From Description to Causation

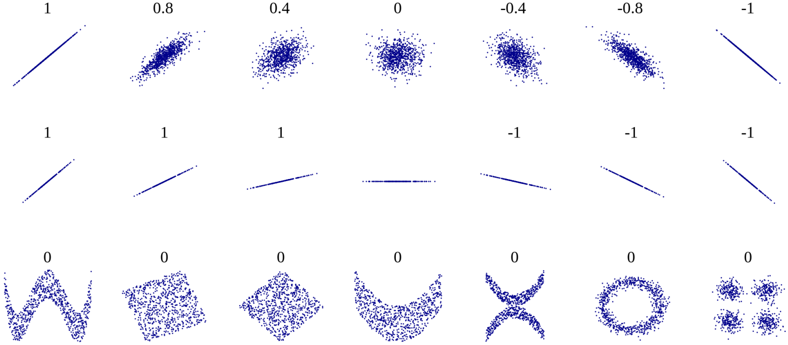
Department of Government
London School of Economics and Political Science

What makes something a *cause*?

Write for 1 minute.

Correlation

Correlation is the *non-independence* of two variables for a set of observations


1
0.8
0.4
0
-0.4
-0.8
-1
1
1
1
-1
-1
-1
0
0
0
0
0
0
0

Correlation

- Synonyms: correlation, covariation, relationship, association
- Any correlation is potentially causal
  - X might cause Y
  - Y might cause X
  - X and Y might be caused by Z
  - X and Y might cause Z
  - There may be no causal relationship

Flashback!

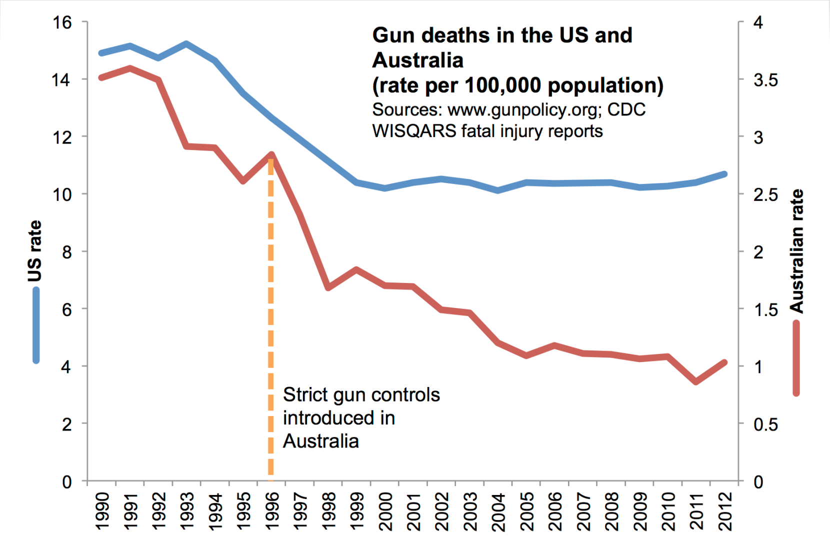
Two Categories of Inference:

1. Descriptive Inference
   - What are the facts?
2. Causal Inference
   - Why does something occur?

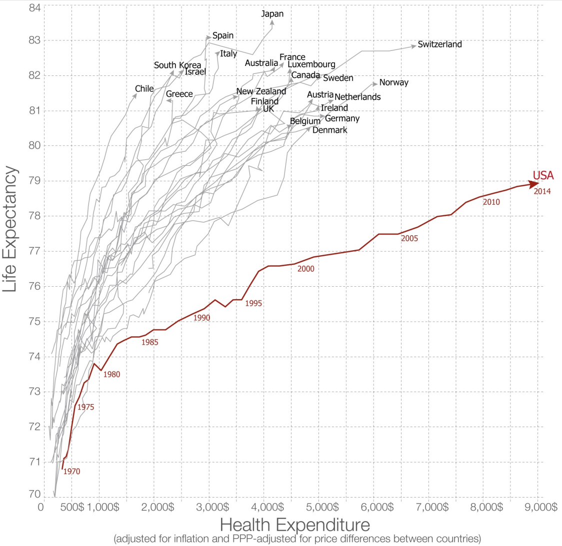
Correlation is Causation?

The mind tends to interpret correlations and patterns as evidence of causal relationships!

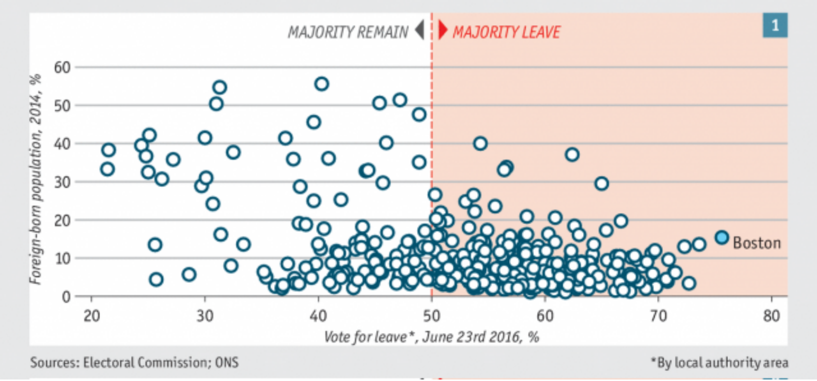
But this is rarely correct!



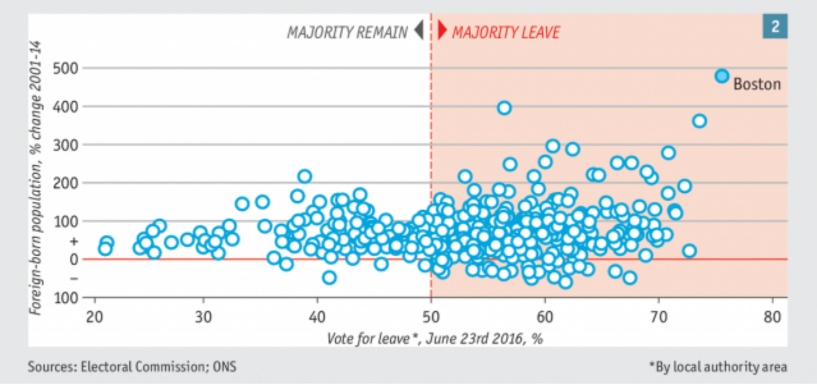
Source: Wikimedia Commons



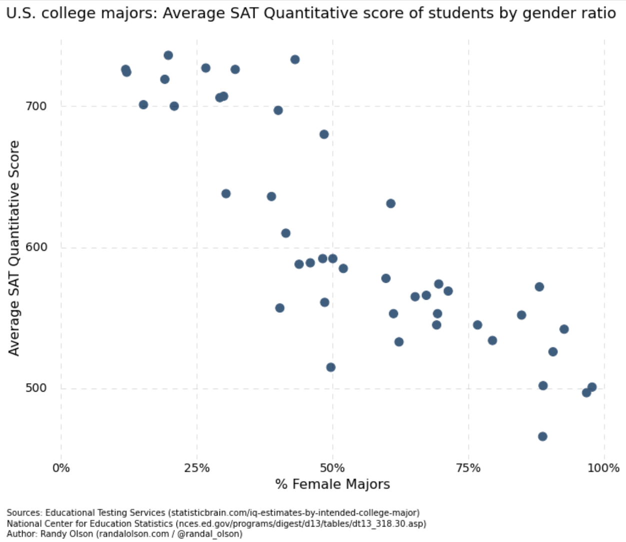
Source: Wikimedia Commons



Source: *The Economist*, 8 July 2016



Source: *The Economist*, 8 July 2016



Source: Randal Olson

Table A: Proportion of individuals at different stages of the CJS process by ethnic group compared to general population, England and Wales

	White	Black	Asian	Mixed	Chinese or Other	Unknown	Total
Population aged 10 or over 2009	88.6%	2.7%	5.6%	1.4%	1.6%	-	**48,417,349**
Stop and Searches (s1) 2009/10	67.2%	14.6%	9.6%	3.0%	1.2%	4.4%	**1,141,839**
Arrests 2009/10	79.6%	8.0%	5.6%	2.9%	1.5%	2.4%	**1,386,030**
Cautions 2010(1)	83.1%	7.1%	5.2%	-	1.8%	2.8%	**230,109**
Court order supervisions 2010	81.8%	6.0%	4.9%	2.8%	1.3%	3.2%	**161,687**
Prison population (including foreign nationals) 2010	72.0%	13.7%	7.1%	3.5%	1.4%	2.2%	**85,002**

Note:
1. Data based on ethnic appearance and therefore do not include the Mixed category.

Source: Ministry of Justice, "Statistics on Race and the Criminal Justice System 2010"

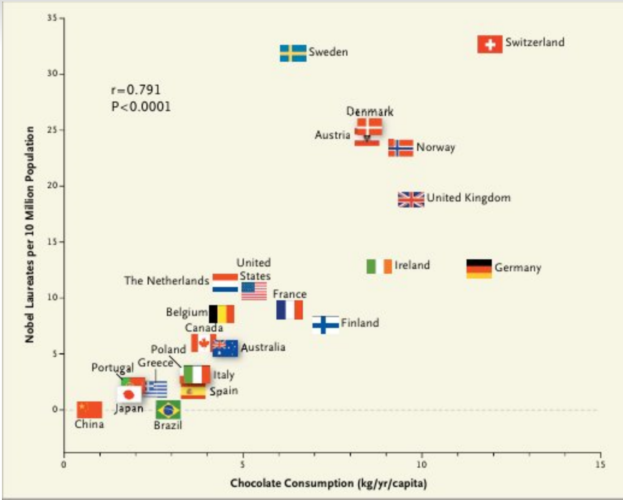
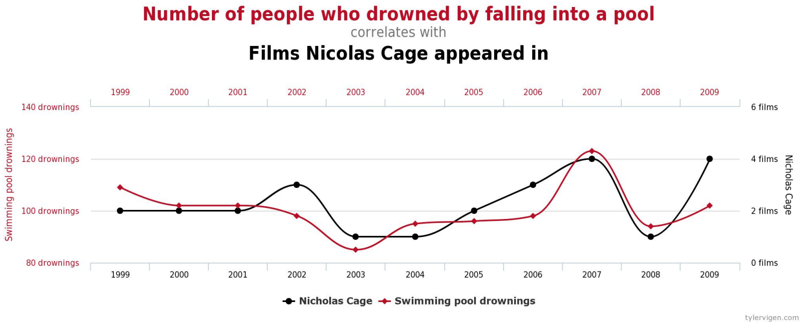


Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.

Source: StackExchange



Source: TylerVigen.com

Naive Causal Inference

- Correlations are not necessarily causal
- Our mind thinks they are because humans are not very good at the kind of causal inference problems that social scientists care about
- Instead, we're good at understanding *physical* causality

Physical causality

- Action and reaction

Physical causality

- Action and reaction
- Example:
  - Picture a ball resting on top of a hill
  - What happens if I push the ball?

Physical causality

- Action and reaction
- Example:
  - Picture a ball resting on top of a hill
  - What happens if I push the ball?
- Features:
  - Observable
  - Single-case
  - Deterministic
  - Monocausal

Pre-Post Change Heuristic

- Our intuition about causation relies too heavily on simple comparisons of *pre-post change* in outcomes before and after something happens
  - No change: no causation
  - Increase in outcome: positive effect
  - Decrease in outcome: negative effect

Pre-Post Change Heuristic

- Our intuition about causation relies too heavily on simple comparisons of *pre-post change* in outcomes before and after something happens
  - No change: no causation
  - Increase in outcome: positive effect
  - Decrease in outcome: negative effect

- Why is this flawed?

Threats to Validity

Campbell and Ross talk about six "threats to validity" (i.e., threats to causal inference) related to time-series analysis

Flaws in causal inference from pre-post comparisons

1. Maturation or trends
2. Regression to the mean
3. Selection
4. Simultaneous historical changes
5. Instrumentation changes
6. Monitoring changes behaviour

Maturation or trends

- Is a shift in an outcome before and after a policy change the impact of the policy or a small part of a longer time trend?
- Case Study: Connecticut crackdown on speeding (1955)

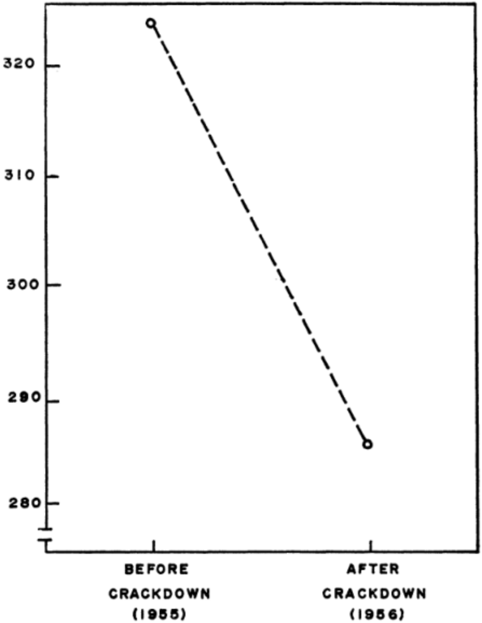


Figure 1. Connecticut Traffic Fatalities, 1955-1956

Regression to the mean

- Is a shift in an outcome before and after a policy change the impact of the policy or a function of statistical variation?
- Case Study: Connecticut crackdown on speeding (1955)

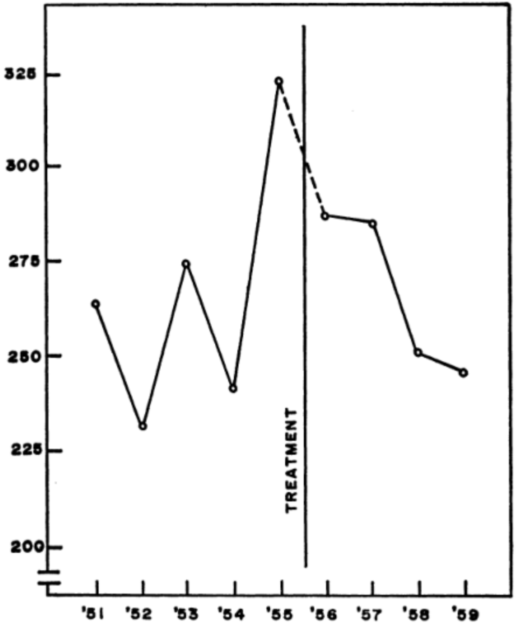


Figure 2. Connecticut Traffic Fatalities, 1951-1959

Selection

- Is a shift in an outcome before and after a policy the impact of the policy or the result of the policy being implemented when outcomes are extreme?
- Case Study: Connecticut crackdown on speeding (1955)

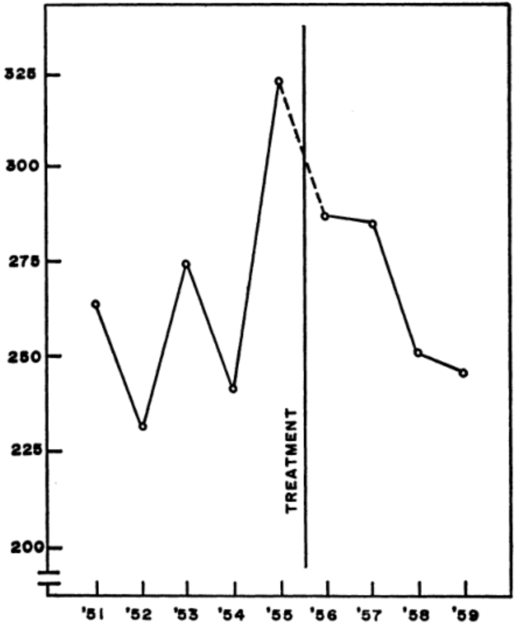
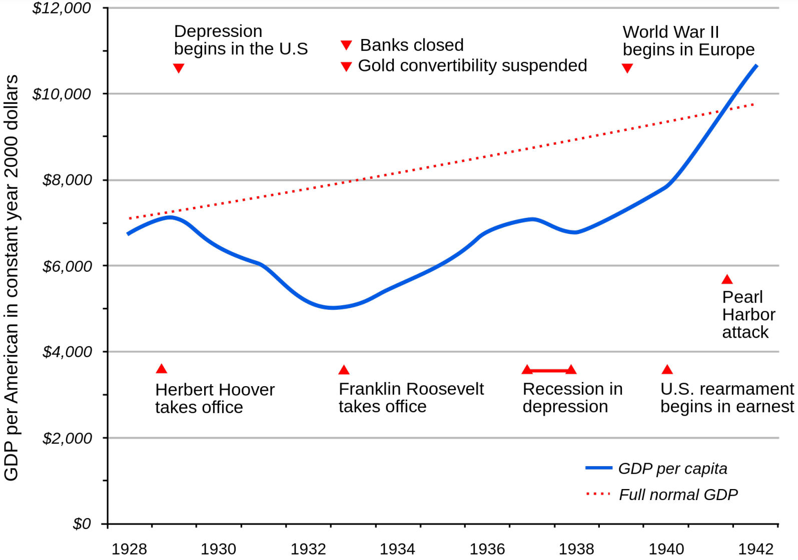


Figure 2. Connecticut Traffic Fatalities, 1951-1959

Simultaneous changes

- Is the shift in an outcome before and after a policy the impact of the policy or the result of a simultaneous historical shift?
- Case Study: US Great Depression Policy


$12,000
$10,000
$8,000
$6,000
$4,000
$2,000
$0
GDP per American in constant year 2000 dollars
1928
1930
1932
1934
1936
1938
1940
1942
Depression begins in the U.S
Banks closed
Gold convertibility suspended
World War II begins in Europe
Pearl Harbor attack
Herbert Hoover takes office
Franklin Roosevelt takes office
Recession in depression
U.S. rearmament begins in earnest
GDP per capita
Full normal GDP

Instrumentation changes

- Is the shift in an outcome before and after a policy the impact of the policy or a change in how the outcome is measured?
- Case Study: Age-adjusted mortality rates

Table 2. Texas 1998 Cancer Mortality Rates (Cases per 100,000), by Cancer Site, Using 1970 and 2000 Standard Populations

	Male			Female		
Cancers	1970	2000	Change (%)	1970	2000	Change (%)
All	202.8	258.9	27.7	131.6	163.7	24.4
Colon and rectum	19.5	25.1	29.2	13.0	17.2	32.9
Lung and bronchus	69.7	85.6	22.7	33.8	40.6	20.1
Prostate	22.4	33.8	50.4			
Breast				22.2	27.0	21.9
Brain, other nervous system	5.1	6.0	18.2	3.5	4.1	16.0

Monitoring changes behaviour

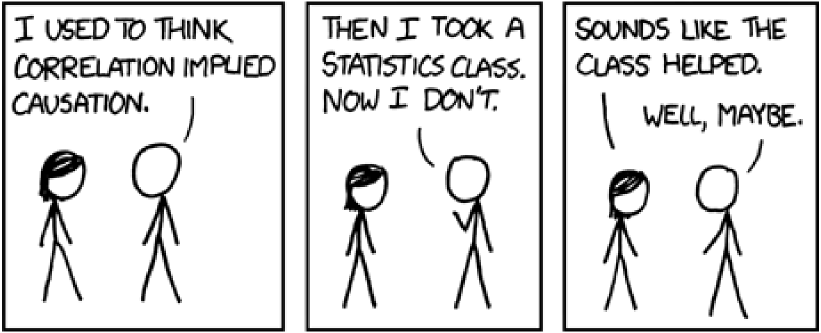
- Is the shift in an outcome before and after a policy the impact of the policy or a change in response to measuring the outcome per se?
- Case Study: Educational testing

The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor.

– Donald T. Campbell (1979)

Flaws in causal inference from pre-post comparisons

1. Maturation or trends
2. Regression to the mean
3. Selection
4. Simultaneous historical changes
5. Instrumentation changes
6. Monitoring changes behaviour


I USED TO THINK CORRELATION IMPLIED CAUSATION.
THEN I TOOK A STATISTICS CLASS. NOW I DON'T.
SOUNDS LIKE THE CLASS HELPED.
WELL, MAYBE.

Directed Acyclic Graphs

- Causal graphs (DAGs) provide a visual representation of (possible) causal relationships

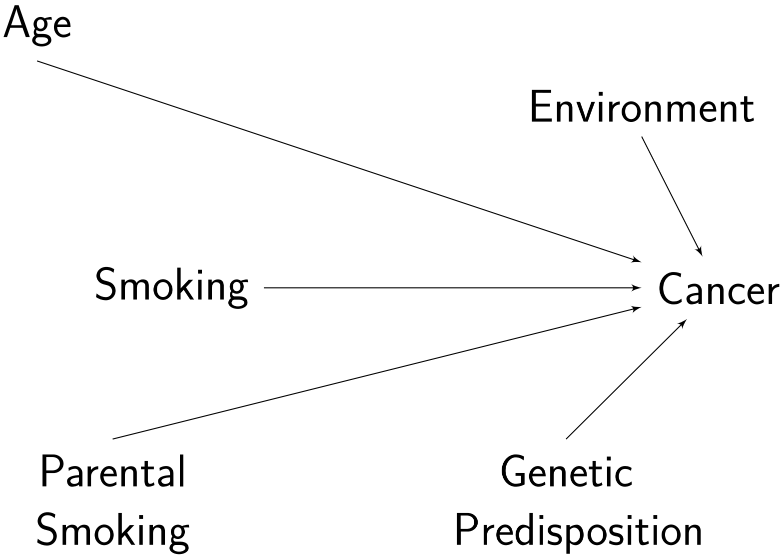
Directed Acyclic Graphs

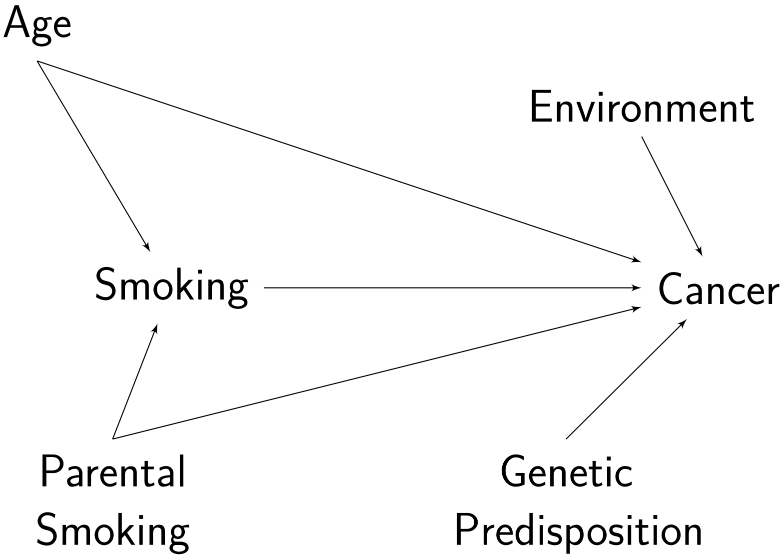
- Causal graphs (DAGs) provide a visual representation of (possible) causal relationships
- Causality flows between variables, which are represented as “nodes”
    - Variables are causally linked by arrows
    - Causality only flows *forward* in time

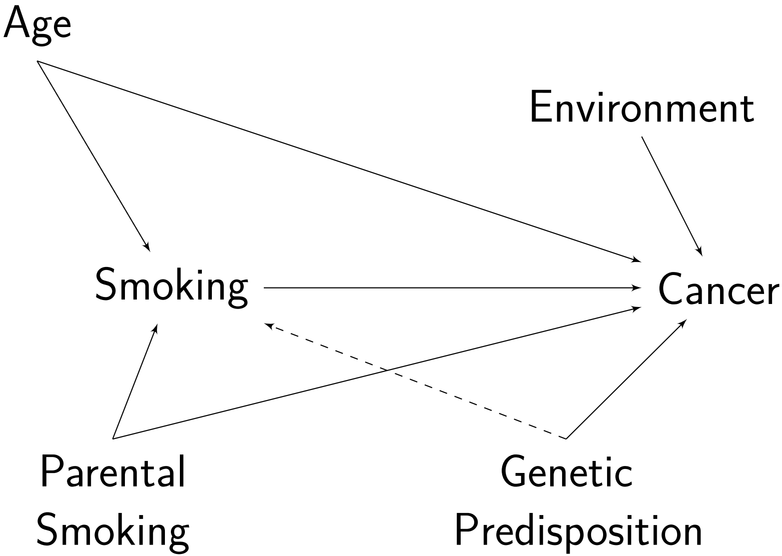
Directed Acyclic Graphs

- Causal graphs (DAGs) provide a visual representation of (possible) causal relationships
- Causality flows between variables, which are represented as "nodes"
  - Variables are causally linked by arrows
  - Causality only flows *forward* in time
- Nodes opening a "backdoor path" from $X \rightarrow Y$ are confounds
  - "Selection bias" or "Confounding"

$$\text{Smoking} \longrightarrow \text{Cancer}$$


Age
Environment
Smoking
Cancer
Parental Smoking
Genetic Predisposition


Age
Environment
Smoking
Cancer
Parental
Smoking
Genetic
Predisposition


Age
Environment
Smoking
Cancer
Parental Smoking
Genetic Predisposition

Causal Inference

Causal inference (typically) involves gathering data in a systematic fashion in order to assess the size and form of correlation between nodes X and Y in such a way that there are no backdoor paths between X and Y by *controlling for* (i.e., *conditioning on*, *holding constant*) any confounding variables, $\mathbf{Z}$.