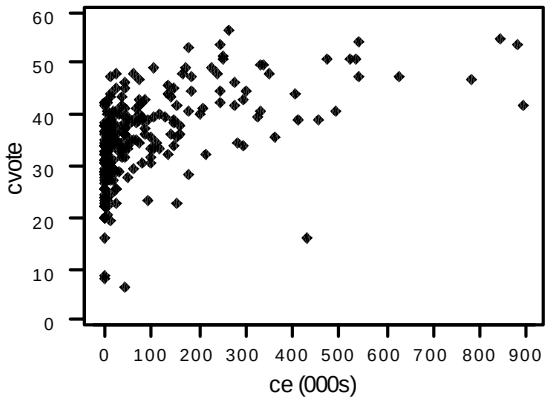
Professor Green
Statistics 102 (Political Science)

Answer sheet for Problem Set #2

---

1. Using the Graph > Plot menu, produce a scatterplot of *cvote* (Y) and challenger in $1000s. Interpret this graph. Is the relationship linear?

The scatterplot is below and does not show a strong linear relationship. It **does** show a strong association, but it's a curvilinear one.



2. Based on this plot, what is your approximate sense of what the average vote is for challengers spending more than $500,000? How does this differ from the average vote for all challengers?

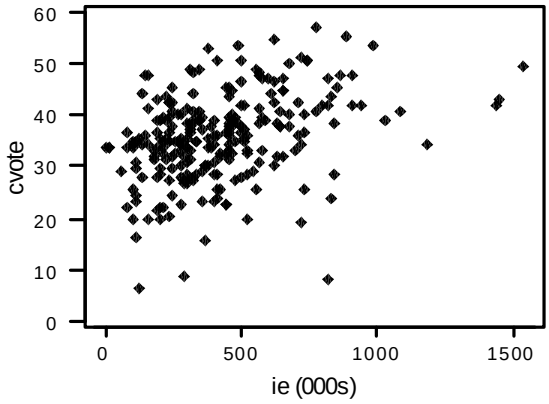
Using this scatterplot as a means to make this approximation, it appears that challengers spending more than $500,000 obtain, on average, nearly 50% of the vote. The mean percentage of votes obtained by all challengers is 35.79%.

3. Calculate the correlation between challenger vote and challenger expenditures in $1000s. How does this correlation change when challenger expenditures are measured in dollars?

The correlation between challenger vote and challenger expenditures is .507. This correlation does not change when challenger expenditures are measured in dollars.

4. Produce a scatterplot of *cvote* and incumbent expenditures (in $1000s), and calculate the correlation between these two variables. What do you infer about the extent to which incumbent campaign expenditures influence the vote?

The scatterplot of these two variables is produced below. The correlation between them is .360. The fact that the correlation is positive suggests the naïve interpretation that the more incumbents spend, the worse they do. The problem with this interpretation, however, is that incumbents spend more as they perceive themselves to be more vulnerable to defeat. Thus, the vote outcome is arguably both a cause and consequence of incumbent spending.



5. Perform a regression in which *cvote* is the response (dependent) variable and challenger expenditures (in $1000s) is the predictor variable. Pull down the Stat > Regression menu. What is the resulting "regression equation"? What do the numbers in this equation mean substantively?

The resulting regression equation is **y [cvote] = 33.2 + .0274 x [ce(000s)].** Substantively, this means that for every $1000 spent by a challenger, his or her share of the vote obtained increases by .0274%. Therefore, for an expenditure of $150,000, this would equation predicts that a challenger would obtain 37.31 percent of the vote.

6. Describe a hypothetical experiment that would, in principle, allow one to estimate the effects of challenger and incumbent expenditures on the vote.

Such an experiment would require that there be **random assignment** of both incumbent and challengers into treatment and control groups. The treatment groups would be

assigned randomly to varying campaign expenditures. This would permit the estimation of the effect of campaign expenditures on the vote obtained.

7. Use the Stats > Tables menu to generate a crosstabulation of *close* and *cpq*. Be sure to percentage your table so as to answer the question: To what extent do more experienced challengers fare better on election day?

The following crosstabulation table can be generated:

Close\CPQ	0	1	2	3	4	5	6	All
1	83.33	81.93	38.10	47.06	61.11	40.00	--	71.07
2	16.67	18.07	33.33	47.06	33.33	40.00	50.00	23.97
3	--	--	28.57	5.88	5.56	20.00	50.00	4.96
All	100	100	100	100	100	100	100	100.00

Generally, the worst results were obtained by the least experienced candidates ["0s" and "1s"], and the best results by the most experienced ["6s"]. However, experience alone may not necessarily be an indicator of a challenger's success in obtaining votes. For example, nearly 29% of challengers in the relatively inexperienced "2" category fell into the most successful category of votes obtained, which is more than the percentage of "3", "4" and "5" candidates who did so. Similarly, candidates in each of these latter experience categories were more likely to fall into the most unsuccessful category than the less experienced "2" candidates.