

Unit 12 Pre-Class Exercise

w203 Instructional Team

```
load("elects.Rdata")
```

The `spending` data frame contains simulated data on election results and campaign spending in the U.S. The `dshare` and `rshare` variables capture the vote shares for the Democratic and Republican parties respectively, while `spent` measures total spending on advertisements in millions of dollars for each race. For this exercise, think of spending as the outcome variable and the vote share variables as predictors.

```
head(spending)
```

```
##      dshare    rshare spent
## 1 0.4573411 0.5426589   3.24
## 2 0.4969932 0.5030068   5.65
## 3 0.4222159 0.5777841   6.84
## 4 0.4220736 0.5779264   8.06
## 5 0.4173364 0.5826636   1.93
## 6 0.4272072 0.5727928   6.82
```

1. Is there a linear relationship between campaign spending and democratic vote share? Generate a scatter plot with a regression line. What does this scatter plot suggest about the appropriateness of the classical linear model assumptions in this case?
2. If we want to improve this model, we have two options - add more variables or transform the variables we have. Let's try to add another variable first. Note that the two vote share variables do not sum to one. This could be due to the existence of a third political party or to error in measuring votes. To begin with, what assumption would be violated if the two variables did sum to one?
3. Since the two variables don't sum exactly to one, we should be able to include republican vote share as a second predictor. But before doing any coding, ask yourself what adding republican voteshare will do to the precision of our estimate on the effect of democratic vote share. Explain why this effect makes sense.
4. Now generate a model that predicts campaign spending using both democratic and republican vote shares. What does the result tell you about the practical implications of very highly correlated predictors?
5. Finally, what is the transformation of variables that might solve the problems identified thus far? Perform that transformation and describe the results.