

## PUBPOL 639: ASSIGNMENT 1

### Winter 2011

Due: Monday, February 7<sup>th</sup> at the start of class

In this assignment you will examine the relationship between school resources and educational outcomes. The data is a sample of 420 school districts in California. It has been posted on the CTools site as well as the online archive. Assume you can treat this data as a random sample of all US school districts. It includes information on average student test scores, expenditure per student, total enrollment, number of computers per district, the average student-teacher ratio, the identity and location of the district (e.g. urban/rural, county), and some socioeconomic measures of students and families in the district (average family income, fraction eligible for free lunch, fraction potentially eligible for TANF).

1. Write down two important causal questions that one could potentially use this data to explore?
2. Write down two important non-causal questions that one could potentially use this data to explore?
3. Pick one of your causal questions from above. What is an ideal experiment that you could use to answer it?

```
sum testscr comp_stu avginc
```

Variable	Obs	Mean	Std. Dev.	Min	Max
-----+-----					
testscr	420	654.1565	19.05335	605.55	706.75
comp_stu	420	.1359266	.0649558	0	.4208333

4. Above is a table of summary statistics for two key variables: test scores (`testscr`) and computers per student (`comp_stu`). Use this table of summary statistics to report or calculate the following:
  - a. Average number of computers per student across all districts
  - b. Standard deviation and variance of computers per student
  - c. Standard error of average computers per student
  - d. Standard error of average test scores
  - e. 95% Confidence Interval for mean number of computers per student. Show your work.

```
by comp_group: sum testscr
```

```
-> comp_group = Low
```

Variable	Obs	Mean	Std. Dev.	Min	Max
testscr	210	649.3207	17.95616	609	700.3

```
-> comp_group = High
```

Variable	Obs	Mean	Std. Dev.	Min	Max
testscr	210	658.9924	18.9309	605.55	706.75

5. The table above summarizes test score data for districts with high levels of computer availability and those with low levels of computer availability. “High” is defined as having computers per student greater than the median.
  - a. Calculate the difference in mean test scores between high- and low-computer use districts
  - b. Calculate the standard error of this difference. Assume that high and low computer-use districts were sampled independently.
  - c. Test the null hypothesis that the mean test scores for the two types of districts are the same. State your null hypothesis, your alternative hypothesis, your test statistic, and your conclusion.
  
6. In this question, you will conduct analysis similar to question 5, but will generate the summary statistics yourself. Create a Stata “do-file” that will enable you to answer the following using the dataset assignment1.dta. Include your do-file and log file as an appendix to your assignment.
  - a. How many total districts are in the dataset?
  - b. How many districts span K-6<sup>th</sup> grade and how many span K-8<sup>th</sup> grade?
  - c. Construct a 95% confidence interval for the mean of district income. Show all work.
  - d. Test the null hypothesis that the mean of district income in the population of all districts is equal to \$16,000, using an alpha level of .05. Specify null and alternative hypotheses, test statistic used, and interpret your result.
  - e. Test the null hypothesis that the mean of district income is the same in high and low computer use districts, where high is defined as having computers per student above the sample median. To do this, you will need to construct a measure of high/low computer use districts as was used in question 5. State your null hypothesis, your alternative hypothesis, your test statistic, and your conclusion.
  
7. In light of your answers to questions 5 and 6, can you conclude that investing in more computers would raise students’ test scores? Explain.