

---

Sociology 350  
Lab Challenge #3

Your Name: \_\_\_\_\_

Subject: Converting ASCII (Raw) Data to SPSS Data

---

—

This lab requires that you begin by first downloading the highly elusive *cuban.dat* file which is contained within the course D2L, but which is not easily copied to your *Personal On-Line File*. It is possible to copy it directly in this way, but I cannot give you the simple and clear directions as to how to accomplish it. So you may attempt it if you dare, or not. Instead I am giving you a CD-ROM with a copy of *cuban.dat* on it and you may download this directly into the SPSS program described below.

To do this you will need to open the CD-ROM file once you have entered the SPSS program at your lab station.

Next, click [OK] to open an existing data source.

Next, go to the bottom of the “Files of Type” listings and click “All files” from the menu. Then download *cuban.dat* with a left double click.

Now at the top of your computer screen it should read “Text Import Wizard- Step 1 of 6.”

Proceed by clicking “Next”

Now in response to Step 2 of 6, click “Next”

Now in response to Step 3 of 6, click “Next”

Now in response to Step 4 of 6, use your computer mouse to insert vertical column dividers between all the variables included in the *cuban.dat* that appears on your screen. You insert vertical column dividers by right clicking your mouse between the two columns that you want to separate.

You should begin this process by first inserting a column divider between column 4 and column 5 in the data file. Next proceed to insert a column divider between every consecutive whole number until you reach column 33. Here you should refer to your “Cuban Refugee Codebook” and specifically Question #25. Notice that in the left-hand column of the page that the data pertaining to this question is stored in columns 33 and 34 of the *cuban.dat* data file. As a consequence, you must put vertical dividers on the inside of column 32 and on the outside of column 34 (in other words, this variable requires two columns to store the data for this one variable). And, the next question #26 asks for the respondent’s year of birth. As a result this variable requires 4 columns to store the data for this one question.

So, the vertical dividers should go before column 35 and after column 38.

[Note: you should check this visually by looking inside these column dividers and making sure that all the data entries begin with the two digits “19”]. And since the next question in this survey asks the respondent to give the interviewer the highest grade level completed, you must put vertical dividers before column 39 and after column 40 as it is indicated on the survey form. To finish this job you will need to continue to click in column dividers between all the remaining columns after 40 since all the remaining survey items have only one column responses.

When you have finished inserting all the column dividers appropriately, click “Next”. Now in response to Step 5 of 6, you are required to insert the appropriate “*Variable Name*” corresponding to that particular data column.

For each of the following variable columns identified below, left double click on that V column number and type in variable name given to you below:

V3	newspap
V4	radio
V5	tv
V6	friends
V7	laxtrib
V8	genimp
V12	perprej1
V13	perprej2
V14	perprej3
V15	cubprej1
V16	cubprej2
V17	cubprej3
V18	cubprej4
V19	cubprej5
V28	criminc
V29	jurybias
V30	yearsres
V31	yearborn
V32	yearsed
V33	married
V35	youngkid
V36	religion
V37	faminc
V38	gender

When you have finished labeling all these variables listed above, click “Next”.

Now in response to Step 6 of 6, click “yes” in response to the question, “Would you like to save this file format?”

Next, click “Save As” and then type in the file name: *cuban.sav*

Then, click “Save”.

And to finish, click “Finish” **Do not forget to do this !!!!!**

---

Now you are ready to add some additional information to your newly created system file called *cuban.sav*. Specifically, what you are going to do next is to add some *Variable Labels* and some *Value Labels*. To do this you must begin by clicking on the “*Variable View*” tab which appears at the bottom left-hand corner of the data grid screen. Next, you must click on the “*Labels*” box corresponding to the variable you wish to work with. Remember from the last lab that this is the action that produces the three dots ... When you click on these three dots the

value labels dialog box opens up and you can then type in the appropriate labels. You must refer to the Cuban Refugee survey codebook to do this however.

To determine what would be an appropriate *Variable Label* you must read the questionnaire item to which the variable refers. For example, the variable called “married” appearing as Question #28 in the survey asks the respondent if he or she has ever been married so you might want to give this a variable label of “Ever-married persons”. The purpose of *Variable Labels* is that they allow you to expand the meaning of the variable name which itself can never be more than 8 characters in length. Good *Variable Labels* will prevent you from making some bad mistakes when you get around to interpreting your data.

Next, you must again refer to Question #28 in the survey to determine what the answer categories are so that you may complete the “*Value Labels*” for this variable. In this instance you should specify that the data code of “1” means “No” and that the data code of “2” means “Yes”. In this instance the words “No” and “Yes” are the “*Variable Labels*” which we are attaching to the variable called “*married*”. Once you have listed all the value categories and the data number that they correspond to in the dialog box, click “*Continue*”. Next you must click “*OK*” to complete the job before moving on to the next variable and repeating the exact same sequence of steps.

Next you are asked to create both “*Variable Labels*” and “*Value Labels*” for each of the following variable names: [You will have to refer to your Cuban refugee codebook to do this.]

Variable Labels

Value Labels

---

— Last      a) *gender*  
Survey  
Item

---

See Q#31    b) *faminc*

---

— See Q#30    c) *religion*

---

See Q#29    d) *youngkid*

---

See Q#24    e) *jurybias*

---

See Q#2     f) *laxtrib*

---

See Q#28    g) *married*

---

One peculiarity of the *cuban.dat* file is that all missing values are coded with the number zero (0) and so you must remember to identify the number (0) as *system missing*. The easiest way to do this for all the variables at the same time is to do as follows:

- a) Click *transform*
- b) Click *recode*
- b) Click *Into Same Variables*
- d) Using your mouse highlight the entire list of variables in the left-hand box.
- e) Click on the right arrow to move all these variables into the center box.
- f) Click *If*
- g) Click *Continue*
- h) Click *Old & New Values*
- i) Keystroke the number "0" into the "Old Value" box.
- j) Click the "system missing" box under the "New Value" side of the screen
- k) Click *Add*
- l) Click *Continue*
- m) Click *OK*

This one operation will convert all the zero codes for this set of variables into missing data.

Once you have completed all the above, it is time to save your work.on your on-line personal file To do this you can't go wrong if you always follow the steps listed below:

First, click **File**

Second, click **Save As**

Third, click **cuban.sav**

[When the computer tells you that you already have this file, do you want to copy over it?  
Answer "Yes".]

Fourth, click **OK**

Your last operation is to run the frequency distributions for the following set of variables:

*married*  
*gender*  
*religion*  
*cubprej5*  
*faminc*  
*jurybias*  
*genimp*

To make this happen you must select the following sequence of menu options ---

- 1) Click *Analyze*
- 2) Click *Descriptive Statistics*
- 3) Click *Frequencies*
- 4) Highlight the variables listed above one at a time and move them into the dialog box to the right side of your screen
- 5) Then, click *OK*

Now compare the results of your frequencies with those of the printed copy of these same frequencies that your instructor ran before class and distributed to you. You should not find any discrepancies (or at least no major ones) between the frequency output generated on your screen and the printed copy of the output distributed to you. If you do find discrepancies you should report these to your instructor to determine what it is that may have caused them.

Remember that you are building a house and this much of what you have done is the foundation. Next lab we will finish the house by actually analyzing the data and testing various research hypotheses.

Also remember that as you exit out of the SPSS program from this point on always answer “**No**” to questions about saving your output file or any other questions about saving your work in this session.

Voila! Job Well Done!

Bring this lab and your Cuban Refugee Survey to the next lab.

Do not turn this lab in to the instructor, please. Hold onto it.

Next week we will fearlessly seek the truth that is contained within this data file!