POS 3713: Assignment 2
Assigned: Monday, 2/7/2000
Due date: In class, Monday, 2/14/2000
Tutorial Session: Thursday, 2/10/2000, 9am-10:45am
Professors Mitchell and Lubell

The purpose of this assignment is to introduce you to various descriptive statistics. You will use SPSS for each part of this assignment. Remember to type up your responses to each question and to either 1) include the output directly in your text, or 2) attach your printed output from SPSS.

We will be using the 1996 National Election Study data, so you should begin by opening the file NES1996Edit.sav in SPSS. To launch the program, click on the Start button, select Programs, and then click on SPSS Windows 9.0. If you saved this file in Assignment 1 on your floppy disk, simply open that saved file on your disk (Click on "File", "Open", and then find the file on your disk). Otherwise, you will need to go back to the web site (ftp://www.coss.fsu.edu) to download the data again. When you are finished with this assignment, be sure to save your working data file again by clicking on "File", and then "Save As".

Preparation: Before you do the analysis in this assignment, you need to change the variable names and labels, in addition to assigning Don't Know/NA categories as missing. This is a standard procedure that we will use for most of our assignments. Saving the data file when you are finished will preserve these changes.

➢ Double click on V960115 and change the variable name to "Region". Then click on "Labels", and change the variable label to "Region". Do the same for V960573 (Variable name = "Prayoftn"; Variable label = "How often R prays"), V960281 (Variable name = "Hillther"; Variable label = "Hillary Clinton Thermometer"), V960272 (Note: this should already be labeled "Clinther" from assignment #1. Change the variable label from "I understand the relabeling procedure" to "Clinton Thermometer".)

➢ You also need to recode the DK and NA categories as system missing values. For each variable that needs re-coding, click on "Transform", "Recode", "Into Same Variable" (see instructions on assignment #1). For example, for the variable "Prayoftn", you need to recode values 8 and 9 as missing. For Hillther (V960281), recode values 997, 998, and 999 as system missing. Finally for Clinther (V960272), recode 998 and 999 as system missing.

Part A: Frequencies and Level of Measurement

One of the most useful pieces of information to obtain in a data set is a frequency distribution of the variables. If you look at the NES Codebook (EditedNES1996Codebook.doc), you will notice that each variable contains the survey question being asked, in addition to each possible response (with the values for each response identified). The numbers to the left of these values are the frequencies for each category of the variable. For example, take a look at the census region variable (Var 960115). You will notice that there are four region categories: 1 = Northeast, 2 = North Central, 3 = South, and 4 = West. The frequencies for these categories are also reported. There are 260 respondents from the Northeast, 458 respondents from the North Central region, 642 respondents from the South, and 354 respondents from the West. You can also obtain this information in SPSS.

Question A1: Produce a frequency distribution of the region variable (v960115). Are the frequencies reported in the codebook the same as those obtained in SPSS for this variable?

➢ Select "Analyze" menu
➢ Select "Descriptive Statistics" and then "Frequencies"
➢ Move Region (v960115) into the Variable box by selecting the variable and using the arrow, or by double clicking the variable
➢ Copy this table into your word processing program. To preserve the formatting of the table, click on the table once in the output window, select Edit and then Copy Objects (Ctrl-K). Then click back into your word processor and select Edit, and Paste Special (or just Paste if this is not an option). If you do a simple Copy and Paste, the information will appear in your word processing program, but the formatting of the output as it appears in SPSS will be lost in the translation.

Question A2: What is the level of measurement for this variable (nominal, ordinal, or interval)? Explain your answer.

Question A3: Produce a frequency distribution for Prayoftn (v960573, How frequently does the respondent pray) using the instructions above. What is the level of measurement for this variable (nominal, ordinal, or interval)? Explain your answer.

Part B: Crosstabulation
As Political Scientists, we are often interested in assessing the relationship between two or more variables, such as party identification and vote, or education and political knowledge. For variables that are measured at the nominal or ordinal level, it is useful to examine the relationship between two or more variables in the form of a contingency table or crosstabulation. One variable is displayed in the rows of the table, while the other variable is displayed in the columns of the table.

For example, in his 1996 article "Charting a Course to Conflict: Territorial Issues and Interstate Conflict, 1816-1992", Dr. Paul Hensel examines the relationship between the issues at stake in militarized disputes and the escalation level of these disputes. He argues that disputes over territorial stakes will be much more likely to escalate to war than disputes over other issues. He tests this relationship with the following data on issues at stake in militarized disputes. Both variables are measured at the nominal level.

Table 2
Militarized Dispute Escalation to War

Territorial Issues at Stake?	Non-War Disputes	Interstate War	Total
No	1420	36	1456
Yes	543	43	586
Total	1963	79	2042

You can see that 2.5% (36/1456) of disputes not involving territorial issues escalate to war, whereas 7.3% (43/586) of disputes over territory escalate to war. He concludes that disputes involving territorial stakes are 3 times more likely to escalate to the level of full-scale war.

Question B1: Produce a contingency table with the variables "Region" and "Prayoftn" using the instructions below. Answer the following questions:

    a) How many respondents in the sample are from the West region and pray once a day?
    b) What percentage of respondents never pray?
    c) What percentage of respondents pray several times a day and live in the North Central region?
    d) What can you conclude in general about the relationship between how often a person prays and the region they live in? Is there any tendency for people in one region of the US to pray more often than people in other regions? If so, what could account for this difference?

➢ Select "Analyze", "Descriptive Statistics", and then "Crosstabs".
➢ Click on the variable "Region" and place it in the "Column" box.
➢ Click on the variable "Prayoftn" and place it in the "Row" box.
➢ Click on "Cells" at the bottom of that window, and then click on the boxes for "Row" and "Column" percentages. Click "Continue", and then "OK".

Part C: Descriptive Statistics for Interval Measures
Finally, we will take a look at some basic descriptive statistics including the mean, median, and mode. These statistics are most appropriate for variables measured at the interval level. You will be comparing people's attitudes about Bill Clinton and Hillary Clinton, using the feeling thermometer scales (Clinther and Hillther).

Question C1: Produce a descriptive statistics table for the variables "Clinther" and "Hillther" using the instructions below.

    a) Compare the mean thermometer levels for Bill and Hillary Clinton. Who do the respondents have warmer feelings for (represented by higher average scores on the thermometer variables)? Do you agree with these averages, i.e., would you rank Bill and Hillary Clinton in the same order based on your feelings about them?
    b) Compare the median and mode thermometer levels for Bill and Hillary Clinton. Is your comparison of these two measures of central tendency similar to what you found in part (a)? Explain how these measures of central tendency differ from the technical standpoint.

➢ Select "Analyze", "Descriptive Statistics", and then "Frequencies". Double click on the variables "Clinther" and "Hillther" to move them into the Variable(s) box.
➢ Next, click on the "Statistics" bar at the bottom of that window. Under the section labeled "Central Tendency", click on the boxes for mean, median, and mode. Click "Continue" and then "OK".