CIV 337 Homework 2

due: **9/11/24**

Instructions: Please submit solutions on canvas. Only a knitted pdf of an `Rmarkdown` file will be accepted.

Problem 1: QSS exercise 1.5.1. Surveys are frequently used to measure political behavior such as voter turnout, but some researchers are concerned about the accuracy of self-reports. In particular, they worry about possible *social desirability bias* where in post-election surveys, respondents who did not vote in an election lie about not having voted because they may feel that they should have voted. Is such a bias present in the American National Election Studies (ANES)? The ANES is a nation-wide survey that has been conducted for every election since 1948. The ANES conducts face-to-face interviews with a nationally representative sample of adults. The table below displays the names and descriptions of variables in the `turnout.csv` data file.

Variable	*Description*
`year`	Election year
`ANES`	ANES estimated turnout (percentage)
`VEP`	Voting Eligible Population (in thousands)
`VAP`	Voting Age Population (in thousands)
`total`	Total ballots cast for highest office (in thousands)
`felons`	Total ineligible felons (in thousands)
`noncitizens`	Total non-citizens (in thousands)
`overseas`	Total eligible overseas voters (in thousands)
`osvoters`	Total ballots counted by overseas voters (in thousands)

a. Load the data into `R` and check the dimensions of the data. Also, obtain a summary of the data. How many observations are there? What is the range of years covered in this data set?

b. Calculate the turnout rate based on the voting age population or VAP. Note that for this data set, we must add the total number of eligible overseas voters since the VAP variable does not include these individuals in the count. Next, calculate the turnout rate using the voting eligible population or VEP. What difference do you observe?

c. Compute the difference between VAP and ANES estimates of turnout rate. How big is the difference on average? What is the range of the difference? Conduct the same comparison for the VEP and ANES estimates of voter turnout. Briefly comment on the results.

d. Compare the VEP turnout rate with the ANES turnout rate separately for presidential elections and midterm elections. Note that the data set excludes the year 2006. Does the bias of the ANES vary across election types?

e. Divide the data into half by election years such that you subset the data into two periods. Calculate the difference between the VEP turnout rate and the ANES turnout rate separately for each period. Has the bias of the ANES increased over time?

f. The ANES does not interview overseas voters and prisoners. Calculate an adjustment to the 2008 VAP turnout rate. Begin by subtracting the total number of ineligible felons and non-citizens from the VAP to calculate an adjusted VAP. Next, calculate an adjusted VAP turnout rate, taking care to subtract the number of overseas ballots counted from the total ballots in 2008. Compare the adjusted VAP turnout with the unadjusted VAP, VEP, and the ANES turnout rate. Briefly discuss the results.

Problem 2: QSS exercise 2.8.1. The STAR (Student-Teacher Achievement Ratio) Project is a four year longitudinal study examining the effect of class size in early grade levels on educational performance and personal development.[1]

A longitudinal study is one in which the same participants are followed over time. This particular study lasted from 1985 to 1989 involved 11,601 students. During the four years of the study, students were randomly assigned to small classes, regular-sized classes, or regular-sized classes with an aide. In all, the experiment cost around $12 million. Even though the program stopped in 1989 after the first kindergarten class in the program finished third grade, collection of various measurements (e.g., performance on tests in eighth grade, overall high school GPA) continued through the end of participants' high school attendance.

We will analyze just a portion of this data to investigate whether the small class sizes improved performance or not. The data file name is `STAR.csv`. The names and descriptions of variables in this data set are:

Variable	*Description*
`race`	Student's race (White = 1, Black = 2, Asian = 3, Hispanic = 4, Native American = 5, Others = 6)
`classtype`	Type of kindergarten class (small = 1, regular = 2, regular with aide = 3)
`g4math`	Total scaled score for math portion of fourth grade standardized test
`g4reading`	Total scaled score for reading portion of fourth grade standardized test
`yearssmall`	Number of years in small classes
`hsgrad`	High school graduation (did graduate = 1, did not graduate = 0)

a. Create a new factor variable called `kinder` in the data frame. This variable should recode `classtype` by changing integer values to their corresponding informative labels (e.g., change 1 to small etc.). Similarly, recode the `race` variable into a factor variable with four levels (White, Black, Hispanic, Others) by combining Asians and Native Americans as the Others category. For the `race` variable, overwrite the original variable in the data frame rather than creating a new one.

[1] This exercise is in part based on: Mosteller, Frederick. 1997. The Tennessee Study of Class Size in the Early School Grades. *Bulletin of the American Academy of Arts and Sciences* 50(7): 14-25.

b. How does performance on fourth grade reading and math tests for those students assigned to a small class in kindergarten compare with those assigned to a regular-sized class? Do students in the smaller classes perform better? Use means to make this comparison while removing missing values. Give a brief substantive interpretation of the results. To understand the size of the estimated effects, compare them with the standard deviation of the test scores. Recall that `na.rm = TRUE` can be added to functions in order to remove missing data (see section 1.3.5).

c. Instead of comparing just average scores of reading and math tests between those students assigned to small classes and those assigned to regular-sized classes, look at the entire range of possible scores. To do so, compare a high score, defined as the 66th percentile, and a low score (the 33rd percentile) for small classes with the corresponding score for regular classes. These are examples of *quantile treatment effects*. Does this analysis add anything to the analysis based on mean in the previous question?

d. Some students were in small classes for all four years that the STAR program ran. Others were assigned to small classes for only one year and had either regular classes or regular classes with an aide for the rest. How many such students of each type are in the data set? Create a contingency table of proportions using the `kinder` and `yearsmall` variables. Does participation in more years of small classes make a greater difference in test scores? Compare the average and median reading and math test scores across students who spent different numbers of years in small classes.

e. Examine whether the STAR program reduced the achievement gaps across different racial groups. Begin by comparing the average reading and math test scores between white and minority students (i.e., Blacks and Hispanics) among those students who were assigned to regular classes with no aide. Conduct the same comparison among those students who were assigned to small classes. Give a brief substantive interpretation of the results of your analysis.

f. Consider the long term effects of kindergarten class size. Compare high school graduation rates across students assigned to different class types. Also, examine whether graduation rates differ by the number of years spent in small classes. Finally, as done in the previous question, investigate whether the STAR program has reduced the racial gap between white and minority students' graduation rates. Briefly discuss the results.

Problem 3: First, describe in your own words the importance of randomization, or random assignment of a treatment, for accurately estimating a causal effect. Second, why is it wrong to estimate the causal effect of COVID lockdowns as the difference in average outcomes for "lockdown states" (states that implemented a lockdown) and "non-lockdown states" (states that did not implement a lockdown?