# POLS 7012: Practice Midterm

## Joe Ornstein

Create an R project, an R script, load tidyverse, and read in the ANES 2019 Pilot data from Problem Set 1.

## Tidy the Data

1. Recode the `vote20jb` variable with the actual responses rather than the numeric codes (see the codebook).

2. Recode gender as "Male" or "Female" instead of 1 or 2.

3. Create a new variable called `age`, equal to 2019 minus the respondent's birth year.

4. Filter out the respondents who say they will "probably not vote".

5. Select only the `caseid`, `gender`, `age`, and `vote20jb` variables.

6. Save the cleaned up dataset to your `data/` folder as a .csv file called `anes_2019_pilot_clean.csv`.

## Visualize and Summarize the Data

1. Visualize the age distribution for Biden vs. Trump voters, faceting by gender.

2. Group by `vote20jb` and `gender`. What is the mean age and number of respondents for each group?

## Make Inferences

1. Compute a 95% confidence interval for the average age.

2. Is the difference in age between Trump and Biden supporters statistically significant? Compute a p-value and 95% confidence interval.

3. Are men significantly more likely than women to support Trump? Compute the associated p-value.