UNIVERSITY OF MARYLAND

MSML610: Advanced Machine Learning

Causal Inference

Instructor: Dr. GP Saggese - `gsaggese@umd.edu`

References:

- Easy:
  - Hurwitz, Thompson: Causal Artificial Intelligence: The Next Step in Effective Business AI, 2024
- Medium / Difficult
  - AIMA
  - Facuce

SCIENCE ACADEMY

- ***Causal AI***
  - Why Causal AI?
  - The Ladder of Causation
  - Correlation vs Causation Models
- Causal Networks
- Business Processes Around Data Modeling

- Causal AI
  - ***Why Causal AI?***
  - The Ladder of Causation
  - Correlation vs Causation Models
- Causal Networks
- Business Processes Around Data Modeling

Big Data and Traditional AI

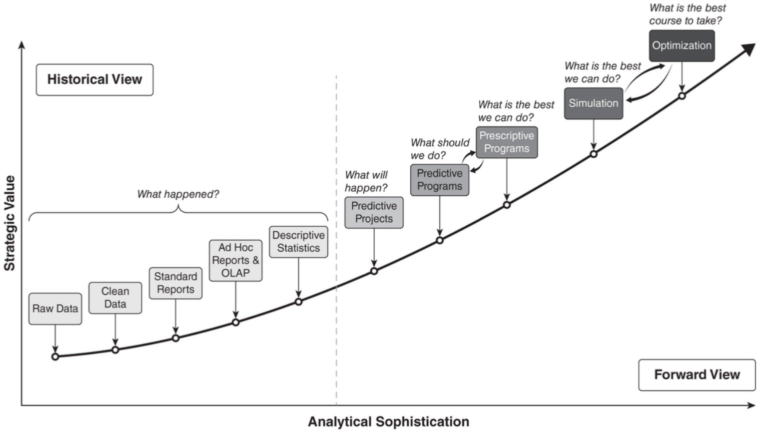
- For the past 10 years, **focus of analytics** on:
  - Organize and analyze massive amount of data
  - Data analytics (dashboards, models, reports)
  - Run machine learning on data
- Problems with **traditional AI**
  - Predicts based on observed correlations
  - Can't explain why an outcome occurred
- **AI in decision making**
  - Understand impact of decisions
  - E.g., *"What happens if a product price is reduced by 10%?"*
    - Will more customers buy?
    - If revenue decreases, what to do?
    - Why are customers leaving? Quality issue? Emerging competitor?

What Are Data Analytics?

- **Collections of data**
  - Aggregated, organized data sets for analysis
  - E.g., customer purchase histories in a CRM system
- **Dashboards**
  - Visual displays of key metrics for insights
  - E.g., dashboard showing quarterly revenue, expenses
- **Descriptive statistics**
  - Summary metrics: mean, median, mode, standard deviation
  - E.g., average sales per quarter to understand trends
- **Historical reports**
  - Examination of past performance
  - E.g., monthly sales reports for past fiscal year
- **Models**
  - Statistical representations to forecast, explain phenomena
  - E.g., predictive model to anticipate customer churn based on behavioral data

Data Analytics Sophistication

Business Question	Methodology
What happened?	Descriptive statistics
What will happen?	Predictive models
What should we do?	Prescriptive programs
What's the best we can do?	Simulation + optimization



Explainability

- **Regulators** require that if you are making decisions using ML / AI, you should be able to defend the results of your analysis
  - E.g., decide who to hire, how to set up a policy
- Organizations can:
  - Be **fined** by regulatory authorities
  - Face **backlash** from customers and activists
- E.g., neural networks are "black boxes"
  - Lack of explainability
    - Humans can't understand how inputs are combined into a conclusion
    - Cannot explain to shareholders why certain decisions were made
  - Bias
    - E.g., using age, race, sex as a feature can introduce bias
- **Explainable AI** allow users to:
  - Comprehend
  - Explain
  - Trust the results by the machine

Correlation is Not Causation!

- **Correlation** is a statistical method for understanding relationships between data
  - Pros
    - Use past outcomes to predict future outcomes by finding patterns and anomalies
  - Cons
    - Doesn't explain the cause
    - Variables may move together due to coincidence or a hidden factor
- **Causation** explains how changing one variable influences the other
  - Cannot be concluded from correlation alone
- **Data does not understand causes and effects**
  - Only humans can identify variables and relationships based on context
  - Without causation, you can't make intelligent decisions

Causal AI

- **Understands the why**
  - Determines cause-and-effect between variables
  - E.g., whether a marketing campaign increased sales
- **Identify interventions**
  - Identifies variables and interventions to change outcomes
  - E.g., which lifestyle changes reduce blood pressure
- **Predicting counterfactuals**
  - Hypothesizes outcomes under different circumstances
  - E.g., student grades if they attended a different school
- **Avoiding bias**
  - Traditional AI biased by training data and ignored variables
  - Ensure fairness by accounting for confounding variables
- **Improving decision-making**
  - Provides understanding of relationships for better decisions
  - E.g., improving supply chain by understanding logistic impact

Causal AI vs Traditional AI

- *"The next revolution of data science is the science of interpreting reality, not of summarizing data"* (Judea Pearl, 2021)
- Current AI uses correlation to:
  - Analyze data
  - Identify patterns
  - Make predictions
- Models depend on data quality
  - Biased or unclean data $\implies$ poor model

- Causal AI
  - Why Causal AI?
  - ***The Ladder of Causation***
  - Correlation vs Causation Models
- Causal Networks
- Business Processes Around Data Modeling

The Ladder of Causation

- Pearl provided a 3-layer framework for understanding causality

Level	Symbol	Activity	Typical Questions
1. Association	$\Pr(Y\|X)$	Observing	What is?
2. Intervention	$\Pr(Y\|do(X), Z)$	Intervening	What if?
3. Counterfactuals	$\Pr(Y_X\|x', y')$	Imagining	Why?

Rung 1: Association

- **Question**: *"How would seeing X change our belief in Y?"*
- **Symbol**: $\Pr(Y|X)$
  - Bayesian update
- **Activity**
  - It is just "passive observation"
  - Determine if two things are related
  - Traditional AI and ML is based on this
- **Example**
  - *"The tree has green leaves during spring"*
  - *"What does a symptom tell you about a disease?"*
  - *"What does a survey tell you about the election results?"*

Rung 2: Intervention

- **Question**: *"What happens to Y if you do X?"*
- **Symbol**: $\Pr(Y|do(X), Z)$
- **Activity**
  - Understand the impact of an action
  - E.g., *"tree has green leaves"* vs *"spring makes tree leaves turn green"*
  - Association is just about observations
  - Interventions involve "doing something" and need a causal model
- **Example**
  - *"Why did the headache go away?"*
    - "Because the pain reliever" or "Because you ate food after skipping lunch"
  - *"If you take aspirin, will your headache be cured?"*
  - *"What if you ban sodas?"*

SCIENCE ACADEMY

Rung 3: Counterfactuals

- **Question**: *"Was X that caused Y?"*
- **Symbol**: $\Pr(Y_X|x', y')$
- **Activity**:
  - Imagine what will happen if facts were different
  - Predicting an outcome is the highest form of reasoning
    - It requires to understand relationships between cause and effect
- **Example**
  - Scientific experiments: *"What if we give a child an adult dose of a drug?"*
  - Litigation: *"What would the jury conclude?"*
  - Marketing: *"Why did my marketing campaign fail to generate sales?"*

- Causal AI
  - Why Causal AI?
  - The Ladder of Causation
  - ***Correlation vs Causation Models***
- Causal Networks
- Business Processes Around Data Modeling

Correlation vs Causation Model

- **Correlation** = identify how variables are related to each other
- **Causality** = determine whether one variable causes another variable
  - Both:
    - Accept inputs and transform them to compute predictions
    - Identify how variables are related to each other
  - Correlation-based AI works well when there is abundant historical and observational data
  - Causal-based AI first creates a business-focused model before integrating data

Correlation-Based Model Process

- **Correlation-based AI** is “data first”
  - The more data collected the better
- **Modeling process**
  - Acquire data
  - Integrate and clean data
  - Exploratory data analysis (EDA)
  - Feature engineering
  - Build and test models
  - Deploy models in production
- **Many AI projects fail because**
  - Cultural and organizational issues
  - Models are opaque and lack explainability
  - Spurious correlations
  - Missing articulating “what’s the goal of doing ML?”

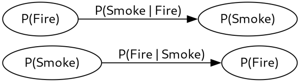
Causation-based Model Process

- **Causal AI** is "model first"
  - Understand business question before ingest and transform the data
- **Modeling process**
  - What is the intended outcome?
  - What is the proposed intervention?
  - What are the confounding factors?
  - What are the effecting factors?
  - Create a model graph or diagram
  - Data acquisition
  - ...

- Causal AI
- ***Causal Networks***
    - Variables
    - Intervention
    - Type of Variables in Causal AI
    - Paths
- Business Processes Around Data Modeling

(Non-Causal) Bayesian Networks

- **Bayesian networks** represent a joint distribution function
  - The direction of the arrow represent **conditional dependence** (not causality)
    - $A \rightarrow B$ requires to estimate $\Pr(A|B)$
  - There are many possible **edges** and **node ordering** for the same Bayesian network
- E.g., a Bayesian network with *Fire* and *Smoke*, which are dependent
  - $\mathit{Fire} \rightarrow \mathit{Smoke}$
    - Need $\Pr(\mathit{Fire})$ and $\Pr(\mathit{Smoke}|\mathit{Fire})$ to compute $\Pr(\mathit{Fire}, \mathit{Smoke})$
  - $\mathit{Smoke} \rightarrow \mathit{Fire}$
    - Need $\Pr(\mathit{Smoke})$ and $\Pr(\mathit{Fire}|\mathit{Smoke})$
  - Networks are equivalent and convey the same information
  - Different difficulties to estimate



- There is an **asymmetry in nature**
  - Extinguishing fire stops smoke
  - Clearing smoke doesn't affect fire

Causal (Bayesian) Networks

- **Causal networks** are Bayesian networks forbidding **non-causal edges**
- Use judgment based on nature instead of just statistics
  - E.g., from *"Are random variables Smoke and Fire correlated?"* to *"What causes what, Smoke or Fire?"*
- **"Dependency in nature" is like assignment in programming**
  - E.g., nature assigns *Smoke* based on *Fire*:
    - ✅ $Smoke := f(Fire)$
    - ❌ $Fire := f(Smoke)$
  - Structural equations describe "assignment mechanism" in causal graphs

$$X_i := f(X_j) \iff X_j \to X_i$$

SCIENCE ACADEMY

Causal DAG

- **Causal DAG**
  - *Directed*: Arrows show direction of cause $\rightarrow$ effect
  - *Acyclic*: No feedback loops
    - Causal relationships assume a temporal order: cause happens before effect
    - A cycle would imply a variable is both a cause and effect of itself (paradox)
- **Benefits**
  - DAGs encode *causal* rather than *associative* links
  - Enables reasoning about interventions and counterfactuals
  - Supports explainable AI models
  - Stability to conditional probability estimation
- **Limitations**
  - Requires domain knowledge to specify structure
  - Assumes all relevant variables are included (e.g., no hidden confounders)

Causal Edges are Stable

- **Causal edge** $X \rightarrow Y$ shows direct causal influence of X on Y, holding other variables constant
  - Captures how manipulating X changes Y, not just their covariance
- Causal edges reflect **stable relationship**
  - Mechanistic stability
    - Causal relationships show system function, not just behavior in one dataset
    - E.g., "Temperature $\rightarrow$ ice melting rate" holds true in Alaska and Arizona
  - Invariance under interventions
    - If X causes Y, intervening on X affects Y consistently, despite confounders or context changes
  - Easier estimation through causal modeling
    - Identifying causal direction focuses estimation on effect size (e.g., regression of Y on X under intervention)
  - Reduced sensitivity to sampling and omitted variables
    - Correlations may change with confounder addition or removal
    - True causal edge persists, stable across model specifications
- **Example**: study *Exercise* $\rightarrow$ *Health*:
  - Correlation may differ in young or elderly populations
  - Causal effect remains stable, as physiological mechanism doesn't change

Structural Causal Model

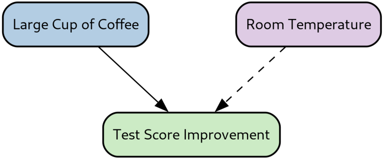
- A **Structural Causal Model** (SCM) translates a causal DAG into mathematical equations
  - DAGs show structure (variables and arrows)
  - SCMs use equations to define how variables interact
- **Structure of SCMs**
  - Variables $X_1, X_2, ..., X_n$ represent quantities in the system
  - Equations model each variable as a function of its direct causes
  - Formally, X_i is modeled as:

$$X_i = f_i(Parents(X_i), \varepsilon_i)$$

  where:
    - $Parents(X_i)$ are direct causes of X_i
    - ε_i is an exogenous (external, unobserved) noise term
- **Properties**
  - Explain causal relationships between variables
  - Provide a foundation for causal reasoning and simulation
  - Describe how the world works, not just variable correlations

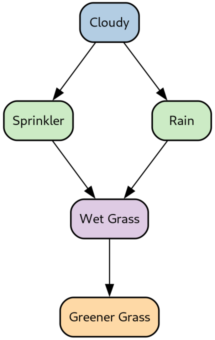
Structural Causal Model: Example

- **Explanatory variables**
  - You can manipulate or observe when changes are applied
  - E.g., *"does a large cup of coffee before an exam help with a test?"*
- **Outcome variables**
  - Result of the action
  - E.g., *"by how much did the score test improve?"*



- **Unobserved variables**
  - Not seen or more difficult to account
  - E.g., *"temperature of the room makes students sleepy and less alert"*

Structural Causal Model: Sprinkler Example



- Structural equations for this model:

$$\begin{cases} C := f_C(\varepsilon_C) \\ R := f_R(C, \varepsilon_R) \\ S := f_S(C, \varepsilon_S) \\ W := f_W(R, S, \varepsilon_W) \\ G := f_G(W, \varepsilon_G) \end{cases}$$

- Unmodeled variables ε_x represent error terms
    - E.g., ε_W is another source of wetness besides *Sprinkler* and *Rain* (e.g., *MorningDew*)
- Assume unmodeled variables are exogenous, independent, with a certain distribution (prior)

- Express joint distribution of five variables as a product of conditional distributions using causal DAG topology:

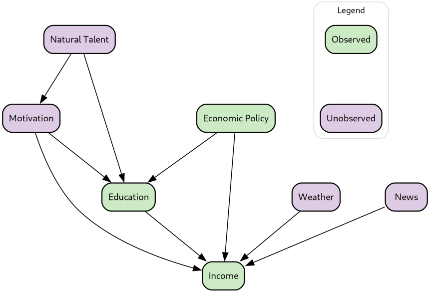
$$\Pr(C, R, S, W, G) = \Pr(C)\Pr(R|C)\Pr(S|C)\Pr(W|R, S)\Pr(G|W)$$

SCIENCE ACADEMY

- Causal AI
- Causal Networks
    - ***Variables***
    - Intervention
    - Type of Variables in Causal AI
    - Paths
- Business Processes Around Data Modeling

Observed Vs. Unobserved Variables

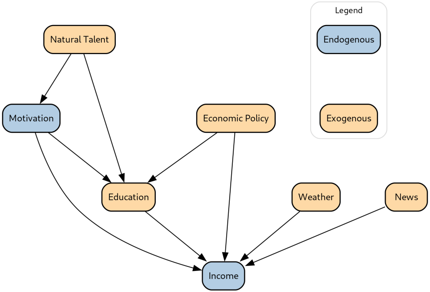
- **Observed variables**
  - Aka "measurable" or "visible"
  - Variables directly measured or collected in a dataset
  - E.g.,
    - Education
    - Income
    - Blood pressure
    - Product price



- **Unobserved variables**
  - Aka "latent" or "hidden"
  - Exist but not measured or included in data
  - E.g.,
    - Natural talent
    - Motivation
    - Company culture
  - Ignoring unobserved variables distorts causal relationships
    - Observed: *IceCreamSales* and *DrowningRates*
    - Unobserved: *Temperature*
    - Misleading conclusion: *IceCream* causes *Drowning*

Endogenous Vs. Exogenous Variables

- **Endogenous variables**
  - Values determined *within* the model
    - Dependent on other variables in the system
  - Represent system's internal behavior and outcomes
  - E.g.,
    - Motivation
    - Income



- **Exogenous variables**
  - Originate *outside* the system being modeled
    - Not caused by other variables in the model
  - Represent background conditions or external shocks
  - E.g.,
    - Natural talent
    - Economic policy
    - Weather
    - News

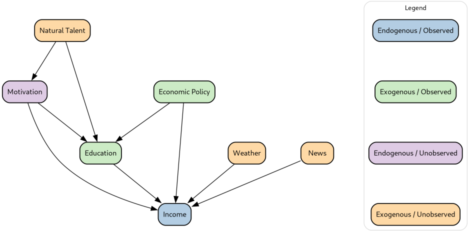
SCIENCE ACADEMY

Endo / Exogenous, Observed / Unobserved Vars

- In **Structural Causal Models**

$$X_i = f_i(Parents(X_i), \varepsilon_i)$$

  where:
  - X_i: endogenous
  - ε_i: exogenous noise
- **Typically**
  - *Endogenous variables*: focus for prediction and intervention
  - *Exogenous variables*: capture randomness or unknown external factors

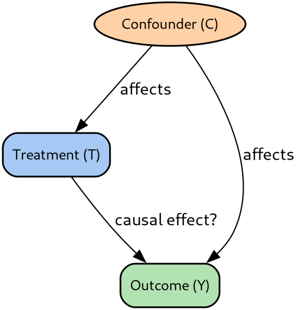


Variable Type	Observability	Example
Endogenous	Observed	Income
Exogenous	Observed	Education
Endogenous	Unobserved	Motivation
Exogenous	Unobserved	Natural Talent

- Causal AI
- Causal Networks
  - Variables
  - ***Intervention***
  - Type of Variables in Causal AI
  - Paths
- Business Processes Around Data Modeling

Estimating Causal Effects

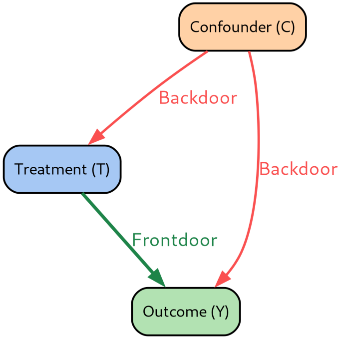
- **Goal**: Determine the causal effect of a treatment variable (aka intervention) T on an outcome Y
- **Example:**
  - T = *"takes drug"*
  - Y = *"recovers"*
  - C = *"overall health"*
- Healthier people may take medicine and recover faster $\Longrightarrow$ correlation without causation
- In **observational data**
  - Confounding variable C affects both treatment T and outcome Y
  - C creates *spurious correlation* between T and Y



- **Problem**
  - There is a "backdoor path" *Treatment* $\leftarrow$ *Confounder* $\rightarrow$ *Outcome*

Frontdoor and Backdoor Paths: Intuition

- A **backdoor path** is any path from T to Y starting with an arrow into T
  - E.g., $T \leftarrow C \rightarrow Y$
  - Interpretation:
    - C is a common cause of T and Y, confounding their relationship
    - Controlling (conditioning) for C blocks the backdoor path, identifying the causal effect of T on Y



- A **frontdoor path** goes directly or indirectly from T to Y through mediators, following causal flow
  - E.g., $T \rightarrow Y$
  - Interpretation:
    - Direct causal path of interest
    - No mediators, so front-door path is direct causal effect of T on Y

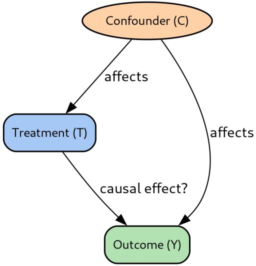
SCIENCE ACADEMY

Randomized Controlled Trials (RCTs)

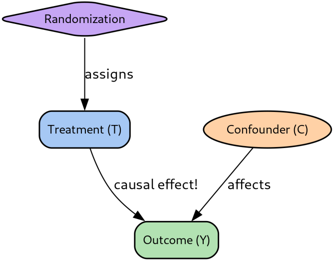
- **Randomized Controlled Trial** is an experimental study to assess causal effect of an intervention or treatment
  - Determine whether an intervention causes an effect, not just associated with it
  - Eliminate selection bias and confounding variables through randomization
- **Key Components**
  - *Randomization*: ensures groups are statistically equivalent at baseline
  - *Control Group*: receives a placebo or standard treatment
  - *Blinding*: participants and/or researchers do not know the assignment to avoid bias
  - *Outcome Measurement*: pre-defined metrics assess the intervention's effect
- **Example**: testing a new drug
  - Treatment group receives the new drug
  - Control group receives a placebo
  - Compare recovery rates after a fixed period
- **Pros**
  - Provides clear causal inference due to randomization
- **Cons**
  - Expensive and time-consuming
  - Ethical or practical constraints may prevent randomization

RCTs Solve the Problem of Confounders

- In **observational data**
  - Confounding variable C affects both treatment T and outcome Y
  - C creates *spurious correlation* between T and Y



- In **experimental settings**
  - Randomization (R) breaks link between C and T
  - Random assignment prevents influence on both treatment and outcome
  - T is independent of C: $T \perp C$
  - Only open path between T and Y is causal path $T \to Y$



Causal Graphs and Interventions

- **Observing correlations** between variables *does not reveal causality*
  - $\Pr(Y|T)$ confounds direct and indirect influences
- **Randomized Controlled Trials** provide the *gold standard* for causal inference
  - Randomization breaks all back-door (confounding) paths
  - RCTs are expensive, slow, or ethically impossible
- **Alternative solution**
  - Can we estimate the *causal effect* from *observational data alone*?
  - Under *what conditions* and using *which variables*?
- **Idea**: Identify and condition on the right *confounders* to:
  - Block spurious associations between T and Y
  - Recover the true causal effect $\Pr(Y|do(T))$

Intervention in Structural Equations

- **Purpose of Structural Equations**
  - Capture causal mechanisms among variables
  - Predict impact of external interventions
- **Effect of Intervention** $do(X_j = x_j)$
  - Original equation:
    $$X_j = f_j(Parents(X_j), \varepsilon_j)$$
  - Modified by intervention:
    $$X_j = x_j \text{ (fixed value)}$$
  - "Mutilate" causal network by *removing incoming edges* to X_j
  - Recompute joint distribution of all variables using modified structure
- **Intuition**
  - *do*-operator enforces variable's value externally, breaking causal dependencies
  - Enables reasoning about "what would happen if...?" scenarios

Adjustment Formula in Causal Networks

- **Goal**
  - Estimate causal effect of intervention $do(X_j = x_{jk})$ on another variable X_i
- **The Adjustment Formula**
  - Derived from the post-intervention joint distribution:

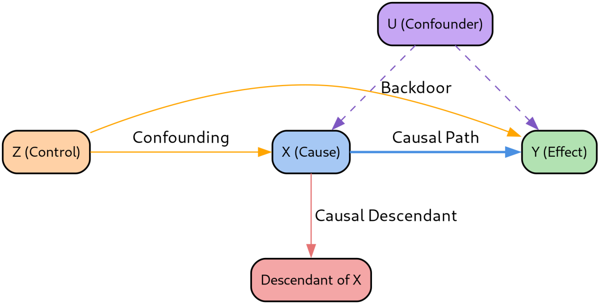
$$\Pr(X_i = x_i | do(X_j = x_j^*)) = \sum_{Parents(X_j)} \Pr(x_i | x_j^*, Parents(X_j)) \Pr(Parents(X_j))$$

  - The mechanism for X_j is *removed*: it is treated as a fixed cause, not a random variable
- **Interpretation**
  - Computes a *weighted average* of effects of X_j and its parents on X_i
  - Weights come from prior probabilities of the parents' values
- **Back-Door Criterion**
  - A set Z is a valid adjustment set if it blocks *all back-door paths* from X_j to X_i
  - Ensures $X_i \perp \text{Parents}(X_j) | X_j, Z$
- **Why It Matters**
  - Enables causal inference from observational data
  - Estimate treatment and policy effects *without randomized trials*

SCIENCE ACADEMY

Backdoor Criterion: Definition

- A set of variables Z satisfies the **backdoor criterion** for variables X (cause) and Y (effect) in a causal graph if:
  1. **No element of Z is a descendant of X**
     - Ensures Z does not "block" part of the causal effect of X on Y
     - Descendants of X may carry information about the causal effect and should not be controlled for
  2. **Z blocks every path between X and Y containing an arrow into X**
     - These paths are *backdoor paths*, representing potential confounding influences
     - Blocking them ensures any remaining association between X and Y is causal, not spurious



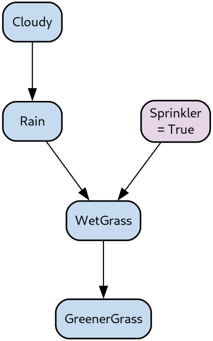
Backdoor Criterion: Intuition

- **Intuition**:
  - The goal is to isolate the causal effect of X on Y by eliminating *confounding bias*
  - Controlling for an appropriate set Z makes the relationship between X and Y as if X were randomly assigned
- **Application**:
  - When Z satisfies the backdoor criterion, we can estimate causal effects from **observational data** (without experiments)
  - The causal effect can be computed using:

$$\Pr(Y|do(X)) = \sum_z \Pr(Y|X, Z = z)P(Z = z)$$

Intervention: Sprinkler Example

- "Intervene" by turning the sprinkler on
  - In do-calculus $do(Sprinkler = T)$
  - Sprinkler variable s is independent of cloudy day c
- Structural equations after intervention:



$$\begin{cases} C := f_C(\varepsilon_C) \\ R := f_R(C, \varepsilon_R) \\ S := True \\ W := f_W(R, S, \varepsilon_W) \\ G := f_G(W, \varepsilon_G) \end{cases}$$

- $\Pr(S|C) = 1$ and $\Pr(W|R, S) = \Pr(W|R, S = T)$ and the joint probability becomes:

$$\Pr(C, R, W, G|do(S = True)) = \Pr(C)\Pr(R|C)\Pr(W|R, S = True)\Pr(G|W)$$

- Only descendants of manipulated variable *Sprinkler* are affected

Intervention vs. Observation in Causal Models

- **Intervention** conceptually *breaks* normal causal dependencies
  - Intervening on *Sprinkler* removes causal link from *Weather* to *Sprinkler*
  - After intervention, causal graph excludes arrow $Weather \rightarrow Sprinkler$
  - *Weather* and *Sprinkler* become independent under intervention
- **Observation vs. Intervention**
  - **Observation**: seeing $Sprinkler = T$
    - Expressed as $\Pr(\cdot | Sprinkler = T)$
    - Reflects *passive observation* — sprinkler on provides information about weather
    - Since *Weather* influences *Sprinkler*, observing $Sprinkler = T$ makes it *less likely* *Weather* is cloudy
  - **Intervention**: forcing $Sprinkler = T$
    - Expressed as $\Pr(\cdot | do(Sprinkler = T))$
    - *Active manipulation* — set sprinkler on regardless of weather
    - Causal link from *Weather* to *Sprinkler* is cut, weather distribution remains unchanged
- **Key intuition**
  - Observation $\rightarrow$ correlation (information flows along causal links)
  - Intervention $\rightarrow$ causation (links into manipulated variable are removed)
  - Thus, $\Pr(Weather | Sprinkler = T) \neq \Pr(Weather | do(Sprinkler = T))$

Controlling for a Variable in Causal Analysis

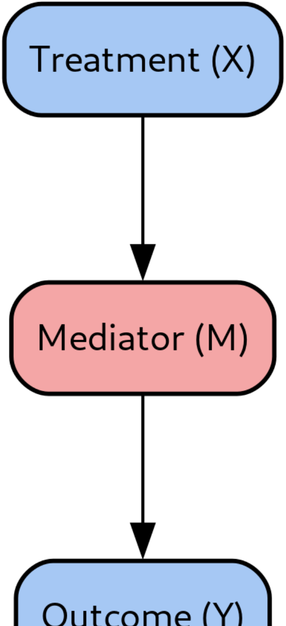
- **Definition**
  - To *control* a variable means to hold it constant (statistically or experimentally) to isolate the causal effect of another variable
- **Example**
  - Does exercise (X) cause weight loss (Y)?
  - Confounder: Diet (Z) affects both exercise and weight
  - By controlling for diet (e.g., comparing people with similar diets), you can estimate the effect of exercise more accurately
- In **regression analysis**
  - Include Z as an additional independent variable
  - E.g., in $Y = \beta_0 + \beta_1 X + \beta_2 Z + \varepsilon$
    - β_1 measures the effect of X *controlling for* Z
    - Coefficient β_1 = *change in* Y with a one-unit change in X_1, *holding* X_2 *constant*
    - Isolates X_1's unique contribution
    - Compares individuals with the same X_2 but different X_1
- In **experiments**
  - Keep Z constant or randomize it

SCIENCE ACADEMY

- Causal AI
- Causal Networks
    - Variables
    - Intervention
    - ***Type of Variables in Causal AI***
    - Paths
- Business Processes Around Data Modeling

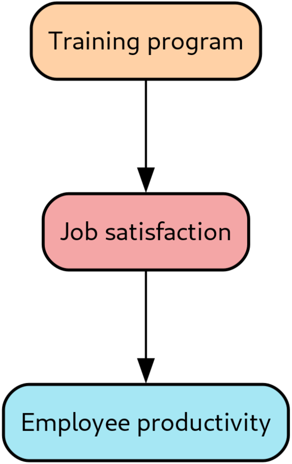
Mediator Variable

- A **mediator variable** M is an intermediate variable that *transmits* the causal effect from X (treatment) to Y (outcome)
  - Lies **on the causal path** between X and Y
  - Captures the **mechanism or process** through which X influences Y



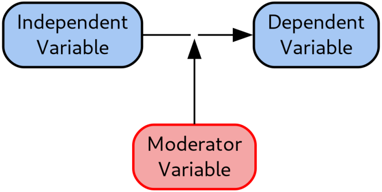
Mediator Variable: Example

- Research question: does a *training program* increase *employee productivity*?
- The causal effect may be **indirect**, operating through a **mediator**
  - The training program might not immediately boost productivity
  - Instead, it could enhance **job satisfaction**, which in turn raises productivity
- **Causal interpretation**
  - X: Training Program (cause)
  - M: Job Satisfaction (mediator)
  - Y: Employee Productivity (effect)
  - Path: $X \to M \to Y$
- **Direct vs. Indirect effects**
  - *Indirect effect* X affects Y through M
  - *Direct effect* X affects Y not through M
  - Controlling for M separates these two effects, clarifying *how* training impacts outcomes



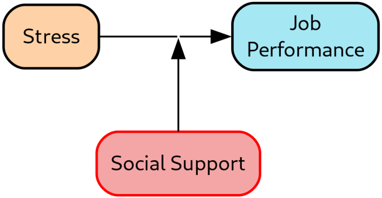
Moderator Variable

- A **moderator variable** changes the *strength* or *direction* of the relationship between an independent variable (X) and a dependent variable (Y)
  - Moderator is not part of the causal chain but conditions the relationship



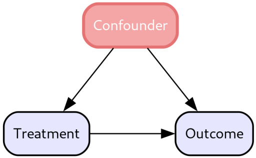
Moderator Variable: Example

- Research question: study relationship between stress X and job performance Y
- Social support M as a moderator
  - High social support weakens stress's negative effect on performance
  - Low social support strengthens stress's negative effect on performance



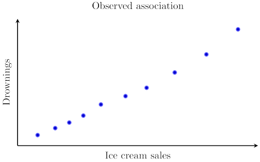
Confounder Variable

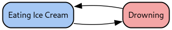
- A **confounder**
  - Influences multiple variables in a causal graph
  - Affects both treatment (cause) and outcome
  - Creates misleading association if not controlled



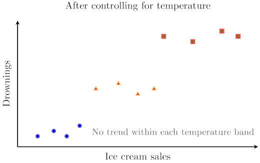
Confounder Variable: Example

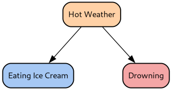
- *IceCreamSales* and *Drowning* move together
  - Correlation-based model claims association, but how to use this relationship?
  - Ban ice cream to prevent drowning?
  - Ice cream maker increase drowning to boost sales?

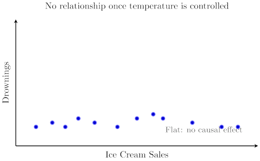




- In reality, no cause-effect between *IceCreamSales* and *Drowning*
  - *Temperature* is a confounder
  - When controlling for season in regression or intervention, association disappears

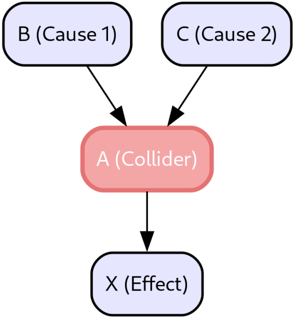






Collider

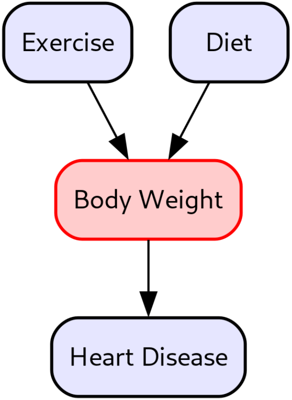
- A **collider** is a variable A influenced by multiple variables
  - In a causal graph A with incoming edges from variables B, C
- A collider complicates understanding relationships between variables B, C and those it influences, X



SCIENCE ACADEMY

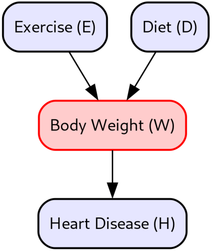
Collider: Examples

- Study the relationship between *Exercise* and *HeartDisease*
  - *Diet* and *Exercise* influence *BodyWeight*
  - *BodyWeight* influences *HeartDisease*
  - *BodyWeight* is a collider



Collider Bias

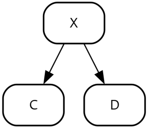
- Aka "Berkson's paradox"
- Conditioning on a collider can introduce a spurious association between its parents by *"opening a path that is blocked"*
- Consider the variables:
  - *Diet* (D)
  - *Exercise* (E)
  - *BodyWeight* (W)
  - *HeartDisease* (D)
- **Without conditioning on W**
  - E and D are independent
    - E.g., knowing someone's exercise level E doesn't give information about diet D, and vice versa
  - The collider W blocks any association between E and D
- **After conditioning on W**
  - E.g., looking for individuals with specific body weight
  - You introduce a dependency between E and D
  - Since W is fixed, any change in E must be balanced by a change in D to maintain the same



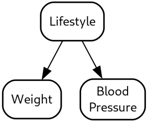
- Causal AI
- Causal Networks
  - Variables
  - Intervention
  - Type of Variables in Causal AI
  - ***Paths***
- Business Processes Around Data Modeling

Fork Structure

- A **fork** occurs when a single variable causally influences two or more variables
  - Formally: $X \to C$ and $X \to D$
- X is a common cause (confounder) of C and D
- Forks induce statistical dependence between C and D
  - Even if C and D are not causally linked
- Conditioning on X blocks the path and removes spurious correlation



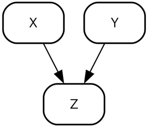
- Example:
  - Lifestyle factors as confounders
  - *Lifestyle* affects both *Weight* and *BloodPressure*
  - These outcomes may appear correlated due to shared cause

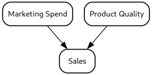


SCIENCE ACADEMY

Inverted Fork

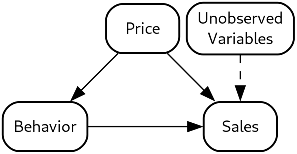
- An **inverted fork** occurs when two or more arrows converge on a common node
  - Also known as a **collider**
- Colliders block associations unless the collider or its descendants are conditioned on
- Conditioning on a collider "opens" a path, inducing spurious correlations
  - This is the basis of selection bias
- Example:
  - Sales influenced by multiple independent causes
  - *MarketingSpend* and *ProductQuality* both influence *Sales*
  - Conditioning on *Sales* can induce false dependence between *MarketingSpend* and *ProductQuality*





Path connecting unobserved variables

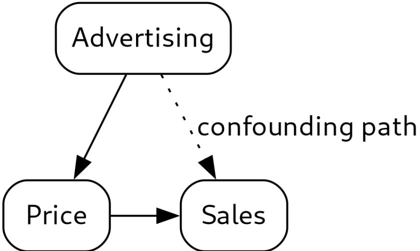
- **Unobserved variables** affect the model but we don't have a direct measure of it
- E.g., consider the causal DAG
  - A retailer does market research, expecting *Price* to influence *Sales* in a predictable way
  - A retailer sets the *Price* of a new product based on market research
  - The retailer can observe and measure *Behavior*, e.g.,
    - Discounts
    - Promotional campaign
  - There are unobserved vars that influence the model, e.g.,
    - Social media buzz
    - Word-of-mouth recommendation



SCIENCE ACADEMY

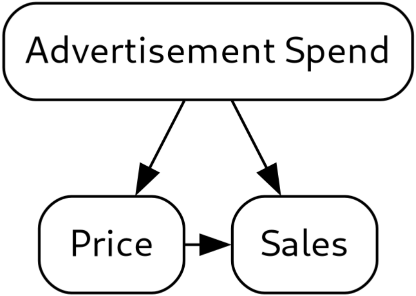
Front-door Paths in Causal Inference

- A front-door path reveals causal influence through an observable mediator
  - The causal effect flows: $A \to P \to S$
- Requirements for identifiability:
  - All confounders of $A \to P$ and $P \to S$ are observed and controlled
  - There are no back-door paths from A to S through unobserved variables
- Enables causal estimation when back-door adjustment is infeasible
- Example:
  - Advertising impacts sales through customer perception of price
  - A: Advertising, P: Price perception, S: Sales
- Pearl's front-door criterion provides a formal method for adjustment
  - Estimate $P(P|A)$, $P(S|P,A)$, and $P(A)$ from data to compute causal effect



Back-Door Paths

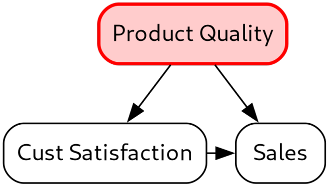
- A company wants to understand the causal effect of price on sales



- `Price` $\rightarrow$ `Sales` is the front-door path
- A confounder is `Advertising spend` since it can affect both:
  - The price the company can set (e.g., the cost increases to cover advertisement costs and the product is perceived as more valuable)
  - The sales (directly)
- The back-door path goes from `Sales` to `Price` via `Advertising spend`

Frontdoor and Backdoor Paths

- Question: *Will increasing our customer satisfaction increase our sales?*
  - Assume that the Causal DAG is



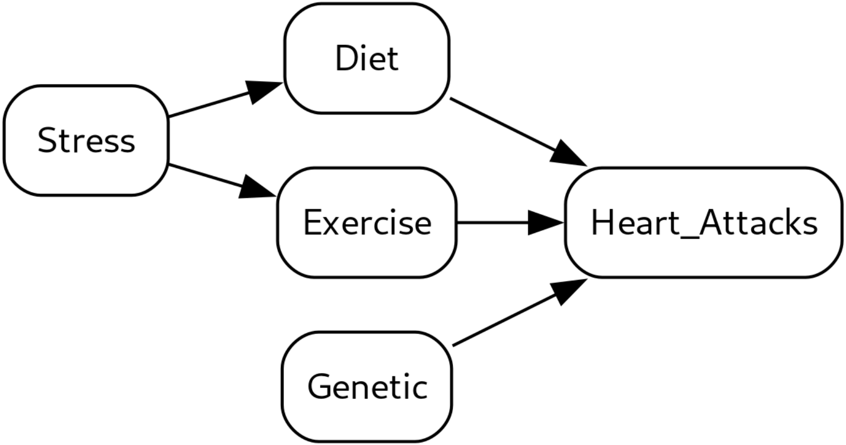
  - **Front-door path** (i.e., a direct causal relationship): *CustomerSatisfaction* → *Sales*
  - **Backdoor path**: *ProductQuality* is a common cause (confounder) of both *CustomerSatisfaction* and *Sales*
  - To analyze the relationship between customer satisfaction and sales, we need to:
    - Control for *ProductQuality* to close the backdoor path
    - Eliminate the confounding effect
  - In reality there are more confounding effects (e.g., price)

Building a DAG

- **Causal models** visually represent complex environments and relationships
- Nodes are like "nouns" in the model:
  - E.g., "price", "sales", "revenue", "birth weight", "gestation period"
  - Variables can be endogenous/exogenous and observed/unobserved
  - Complex relationships between variables:
    - Parents, children (direct relationships)
    - Descendants, ancestors (along the path)
    - Neighbors
- **Iterative Refinement**:
  - Models are continuously updated with new variables and insights
- **Modeling as a Communication Tool**:
  - A shared language that bridges gaps between technical and non-technical team members
- **Unobservable Variables**:
  - Supports inclusion of variables not empirically observed but known to exist
  - E.g., trust or competitor activity can be modeled despite lack of direct data

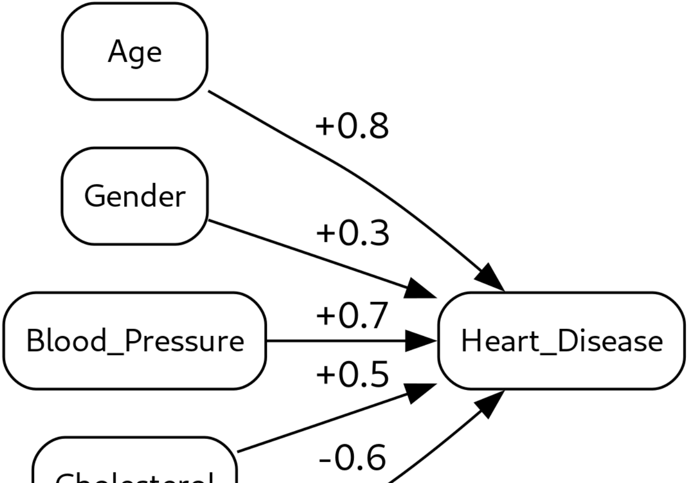
Heart Attack: Example

- What's the relationship between stress and heart attacks?
  - Stress is the treatment
  - Heart attack is the outcome
  - Stress is not a direct cause of heart attack
    - E.g., a stressed person tend to have poor eating habits



Weights

- Weights can be assigned to paths to represent the strength of the causal relationship
  - Weights can be estimated using statistical methods
- Sign represents the direction



Counterfactuals

- A **counterfactual** describes what would have happened under a different scenario
  - *"What would the outcome have been if X had been different?"*
  - *"If kangaroos had no tails, they would topple over"*
  - *"What if we had two suppliers of our product, rather than one? Would we have more sales?"*
  - *"Would customers be more satisfied if we could ship products in one week, rather than three weeks?"*
- **Causal reasoning**:
  - Goes beyond correlation and association
  - Requires a causal model (like an SCM) to simulate alternate realities
  - E.g.,
    - Actual: A student received tutoring and scored 85%
    - Counterfactual: What if the student didn't receive tutoring?
    - Causal model estimates the alternative outcome (e.g., 70%)
- **Challenges**:
  - Requires strong assumptions and accurate models
  - Difficult to validate directly since counterfactuals are unobservable

- Causal AI
- Causal Networks
- ***Business Processes Around Data Modeling***
  - Modeling Processes
  - Roles

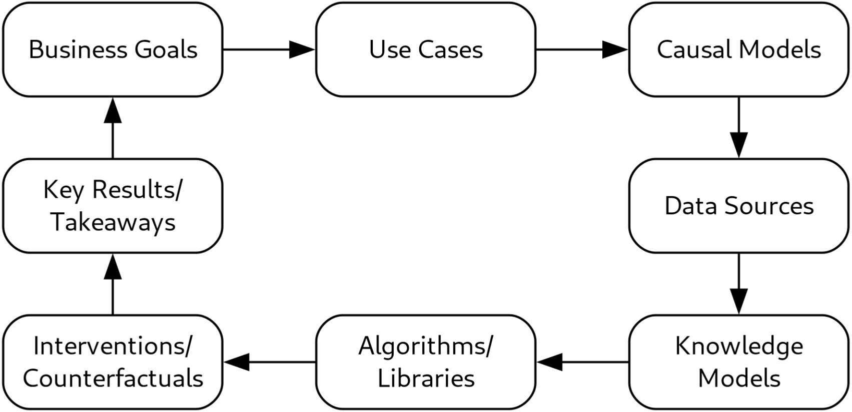
- Causal AI
- Causal Networks
- Business Processes Around Data Modeling
    - ***Modeling Processes***
    - Roles

Digital Transformation

- Integration of Digital Technology
  - Embed digital tools (AI, cloud, IoT, automation) into all business areas to enhance efficiency and value delivery
- Cultural & Organizational Change
  - Encourage innovation, agility, and a data-driven mindset to adapt to new digital workflows and business models
- Customer-Centric Approach
  - Use digital solutions (e.g., personalized experiences, AI-driven insights) to enhance customer engagement and satisfaction
- Process Automation & Optimization
  - Streamline operations through AI, robots, and analytics to reduce costs and improve decision-making
- Data-Driven Decision Making
  - Leverage big data, machine learning, and real-time analytics to make smarter, faster, and more strategic business decisions

Causal Modeling Process

- The overall modeling process looks like:



Step 1: What Are the Intended Outcomes?

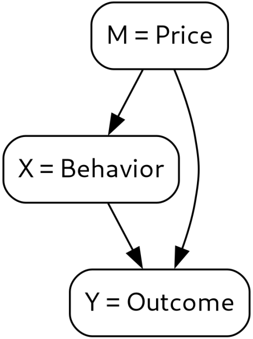
- What is the process/environment we are interested in analyzing?
- What will happen if a course of action is (or not) taken?
- What outcomes are positive, negative, unacceptable, optimal?
- What are the possible / feasible interventions?
- What confounding factors might be correlated with outcomes and treatments?
- What factors exist but we cannot accurately measure?
- What related data sets can be combine / leverage?

Step 2: What Are the Proposed Interventions?

- We will make reference to a use case of customer marketing
- Can we introduce a new product?
- Should we buy one or more competitors?
- Does bundling multiple products improve sales?
- Does bundling multiple products inhibit long-term sales?
- Should advertisement focus on quality of our product vs other options?
- Should we divest the product line?
- Should we discontinue the product?
- Should we add more variations of the same product?

Marketing example: price intervention

- Assume price is our intervention and the Causal DAG is:

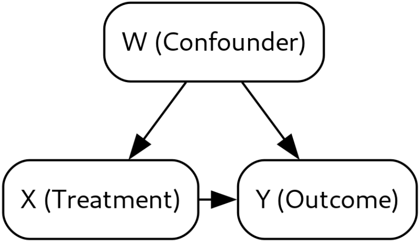


- What happens to sales when we change the price often?
- What pricing interventions are optimal?
  - Should we increase the price and how much?
  - Should we decrease the price and how much?
  - Should price change in one-time or over time?
- Should we adopt a dynamic pricing model?
- Should we develop individual pricing model for each customer?

SCIENCE ACADEMY

Step 3: What are the confounding factors?

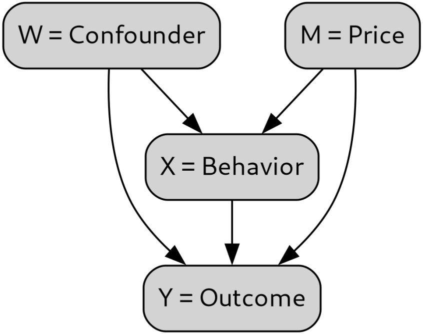
- There can be a variable W that affects both X and Y



- E.g.,
  - Competitive offers
  - Distance to store
  - Amount of product
  - Time to consume product
- A confounder can:
  - Make it difficult to understand the relationship between variables
  - Mute or inflate a relationship

Marketing Example: Effect of Confounder

- E.g.,
  - Intervention = a marketing campaign to sell winter jackets
  - A confounding variable can be "running the campaign in the middle of the winter, after customers have already purchased their jackets"
  - A confounding var can be "a warm winter"



SCIENCE ACADEMY

Step 4: What Are the Factors Creating the Effects and Changes?

- Total causal effect
  - Effect of all factors in the environment or model that modify the outcome
- Direct effect
  - Effect introduced through an intervention
- Indirect effect
  - Effect introduced by environment or it is a byproduct of the intervention in a way that was not planned

Step 5: Build Causal DAG

- Causal models
  - Simplify complex systems without losing key relationships
  - Focus on essential variables and their interactions
- Visual models (e.g., DAGs) help abstract complexity into interpretable formats
  - Highlight direction and strength of influence between variables
- Simplicity
  - Aids communication between technical and non-technical stakeholders
  - Promotes shared understanding and collaborative refinement
  - Reduces cognitive overload by excluding irrelevant details or noise
  - Guides data collection by identifying the most impactful variables
  - Supports hypothesis testing through counterfactual and intervention scenarios
- Balance is key
  - Too much simplicity loses insight
  - Too much complexity loses clarity

Step 7: Data Acquisition and Integration

- You can use the data collected for correlation-based ML
- Data collection can be done specifically for causal AI
  - Treating, conditioning, transforming data

Step 8: Model Modification

- Once the DAG is designed, use software packages to build models
  - Refine the initial DAG and causal model to reduce bias and improve reliability
  - Clarify variables as confounders, mediators, or outcomes
- Avoid common pitfalls:
  - Do not control for mediators or effects, which can distort results
  - Control for direct and indirect confounders to prevent biased estimates
- Implementation tools:
  - Use libraries to operationalize models
  - Test models against technical and business objectives
- D-separation:
  - Identifies conditional independence relationships in a DAG
  - Determines necessary controls to isolate causal effects
  - Based on Judea Pearl's definition: independence = separation in the graph
  - Prevent unintentional inclusion of bias
  - Ensure causal assumptions align with data and domain logic
  - Improve model interpretability and predictive power
- Goal: Ensure final models are technically valid and business-relevant before proceeding to data transformation and testing stages

Step 10: Data Transformation

- Prepare the data to match the refined causal model
  - Clean, normalize, and align data with model assumptions
- Transformations include:
  - Mapping observed variables to nodes in the DAG
  - Encoding categorical variables appropriately
  - Handling missing or unobserved data (e.g., imputation or exclusion)
  - Normalizing or scaling values to align with model expectations
- Control for bias and confounding:
  - Apply methods like propensity score matching or stratification
  - Exclude or adjust for variables that introduce bias per d-separation insights
- Goal: Ensure the data structure supports causal estimation
  - Consistent with assumptions made in model refinement
  - Aligned with theoretical model
  - Fit for downstream tasks like estimation, inference, and simulation

Step 11: Preparing for Deployment in Business

- Operationalize the causal model within a business context
  - Transition from experimentation to integration with decision-making processes
  - Validate the model against real-world business data and outcomes
  - Ensure stakeholders understand and trust the causal logic and assumptions
- Model packaging:
  - Develop user-friendly interfaces or dashboards for business users
  - Automate data pipelines for timely updates and monitoring
  - Embed the model within decision-support tools or policy engines
- Governance and monitoring:
  - Establish metrics for performance tracking and drift detection
  - Create feedback loops to refine and improve models post-deployment
- Documentation and training:
  - Provide clear model documentation for auditors and users
  - Train stakeholders on interpreting causal results and making informed decisions
- Goal: A deployable causal AI solution that supports strategic decisions and delivers measurable business value

- Causal AI
- Causal Networks
- Business Processes Around Data Modeling
    - Modeling Processes
    - ***Roles***

Why ML / AI Projects Fail?

- AI projects fail because they approach problems only from a ML perspective
  - Data scientists:
    - Use data to create models
    - Work in isolation from business users and internal data teams
  - Black-box models unable to produce solutions to real-world problems

How to Make ML / AI Project Succeed

1. Create a hybrid team
   - Organizations are complex in structure and offerings
   - A single group lacks the knowledge / skills to tackle difficult problems
   - Need an hybrid team:
     - Represents all aspects of the business problem
     - Uses a collaborative framework
     - Communicates with a single language (e.g., through DAGs)
     - Team size depends on company size and project complexity
2. Meet regularly to ensure project continuity
3. Find an executive sponsor for the project
   - Someone who understands the project's goals and potential
4. Initial pilot
   - Small team for a targeted problem
   - Demonstrate the merit of the AI approach

Roles in Hybrid Teams

Role	Responsibilities
Business Strategists	Align modeling with business goals
	Sponsor projects
	Communicate insights to stakeholders
Subject-Matter Experts	Provide domain expertise
	Identify relevant variables and assumptions
	Validate DAGs
Data Experts	Source and clean data
	Map data to model variables
	Handle missing values
Data Scientists	Construct and validate DAGs
	Apply causal inference methods
	Simulate decisions
Software Developers	Build tools and interfaces
	Create data pipelines
IT Professionals	Provide infrastructure and governance
	Ensure model execution
	Integrate with enterprise systems
Project Managers	Coordinate collaboration and timelines
	Manage documentation
	Ensure alignment with strategic goals

Steps for a Hybrid Team Project

- **Establish a Phased, Collaborative Approach**
  - Align strategic goals with technical efforts
  - Emphasize early stakeholder engagement and shared ownership
- **Strategic Kickoff Meeting**
  - Unite business, technical, and operational roles
  - Clarify problems, outcomes, responsibilities, success metrics
- **Define Team Goals**
  - Use SMART objectives aligned with business strategy
  - Focus on outcomes and supported decisions, not tools
- **Target a Project**
  - Choose a bounded, feasible, high-value use case
  - Prioritize early wins for trust and momentum
- **Define the Hypothesis**
  - Translate business problems into testable causal assumptions
  - Build a preliminary DAG with experts' input
- **Incremental Model Development**
  - Build the model in small, reviewable stages
  - Iterate with regular feedback, refining scope and variables
- **Embrace Iteration and Continuous Refinement**
  - Keep progress collaborative and aligned with business needs
  - Add complexity gradually to manage risk and enhance understanding

The Importance of Explainability

- Managers rely on AI systems to automate decision-making
  - Decisions rely on complex algorithms and data
- Understanding AI-based models is growing in importance
  - How do ML models make decisions?
  - How can they be trusted?
  - Are they biased?
- Management often faces demands to prove code validity
  - Loss of trust, regulation violations, fines, additional development costs, lawsuits
  - E.g., a false negative in a medical screening for cancer
- Well-designed AI systems must foster trust, transparency, and user confidence
- **Humans in the loop**
  - AI systems lack true reasoning and contextual understanding
  - Human involvement ensures interpretation and context are considered

Techniques for Interpretability

- Local Interpretable Model-agnostic Explanations (LIME)
  - Focuses on a single prediction (local fidelity)
  - Approximates the model locally with an interpretable model
  - Perturbs input data and observes changes in predictions
- Partial Dependence Plots (PDP)
  - Show the marginal effect of one or two features on the predicted outcome
  - Vary the value of one feature while keeping others constant
  - Plot feature values against the average predicted outcome
- Individual Conditional Expectation (ICE)
  - Show the relationship between a feature and the prediction for individual instances
- SHapley Additive exPlanations (SHAP)
  - Quantify the contribution of each feature to a specific prediction

Causal AI in Interpretable AI

- Causal AI helps understand causes, effects, and potential solutions
  - Uses causal graphical models to present variables, relationships, and strengths
  - Counterfactual analysis predicts outcomes of different actions or policies before deployment
  - Model output is understandable to humans and non-experts
  - Removing confounding variables prevents skewed causal estimates due to hidden influences
  - Hybrid teams (technical + domain experts) enhance context awareness and reduce blind spots

The Future of Causal AI

- Causal AI:
  - Is moving out of academia into the commercial world
  - Is a departure from the 2000 approach of a purely data-driven AI systems
  - Is a reflection of the reality of how humans think, analyze, and make decisions and how the real world works
- Causal, traditional AI, deep learning, and generative techniques will merge