Course Project: Task 1 Kick-off

INFO/STSCI/ILRST 3900: Causal Inference

25 Sep 2024

Course Project

- Opportunity to engage with the course content via a real-world example
- Students in class come in with different backgrounds and skills
- The project will require various skill sets
  - Asking a causal question
  - Reasoning about assumptions
  - Data Analysis
  - Communicating Results

Overview

Pieces of the project will be completed throughout the semester

Individual Tasks

- Task 1: Define a causal question
- Task 2: Draw and reason about a causal graph

Overview

Pieces of the project will be completed throughout the semester

Individual Tasks

- Task 1: Define a causal question
- Task 2: Draw and reason about a causal graph

Group Tasks

- Task 3: Gather and process relevant data
- Task 4: Select an identification strategy and analyze the data
- Task 5: Communicate the results with short report and presentation video

Details for each task are in the Project Overview pdf on the course website

Task 1: Define a causal question

Details are in the Task 1 Details pdf on the course website

Task 1: Define a causal question

Details are in the Task 1 Details pdf on the course website

Main pieces

- What is a causal question that you are interested in?
- Carefully define a treatment and outcome of interest.
- Make sure a data set exists which at least includes observations of both treatment and outcome.

Data sources

- Default option of using National Longitudinal Study of Adolescent to Adult Health (Add Health)
- Many other interesting data sets
- Proposing a causal question which uses a data set aside from Add Health can give up to 5% extra credit[1]

[1] If it gets selected; see project overview for details

Data sources

- Default option of using National Longitudinal Study of Adolescent to Adult Health (Add Health)
- Many other interesting data sets
- Proposing a causal question which uses a data set aside from Add Health can give up to 5% extra credit[1]
- Be creative in thinking about a question!
- Could end up being the start of a senior thesis

[1] If it gets selected; see project overview for details

Explore other data sets as a group

- Select a dataset other than Add Health described in the Task 1 Details details
- Take 7-10 minutes to first poke around and see what is in the data set
- What is a causal question which could potentially be answered using this data? What is the treatment? What is the outcome?
- No need to verify that the appropriate data actually exists, so be creative!
- How would having a good understanding of the causal effect allow for better decision making?

Add Health Exercise

Assessing Causality and Persistence in Associations Between Family Dinners and Adolescent Well-Being by Musick and Meir use Add Health data to examine the causal effect of eating family dinners on a variety of outcomes

As a group

- Read the Abstract
- Read first three subsections of Methods section
  - National Longitudinal Survey of Adolescent Health
  - Outcomes
  - Family Dinners
- Use Add Health Codebook Explorer to find the questions in the survey corresponding to the treatment and at least one of the outcomes
- Answer the question on the next slide

Add Health Exercise- In Class Assignment

As a group

- Read the Abstract and first three subsections of Methods section
- Use Add Health Codebook Explorer to find the questions in the survey corresponding to the treatment and at least one of the outcomes
- Answer the questions and submit your answers
  1. What is the most common response to the question corresponding to treatment?
  2. How many respondents buy recorded music that is not in English?
  3. Did the month in which most respondents were born change between the waves? (*Hint: birth date*)