Causal Inference: Going beyond Associations

Chair of Business Analytics

What is causality?

- **Question**: How do we distinguish cause and effects?

Chocolate consumption vs number of Nobel prizes

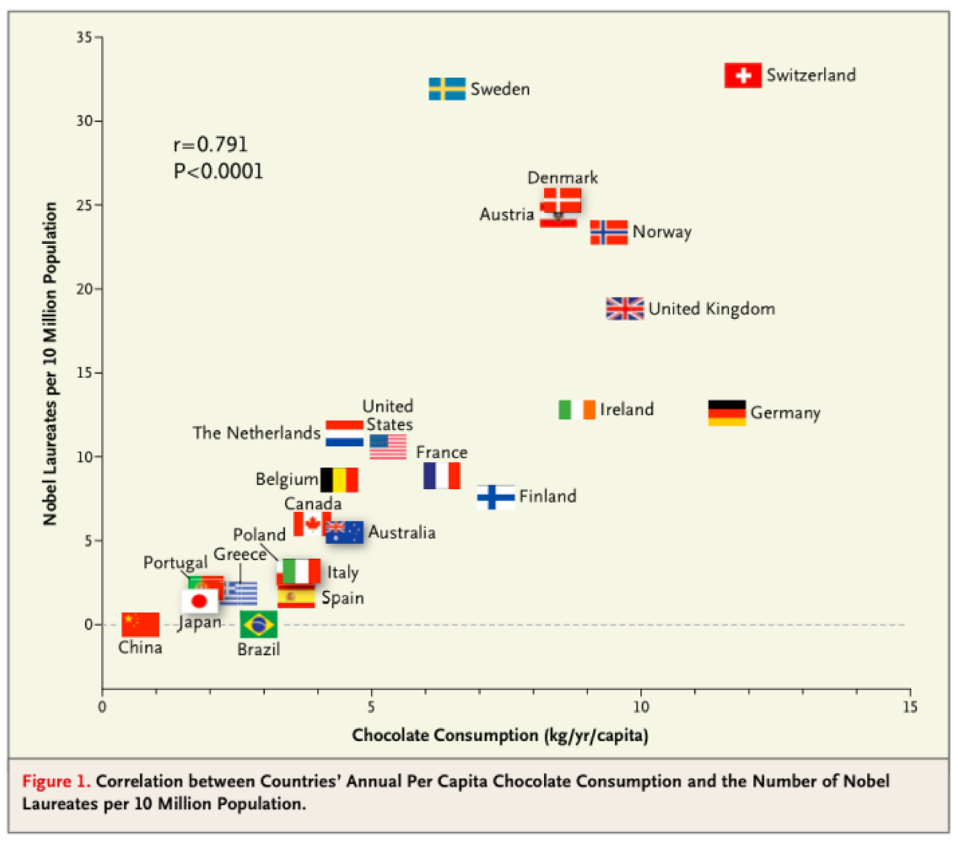


Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.

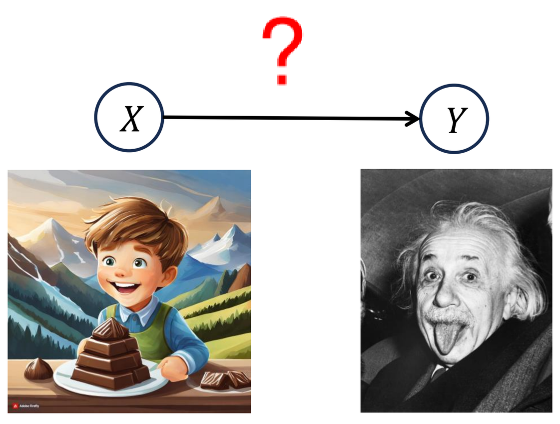
F. H. Messerli: Chocolate Consumption, Cognitive Function, and Nobel Laureates, N Engl J Med 2012

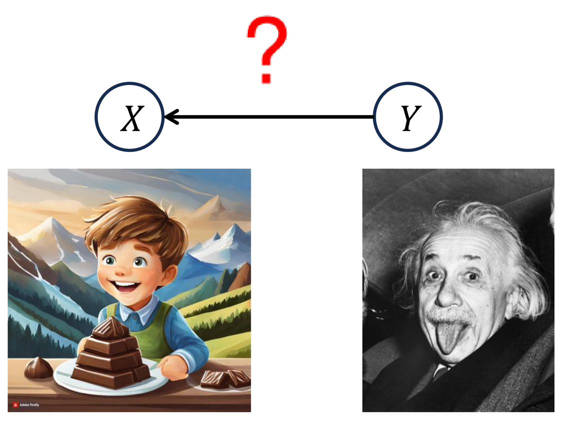
CORRELATION≠CAUSATION

- **Problem**: *Probability theory has an associational, rather than a causal nature*

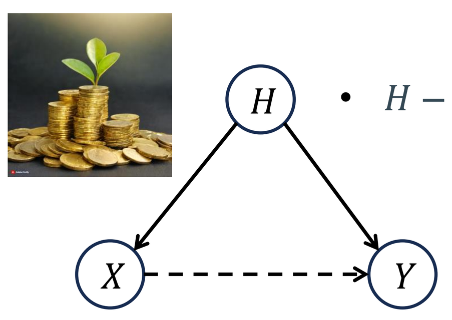
POSSIBLE THEORIES

- X – chocolate consumption
- Y – obtaining Nobel prize





OR



- H – "wealth"

Well... you may have your own theories… :)

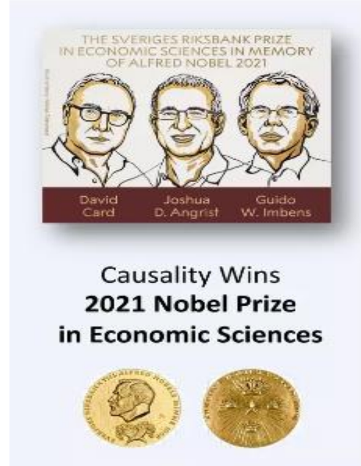
This operation, called intervention, changes the distribution of the variable regardless of its previous value or other circumstances (covariates). In our example it means forcing people to eat chocolate regardless of everything else.

- ***Randomized Controlled trial (RCT)***
  - Form two groups at random.
  - Force one group to eat lots of chocolate. Observe: $P(Y|do(X=1))$.
  - Ban the other group from eating chocolate at all. Observe: $P(Y|do(X=0))$.
  - Compare: $P(Y|do(X=1))$ and $P(Y|do(X=0))$.

Why causality?

Trending applications of causality:

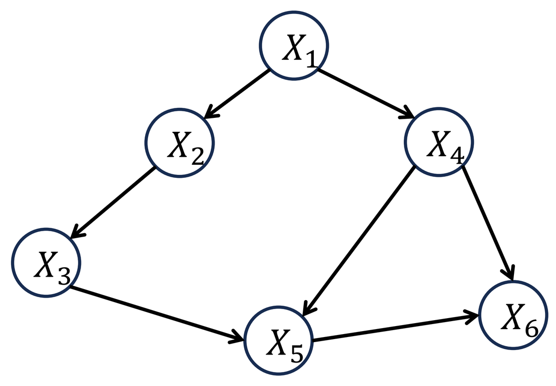
- Artificial intelligence and computer science
- Biology and genetics
- Epidemiology, public health and medicine
- Management and business
- Economics, education, psychology, social sciences



Fundamental question of causality

1. Causal Discovery

Data + Assumptions →



Observations and/or interventions

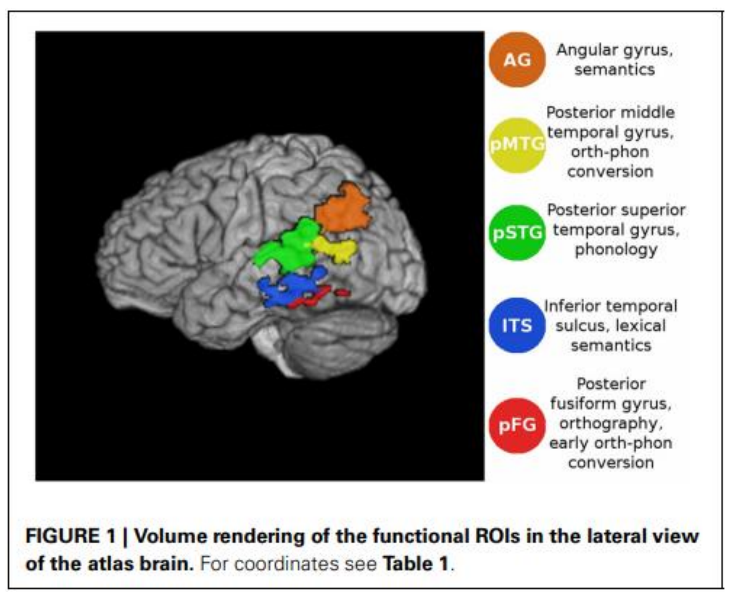


FIGURE 1 | Volume rendering of the functional ROIs in the lateral view of the atlas brain. For coordinates see **Table 1**.

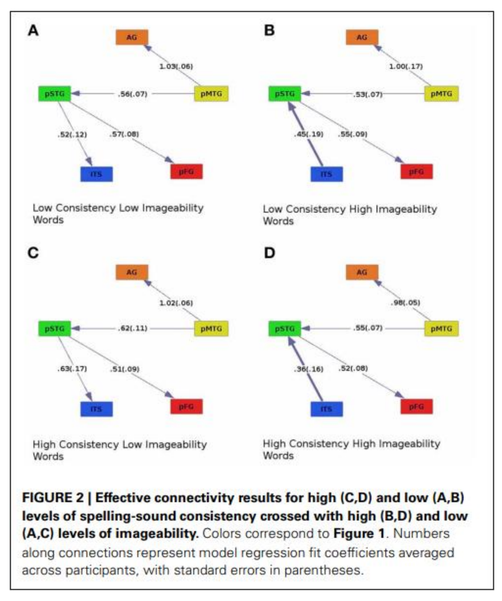


FIGURE 2 | Effective connectivity results for high (C,D) and low (A,B) levels of spelling-sound consistency crossed with high (B,D) and low (A,C) levels of imageability. Colors correspond to **Figure 1**. Numbers along connections represent model regression fit coefficients averaged across participants, with standard errors in parentheses.

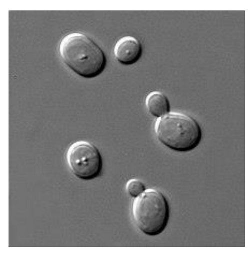
Boukrina & Graves, 2013

2. Causal Effect Estimation\Identification

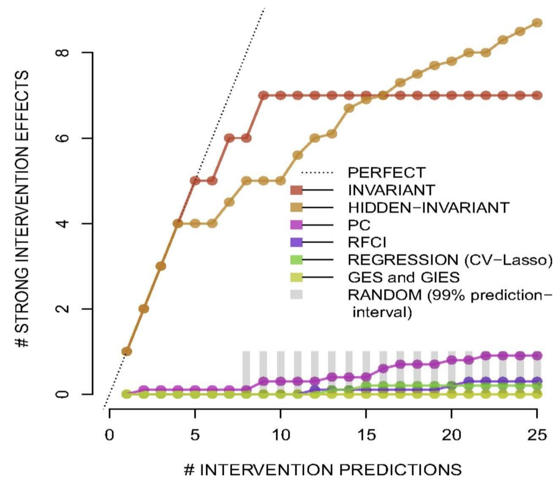
Quantify\identify how interventions impact different groups within your data

Gene deletion experiments in yeast

- Y – expression of a target gene
- X – gene expressions of all other 6169 genes



- **Observation:** 160 observational data points from wild-type individuals and 1,479 interventional data arising from single gene deletions
- **Goal:** estimate\predict the effects of unseen gene deletion on all other genes

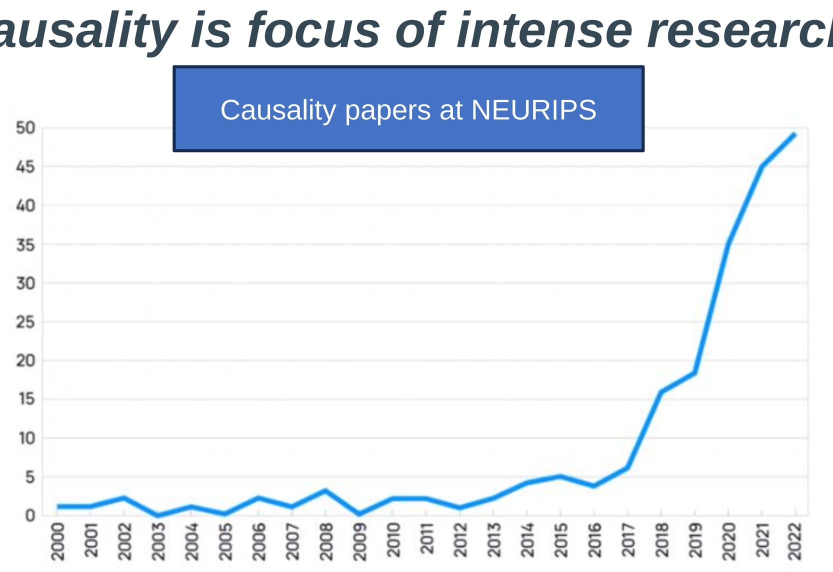


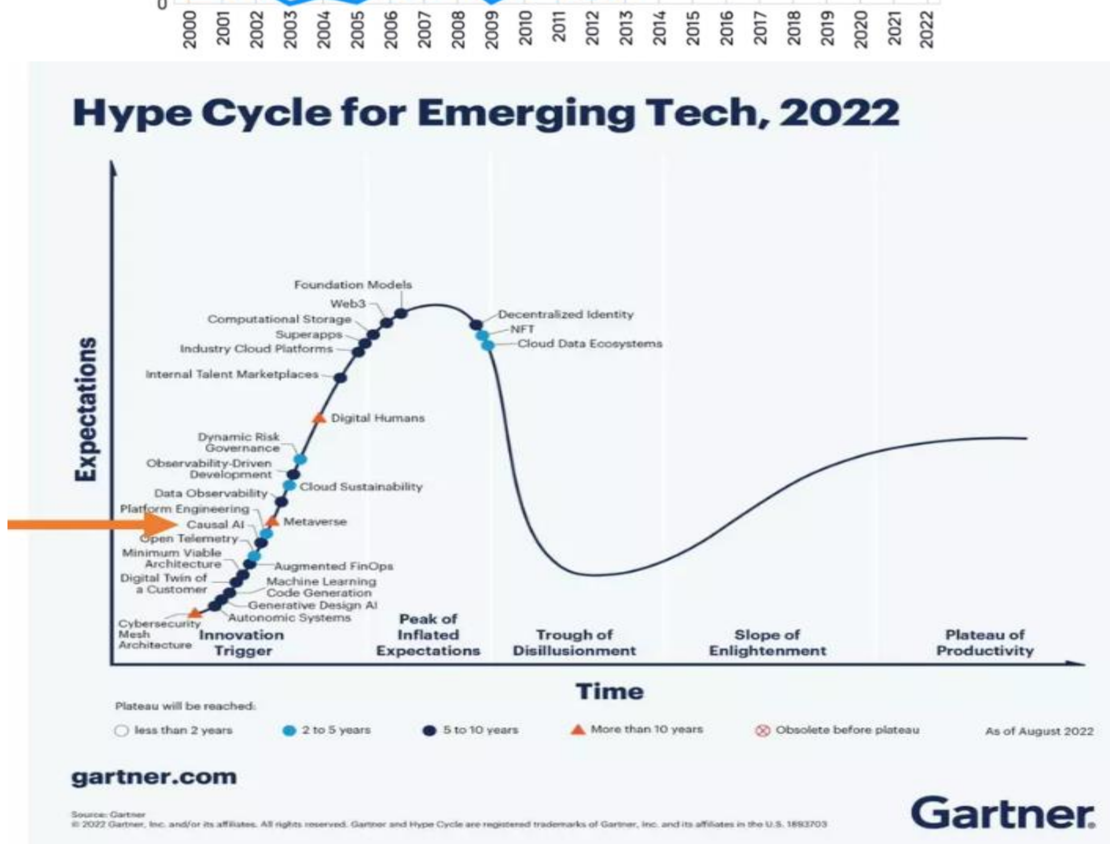
Meinshausen, Nicolai, et al., *2016*

Causality: growing trends in AI

- Current AI models: black box correlation-based algorithms
- As a result AI decisions lack of explainability
- Humans beat AI systems in part because they understand causal/effect relations

Causality is focus of intense research!

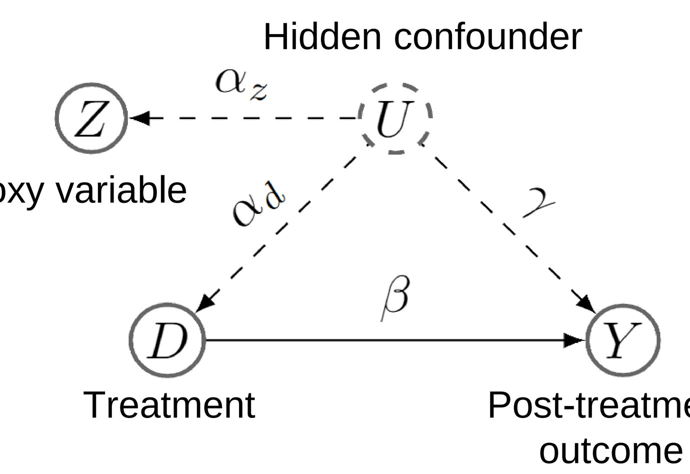




Selected samples of work in our group:

- **A Cross-Moment Approach for Causal Effect Estimation (NeurIPS 2023)**

Propose a statistical solution how to estimate the causal effect of the treatment on the post-treatment outcome



- **A Unified Experiment Design Approach for Cyclic and Acyclic Causal Models (JMLR 24.354 2023)**

Propose an experimental design algorithm for learning a causal graph. This framework is the first unified algorithm for experimental design for cyclic and acyclic graphs.

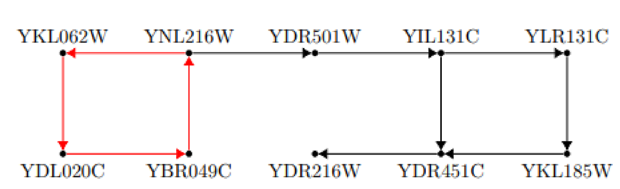


Figure 2: A sub-network of Yeast's gene regulatory network that contains a directed cycle of length 4 (the edges in red).

References:

- Mokhtarian, Ehsan, et al. "A unified experiment design approach for cyclic and acyclic causal models." *Journal of Machine Learning Research* 24.354 (2023): 1-31.
- Kivva, Yaroslav, Saber Salehkaleybar, and Negar Kiyavash. "A Cross-Moment Approach for Causal Effect Estimation." *Advances in Neural Information Processing Systems* 36 (2024).
- Meinshausen, Nicolai, et al. "Methods for causal inference from gene perturbation experiments and validation." *Proceedings of the National Academy of Sciences* 113.27 (2016): 7361-7368.
- Boukrina, Olga, and William W. Graves. "Neural networks underlying contributions from semantics in reading aloud." *Frontiers in human neuroscience* 7 (2013): 518.
- Messerli, Franz H. "Chocolate consumption, cognitive function, and Nobel laureates." *N Engl J Med* 367.16 (2012): 1562-1564.

Find out more of our work here: