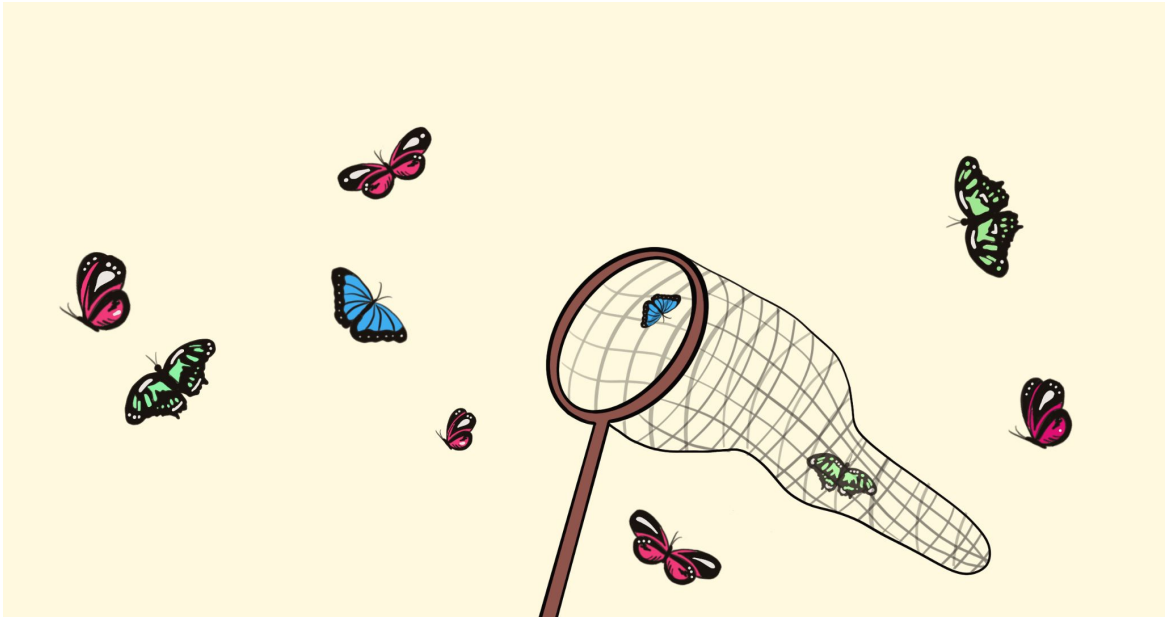
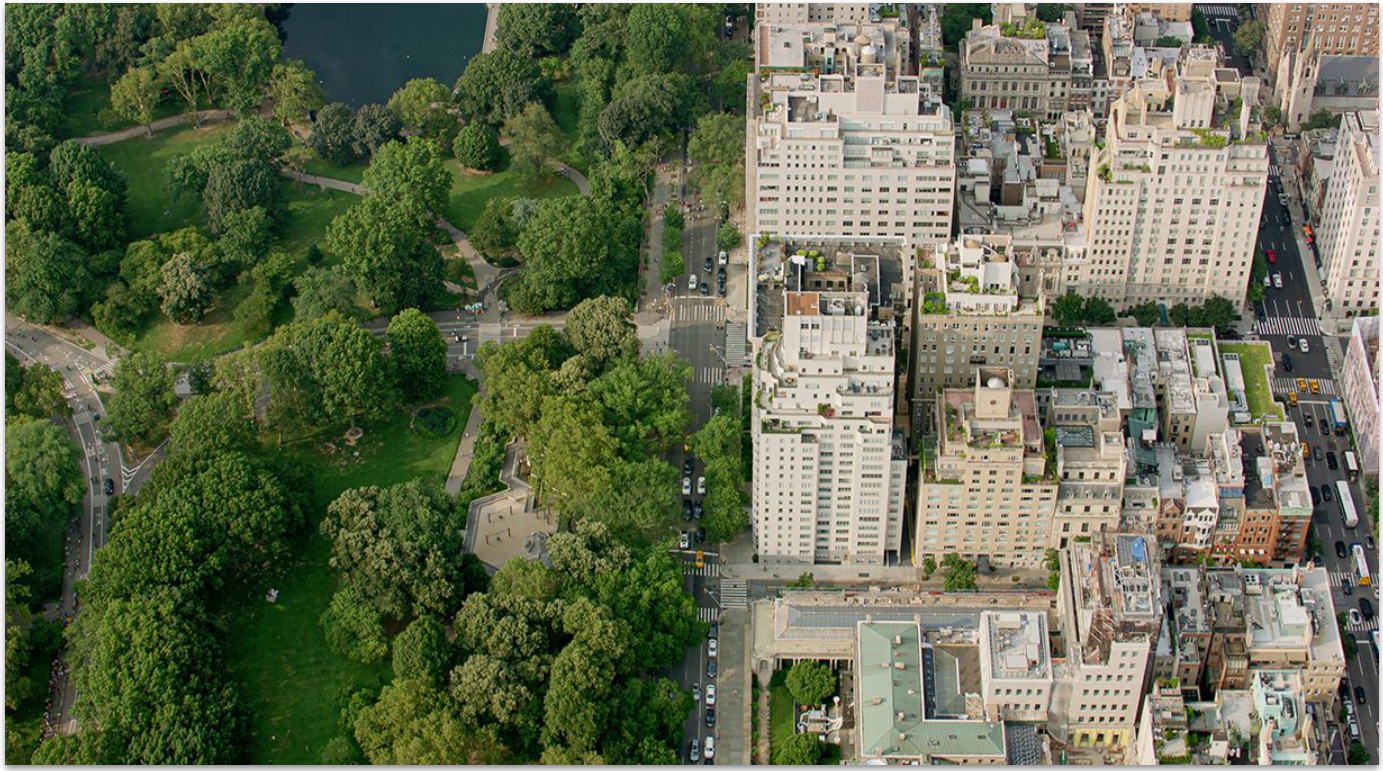
Week 6 Lecture 1:
Forks, pipes, and colliders

EDS 222: *Statistics for Environmental Data Science*



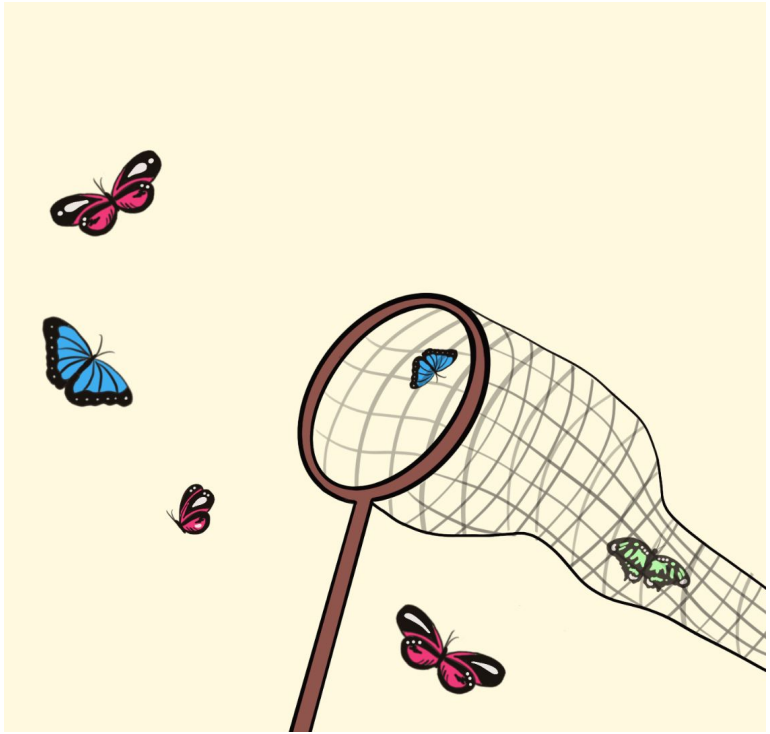
Urban heat islands



EPA

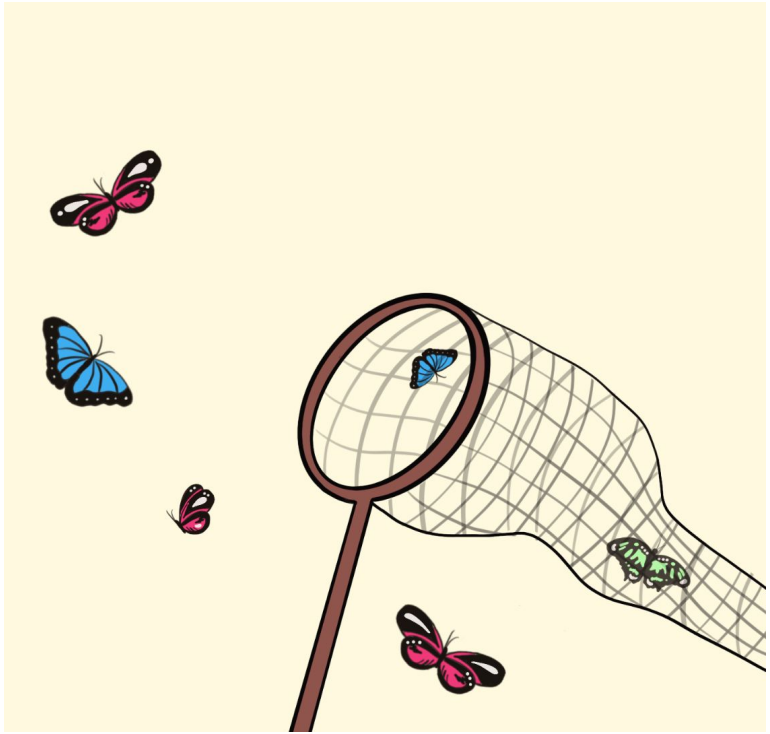
Today’s agenda

- ➔ DAGs and causal relationships
- ➔ Choosing variables
- ➔ Your final project



Today’s agenda

- ➔ **DAGs and causal relationships**
- ➔ Choosing variables
- ➔ Your final project



Density and UHIs

Think-pair-share

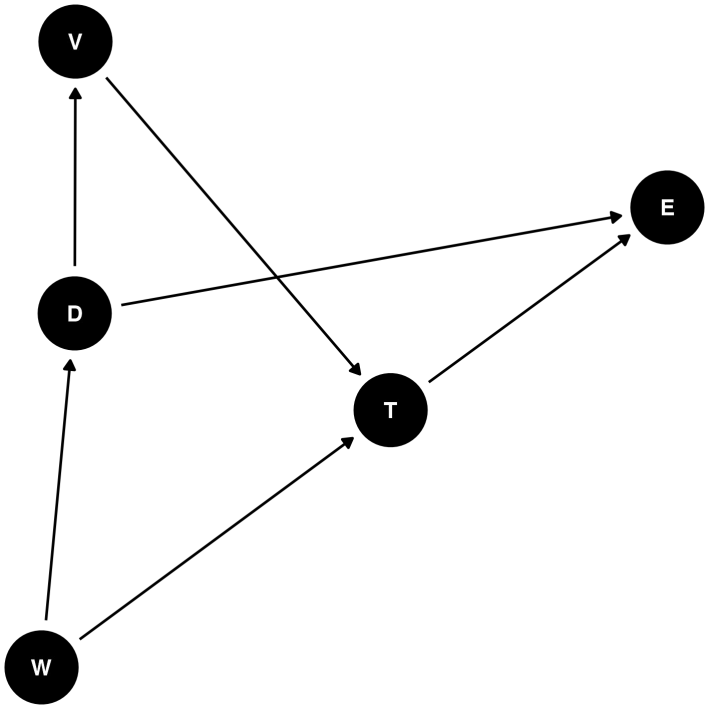
Does urban density increase the urban heat island effect?

1. What **response** variable would you collect?
2. What **predictor** variables are relevant?
3. What **causal relationships** exist between the variables?

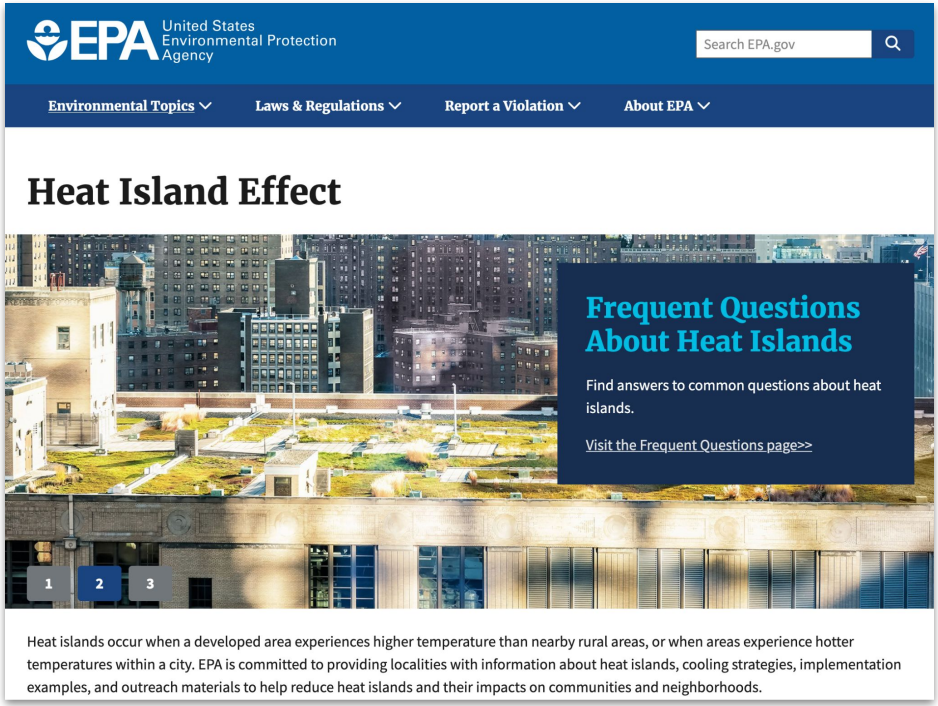
Regression = association

Simple DAGs

State your assumptions

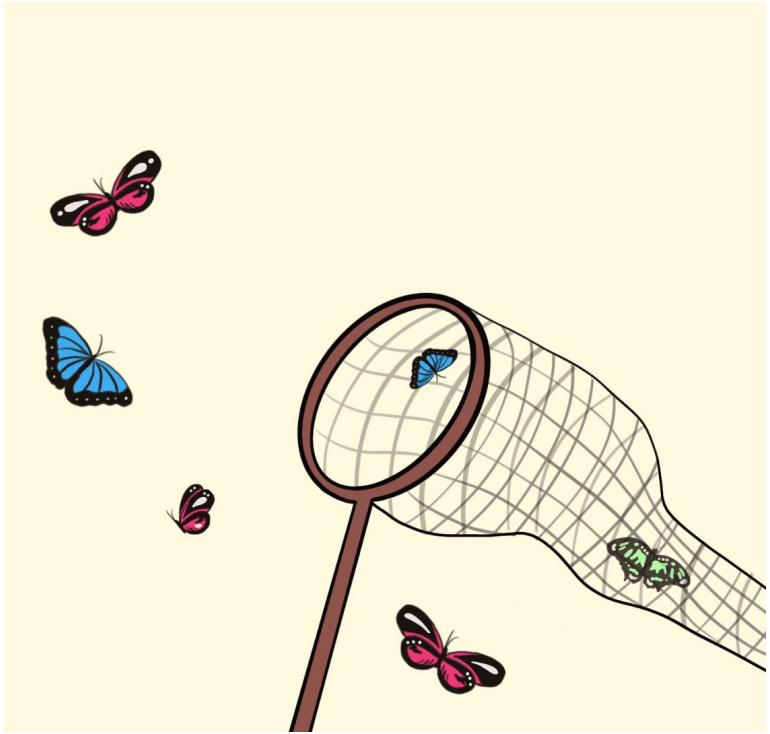


Where do relationships come from?

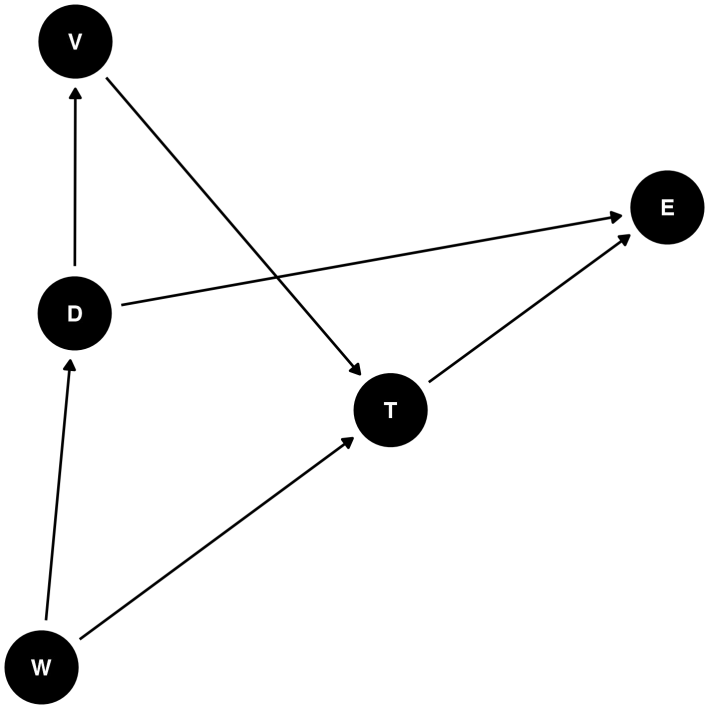


Today’s agenda

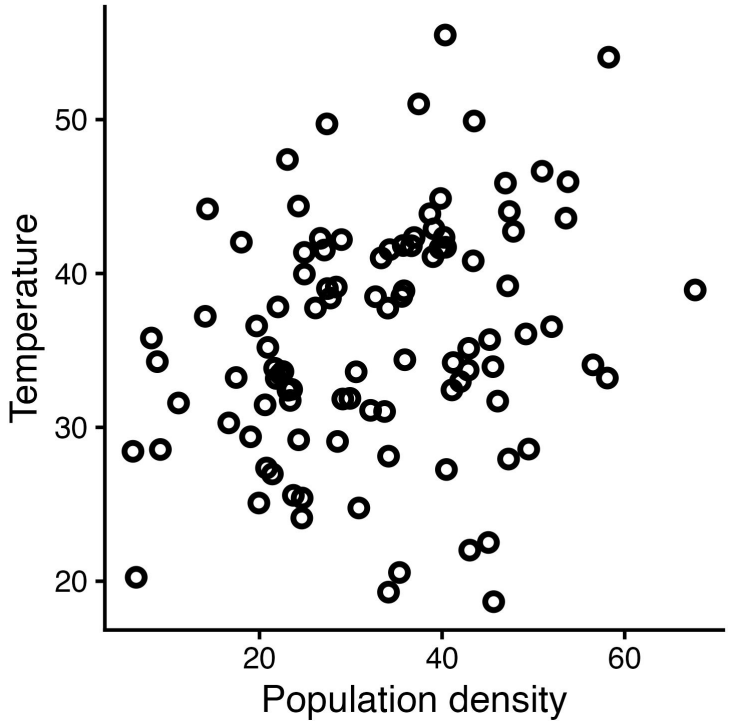
- ➔ DAGs and causal relationships
- ➔ **Choosing variables**
- ➔ Your final project



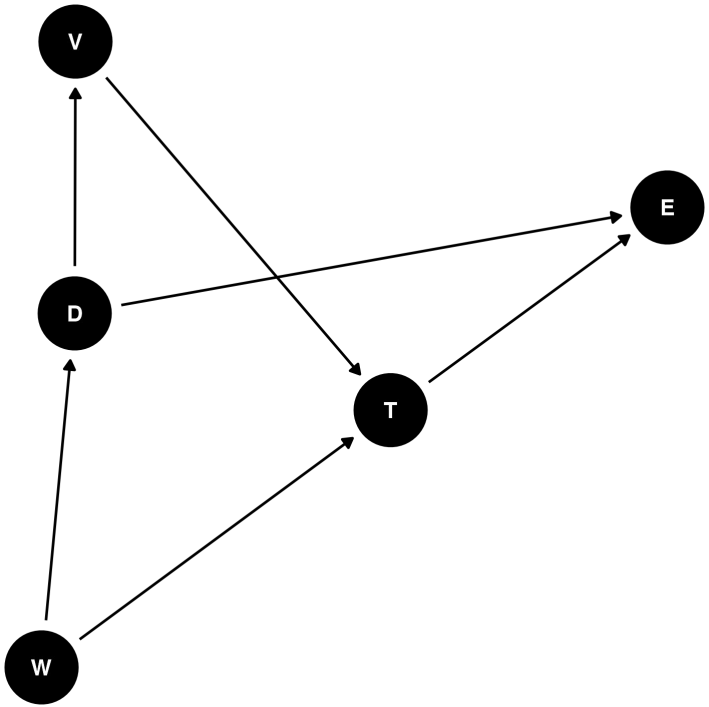
Forks, pipes, and colliders



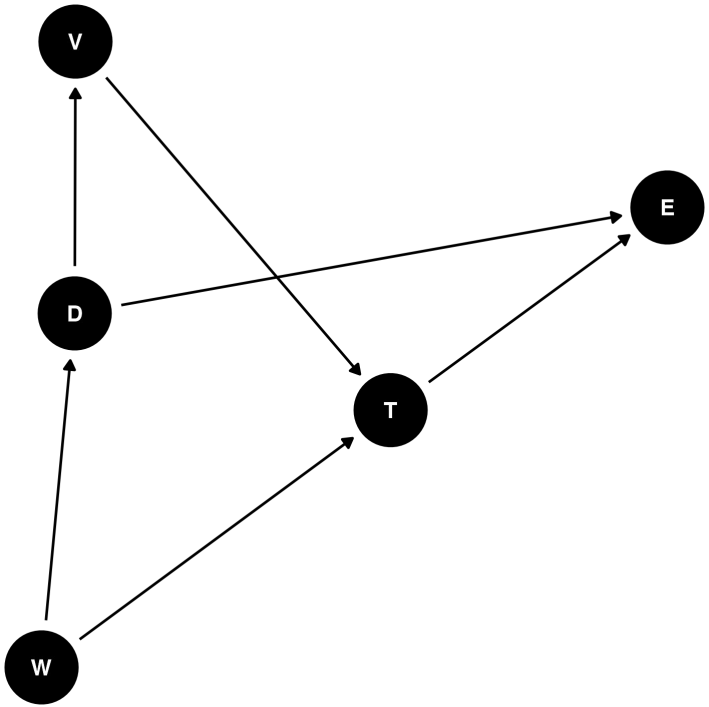
The data generating process



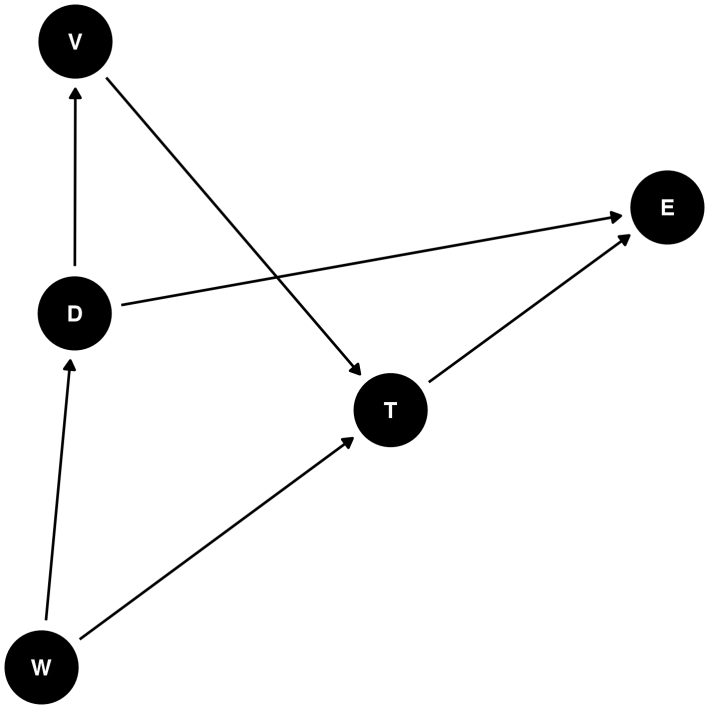
The fork



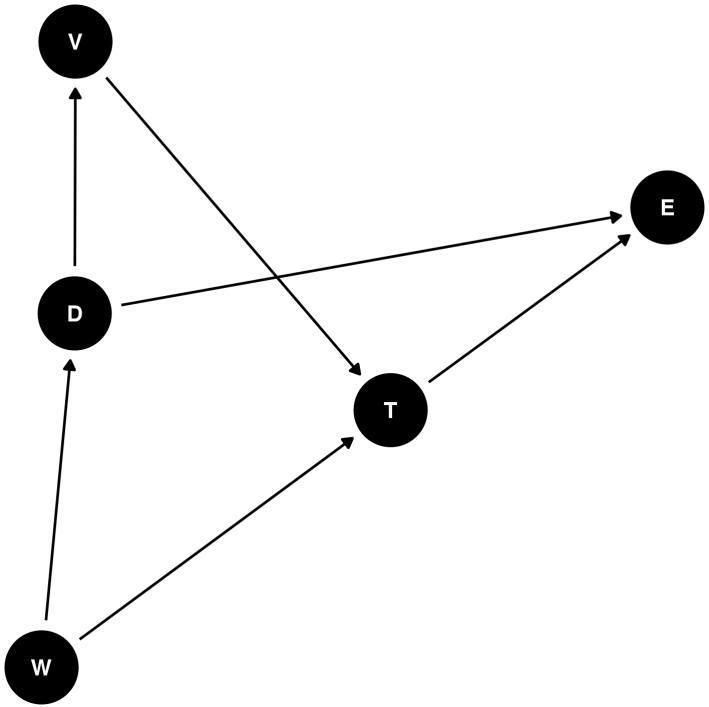
The pipe



The collider

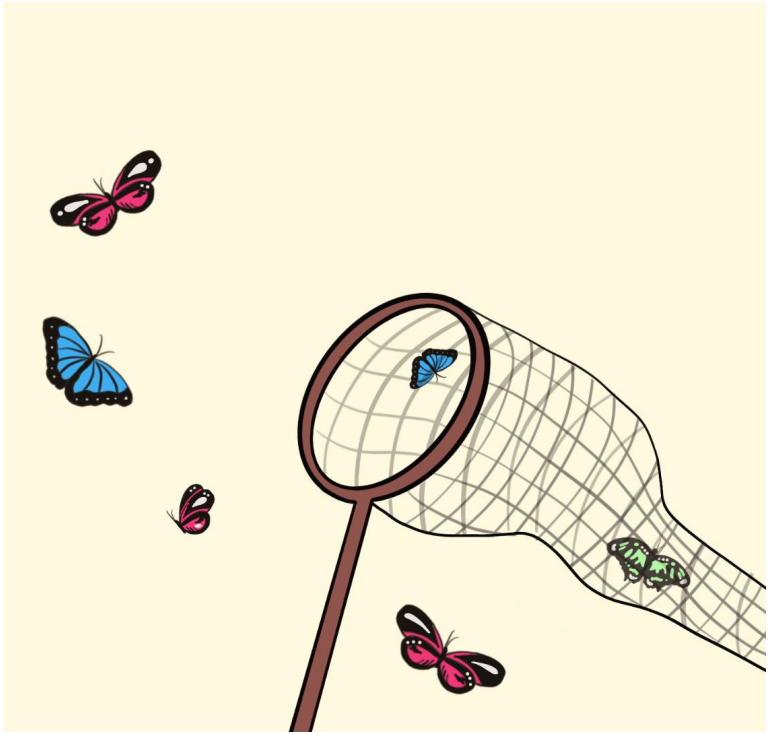


The “backdoor path” rule



Today’s agenda

- ➔ DAGs and causal relationships
- ➔ Choosing variables
- ➔ **Your final project**



Final project DAG

Consider your final project (8 minutes)

What's the research question?

1. List the variables you have or plan to collect. Include at least 3 predictors.
2. Draw a DAG explaining the relationships between the variables.
3. Are there any confounds? If so, where are the backdoor paths?

Final project DAG

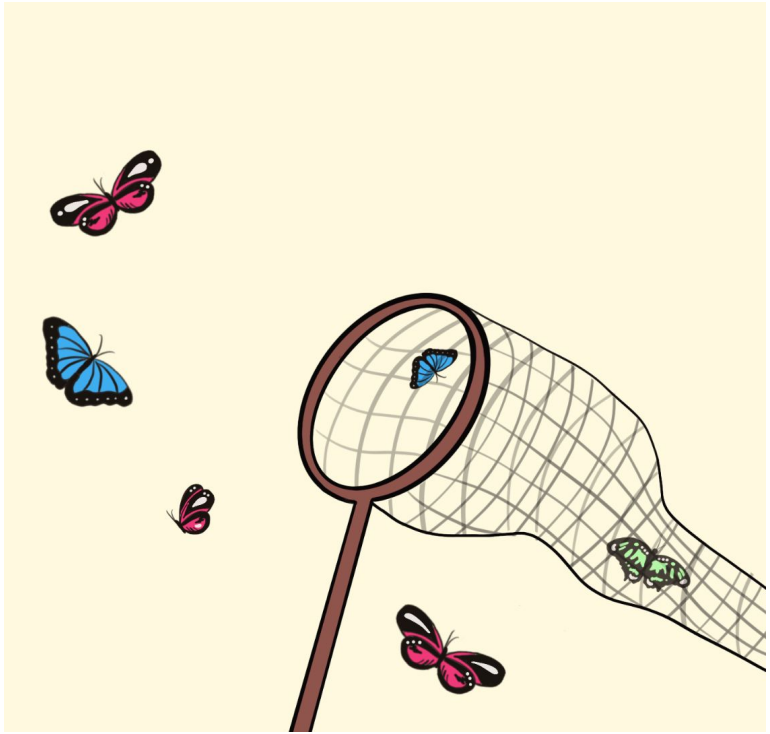
Peer feedback (6 minutes)

Share your DAG with a peer and give each other feedback.

1. Are the causal relationships reasonable? Do any require clarification?
2. Do you agree about the confounds?
3. Together, think of an example of a collider variable. This doesn't have to be in the actual dataset.

Today’s agenda

- ➔ DAGs and causal relationships
- ➔ Choosing variables
- ➔ Your final project



Midterm 2

Midterm preparation

- Practice midterm available through Slack
- Attempt the questions before lab tomorrow
- Tomorrow's quiz: write down one *muddy point* you have about the practice midterm or DAGs