

# Week 6 Lecture 1:

## Forks, pipes, and colliders

*EDS 222: Statistics for Environmental Data Science*



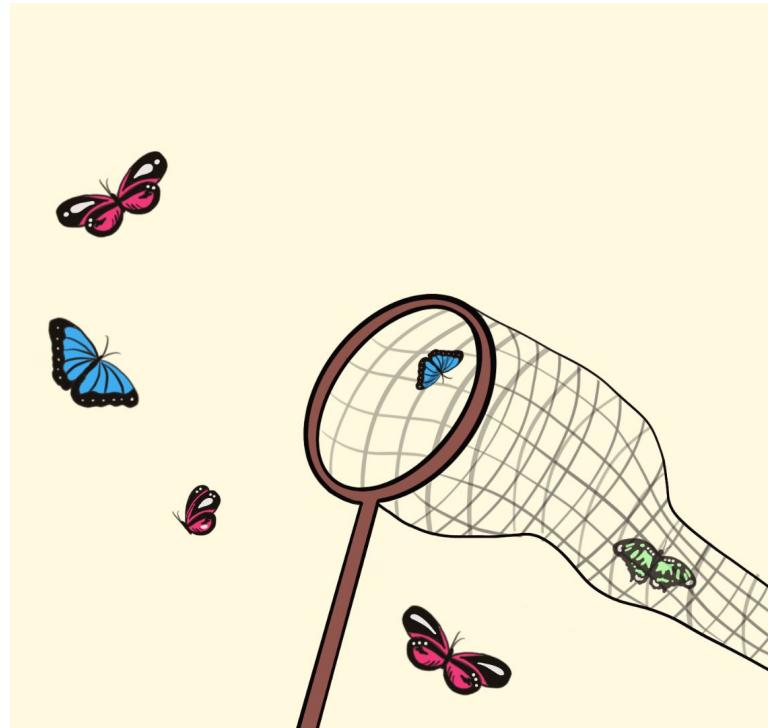
# Urban heat islands



EPA

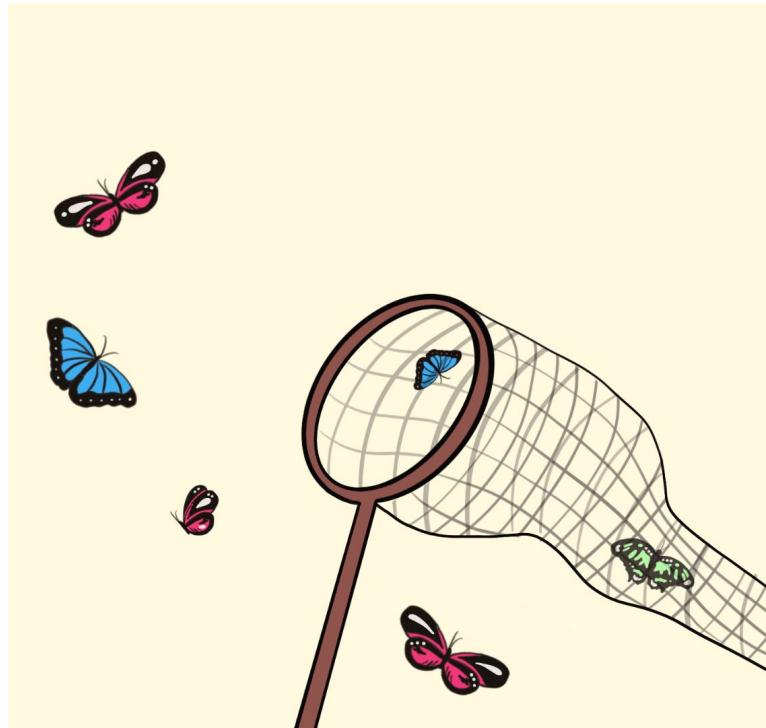
# Today's agenda

- 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

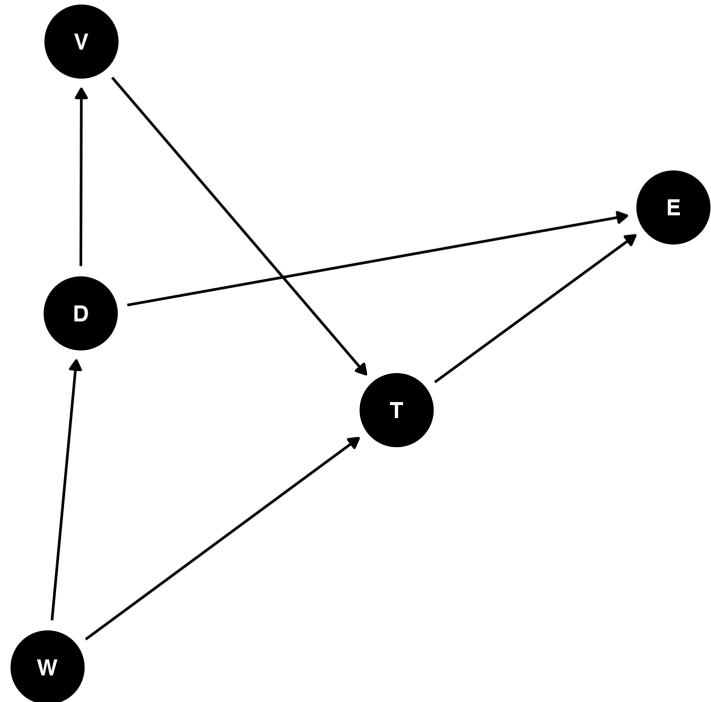
*Does urban density increase the urban heat island effect?*

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?

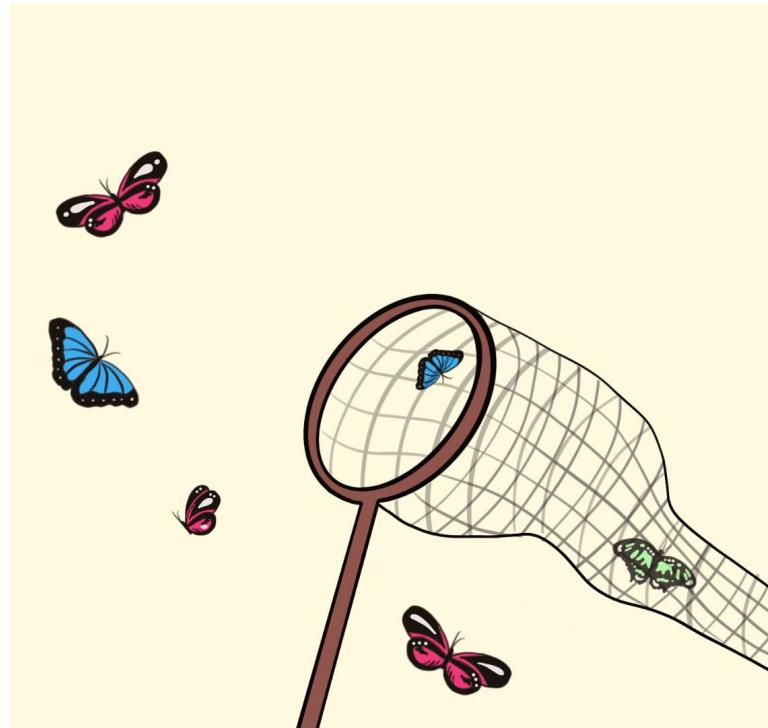


The screenshot shows the EPA website's header with the logo and navigation links: Environmental Topics, Laws & Regulations, Report a Violation, and About EPA. Below the header, a large image of a city skyline with a green roof in the foreground is displayed. A dark blue overlay box contains the text "Frequent Questions About Heat Islands", "Find answers to common questions about heat islands.", and a link "Visit the Frequent Questions page>>". At the bottom left of the main image, there are three numbered buttons: 1, 2 (which is highlighted in blue), and 3.

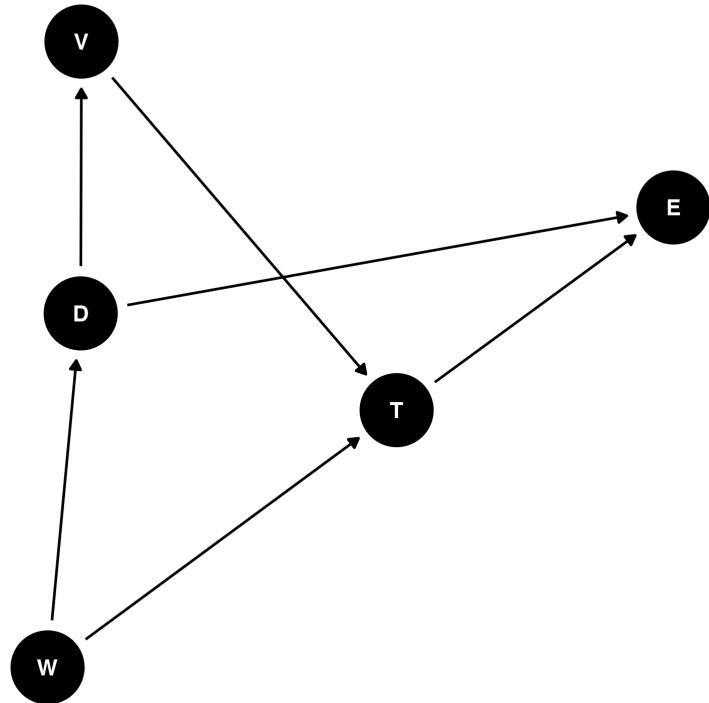
Heat islands occur when a developed area experiences higher temperature than nearby rural areas, or when areas experience hotter temperatures within a city. EPA is committed to providing localities with information about heat islands, cooling strategies, implementation examples, and outreach materials to help reduce heat islands and their impacts on communities and neighborhoods.

# Today's agenda

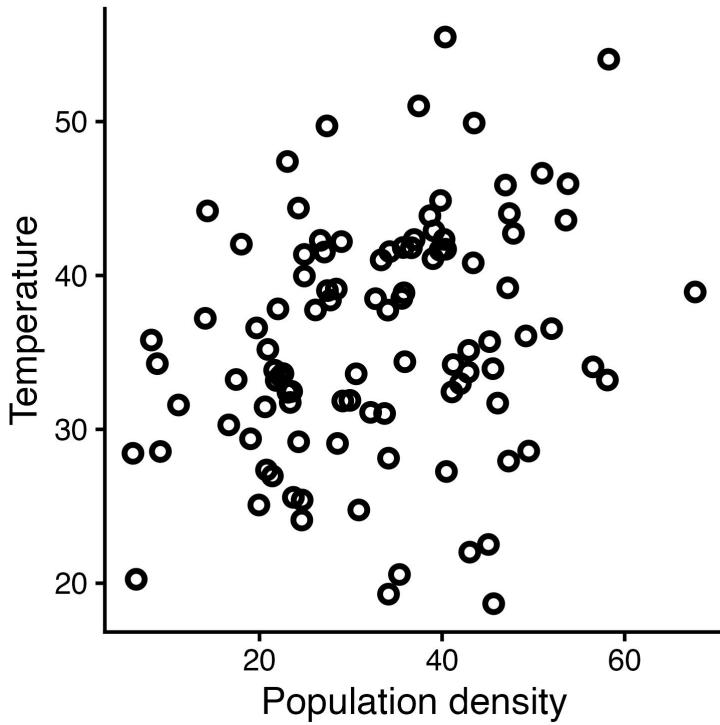
- DAGs and causal relationships
- **Choosing variables**
- Your final project



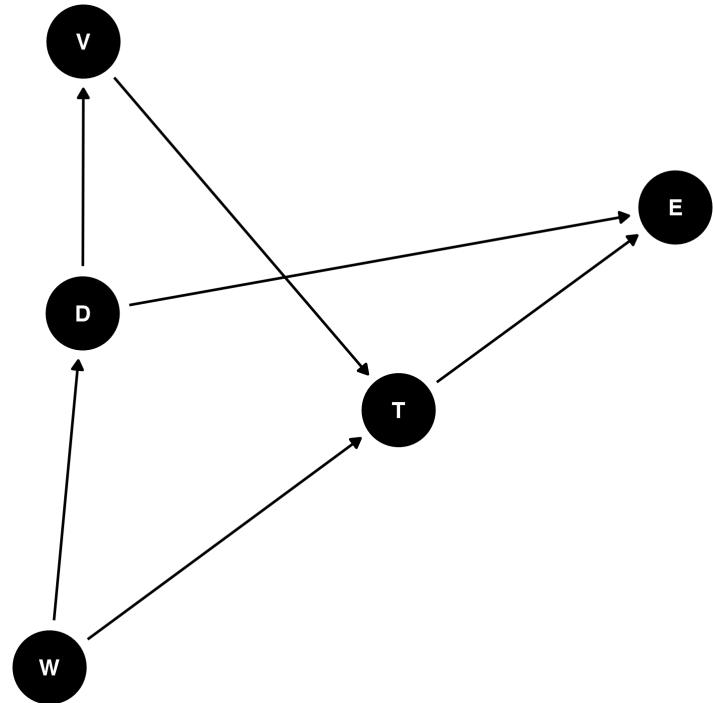
# Forks, pipes, and colliders



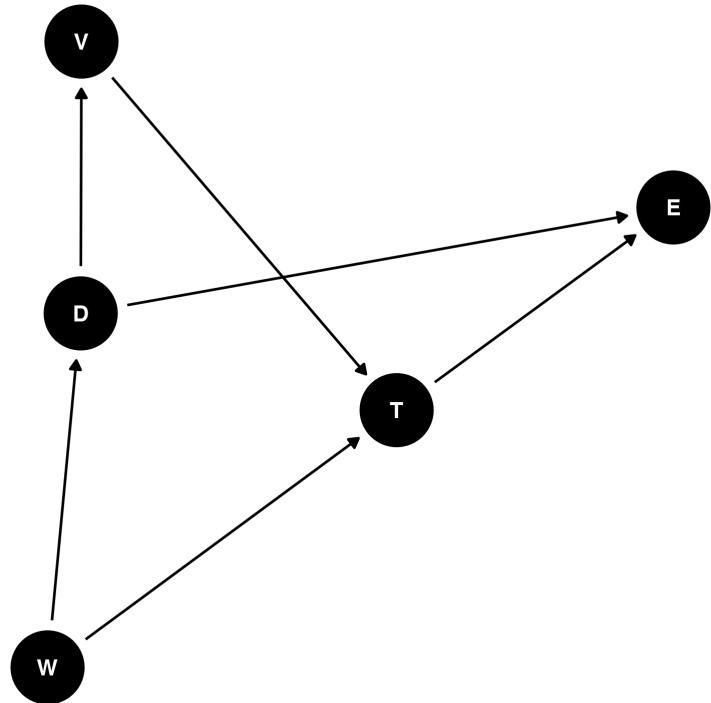
# The data generating process



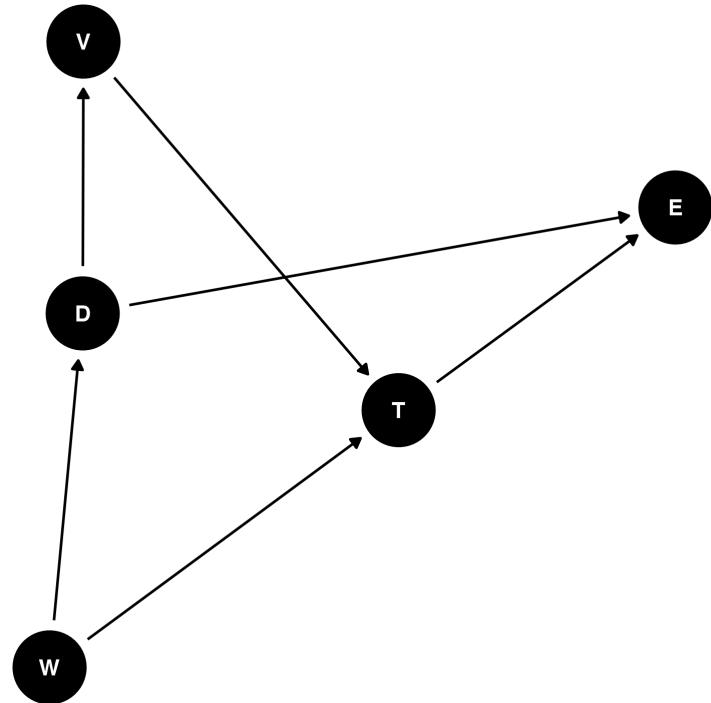
# The fork



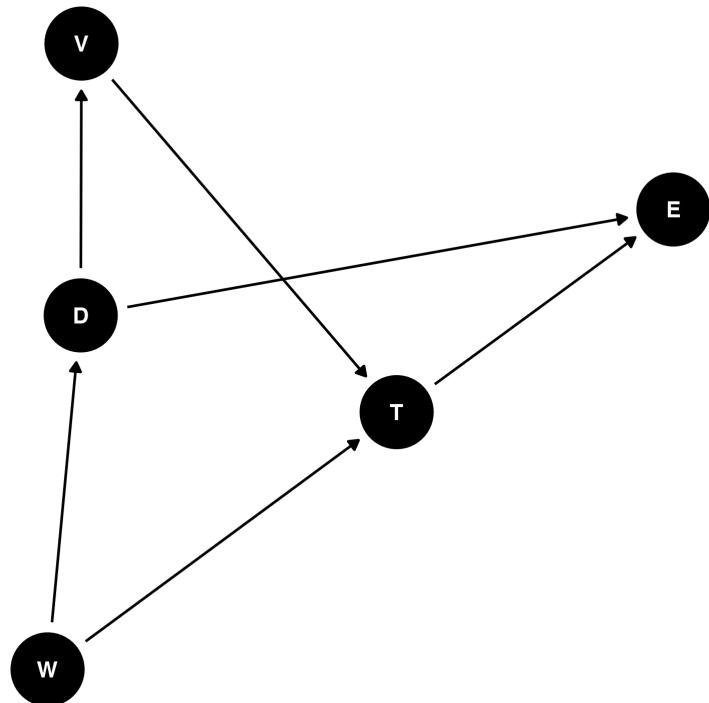
# The pipe



# The collider

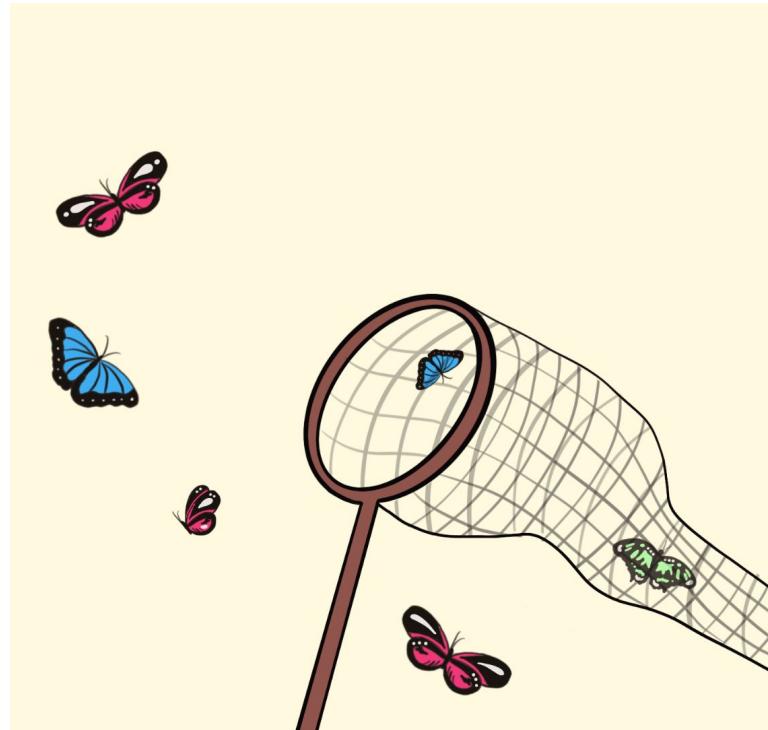


# The “backdoor path” rule



# 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

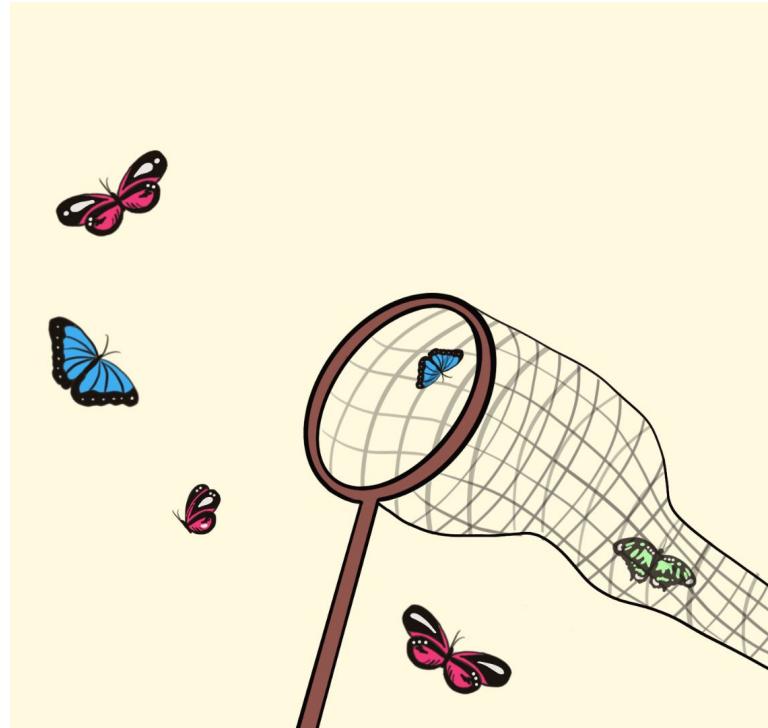
## Peer feedback (6 minutes)

*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