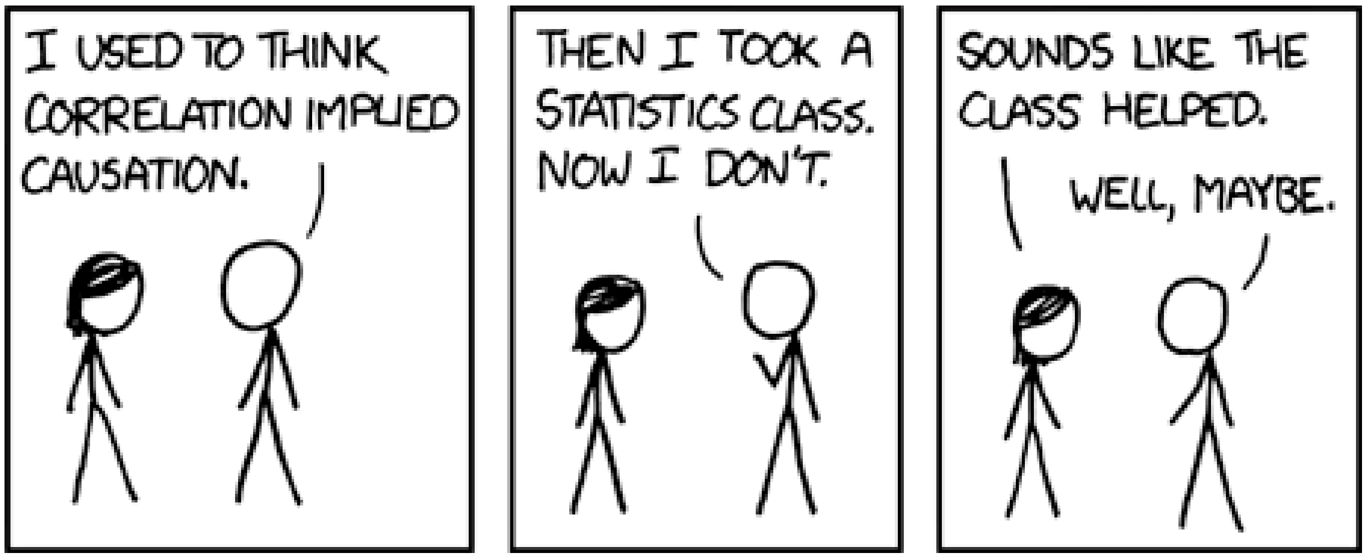
# Causal Inference



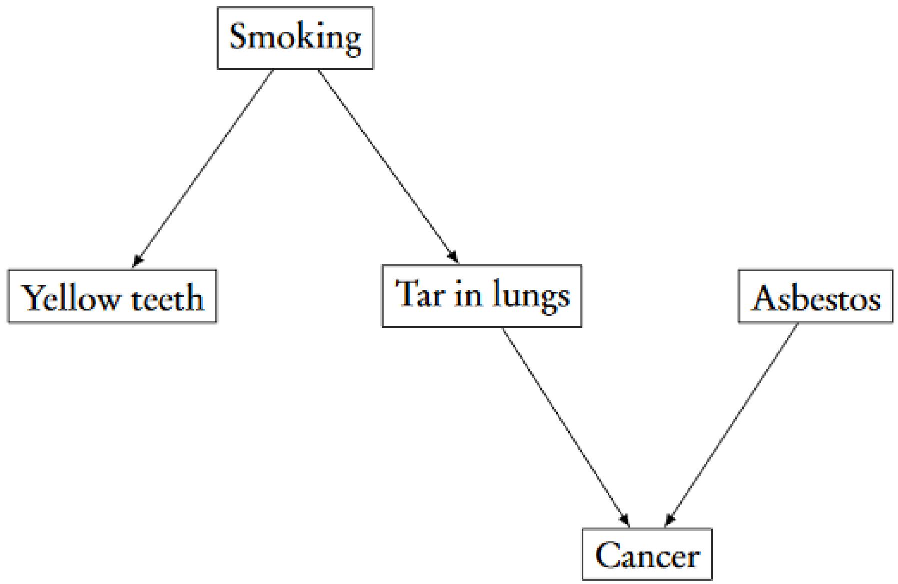

https://xkcd.com/552/

ECE324, Winter 2022

Michael Guerzhoy

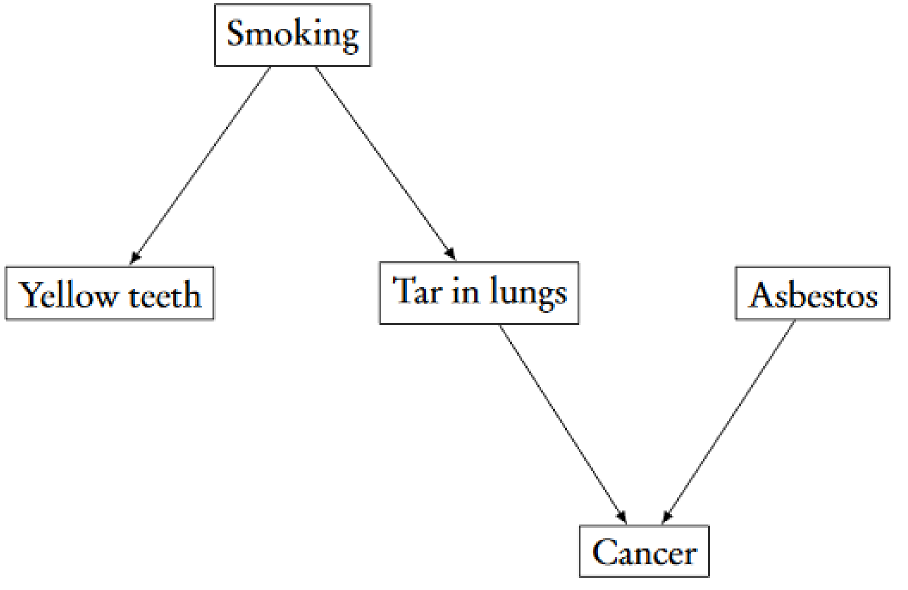
Content from Cosma Shalizi,
*Advanced Data Analysis from an Elementary Point of View*

# Directed Acyclical Graphs (DAG) for representing causal structure



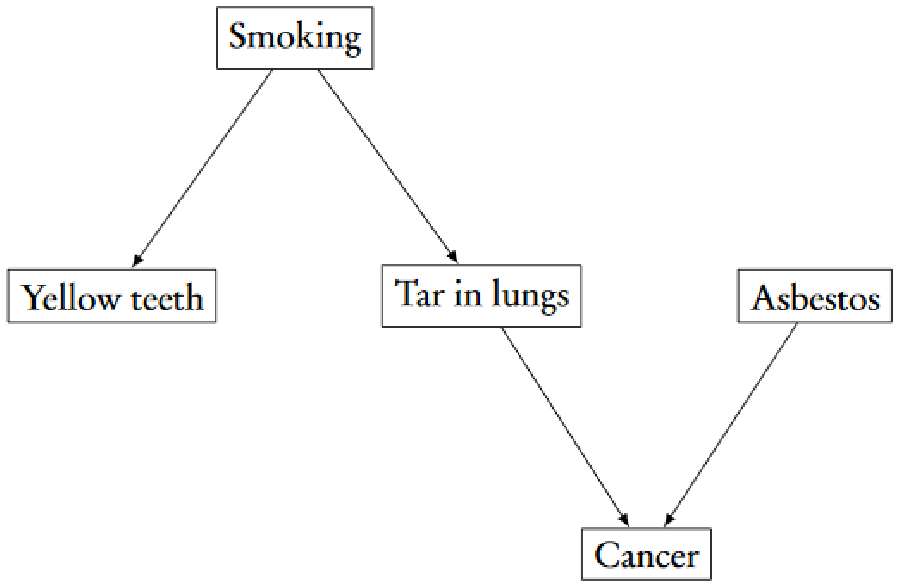

- If we know the value of *Smoking (0 or 1)*, we can generate the value of *Yellow teeth* and *Tar in lungs*
- If we know the value of *Tar in lungs* and *Asbestos*, we can generate the value of *Cancer*
- A datapoint is generated by first generating *Smoking* and *Asbestos,* then *Yellow teeth* and *Asbestos*, then *Cancer*

# Directed Acyclical Graphs (DAG) for representing causal structure



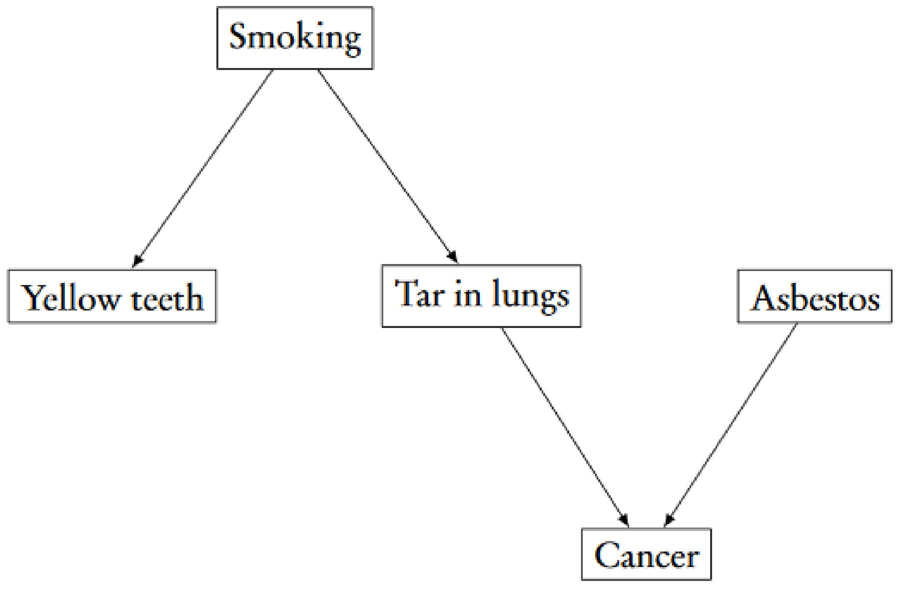

- The graph encodes our knowledge (or assumptions) about the causal structure of the data
- Can help with inference

# Directed Acyclical Graphs (DAG) for representing causal structure



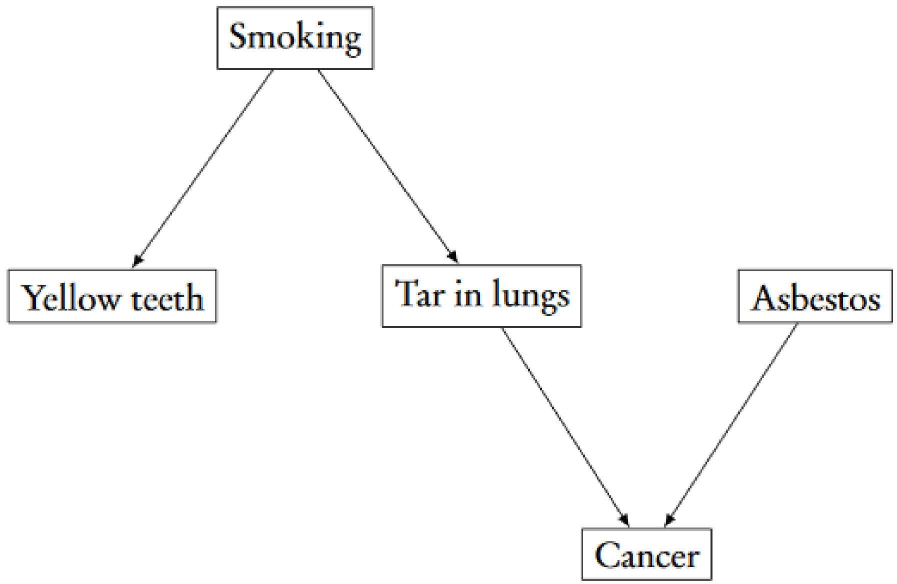

- *Tar* is independent of *Asbestos*
  - They are independently generated
- *Tar* is *not* independent of *Asbestos* given *Cancer*
  - Intuition: if $Cancer = 1$ and $Asbestos = 1$, then $Tar = 1$ is less likely than otherwise, since the cancer is already explained
    - This called "Explaining away"

# Directed Acyclical Graphs (DAG) for representing causal structure




- *Yellow teeth* is not indep. of *Tar*
  - Both caused by *Smoking*
- *Yellow teeth* is indep. of *Tar* conditioned on *Smoking*

# Directed Acyclical Graphs (DAG) for representing causal structure




- The right way to think of the arrows: *Tar might* cause *Cancer*, or have no effect
- *Smoking* is *definitely* independent of *Asbestos*

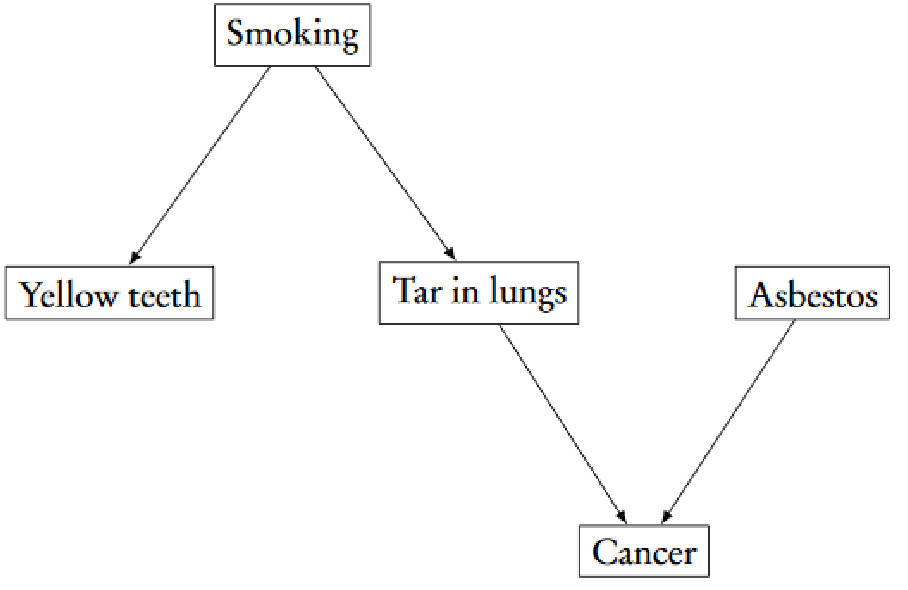
# Causation

- Define "A caused B"

# Causation

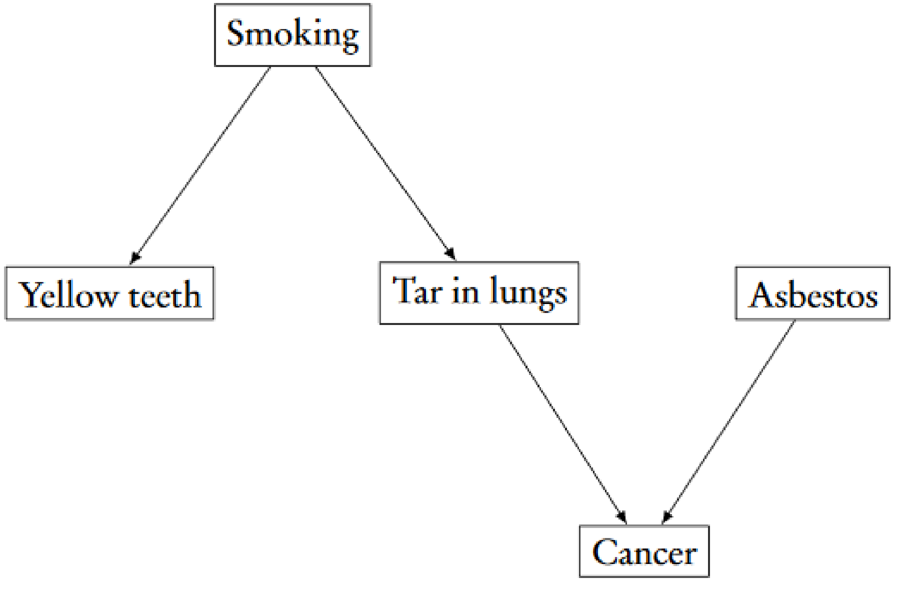
- Counterfactual: we say that *A* causes *B* if *B would not have happened* if *A* had not happened
- Causal inference: trying to answer causal questions from empirical data
    - Difficult to derive counter-factual conclusions from factual premises

# Effect of causes




- *What is the causal relationship between exposure to asbestos and yellow teeth?*
  - There is none!
  - $Yellow$ and $Asbestos$ are not indep. conditioned on Cancer
    - Explaining away phenomenon
    - Can predict $Yellow$ from $\{Asbestos, Cancer\}$

# Effect of causes



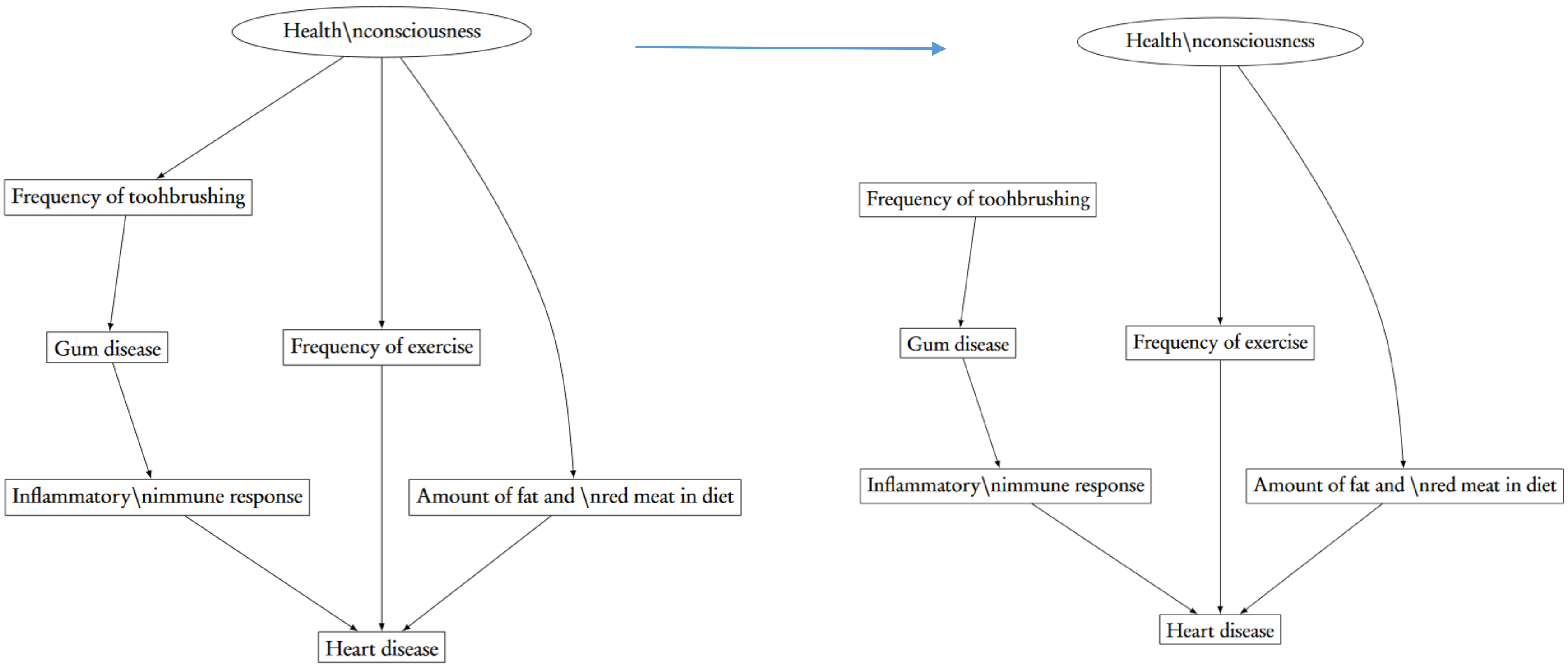

- A way of thinking about this:

$$P(Yellow|Asbestos = 1, Cancer = 1) \neq P(Yellow|Cancer = 1)$$

$$P(Yellow|Asbestos = 1, do(Cancer = 1)) = P(Yellow|do(Cancer = 1)) = P(Yellow)$$

- $do(Cancer = 1)$ sets $Cancer$ to 1, and changes the causal graph eliminating the mechanism that generates *Cancer*

# do(brushing)




Again, $\mathrm{P}(Heart\ disease|Brushing = b) \neq \mathrm{P}(Heart\ disease|do(Brushing = b))$

# Identifying Causal Effects from Observations

- The most straightforward way to compute $P(Y|do(X = x))$ is to manipulate $x$ physically and see what happens to $Y$
  - Run an experiment
    - Hold all other variables constant

      Or
    - Randomize all other variables

# Identification

- Want to calculate the causal effect of $X$ on $Y$ (i.e., $P(Y|do(X = x))$, but can't run an experiment.
- Can do this if we have the causal graph and observe all the variables
    - Saw this in the toothbrushing example
- Can *sometimes* do this if not all variables are observed
    - Need to carefully look at the graph