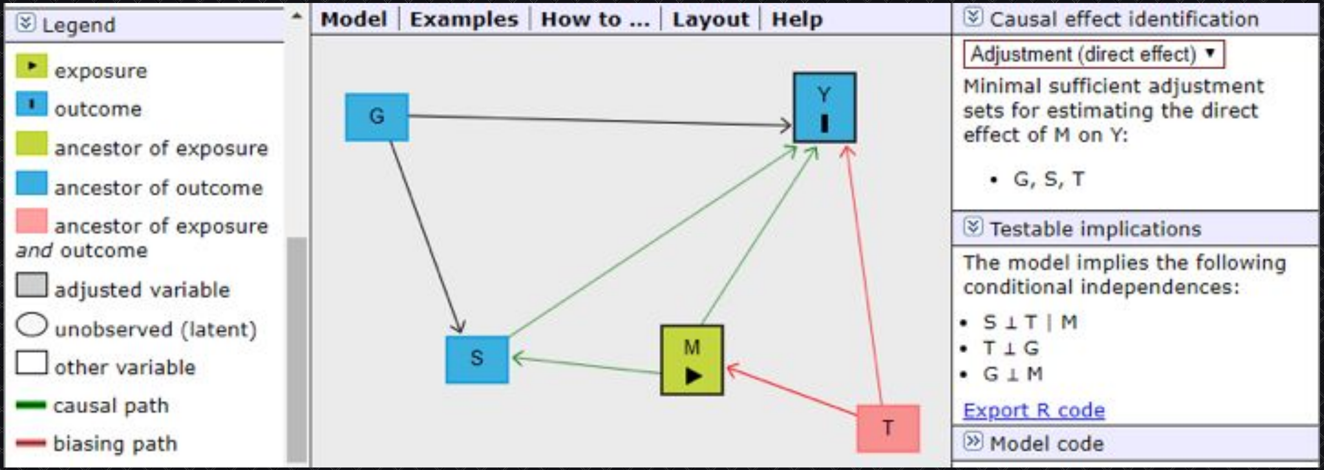
TRIPTEASE

Diagram: A/B test



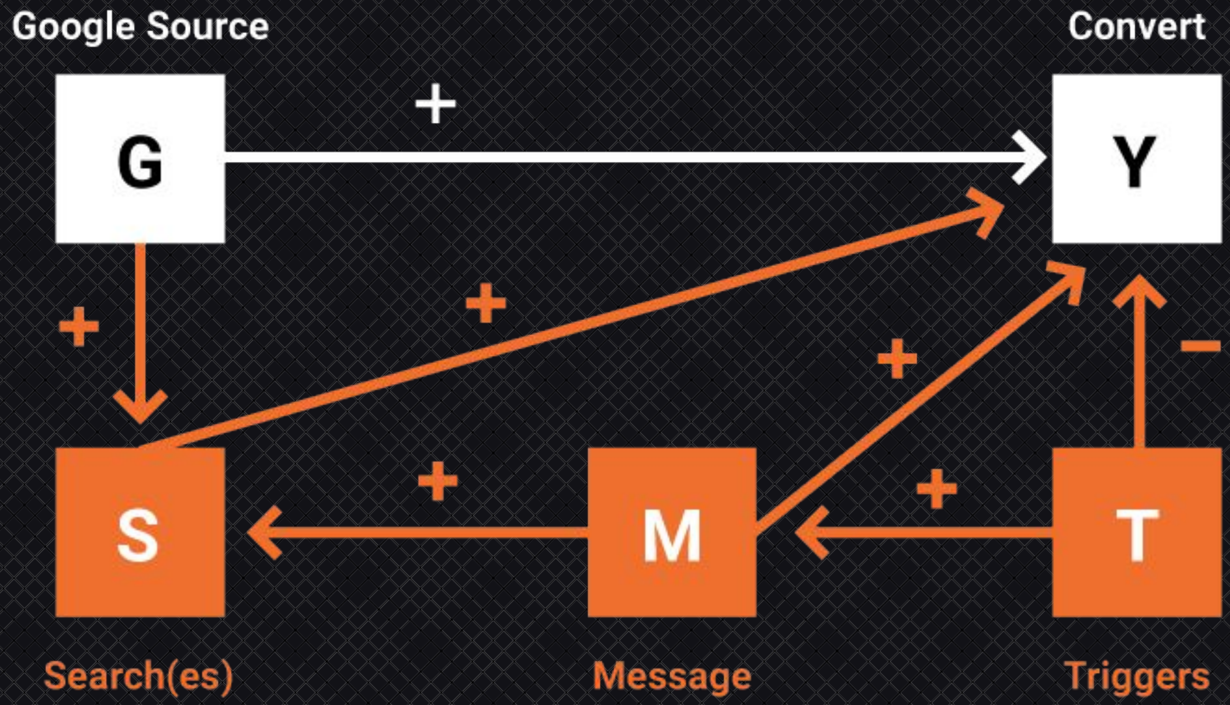
TRIPTEASE

Tooling



- → Dagitty (above)
- → Python packages, e.g.: `causality; causalgraphicalmodels; dowhy`

Diagram: effect of *Google source*

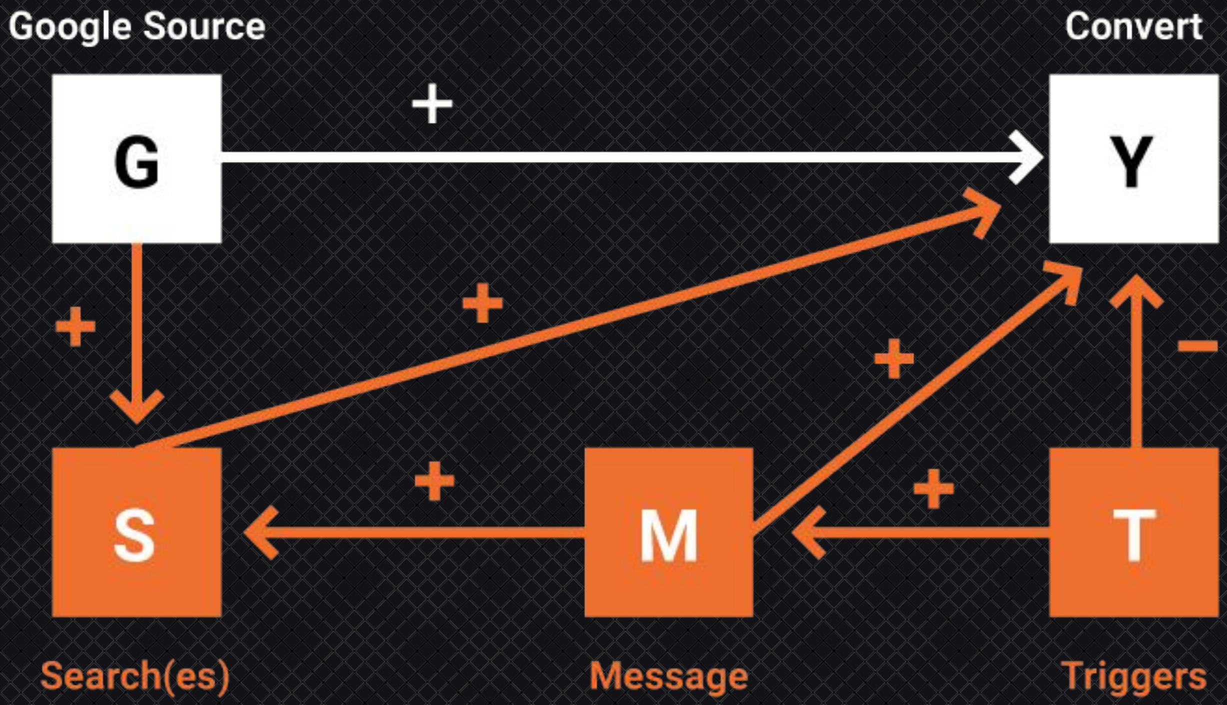


Confounders: M & S

Control for {M, S}

TRIPTEASE

Diagram: effect of *Google source*



$G \rightarrow S \leftarrow M \leftarrow T \rightarrow Y$

S is collider here!

$G \rightarrow |S| \leftarrow M \leftarrow T \rightarrow Y$

Block with {M}

Why causal diagrams are great

- Assumptions for a model are explicit
- A toolkit for identifying appropriate control variables
- Diagrams can be refined iteratively

TRIPTEASE

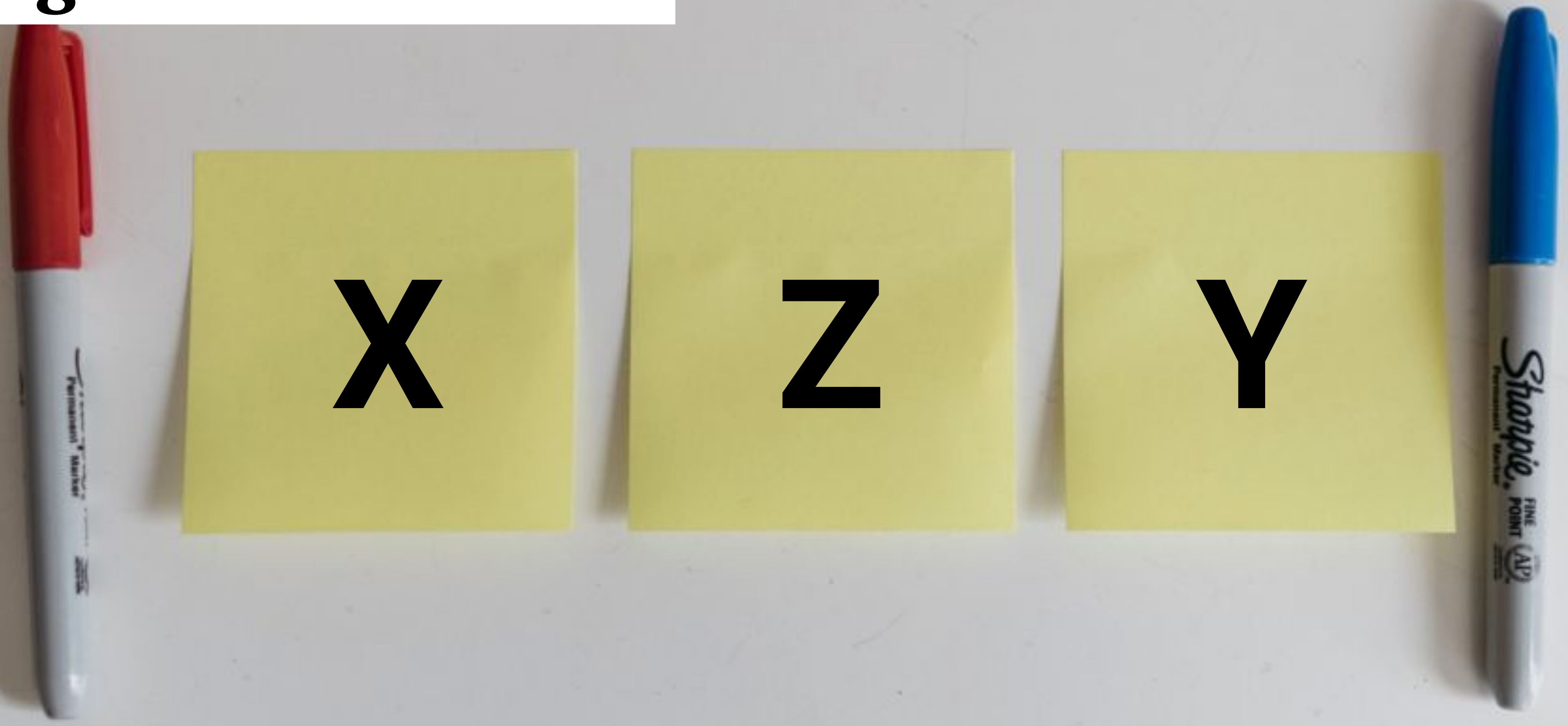
On the Path to Causal Inference

Types of bias

TRIPTEASE

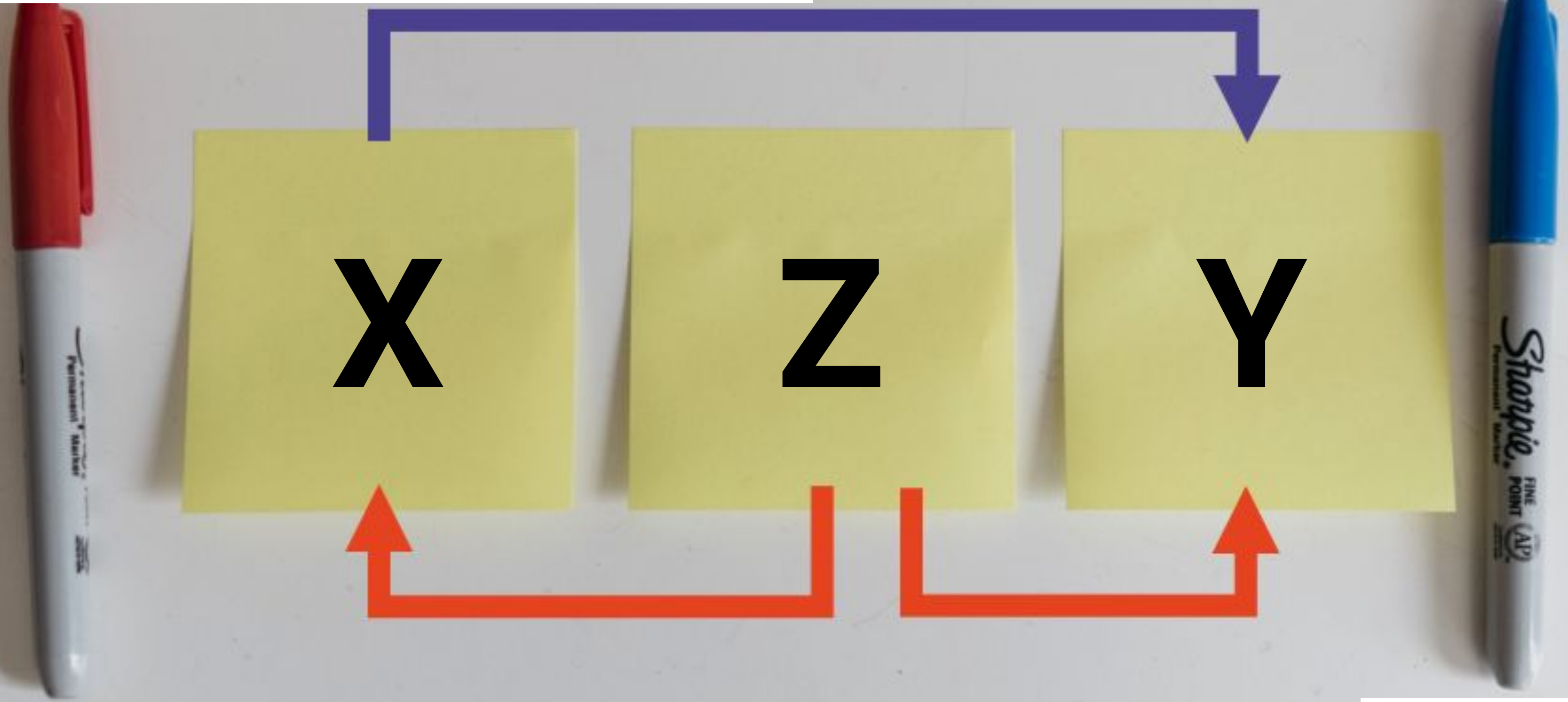
On the Path to Causal Inference

Regress Y on X & Z



TRIPTEASE

Z: confounding variable



TRIPTEASE

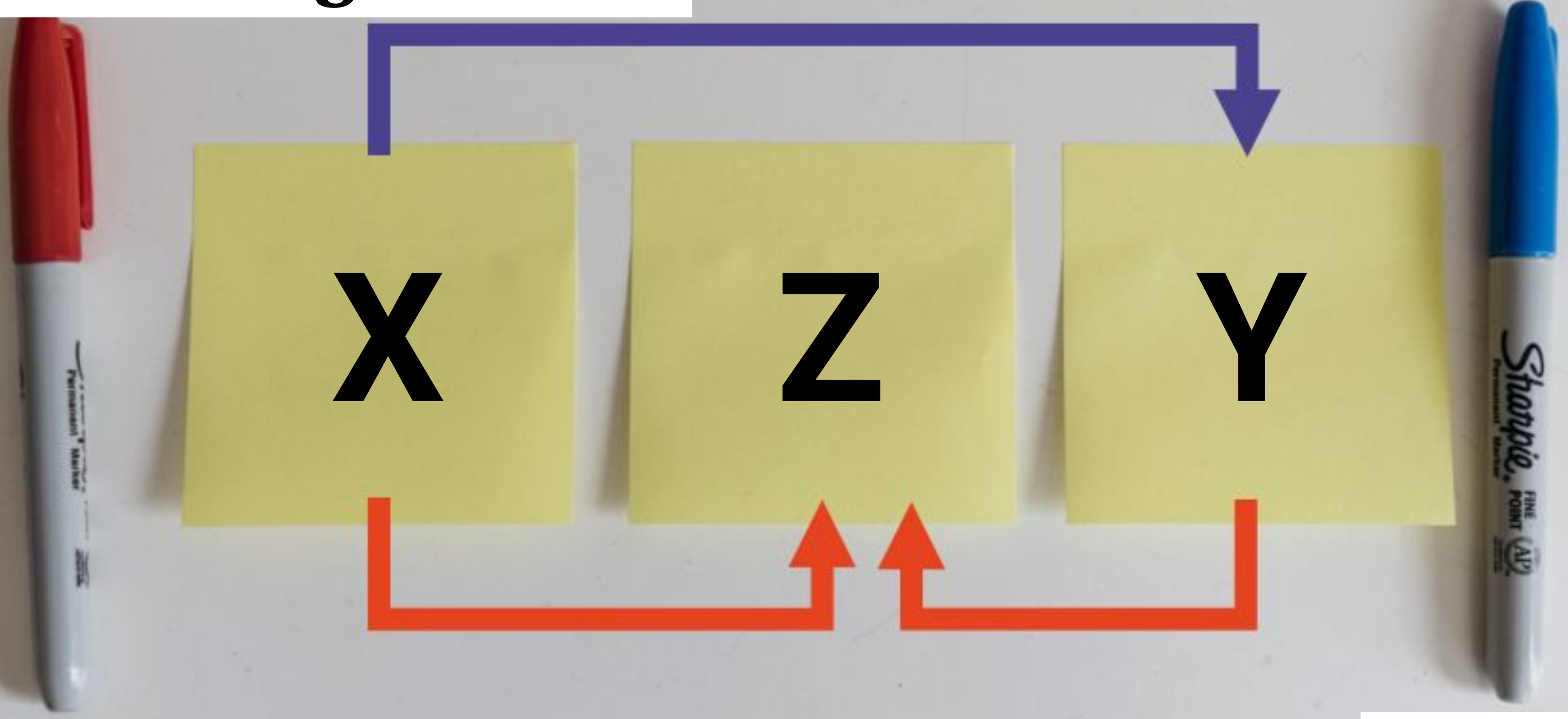
$X \leftarrow Z \rightarrow Y$

Confounders

- Control for confounders – they help your model’s estimates by:
  - Increasing their precision;
  - Reducing their bias.

TRIPTEASE

Z: colliding variable



$X \rightarrow Z \leftarrow Y$

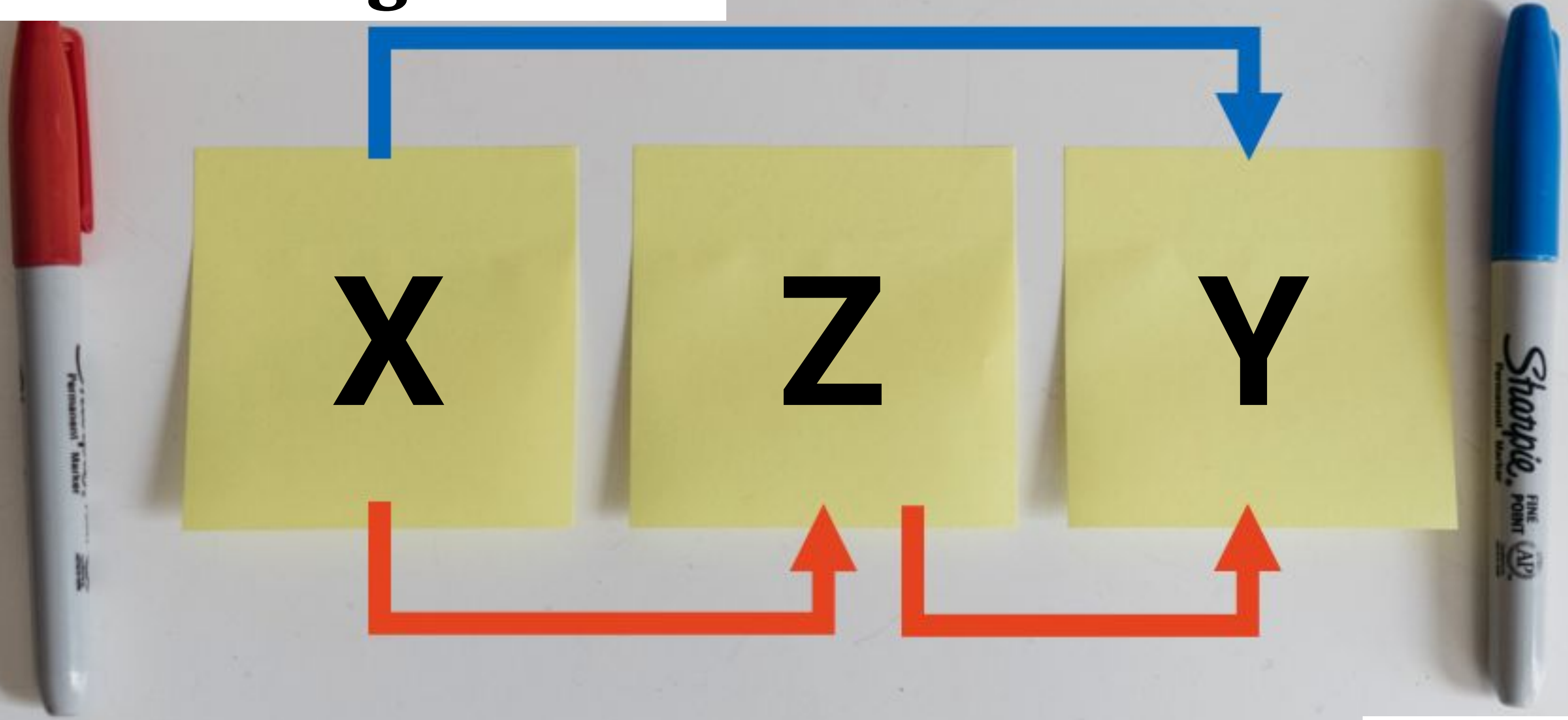
TRIPTEASE

Colliders

- Collider automatically blocks backdoor path
- Common effect of X & Y – beware
- Sometimes a variable can be a collider & a confounder on different paths!

TRIPTEASE

Z: mediating variable



$X \rightarrow Z \rightarrow Y$

Mediators

- A mediator (M) introduces an indirect effect of X on Y.
- If we care only about the total effect of X on Y, then there's no need to control for M.

Summary

- Causal questions are ubiquitous
- Causal effects can be estimated without randomised experiments (fortunately)
- The causal diagram is useful for modelling
- Control for confounders & beware of colliders
- Tools can assist with the choice of controls

Further resources

- Pearl: The Book of Why
- Angrist & Pischke: Mastering ‘Metrics
- Python libraries: `statsmodels`; `linearmodels`; `appelpy`

mfarragher

TRIPTEASE

TRIPTEASE