**POL 199: SPRING 2009**

**PROBLEM SET #1**
**DUE IN-CLASS, THURSDAY FEBRUARY 19**

**DIRECTIONS**
Answer each of the questions below. You should submit your answers in the following form (points will be deducted for improper form):

1. Type up all of your answers, using a reasonable font and margins.
2. Use complete sentences in answering all of the questions.
3. Properly label your problem set with (a) your name, (b) the number of the question you are answering. If it makes sense to do so, you can combine parts of a question together in your answer (i.e. do parts 1a, 1b, and 1c together, etc.). Just be sure to label your answers appropriately.
4. Simply cutting and pasting output from STATA into Word is not enough to answer the question (unless the question specifically specifies otherwise). You should run the analysis in STATA, then pick out the information you need to answer the question and write it up in your problem set. If you are including a table of data in your answers, then you should copy it from STATA into Excel, then properly format it and keep only the information you need to answer the question.

See the example of how to properly submit a problem set.

**QUESTION 1**
The table below contains some data that might come up in a discussion of gun control.

<table>
<thead>
<tr>
<th>Country</th>
<th>Death Rate from Firearms</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>13.7</td>
</tr>
<tr>
<td>Australia</td>
<td>2.9</td>
</tr>
<tr>
<td>Canada</td>
<td>3.9</td>
</tr>
<tr>
<td>Denmark</td>
<td>2.1</td>
</tr>
<tr>
<td>England and Wales</td>
<td>0.4</td>
</tr>
<tr>
<td>France</td>
<td>6.3</td>
</tr>
<tr>
<td>Israel</td>
<td>2.8</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.5</td>
</tr>
<tr>
<td>New Zealand</td>
<td>3.1</td>
</tr>
<tr>
<td>Norway</td>
<td>4.3</td>
</tr>
<tr>
<td>Scotland</td>
<td>0.6</td>
</tr>
</tbody>
</table>

a. Enter the data in STATA. Create a copy of your table using the list command (use the help menu if needed). Using the list command, create a table that lists all the countries and death rates from firearms (i.e. it should duplicate the table above). Use Excel to properly format the table and paste it into your problem set.
b. What is the mean death rate from firearms?
c. What is the median death rate from firearms?
d. What does mean and median tell you about the distribution of the data? Explain your answer.
e. Imagine that you are working with policymakers on gun control policy. They want to know how the United States compares to the middle 50% of countries in terms of death rates from firearms. In other words, there will obviously be some countries with relatively high and relatively low death rates from firearms. What is the range of death rates for the 50% of countries in the middle? As you report your answer, and indicate how the US compares.

f. What is the standard deviation of death rates? Interpret what this means.

g. What is the z-score of the United States? Interpret what this means.

h. Use STATA to draw a histogram to describe the shape of the distribution. How would you describe the distribution’s shape? (Normal vs. Skewed?) Properly format the histogram and paste it into your problem set.

**QUESTION 2**

The purpose of this question is to give you practice recoding data and constructing scales. For this question, download the American National Elections Study 2004. Download the dataset and the codebook from the course conference. Let’s say we want to know how much respondents trust government.

a. Identify the construct you are trying to measure.

b. What are the indicators of that construct? Using the ANES 2004 codebook, find questions that are good indicators of whether or not the respondent trusts government. In particular, there are four questions that the ANES identifies as being part of the “Trust in Government” index. Indicate what those four questions are, along with their variable names (NOTE: do NOT use v141, which is a different scale that has already been created).

c. For the first indicator that you find, write out the original question and response categories, and indicate how you will recode it for scale construction. Then, do the recoding in STATA, making sure to properly label the recoded variables. Then run a frequency table of your new recoded variable and cut and paste that frequency table from STATA into your problem set. Recode each item so that higher numbers mean more trust.

d. For the second indicator that you find, do the same stuff that you did in part c.

e. For the third indicator that you find, do the same stuff that you did in part c.

f. For the fourth indicator that you find, do the same stuff that you did in part c.

g. Create a z-score for the first indicator. Indicate what the minimum value, maximum value, mean, and standard deviation of the z-score is.

h. Create a z-score for the second indicator. Indicate what the minimum value, maximum value, mean, and standard deviation of the z-score is.

i. Create a z-score for the third indicator. Indicate what the minimum value, maximum value, mean, and standard deviation of the z-score is.

j. Create a z-score for the fourth indicator. Indicate what the minimum value, maximum value, mean, and standard deviation of the z-score is.

k. Using the four z-scores that you just created, create an aggregate scale for trust in government. The scale should be the average of each respondent’s score on the four indicators. Indicate what the minimum value, maximum value, mean, and standard deviation is for the scale.

**QUESTION 3**

The purpose of this question is to give you practice recoding variables. Download Fearon.dta from the course conference. Let’s say you want to examine the relationship between ethnic fractionalization and wars in your dataset.

a. Create a joint distribution table of the variable wars and ethnic fractionalization.

b. Let’s say you want to divide ethnic fractionalization into 4 categories. Identify the cutpoints for the 25th, 50th, and 75th percentile of ethnic fractionalization.

c. Using those cutpoints, create a new variable (called ethfrac4) that divides the ethnic fractionalization variable into 4 categories: (1) less than or equal to the 25th percentile, (2) less
than or equal to the 50th percentile, (3) less than or equal to the 75th percentile, (4) the top 25th percentile. Run a frequency table of your variable and show your results.

d. Create a crosstab between ethfrac4 and wars, with wars in rows and ethfrac4 in columns. Show your results.

e. Based on the work you did, describe the relationship between ethnic fractionalization and war.

f. Create a simple bar chart in Excel that shows the percentage of cases in each region of the world (use the region variable).