

POS 3713: Assignment 1

Assigned: Monday, 1/31/2000

Due date: In class, Monday, 2/7/2000

Tutorial Session: Thursday, 2/3/2000, 9am-10:45am

Professors Mitchell and Lubell

Part A: Finding the Data

We will be accessing the data for our laboratory assignments from the computing lab File Transfer Protocol (FTP) site on the Internet. To access the data, open up Netscape or Internet Explorer and type <ftp://www.coss.fsu.edu> in the URL address field. You will then have access to the files for the class (in the POS3713 directory). When you click on the files, you will have the option of opening them or saving them to disk. We recommend keeping a copy of each file on your own floppy disk. Note that you can access these files from any computer with Internet access, including your home computer. However, to work on the data you must have SPSS installed on the computer. Click on **POS3713** and you will have access to four files:

1) *EditedNES1996CodeBook.doc*

This Microsoft Word file is the codebook for each of the variables in the 1996 National Election Study (NES) dataset. The codebook contains the variable names and definitions for each of the values in the dataset. You should always look at this codebook to understand the meaning of the specific variables you are working with. We do not recommend printing this file due to its length.

2) *EditedNESVariableList.doc*

This Microsoft Word file is the list of variables included in the NES dataset. The list of variables includes the variable name and a brief description. We recommend printing this file and keeping notes about the various items on this list.

3) *NES1996Edit.sav*

This is the SPSS dataset containing the 1996 Election Study Data. The data was collected prior to the 1996 election. The units of analysis are individual citizens randomly selected from the United States population.

4) *World95.sav*

This is the SPSS dataset containing data for a 1995 world survey collected from 109 countries. It includes information such as Gross National Product, life expectancy, and literacy.

All of your laboratory exercises should be completed using Microsoft Word (or any other word processor) and then printed out in the computer lab or on some other computer. When appropriate, the Microsoft Word document can include SPSS output pasted into the Word document from the SPSS output screen (see exercise B below). For some exercises, it is more efficient to turn in the edited SPSS output. In all cases, choose the method that saves the most space and paper. Turn in the printed results of your work to your professor; we will NOT accept handwritten homework. You should always save your homework on a 3.5" floppy disk. Erasure of homework from the C:/ drive of the data lab computer is NOT an acceptable excuse. If you lose your copy of the homework assignment, it is also available on Professor Mitchell's homepage, <http://www.geocities.com/CollegePark/Plaza/3094/>.

Saving the Data

The files on the website are read-only files, so you cannot save any modifications you make to these files on the FTP site. To save your work, you must save the file at a different location. You can save the file on the C: drive of the data lab computer, but this is a risky strategy because the data lab staff eliminates student files from the computer once per week. Overall, the best strategy is to save the file on the C: drive of the computer when working in the data lab, then save your final work on a 3.5" floppy disk placed in the A:/ drive of the computer. A second option is to save the SPSS file directly on a 3.5" floppy disk and work from the A:/ drive, but this will be significantly slower than working from the C: drive. To save your file with a name of your choosing at a specific location, use the mouse to choose the following menu options:

- Select "File" menu
- Select "Save as"
- Enter the name of your file in the "File name:" text box, and choose the appropriate location from the "Save In:" drop-down menu.

Part A Exercises

Exercise A1. Find the following variables from the NES code book: v960272, v960417, v960548. What survey questions are these variables responses to? What are the possible values for each variable, and what does each value mean? What concept would you say each of the variables is measuring?

Part B: Renaming and Labeling

It is often useful to rename and label a variable and the variable's values using words and phrases that suggest the meaning of the variable to the researcher.

Part B Exercises

Exercise B1. Rename the variables from above with the following names: v960272= "ClinTher", v960417 = "Party Id", and v960548= "VotePref". Note that variable names cannot exceed 8 characters, and must begin with a letter. To rename the variables:

- Double click on variable heading
- Enter the new variable name in the "variable name" text box
- Click "Ok"

Exercise B2. Add the phrase "I understand the relabeling procedure" to the variable label for *ClinTher*.

- Double click on variable heading
- Click "Labels"
- Enter phrase at the end of current variable label
- Click "Continue" and then "Ok"

Exercise B3. Add 3 new value labels to the *ClinTher* variable. Label 0= very cold, 100 = very warm, and 50= neutral. These are the endpoints and the middle of the feeling thermometer scale.

- Double click on variable heading
- Click "Labels"
- Enter the value and label in the appropriate text boxes
- Click "Add", and repeat this step and previous step for each value
- Click "Continue" and then "Ok"

To turn in the results of this exercise, click on the "Utilities" menu and then choose "File Info". This will produce a long SPSS output file, which lists all variables and labels. Double click the table containing all the variable names and information, so that you can navigate within the table using the side bar. Then, find each of the three variables manipulated in this exercise. Using the mouse, highlight all of the information for each variable and hit "Ctrl-C" (or use the Edit menu) to copy the information into memory. Next, paste (ctrl-V, or use the edit menu) the information for each variable into the Word document you are using for this homework. You should then have information for all three variables in your homework Word document.

Part C: Recoding and Computing New Variables

Survey companies often use values to represent answers to survey questions like “don’t know” or “not applicable”. While it is important for descriptive purposes to know how many respondents choose these values, these values often don’t make sense for the purpose of more complicated statistical analyses. A common strategy for analysis is to transform these variables into missing data. Often times, analysts want to create a new variable based on the values of a different variable for the purposes of representing the data in a different format.

Part C Exercises

Exercise C1: Recode unknown responses of *PartyID* into missing data. In other words, transform the values [4,5,8,9] into missing data.

- Select “Transform” Menu
- Select “Recode Into Same Variable” (note that instead of transforming values within the same variable, you can choose to create a new variable containing the transformed values)
- Move *PartyID* into the “Numeric Variable” text box by selecting the variable and using the arrow, or double clicking the variable
- Click “Old and New Values”
- Enter each value you want to change in the “Old Values” text box and click “System-missing” in the new variables text box
- Click “Add” for each value change
- Click “Continue” and then “Ok”

Exercise C2. Compute a new, 4-category feeling thermometer scale using *ClinTher* as a basis. The categories should be 1= [0,25], 2 = [26-50], 3= [51-75], and 4= [76-100]. Name the variable *NewTherm*, label the variable “Collapsed Clinton feeling thermometer”, and label the values 1= Cold, 2=Luke Warm, 3=Warm, 4= Hot. Use the following steps to computer the new variable:

- Select the “Transform” menu
- Select “Compute”
- Type in name of new variable in “Target Variable” text box
- Enter new first new value (“1” in this case) into “Numeric expression” text box
- Click “If” and enter appropriate conditions for category 1 ($ClinTher >= 0$ and $Clinther <= 25$)
- Click “Continue” and then “Ok”
- Repeat for each of the categories (2,3,4 still need to be done!), and choose “Ok” when prompted to replace existing variable

To produce the needed output for the Part C exercises, select the “Analyze” menu and then click “Frequencies”. Then select both *PartyID* and *NewTherm* and click “Ok”. The output should then give you the correct frequencies for the recoded and new variable. Note that you can print this output page directly, since it is very small, or cut and paste the frequency tables into Microsoft Word.

To check your work for the *NewTherm* variable, select the “Analyze” menu, click “Descriptive statistics” and then “Crosstab”. Enter *ClinTher* as the Row and *NewTherm* as the column and click “OK”. The categories of *NewTherm* should correspond to the appropriate value of *ClinTher*. You don’t have to turn in the cross-tabulation; just make sure the frequency is correct.