Problem Set 0

Bias in Self-reported Turnout

[YOUR NAME]

Due Date: 2025-01-17 by 11:59PM ET

If you utilized any assistance to complete this problem set (e.g., classmate, tutor, AI), you should cite it here: For example, I used ChatGPT for this assignment to help with an error I received on Question 2. The log of the session can be found at this link: chatgpt.link.com. I used Microsoft Copilot for an error I received on Question 3. I have included a printout of that session on Blackboard along with this assignment. Finally, I worked with Classmate O'Classmate on the entire problem set; however, these answers are my own.

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.

Name	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)

Question 1

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?

Answer 1

```
# Write your code here
```

Write your answer here.

Question 2

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?

Answer 2

```
# Write your code here
```

Write your answer here.

Question 3

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.

Answer 3

```
# Write your code here
```

Write your answer here.