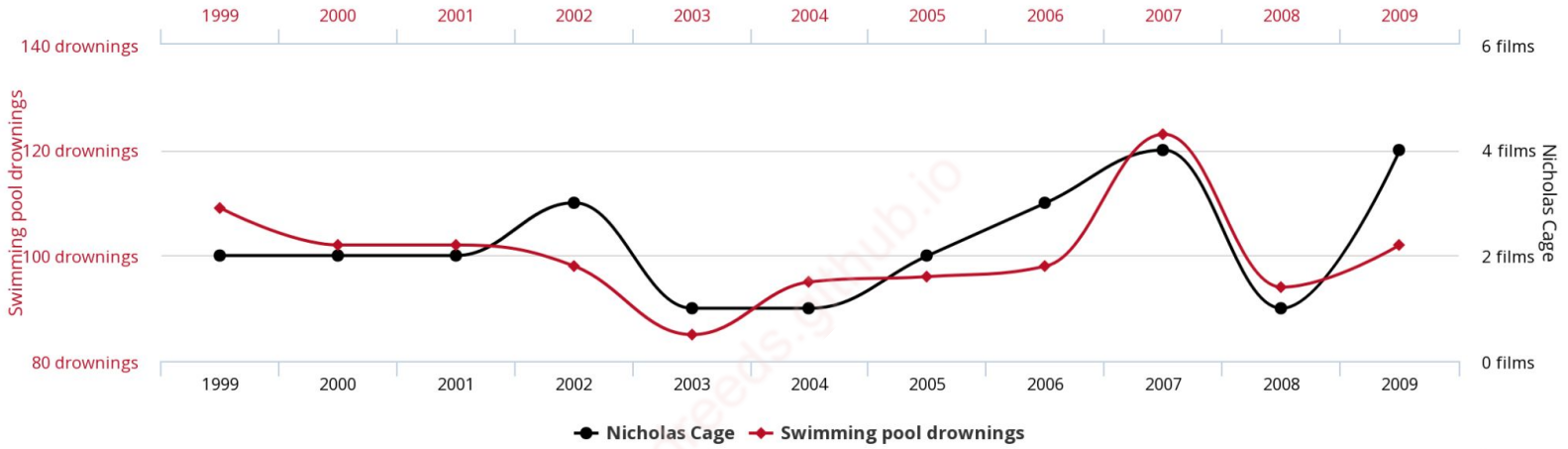
Number of people who drowned by falling into a pool

correlates with

Films Nicolas Cage appeared in

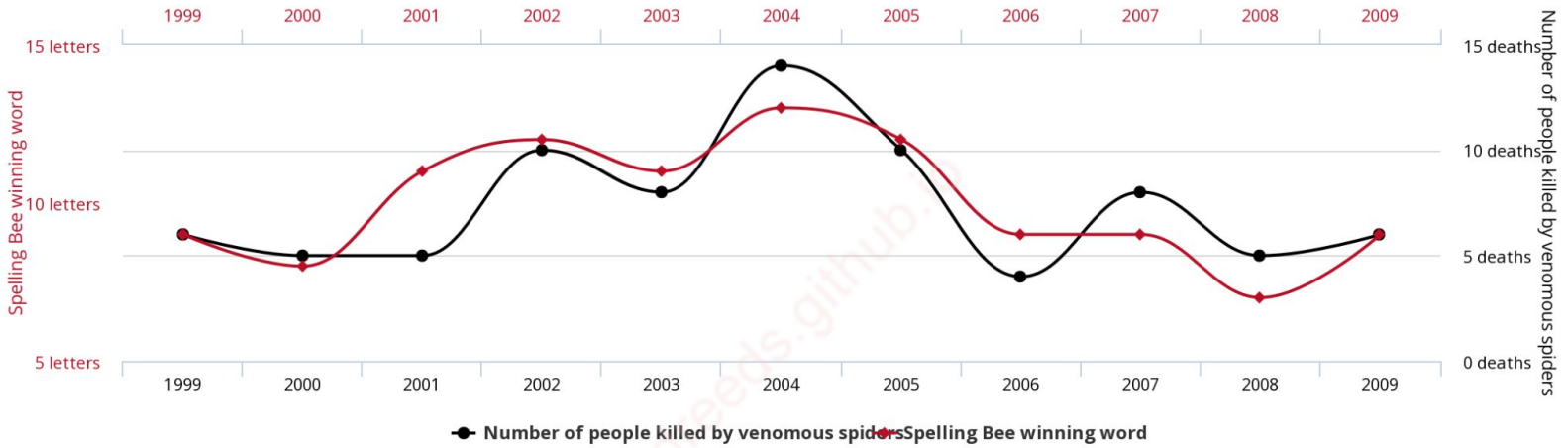


tylervigen.com

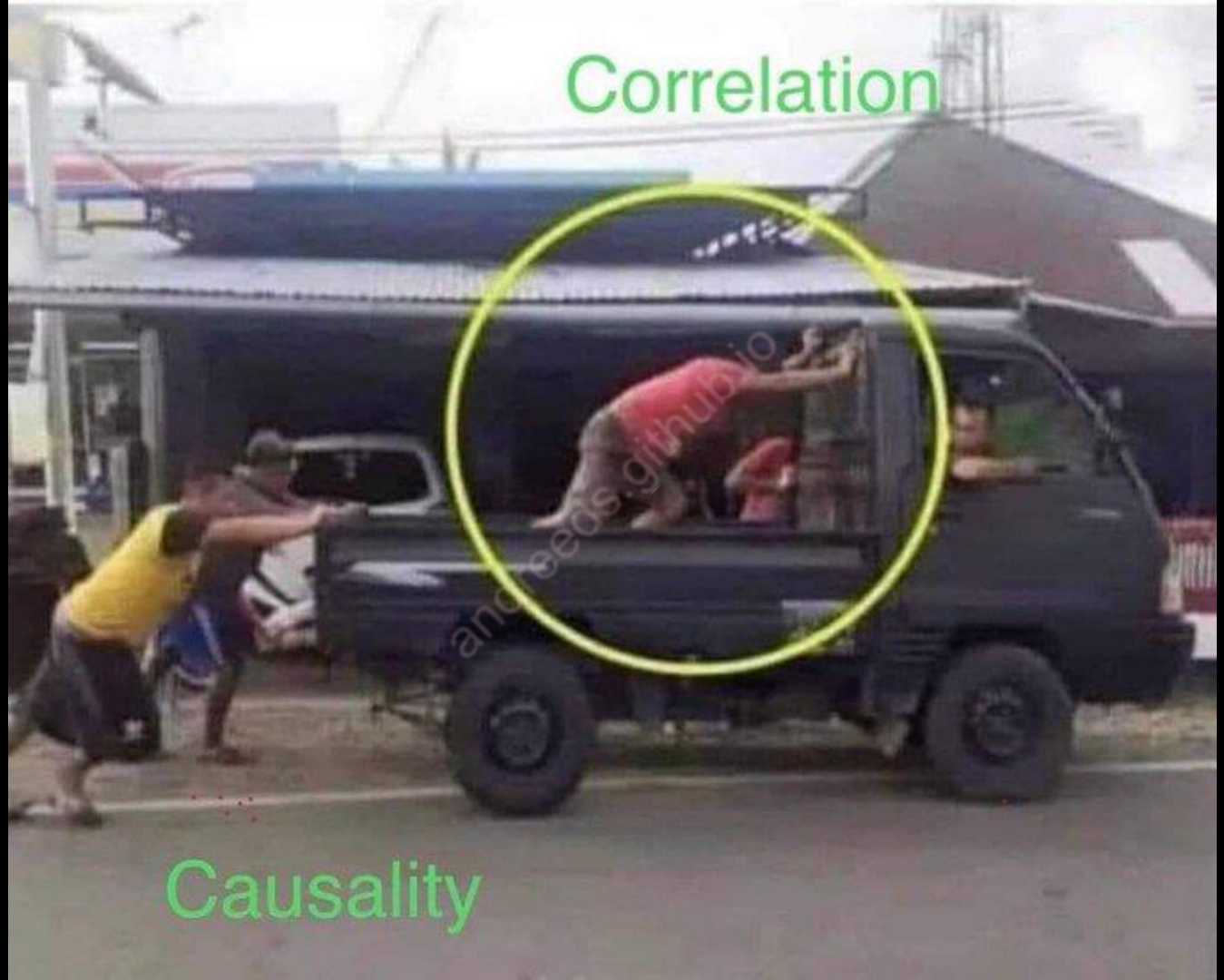
Letters in Winning Word of Scripps National Spelling Bee

correlates with

Number of people killed by venomous spiders



tylervigen.com


Correlation
Causality

CAUSALITY

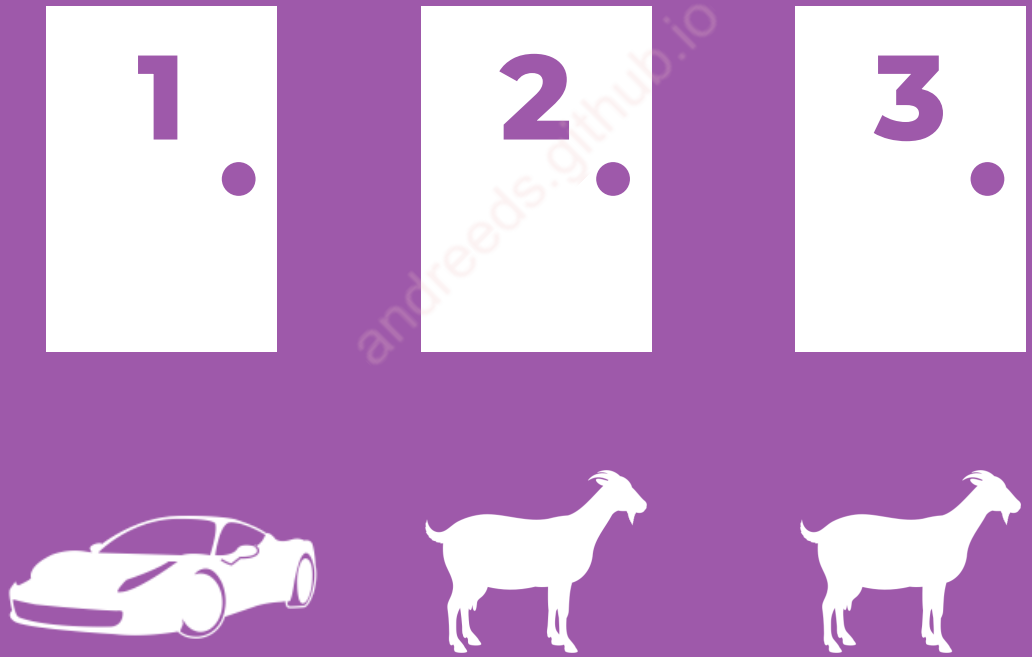
PART I - CONFOUNDING

André dos Santos, *Ph.D.*


Have you ever noticed that,
among the people you date
the attractive ones tend to be jerks?
Frozen (2013)

The Monty Hall problem

"Let's make a deal"



The Monty Hall problem

"Let's make a deal"

The Monty Hall problem

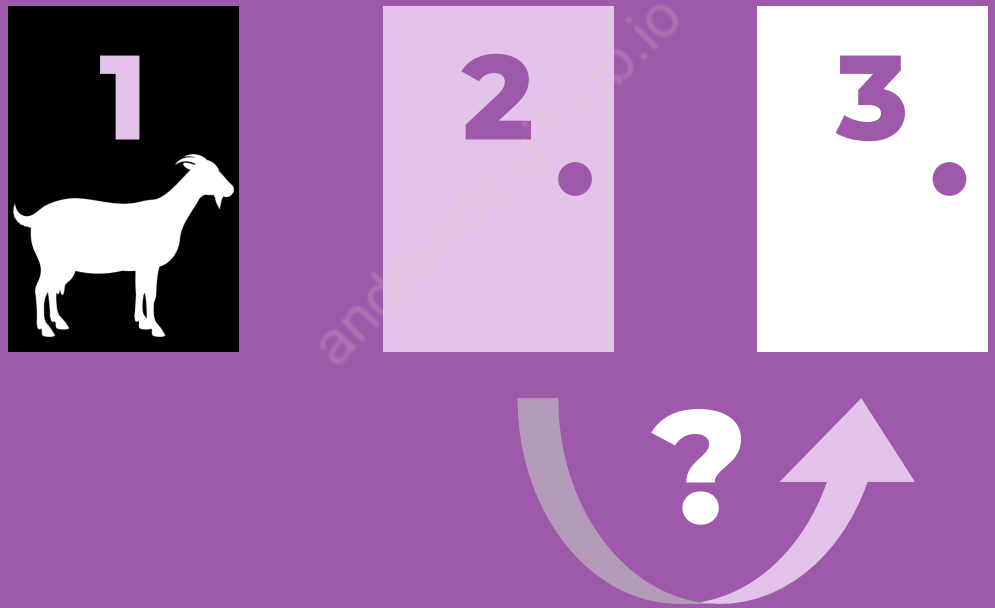
"Let's make a deal"

2.

3.

The Monty Hall problem

"Let's make a deal"



The Ladder of Causation

1

SEEING

ASSOCIATION

2

DOING

INTERVENTION

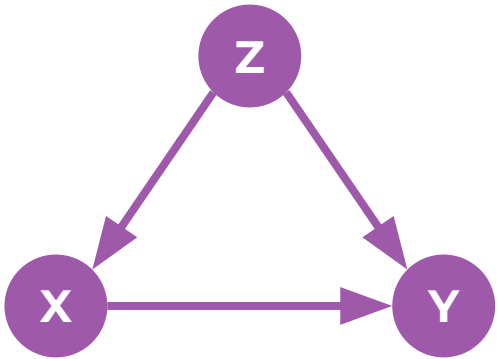
3

IMAGINING

COUNTERFACTUALS

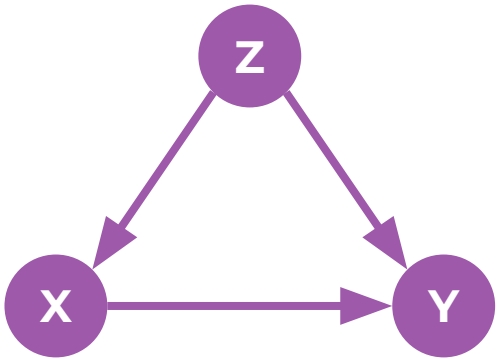
Confounding

- Not a statistical notion
- P(Y|X) ≠ P(Y|do(X))
- **Discrepancy** between what we want to a asses (the causal effect) and what we actually do assess using statistical method



Backdoor path is any path from X to Y that starts with an arrow pointing into X.

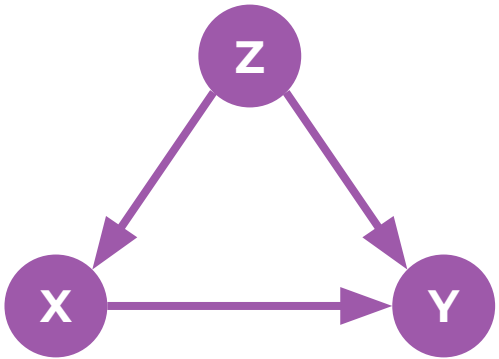
$$X \leftarrow Z \rightarrow Y$$



X and Y will be **deconfounded** if we block every backdoor path.

Backdoor path is any path from X to Y that starts with an arrow pointing into X.

$$X \leftarrow Z \rightarrow Y$$



X and Y will be **deconfounded** if we block every backdoor path.

Flow of Information In Causal Graph

CHAIN

FORK

COLLIDER

Flow of Information In Causal Graph

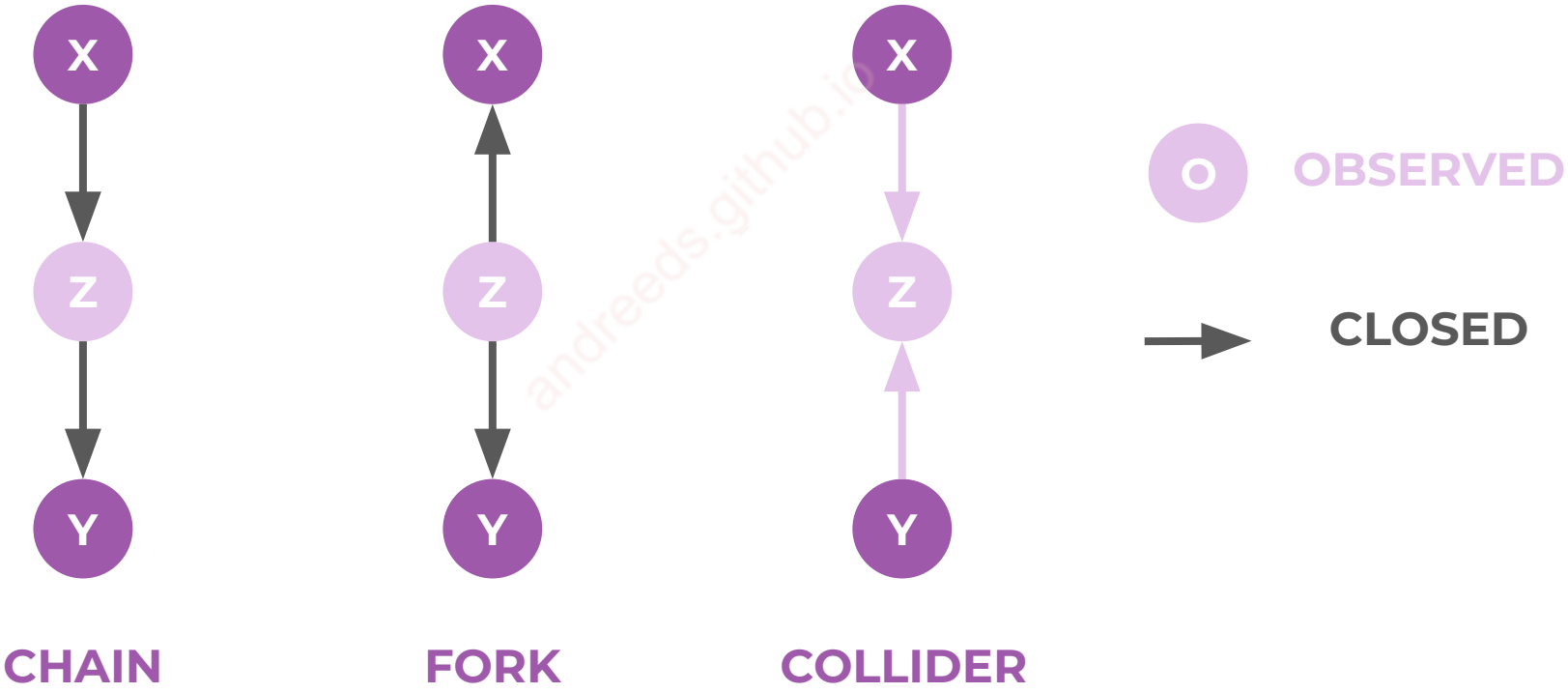
CHAIN

FORK

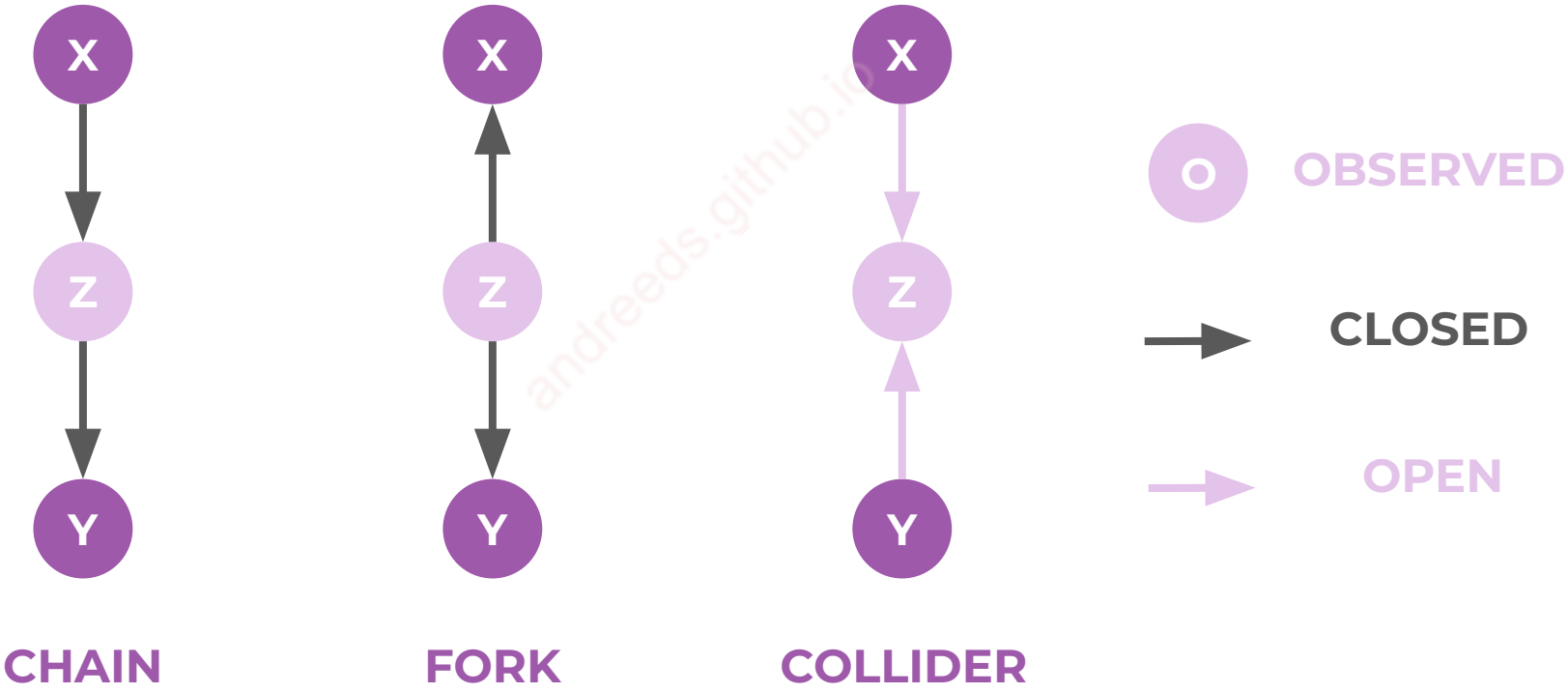
COLLIDER

OBSERVED

Flow of Information In Causal Graph



Flow of Information In Causal Graph

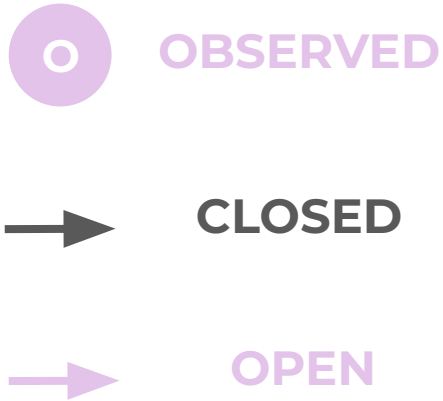


Flow of Information In Causal Graph

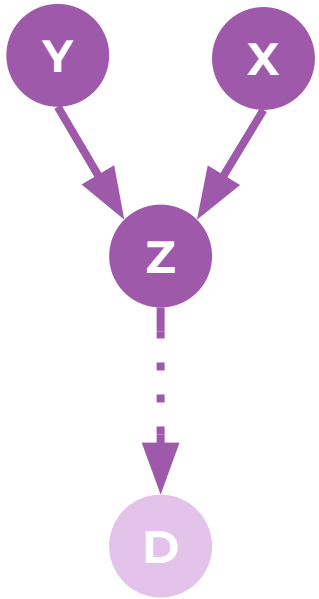
CHAIN

FORK

COLLIDER

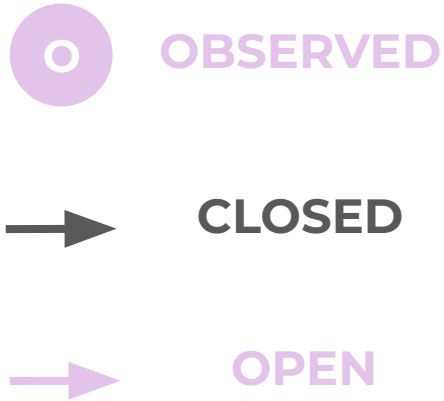


Flow of Information In Causal Graph

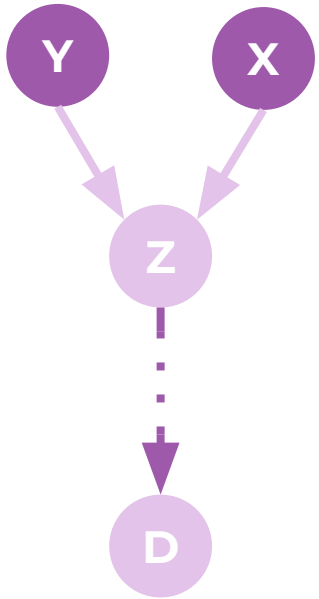


*DESCENDENTS

COLLIDER

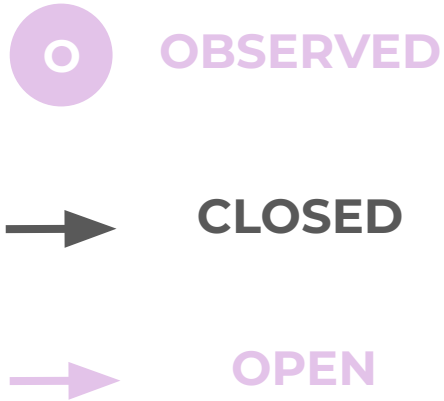


Flow of Information In Causal Graph



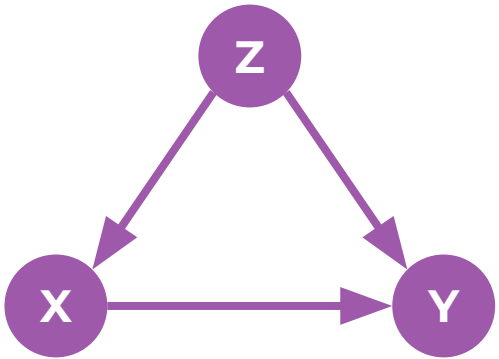
*DESCENDENTS

COLLIDER

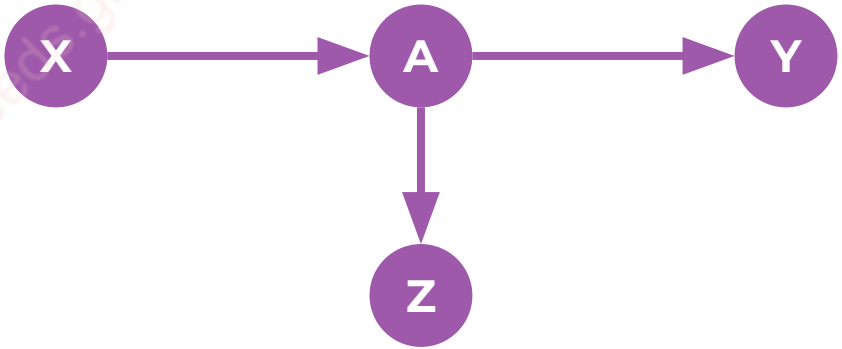


Backdoor path is any path from X to Y that starts with an arrow pointing into X.

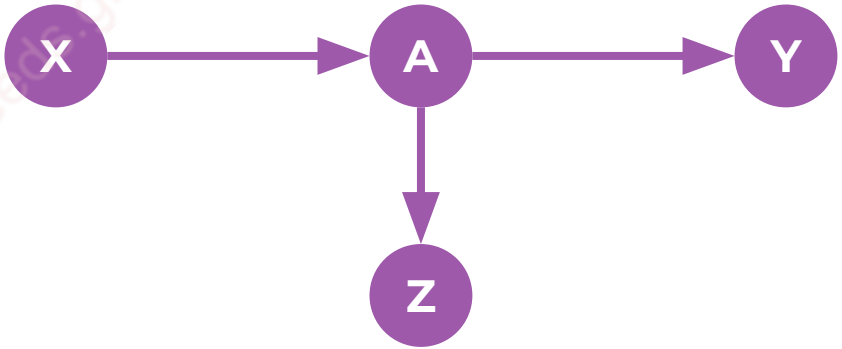
$$X \leftarrow Z \rightarrow Y$$

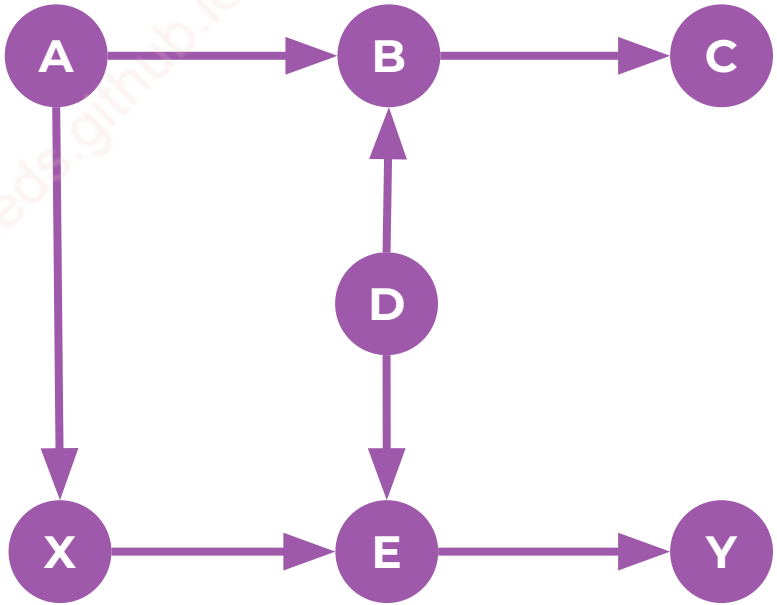


X and Y will be **deconfounded** if we block every backdoor path.

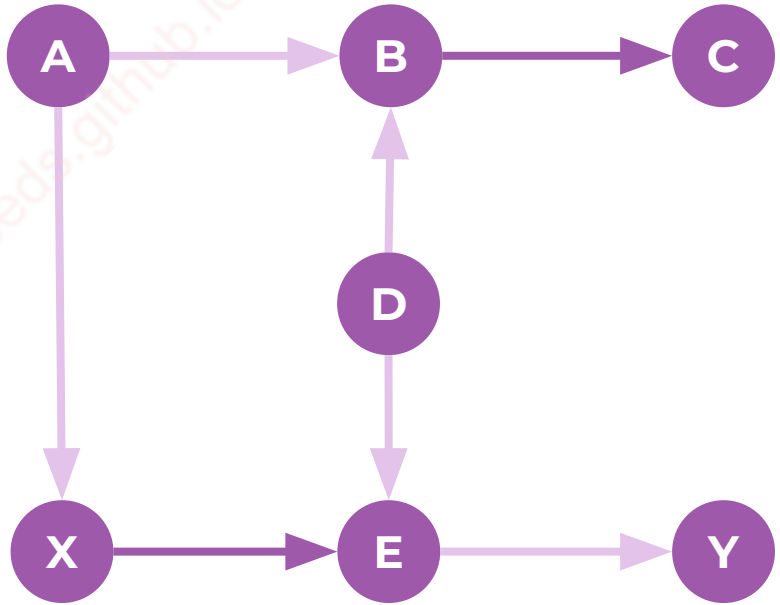


- No backdoor

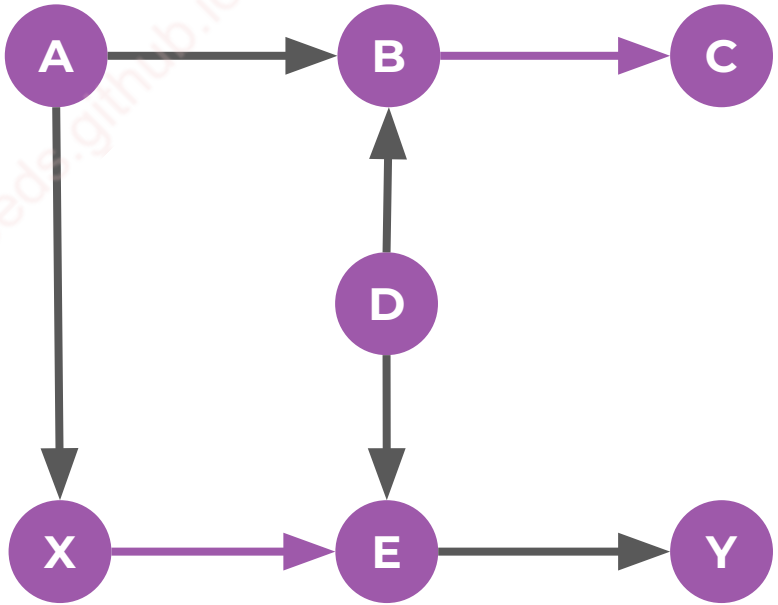




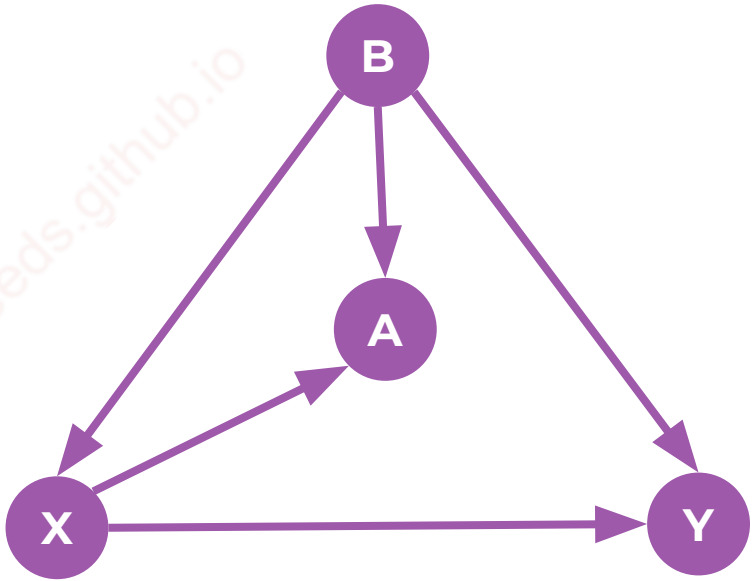
- One backdoor path
- X ← A → B ← D → E → Y



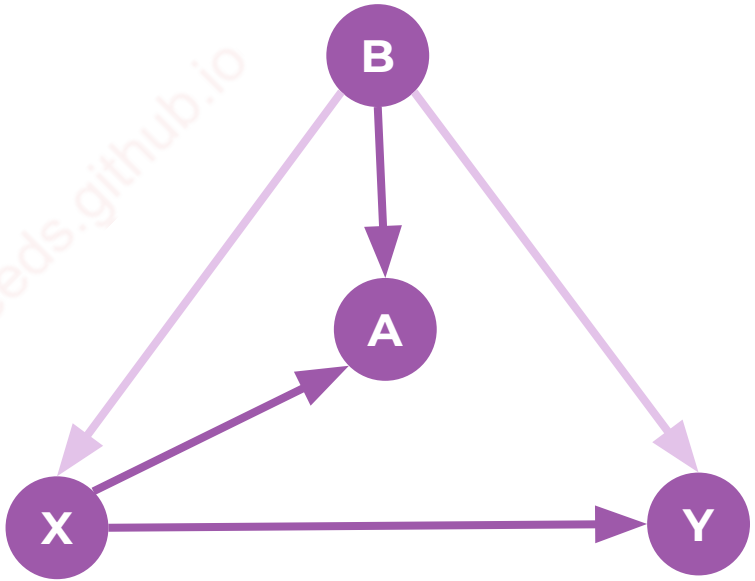
- One backdoor path
- $X \leftarrow A \rightarrow B \leftarrow D \rightarrow E \rightarrow Y$
- Control $\varnothing$



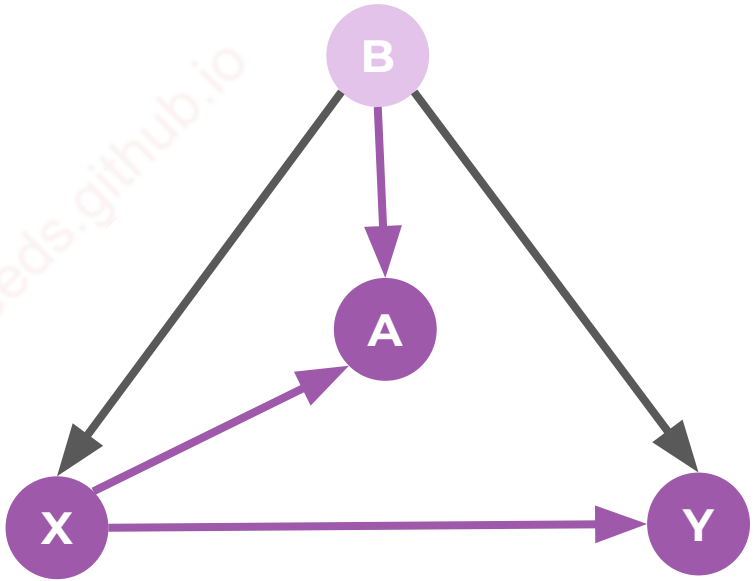
- One backdoor path
- X ← B → Y

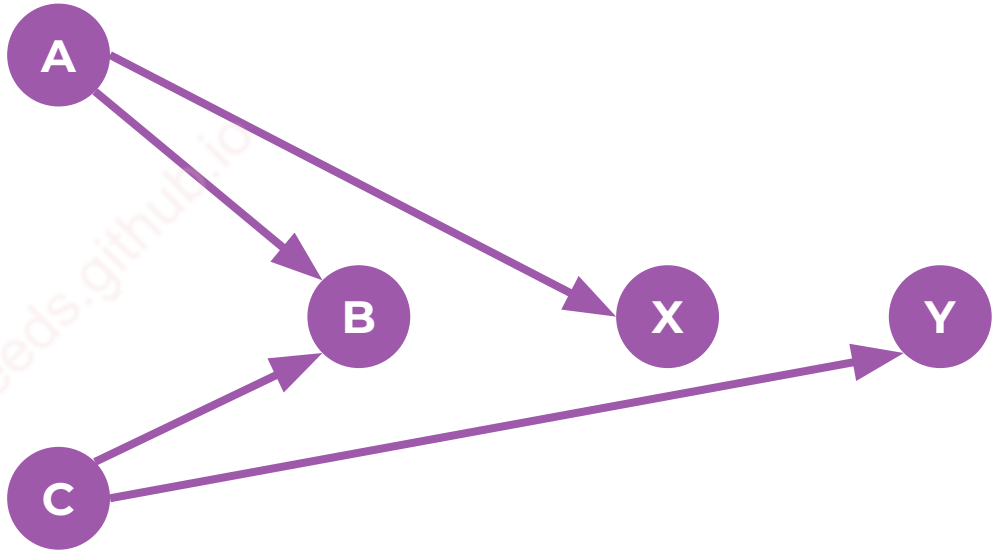


- One backdoor path
- X ← B → Y

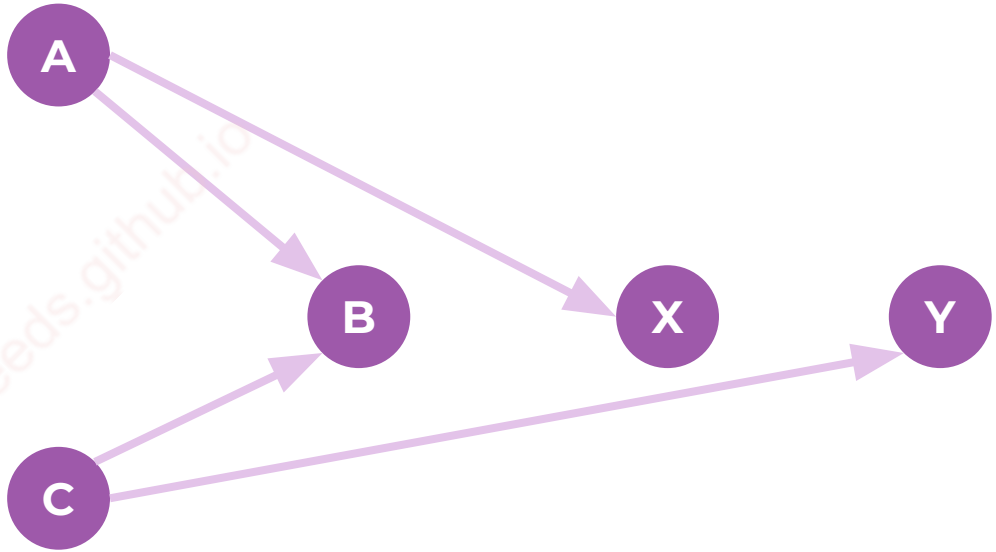


- One backdoor path
- X ← B → Y
- Control B

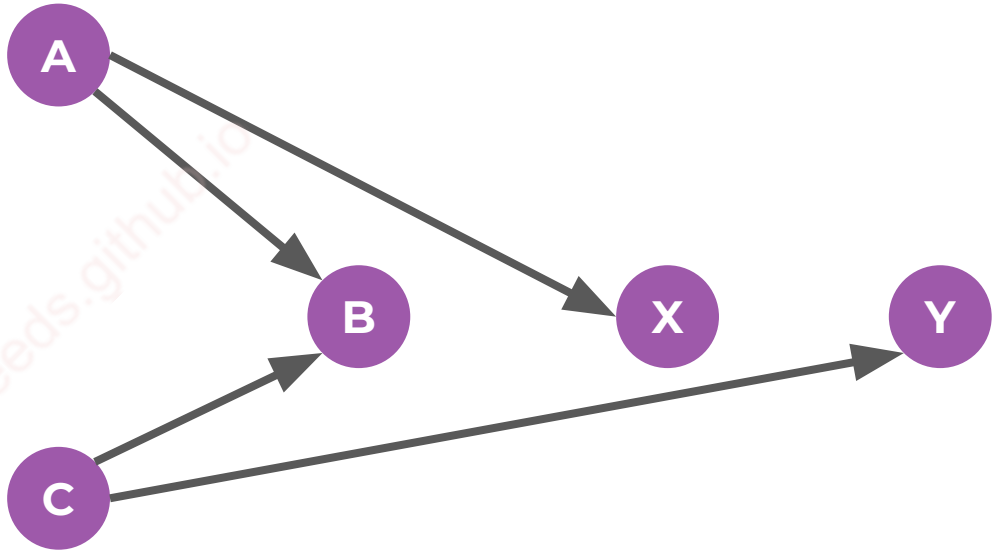




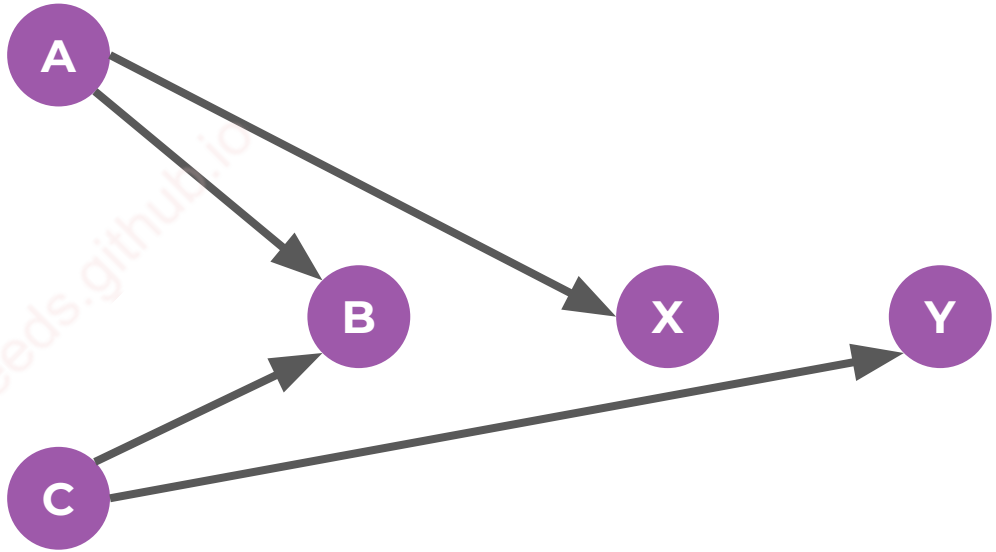
- One backdoor path
  - X ← A → B ← C → Y



- One backdoor path
  - X ← A → B ← C → Y
- Control Ø

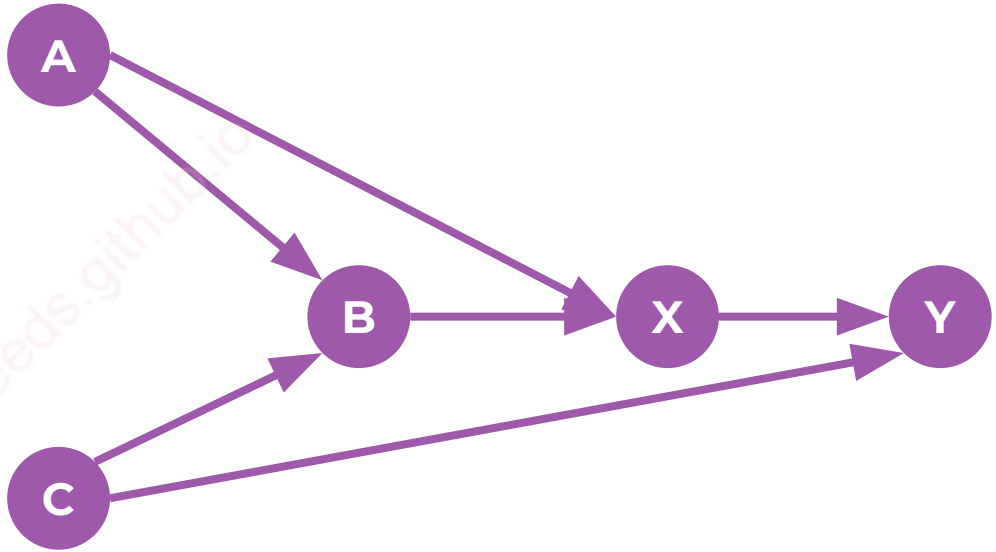


- One backdoor path
  - X ← A → B ← C → Y
- Control ∅

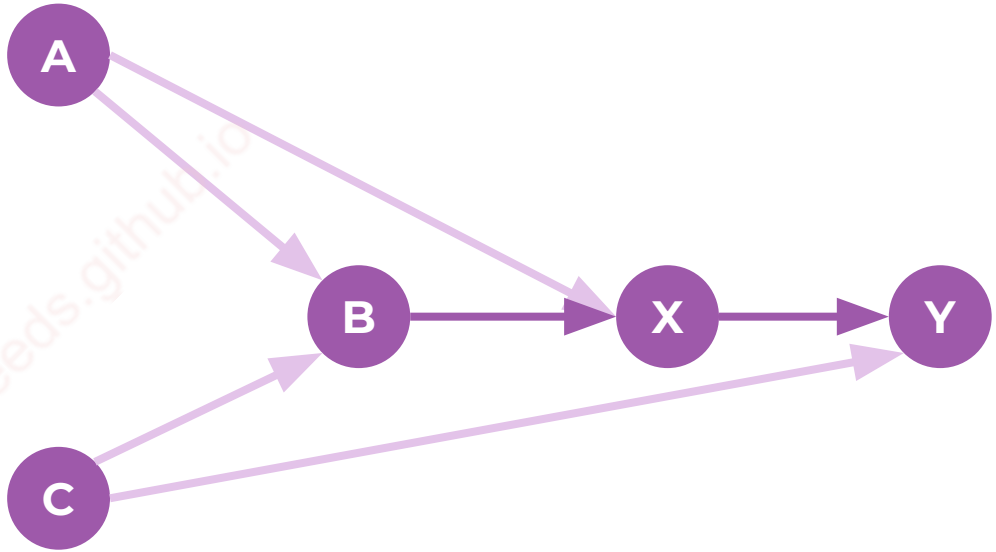


A := Societal norms
B := Seat belt usage
C := Safety and health related measures

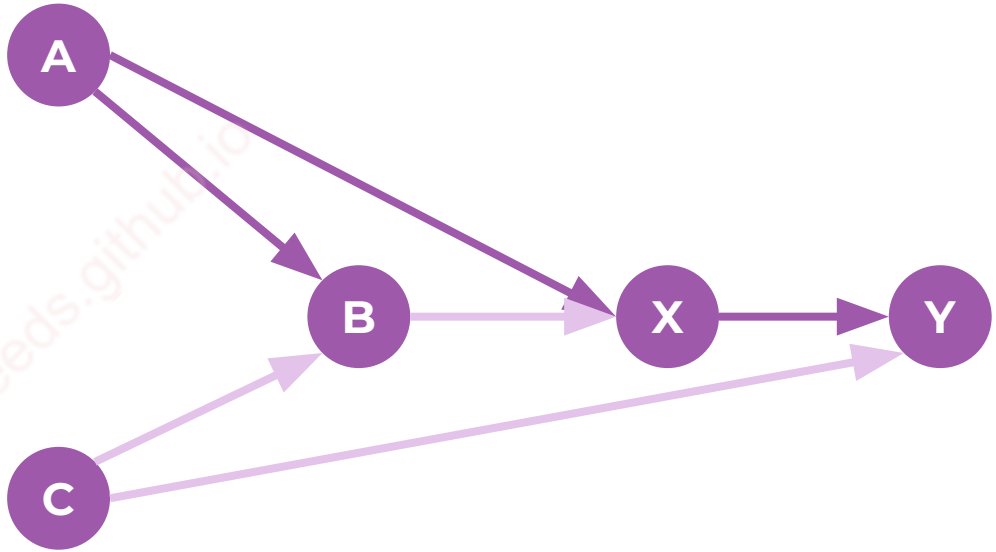
X := Smoking
Y := Lung disease



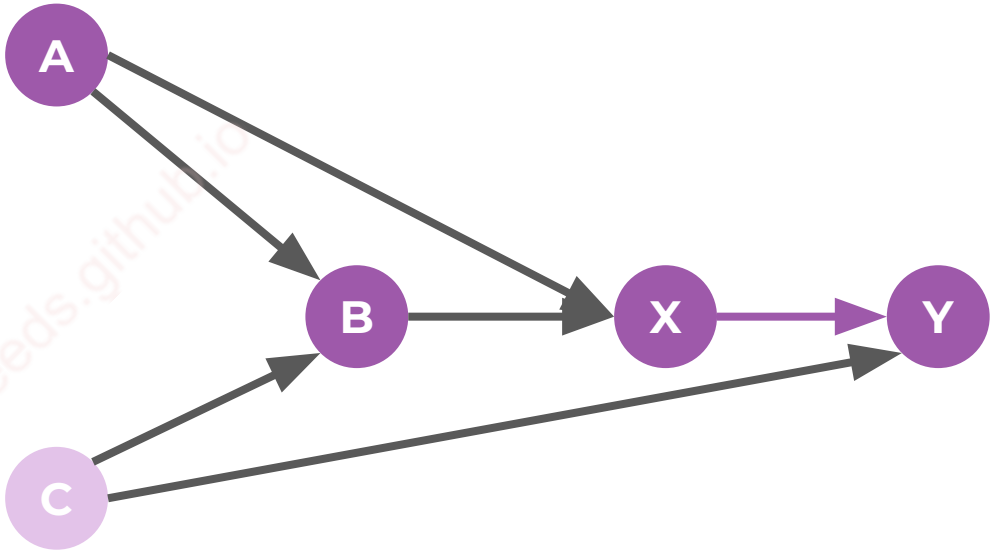
- Two backdoor paths
  - X ← A → B ← C → Y

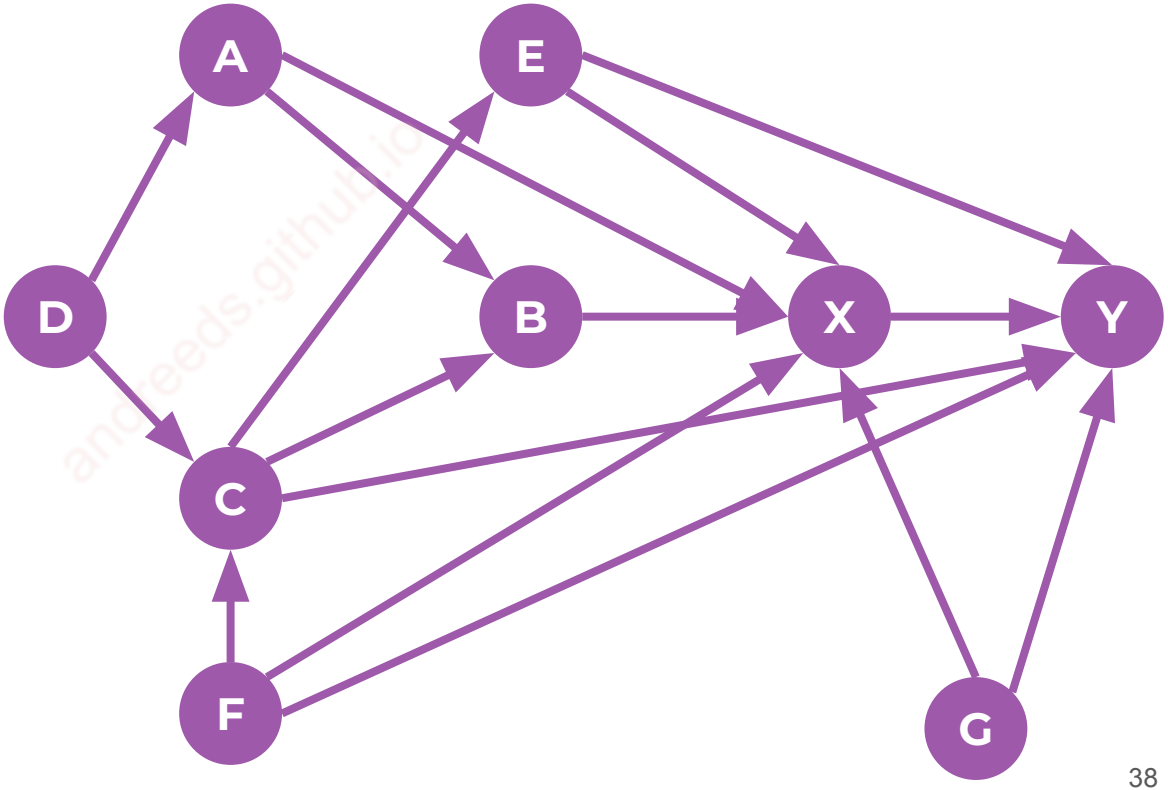


- Two backdoor paths
  - X ← A → B ← C → Y
  - X ← B ← C → Y



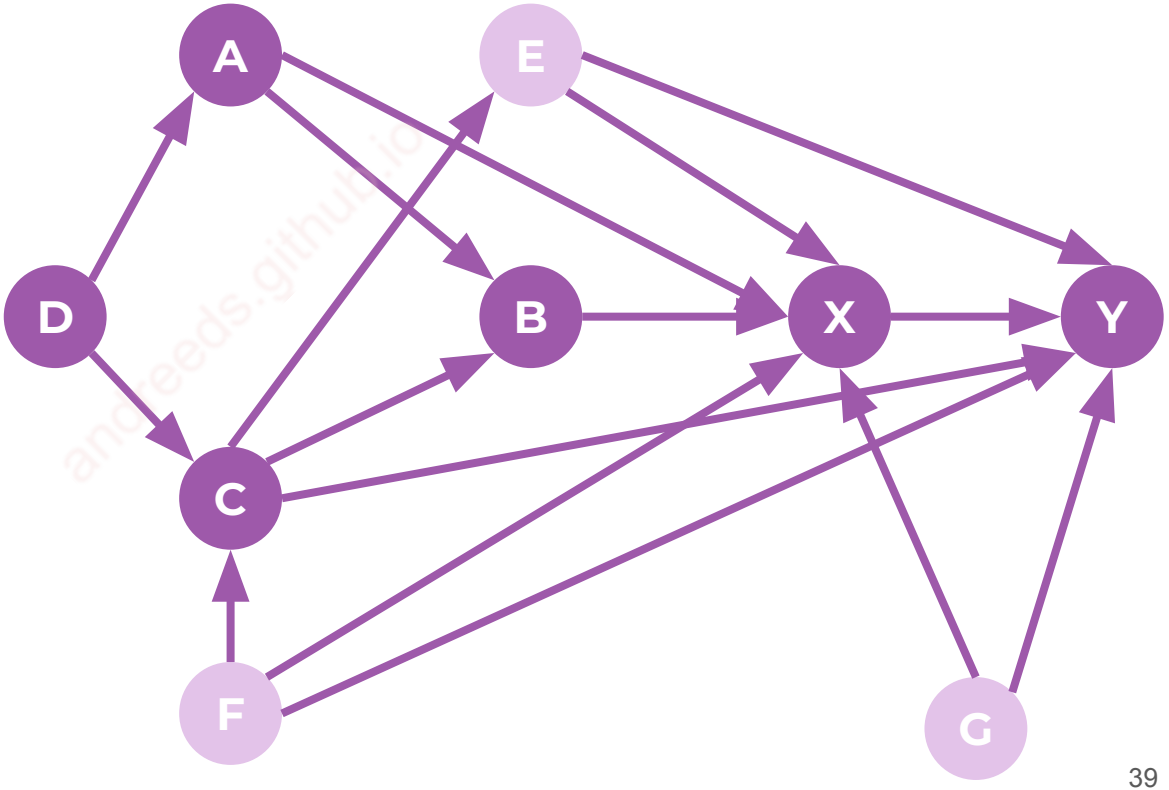
- Two backdoor paths
  - X ← A → B ← C → Y
  - X ← B ← C → Y
- Control C





A := Societal norms
B := Seat belt usage
C := Safety and health related measures
D := Parental asthma
E := Chronic bronchitis
F := Sex
G := Socioeconomic status
X := Smoking
Y := Lung disease

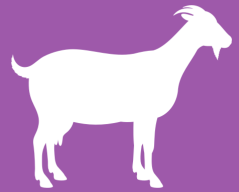
- Control E, F, and G



A := Societal norms
B := Seat belt usage
C := Safety and health related measures
D := Parental asthma
E := Chronic bronchitis
F := Sex
G := Socioeconomic status
X := Smoking
Y := Lung disease

The Monty Hall problem

"Let's make a deal"

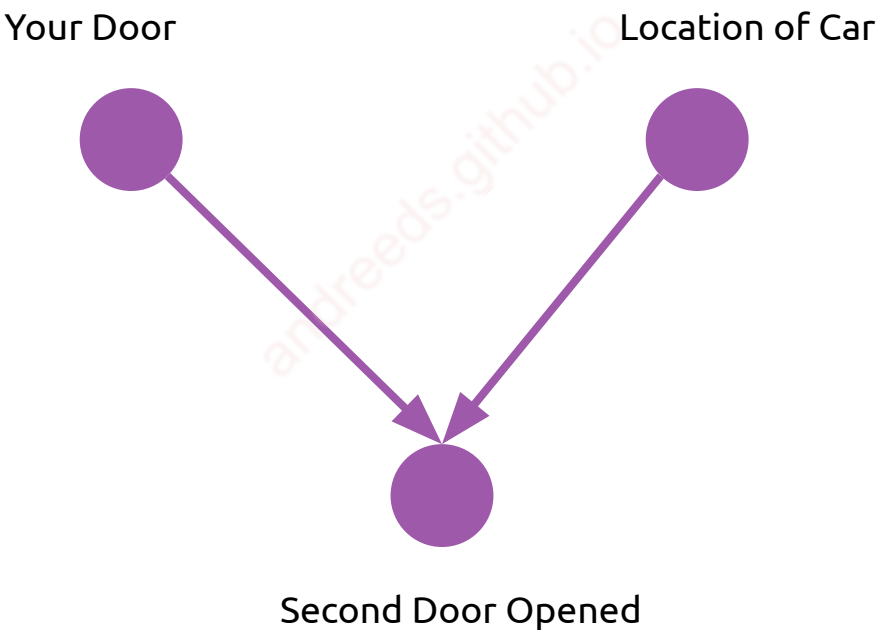


The Monty Hall problem

Chosen

Door 1	Door 2	Door 3	Outcome if **switch**	Outcome if **stay**
			Lose	Win
			Win	Lose
			Win	Lose

The Monty Hall problem



The Monty Hall problem 2.0

"Let's **fake** a deal"

1. 2. 3.

The Monty Hall problem 2.0

"Let's **fake** a deal"

Door You Choose	Door with Car	Door Opened	Outcome if switch	Outcome if stay
1	1	2 	Lose	**Win**
1	1	3 	Lose	**Win**
1	2	2 	Lose	Lose
1	2	3 	**Win**	Lose
1	3	2 	**Win**	Lose
1	3	3 	Lose	Lose

The Monty Hall problem 2.0

"Let's **fake** a deal"

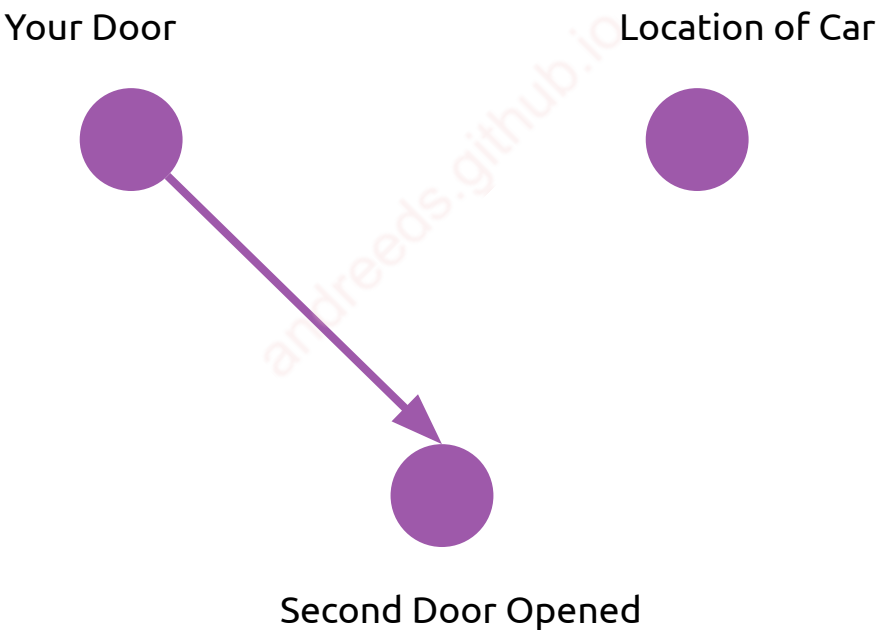
Door You Choose	Door with Car	Door Opened	Outcome if switch	Outcome if stay
1	1	2	Lose	**Win**
1	1	3 	Lose	**Win**
1	2	2 	Lose	Lose
1	2	3 	**Win**	Lose
1	3	2	**Win**	Lose
1	3	3 	Lose	Lose

The Monty Hall problem 2.0

"Let's **fake** a deal"

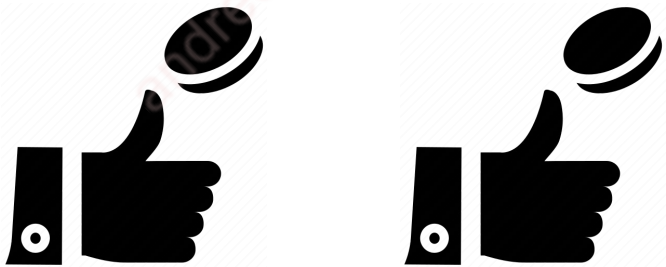
Door You Choose	Door with Car	Door Opened	Outcome if switch	Outcome if stay
1	1	2 	Lose	Win
1	1	3	Lose	Win
1	2	2 	Lose	Lose
1	2	3	Win	Lose
1	3	2 	Win	Lose
1	3	3 	Lose	Lose

The Monty Hall problem



2 coins and 1 date

Berkson's Paradox



Write down the results only when **at least** one of them comes up **heads**

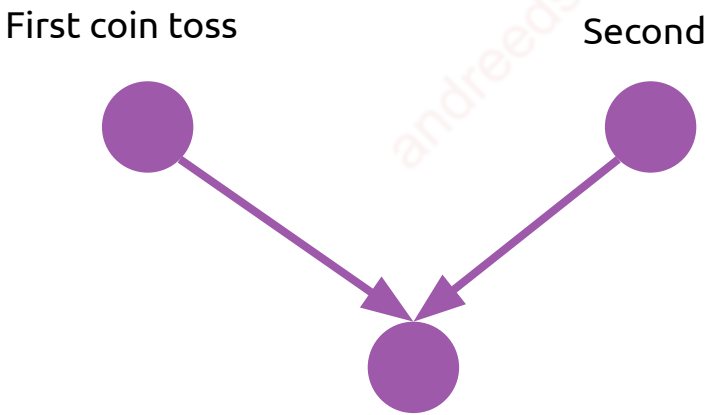
2 coins and 1 date

Berkson's Paradox

Coin 1	Coin 2	Ω
Heads	Heads	25
Heads	Tails	23
Tails	Heads	27

2 coins and 1 date

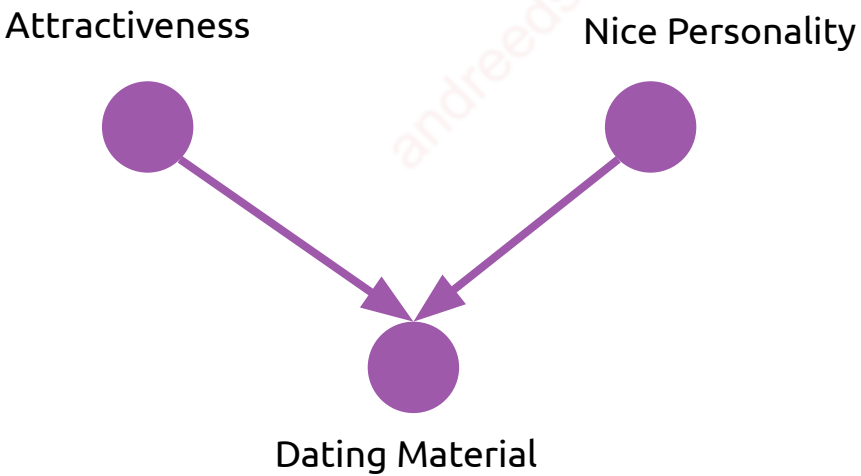
Berkson's Paradox



Coin 1	Coin 2	Ω
Heads	Heads	25
Heads	**Tails**	**23**
Tails	Heads	27
Tails	Tails	X

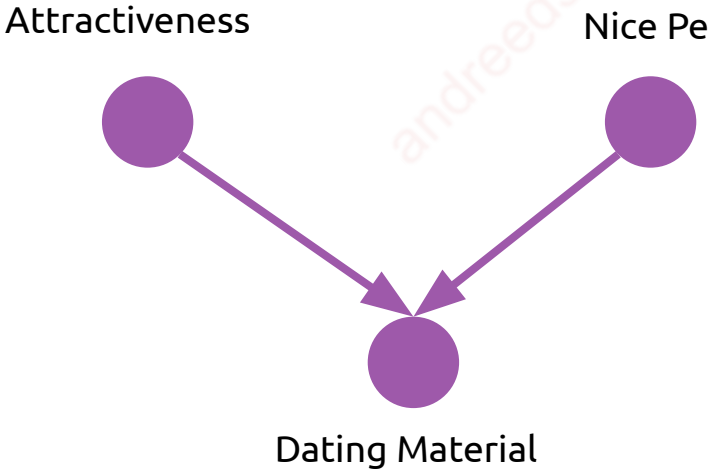
2 coins and 1 date

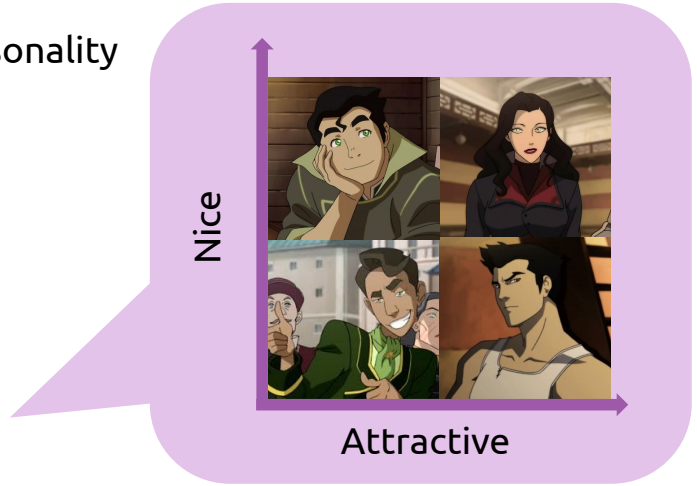
"How Not to Be Wrong" by Jordan Ellenberg



2 coins and 1 date

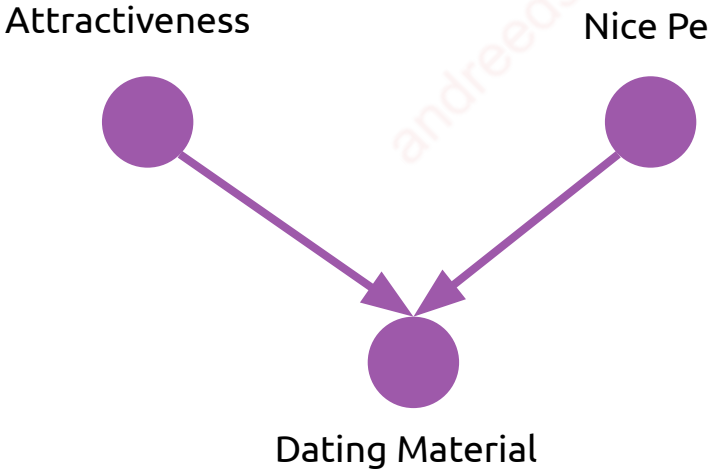
"How Not to Be Wrong" by Jordan Ellenberg

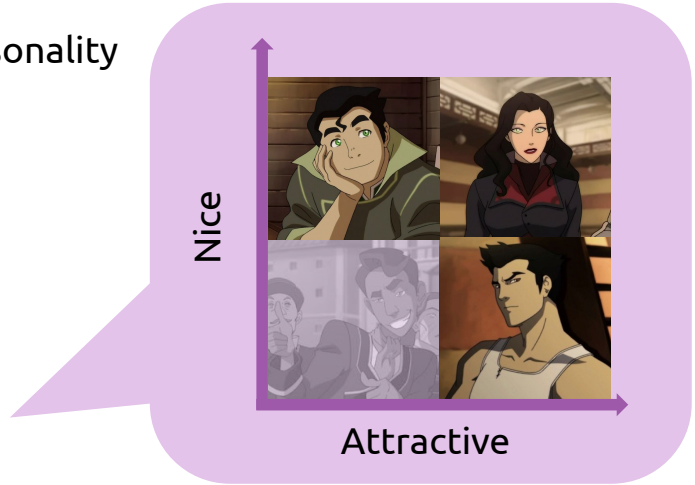




2 coins and 1 date

"How Not to Be Wrong" by Jordan Ellenberg





CAUSALITY

PART I - CONFOUNDING

André dos Santos, *Ph.D.*