Econ 5100 - Quantitative Methods and Applications
In-class problem set 1

Alumni Giving

Alumni donations are an important source of revenue for colleges and universities. If administrators could determine the factors that influence increases in the percentages of alumni who make a donation, they might be able to implement policies that could lead to increased revenues. The data set Alumni.csv on Canvas contains information on alumni giving rate (almunigivingrate) and two possible determinants, percent of classes with less than 20 students (classlt20) and student/faculty ratio (sfratio).

1. Use ggplot to draw scatter plots of the data, i.e. alumni giving rate vs each of the two other university characteristics (clssslt20 and sfratio). Try to experiment with jitter. Comment on whether it appears that linear models might be appropriate.

2. Use tidyverse to find the means, standard deviations, and min/max for each of the three variables. This is more cumbersome than it sounds. If you want a challenge, figure out how to write this with a function and call it for each variable.

3. Run a regression model with alumni giving rate as the dependent variable and the percent classes with less than 20 students as the explanatory variable.
   a. What is the interpretation of the estimated coefficients? Make sure to discuss whether it is statistically significant and, if so, at what level.

b. What is R-square for this model, and what is the interpretation?

4. Run a regression model with alumni giving rate as the dependent variable and the student/faculty ratio as the explanatory variable.
   a. What is the interpretation of the estimated coefficients? Make sure to discuss whether it is statistically significant and, if so, at what level.

   b. What is R-square for this model, and what is the interpretation?

5. Discuss to what extent you can think of the explanatory variables as having a **causal** relationship with the dependent in these two cases.