

Quantitative Analysis in International Affairs
SIS 600 Prof. Powner Fall 2009

Problem Set 1: Descriptive Statistics

Due Week 3: Sept 14/17

General policies on problems sets: You may collaborate with others in this class (from either of my sections). Each person, however, *must* submit his/her own set of answers and generate his/her own graphs and results. You may print graphs and figures in black and white; if you do so, please be sure that any markers on the graph that depend on color (and not just shape) for differentiation are adjusted. Always provide a key on graphs unless the graph shows a single piece of information that you can capture in the title (i.e., “Economic Growth in Thailand, 1970-2005”). Any “math” can be done by hand, but verbal responses more than a couple sentences long should be typed. It’s ok to go back and forth between typed sections and handwritten ones on your problem sets; always identify the question part to which you’re responding, no matter the format. “Briefly discuss” generally means 2-5 sentences. Please track how long it takes you to complete each problem set and note this at its end.

1. Use the “Econ Dataset” you made in class.
 - a) Use Excel to calculate the average growth across these states in the studied years.
 - b) Which state has the highest average growth? Which state has the lowest? Which state’s growth rates span the largest range (i.e., the largest difference between its highest and lowest rates)?
 - c) Go back to WDI and get the world’s GDP growth rate during the studied years. (*Hint*: “World” is one of the entries in the “Countries” list.) Add this data to your dataset. Is “World” a new case, or a new variable? Briefly defend your answer.
 - d) Create a new time-series graph showing the three states you identified in part B, the sample average, and the world average.
 - e) Briefly describe highlights of this graph. (3-4 sentences)

2. Use the “Gender Dataset” you made in class.
 - a) Use the formula functions in Excel to calculate percentage change for birth rate and your additional variable for each country.
 - b) Create a column graph for birth rate showing the countries in ascending order of percentage change. (*Hint*: Excel graphs observations in the order they appear in the dataset.)
 - c) Create a bar graph for your additional variable showing the countries in descending order of percentage change on the additional variable.
 - a) Calculate average percentage change for each observed year across countries.
 - b) Make a graph (your choice of type) showing average percentage change across the two observed years.

3. Think of another variable that should have a direct relationship with the crude birth rate, that is, one that you would expect to move up (down) as the birth rate increases (decreases).
 - a) Identify this variable, and briefly describe the logic of why you expect it to trend with the crude birth rate.
 - b) Use the WDI to capture your new variable and the crude birthrate for all countries for the year 2000. Export your dataset to Excel. (*Hint:* Exclude “World” and any regions or country groupings.)
 - c) Create a scatterplot of your two variables. Do you find the expected relationship? Just from eyeballing it, how strong is the relationship: Are you totally convinced that the two variables trend together, or is the relationship more iffy? On what feature(s) of the graph do you base your judgment?
 - d) Use the WDI to create a time series dataset for two countries, one developed and one developing, that includes crude birth rate and your new variable for all available years.
 - e) Consider the variables you chose. Do you expect the time-series graph for the developed country to look the same as the one for the developing country? Why or why not? (Think theoretically.) How do you expect the graphs to differ?
 - f) Create a time series graph of your two countries.
 - g) Does your graph in part f match your expectations from part e? Why or why not? Briefly discuss your graph and its congruence with your expectations. (3-4 sentences)

4. Select any interval-ratio level variable and region of interest to you in the WDI dataset, and generate a cross-sectional dataset for the year 1994.
 - a) What is the maximum value in your dataset? What is the minimum value? What is the range?
 - b) Find the mean value of the variable.
 - c) Find the median value of the variable.
 - d) Use SPSS to generate a histogram of the variable.
 - e) Convert your interval-ratio variable into an ordinal one. Use common sense or the histogram to establish the class intervals for your new variable, then generate a new variable to capture the ordinal data.
 - f) Use SPSS to generate a histogram of the transformed variable.
 - g) Compare your new histogram to the one you produced in part d. What similarities and differences do you notice?
 - h) Think of a research question or research use for which you would probably prefer to use the original interval-ratio level data. Identify a research question or research use for which you would probably prefer to use the ordinal data.

5. Get some practice using the MS Word Equation Editor.
 - a) Type the formula for finding the mean.
 - b) Type the formula for the standard deviation (from ch 5, since you’re reading it anyway).
 - c) Type the formula for for $Z(\text{obtained})$. (See the inside cover of Healey or ch 8.)

6. **Bonus:** Time for a bit of fun. Grossly misrepresent some piece of statistical data from the WDI. You might display changes in the value of some variable for some country across several years; you might choose to exaggerate the disparity of two countries (or regions) on some variable. You might choose to do any number of things. The key thing here is to downplay or overplay differences so as to make a wildly exaggerated picture of the true situation.