

Name(s): _____

DATA 101 Assignment 4: ANSWER KEY

Work in teams of 2

Instructions. Answer the following questions about experiments and causal designs. You may use the information from class and the lecture slides. Write your responses clearly in the spaces provided.

1. Suppose Dickinson wants to know whether weekly yoga classes reduce student stress.

(a) Propose a treatment and control group.

Treatment: A random group of students attend yoga classes 5 days a week.

Control: A random group of students who don't attend any yoga classes.

(b) Define the outcome variable(s). Stress, which could be on a scale from 0-10.

(c) Explain how you would implement random assignment in practice. You can get a roster of all Dickinson students and randomly select eg 100 students to be included in the treatment group and 100 for the control group.

(d) Identify one possible source of bias and how you would address it. If number of students who are supposed to be included in the treatment group skip yoga classes that could limit the treatment's efficacy. The researchers should monitor attendance and take it into account when measuring treatment's efficacy.

2. Read the following proposal:

Treatment: Students asked to give up their smartphones for 48 hours.

Control: Students continue normal use.

Outcome: Daily mood logs and GPA.

(a) Would you approve or reject this as an IRB member? Why? I would conditionally approve this study. One concern is, if the students stop using their smartphones for 48 hours, their families might be concerned because they can't reach out to them. If the experimenters are okay with it, they could assign friends who could personally check up on them to see if they need anything, or even more simply, allow the participants to use non-smarthpone means of communication such as a flip phone (provided by the researchers.)

(b) Suggest one modification that could make this study more ethical. The researchers can make sure that the 48-hour restriction does not overlap with midterms, finals, or major assignment deadlines because the transition from smartphone to no smart-phone could exacerbate stress during busy periods.

3. Imagine you have 12 students and need to split them into treatment and control groups.

- (a) Write down one unfair way of assigning them to treatment that would be biased. *Sorting students into groups alphabetically.*
 - (b) Write down one method that would be considered random and fair. *Random assignment of students to groups.*
 - (c) Explain why randomization is crucial in an experiment. *Randomization ensures that we get two comparable groups who have, on average, identical characteristics except for the treatment status.*
4. In 2007, Washington state became the first state to pass a law banning all drivers from sending text messages while driving. The law went into effect on January 1, 2008. Oregon did not adopt a texting ban until later, so it can serve as a reasonable control state in this comparison.

The table reports fatal crashes per 100,000 licensed drivers for Washington (treatment) and Oregon (control), before and after the ban:

	Pre (2007)	Post (2009)
Washington (treatment)	11.64	9.79
Oregon (control)	16.09	13.27

- (a) Compute the DiD estimate by hand. Show your steps.

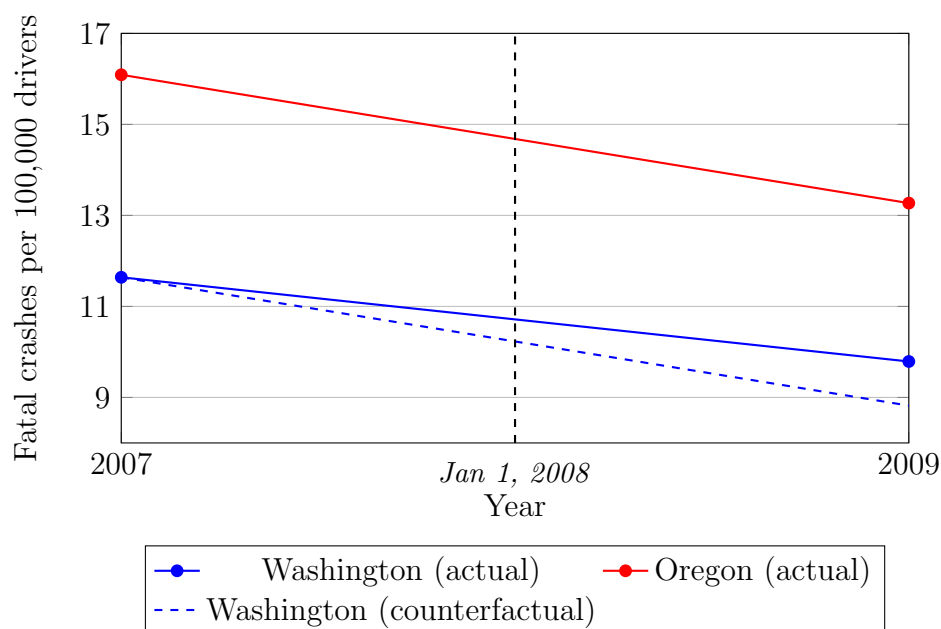
$$\text{Change in Washington (treatment)} = 9.79 - 11.64 = -1.85$$

$$\text{Change in Oregon (control)} = 13.27 - 16.09 = -2.82$$

$$\begin{aligned}
 \text{DiD estimate} &= (\text{WA change}) - (\text{OR change}) \\
 &= (-1.85) - (-2.82) \\
 &= -1.85 + 2.82 \\
 &= 0.97 \quad \text{fatal crashes per 100,000 licensed drivers.}
 \end{aligned}$$

- (b) Interpret what the sign and magnitude of the estimate mean in context.

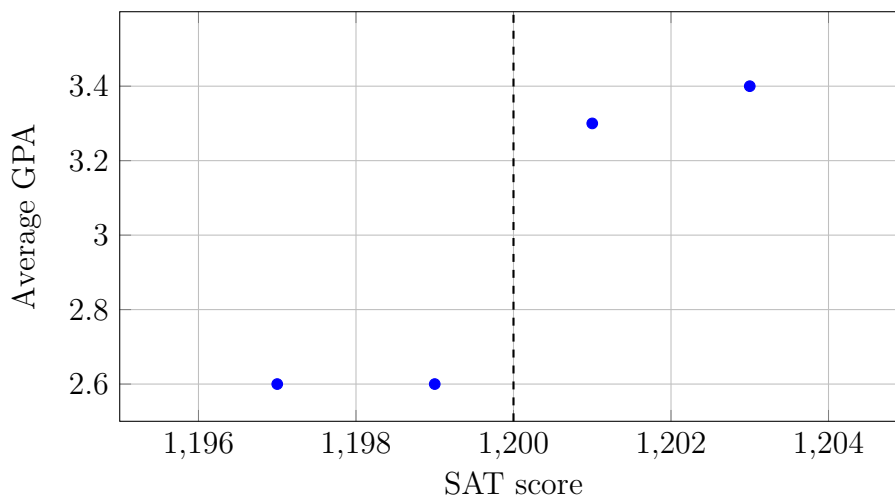
The DiD estimate is positive (0.97), which means that although fatal crashes declined in Washington after the texting ban, they declined by less than in Oregon. Under the parallel trends assumption, this suggests the texting ban is associated with an *increase* of about 0.97 fatal crashes per 100,000 licensed drivers relative to the control state. In other words, Washington improved less than Oregon in the post period.



5. Suppose a scholarship awarded if $SAT \geq 1200$. Below are average college GPA outcomes:

SAT score	Average GPA
1197	2.6
1199	2.6
1201	3.3
1203	3.4

(a) Plot by hand these four points on graph paper (SAT on x-axis, GPA on y-axis).



(b) Does there appear to be a jump in GPA at the cutoff? **Yes.** The GPA rises from about 2.6 just below 1200 to above 3.3 just above 1200, suggesting a discontinuity at the cutoff.

(c) In plain words, what does that jump tell us about the effect of getting the scholarship?

The jump suggests that students who barely qualified for the scholarship (scoring just above 1200) went on to earn significantly higher GPAs than students who barely missed it. Interpreted causally, this means the scholarship likely improved academic performance for students near the cutoff.

6. Suppose Spotify wants to test a new playlist interface.

- (a) Sketch an experimental design (treatment, control, outcome).

Spotify could randomly assign users into two groups at the account level. The treatment group receives the new playlist interface, while the control group continues with the old design. The main outcome would be user engagement, measured by variables such as minutes listened per day, tracks completed, skips, playlist saves, or retention rates. By comparing the two groups, Spotify can estimate the causal effect of the new interface on usage.

- (b) Spotify runs the test for one week and finds the new design “wins.” A month later, usage drops back to normal. What might explain this?

A likely explanation is a novelty effect. Users initially engage more with the new interface out of curiosity, but once the design becomes familiar, their long-run behavior returns to baseline. Another possibility is that the one-week testing window captured a temporary spike that was not representative of sustained behavior.

- (c) Spotify changes the design at the same time Beyoncé drops a surprise album. How could that confuse the results?

The album release is a confounding factor. It could increase music consumption independently of the new interface, making it difficult to know whether higher usage is due to the design or to excitement about new content. If Beyoncé fans are unevenly distributed between treatment and control, or if the new interface promotes the album differently, the estimated treatment effect would be biased.