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.

Answer: See figures at end. In both cases, a linear models seems 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.

Answer: for sfratio, classeslt20, and alumnigivingrate:

	mean	sd	min	max
	<dbl>	<dbl>	<dbl>	<dbl>
1	11.5	4.85	3	23
2	55.7	13.2	29	77
3	29.3	13.4	7	67

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.

      Answer: For a one-percentage point increase in the number of class with less than 20 students the percentage of alumni that donate increases by 0.66 percentage points. This is statistically significant at the 0.001 percent level.

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

Answer: The variation in the percent of classes with less than 20 students explain 41.7% of the variation in the alumni giving rate.

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.

      Answer: As the student/faculty ratio increases by one (one extra student per faculty) the percentage of alumni that donate decreases by 2.06 percentage points. This is statistically significant at the 0.001 percent level.

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

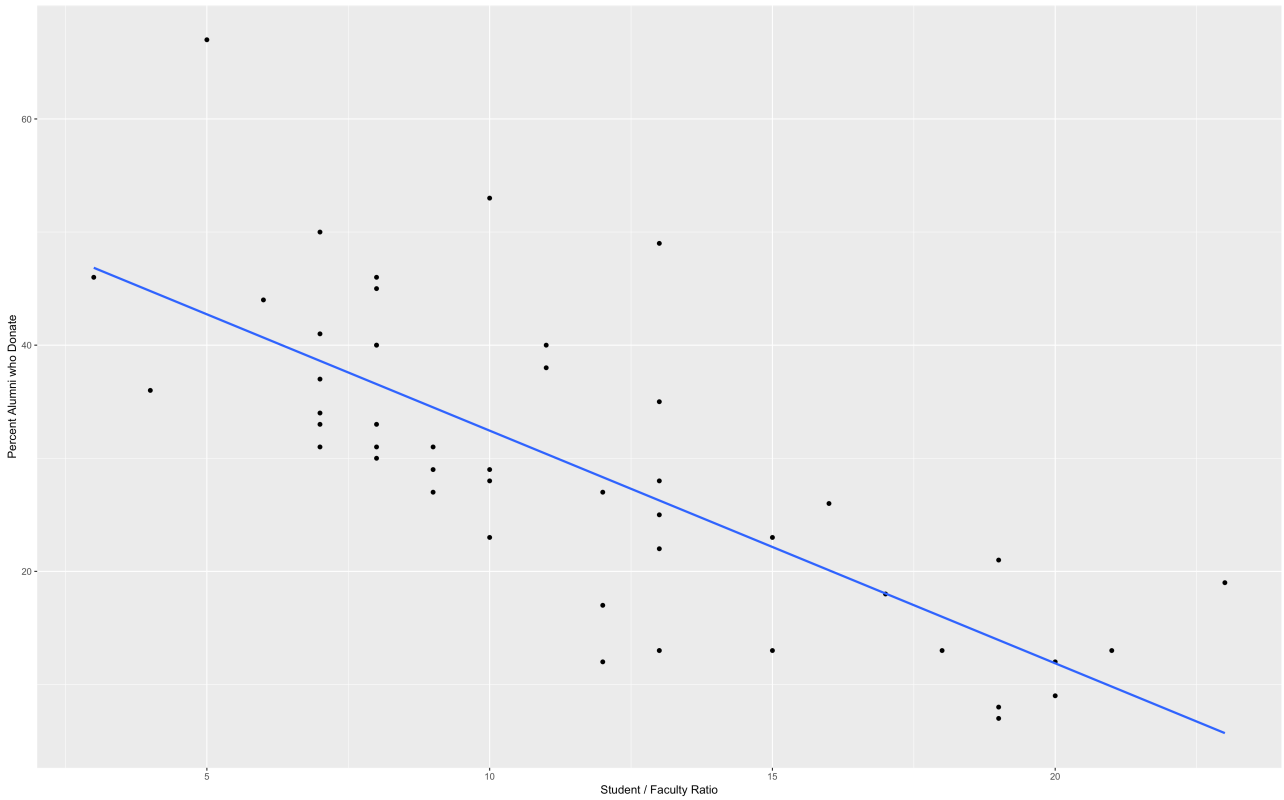
Answer: The variation in the student / faculty ratio explain 55.1% of the variation in the alumni giving rate.

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

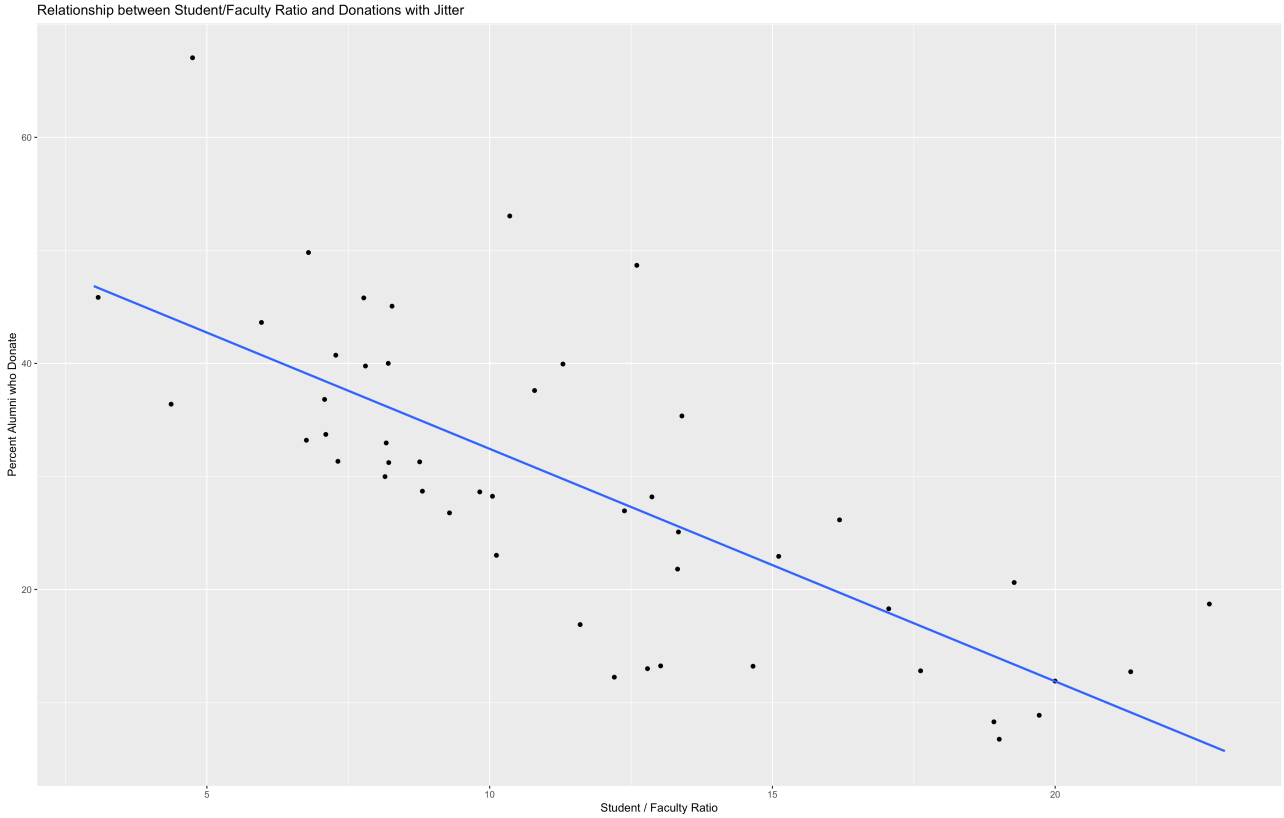
Answer: Thinking of both variables as causal might be problematic. First, it is likely that it is the student/faculty ratio and number of class with less than 20 students when the alumni were in school that matter rather than the current numbers. Secondly, a school with a high alumni giving rate is more likely to be able to afford lowering the student/faculty ratio and increasing the number of class with less than 20 students.

Figures

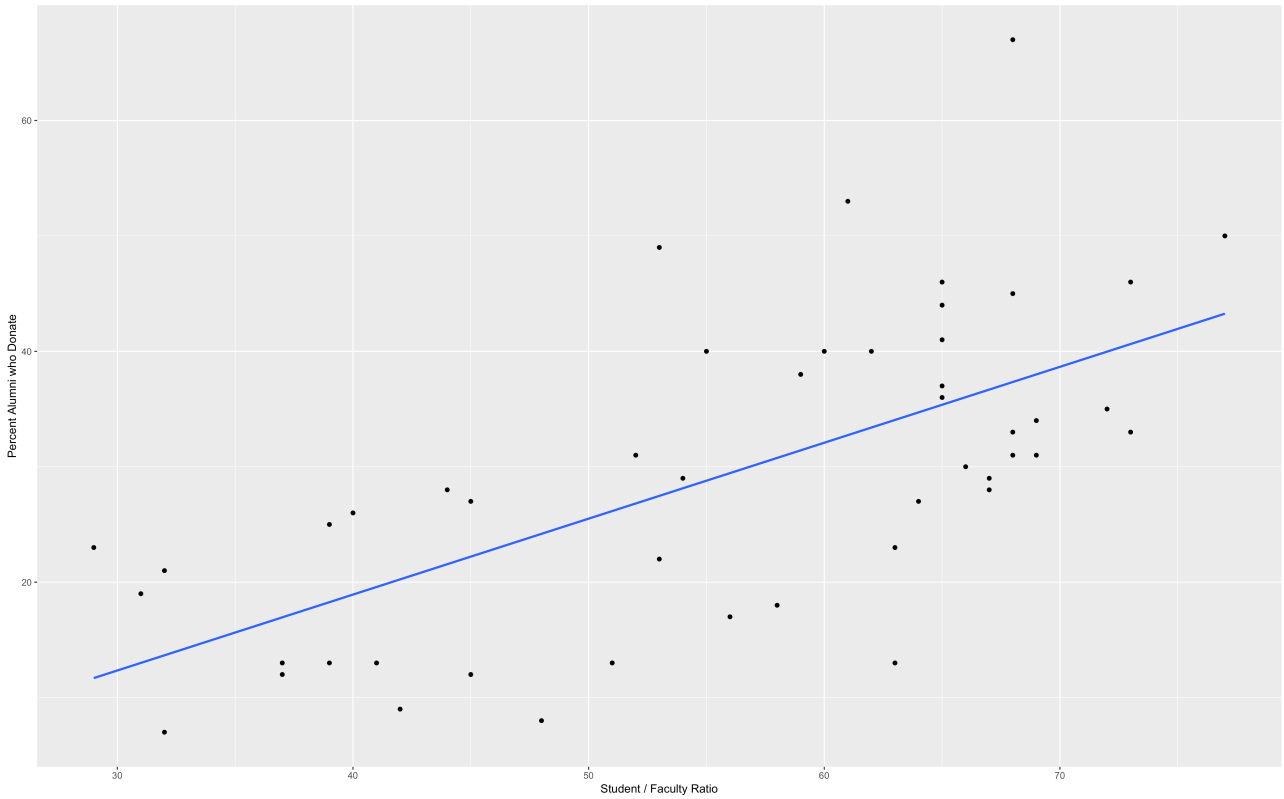
Relationship between Student/Faculty Ratio and Donations



Relationship between Student/Faculty Ratio and Donations with Jitter



Relationship between Percent Classes with less than 20 Students and Donations



Relationship between Percent Classes with less than 20 Students and Donations with Jitter