POL 212
Winter 2024
Midterm exam

1. Load the Human Rights dataset and README file: `humanrights.dta` (A Stata file) and `humanrights_README.txt`. To confirm you have successfully loaded the data in R, print the first six rows.

2. What is the (Pearson) correlation between the `time` and `humanrights` variables? Provide a scatterplot (with lowess smoother) between these two variables.

3. Provide a histogram showing the univariate distribution of the `kingsfans` variable. Also provide a scatterplot (with lowess smoother) between it and `humanrights`. Based on this, what might be an appropriate transformation for `kingsfans`?

4. Consider the `tort` variable: in the next questions, you will be asked to include it as an independent (explanatory) variable in a linear regression model. Here, decide how you want to treat its level of measurement, either nominal/ordinal OR continuous. There is no right or wrong answer, but justify your choice. How will you code its class in R?

5. Fit a bivariate (simple) linear regression in R where `humanrights` is the dependent variable (y) and `tort` is the independent variable (x).

    a. Report and interpret the estimated regression coefficient(s).
    b. What if you first standardize `humanrights` (y)? Report and interpret the regression coefficient(s) now.

6. Fit a multiple linear regression in R where `humanrights` is the dependent variable (y) and `tort`, `genocide`, and `time` are the independent variables (x).

    a. Which coefficients are statistically significant at the $p < 0.05$ (two-tailed) level?
    b. What is the sign (positive or negative) on the regression coefficient for `time`? Substantively, what does this mean (in informal terms)?
    c. Report and interpret the R^2 value for the estimated regression model.
    d. Bootstrap 500 R^2 values and describe the simulated distribution (show a histogram and report its mean and standard deviation).