Homework # 1

A/B Testing
Fall 2024

NAMES: names of TWO students here
ANDREW IDs: Andrew IDs of TWO students here

One student submits homework on Gradescope
Then add other student as team member

DUE November 4, 11.00pm EST

Show your work as part of your answers.
Include code and screen captures of the generated output.
Cite all your sources human generated and otherwise.
If you use a LLM to help with your homework
you are required to list all prompts that you used to do so.

Question 1 [60 points]

Define the following concepts using your own words: **[10 points each]**

a) Counterfactual

b) Selection

c) Causal effect

d) Back door

e) Randomized Controlled Trial

f) Pre-experimental balance

Type your answer here!

Question 2 [90 points]

A team on analysts are trying to find the effect of allowing smartphones into the classroom on the performance of students. For this purpose, a school shared data with this team, available to you for the purpose of this homework in file HW-1-a.csv. Several students in the school were subject to different conditions during one lecture, namely:

B: smartphones banned from the classroom

ANL: smartphones allowed into the classroom but not used to assist instruction

AML: smartphones allowed into the classroom and used to assist instruction

Using the smartphones to assist instruction was accomplished by having the teacher actively asking the students to use the devices to complement learning while in the classroom. The following covariates are included in this file:

stu_id: an anonimized identifier for each student

treatment ("B","ANL","AML"): experimental condition as described above

prescore (0-100): the grade students had in a test given at the beginning of the lecture

postscore (0-100): the grade students had in a test given at the end of the lecture

female: an indicator for gender (TRUE-female, FALSE-otherwise)

Use this dataset to answer the following questions:

a) How many students are included in this dataset and in each condition? **[2+2 points]**

b) Show descriptive statistics for the improvement in test scores from the pre-test to the post-test **[16 points]**

c) Show whether the improvement in the scores (as defined above) reduced when smartphones were allowed into the classroom and not used to assist instruction (compared to when they were banned from

the classroom) **[20 points]**

d) Show whether the improvement in the scores (as defined above) increased when smartphones were allowed into the classroom and used to assist instruction (compared to when they were banned from the classroom) **[20 points]**

e) Are the effects computed in d) and e) causal? Why or why not? Show suggestive empirical evidence to support your conclusion **[5+10+15 points]**

(note: if you are running regressions (and depending on the software you use) it may be helpful to transform categorical variables (like treatment and female) into numerical types).

Type your answer here!

Question 3 [100 points]

Later on, the team of analysts arranged to run an experiment with this school to find the causal effect discussed in the previous question. The school agreed to use a subset of its students to participate in the experiment. Students were randomly assigned to the 3 experimental conditions referred above during one lecture. The school shared the dataset in file HW-1-b.csv available to you for the purpose of this homework. Use this dataset to answer the following questions:

a) How many students participated in the experiment in each condition? How many of them are female? **[2+2 points]**

b) Show suggestive evidence that this data came from a well executed RCT **[15 points]**

c) Show descriptive statistics for the improvement in test scores from the pre-test to the post-test **[16 points]**

d) Show whether the improvement in the scores (as defined above) reduced when smartphones were allowed into the classroom and not used to assist instruction (compared to when they were banned from the classroom)? **[20 points]**

e) Show whether the improvement in the scores (as defined above) increased when smartphones were allowed into the classroom and used to assist instruction (compared to when they were banned from the classroom)? **[20 points]**

f) Show whether the results obtained in d) and e) above are different for female and male students **[25 points]**

Type your answer here!