INFO371 Lab 04

Your name:

Deadline: Thursday, Feb 7, 11:59pm

Introduction

Please submit the completed lab by end of Tuesday, Feb 4th. You should submit a) your code (rmd, notebooks, whatever you use) and b) the lab in a directly readable form (html or pdf).

Please do not just provide computer output. Always comment on your main findings. Include any substantial comments as a separate text blocks. Also limit your output: do not submit pages and pages of whatever your code spits out.

Note: you may want to do some of it on paper instead of computer. You are welcome to do it but please include the result as an image into your final file.

Working together is fun and useful but you have to submit your own work. Discussing the solutions and problems with your classmates is all right but do not copy-paste their solution! Please list all your collaborators below:

1.

2. . . .

1   Was Montana Meth Prevention Project Effective?

Your task is to determine if Montana Meth Prevention Project (MMP) was effective in decreasing the methamphetamine use among Montana teens.

Background and Data

2005-2007 Montana conducted an agressive ad campaign to decrease the meth abuse by the youth.

Your task is to analyze the efficacy of the campaign using Youth Risk Behavior Surve (YRBS) data (dowload if from canvas: files - yrbs.tsv). You will implement a) cross-sectional estimator; b) before-after estimator; c) differences-in-differences (DiD) estimator. The sample, necessary for this lab, is on canvas. The full data and documentation can be obtained from `https://www.cdc.gov/healthyyouth/data/yrbs/data.htm`.

The variables in the sample are:

year survey year (2003-2009)

before engineered variable, TRUE if before the ad campaing (i.e. 2003 and 2005).

age age, 14-17 years

sex M,F

**meth** 1 = has used methamphetamines in life

tv 1 = watches more than 1hr TV on average schoolday (a large number of the project ads were displayed on TV).

state MT = Montana, XX = national sample

Here is a small sample of the data:

```
##      year age sex meth tv state before
## 1: 2003   15   M    0  1    MT   TRUE
## 2: 2005   15   F    0  1    MT   TRUE
## 3: 2005   15   M    0  1    MT   TRUE
## 4: 2009   14   M    0  1    XX  FALSE
## 5: 2009   14   M    0  1    XX  FALSE
## 6: 2007   14   M    0  0    XX  FALSE
## 7: 2007   17   M    0  1    XX  FALSE
## 8: 2003   17   M    0  1    XX   TRUE
```

And here is how the average meth use data looks on a graph for different years and states:
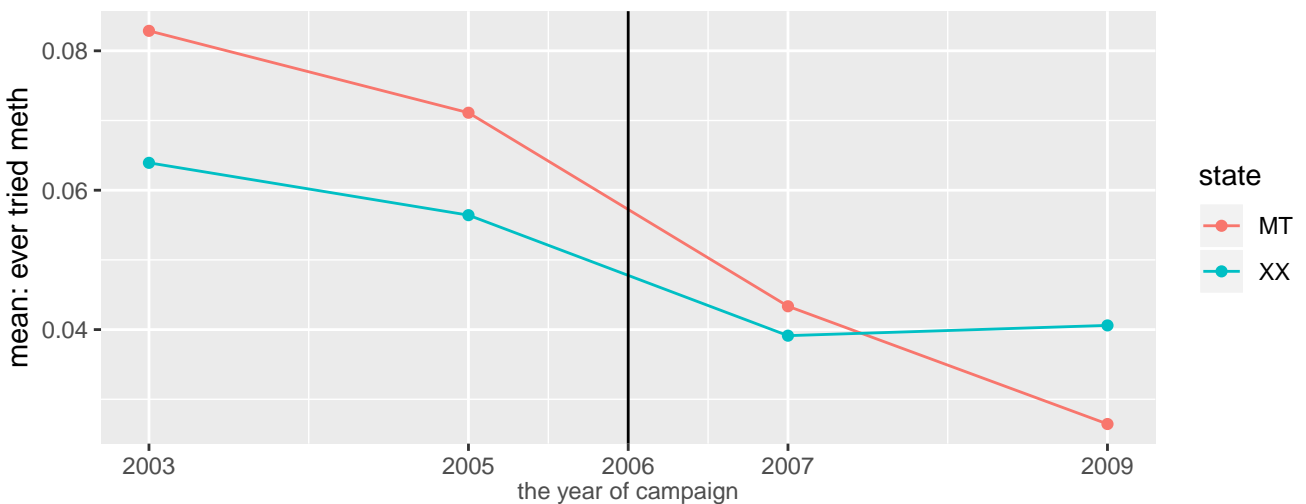
1.1 Your Tasks

Load the data

1.1.1 Before-After Estimator

If the campaign had any effect, it should have decreased the amphetamine use in Montana. Let's compare the meth use (mean value) in Montana before and after the campaign. In each case interpret the most important outcomes, and compare the results for all three approaches.

1. Do it as a simple table: just present the average meth use in MT before and after the campaign.

2. Do it using regression without any other controls (no age, sex, tv).

   Hint: you have to focus on a single state only (which?). Note: you can select only a subset with `subset=<a logical vector>` argument in `lm`.

3. Run the same regression, but now add full controls (age, sex, tv).

1.1.2   Cross-Sectional Estimator

Alternatively, if the project had an effect, Montana should have lower rate of meth use afterwards. Now compare the meth use (mean value) in Montana and nationally after the project. Interpret the most important outcomes, and compare the results for all three approaches.

- Do it as a simple table: compare the average post-project meth use in MT and nationally.

- Do it using regression without any other controls

- Run a regression with full controls.

1.1.3   5 Differences-in-Differences Estimator

However, both the long-time trends and levels in meth use in Montana and elsewhere may be different (this is also suggested by the figure). Now let's see how much more will meth use decline in Montana compared to the national average. Compare the differences in trends in meth use in Montana and nationally between pre-project and post-project years. In each case interpret the most important outcomes, and compare the results for all three approaches.

- Do it as a simple table: report average use before and after the campaign in both Montana and nationally. Compute the effect by calculating the trend difference.

- Do it using regression without any other controls.

  Hint: you need a regression that can predict four different values for MT before, MT after, national before, national after. What kind of categories you have to include in the model?

- Run a regression with full controls.