

MAX 201

Lab #1: Introduction to the General Social Survey (GSS) and SPSS

Note: Please read through the entire assignment before you start the lab. You will get a sense of what to do, and how to proceed—this will save you time in the end.

Introduction:

This assignment introduces you to the General Social Survey data we will be using all semester, and familiarizes you with some of the features of SPSS.

- If you want some extra help with SPSS, I use the recommend the book from the syllabus, “An IBM SPSS Companion to Political Analysis” by Philip Pollock
- Or come to office hours!

Files you need to start:

1. GSS DATASET: can be found on Blackboard under the tab “GSS Data and Codebook”.
 - When in the “data view” tab (found towards the bottom of your SPSS window)...
 - Each row of data in the GSS file represents an individual respondent.
 - Each column represents one of the questions a respondent was asked in the survey.
 - Each cell contains the answer of a particular respondent to a particular question.
2. GSS CODEBOOK: also on Blackboard under the “GSS Data and Codebook” tab.
 - This is a pdf file containing the codebook for the GSS. The codebook is the key to making sense of the data. In this file, you will find the variables in numerical order, with the question wording and response categories.

Save a copy of the GSS:

1. A full (and clean) copy of the GSS will always be available on Blackboard.
 - a. If you want to have a copy in your own files, you will need to save on OneDrive. Files saved on the hard drive on cluster computers are not secure and may be deleted or changed by anyone with access to that computer. Therefore, it is very important that you save all of your work in your folder on the OneDrive or on a thumb drive.

Note: This week we are using these six variables from the GSS:

1. age, 2. educ, 3. health, 4. reqmasks, 5. partyid, 6. class

2. Save any changes you make to the dataset to a new file with a new name that describes what you are saving.
 - a. Although you can always download a fresh copy of the GSS from Blackboard, any changes that you make to the dataset (recoding, creating indices, etc.) must be saved as an SPSS file with a different name to your own directory on OneDrive or a thumb drive.

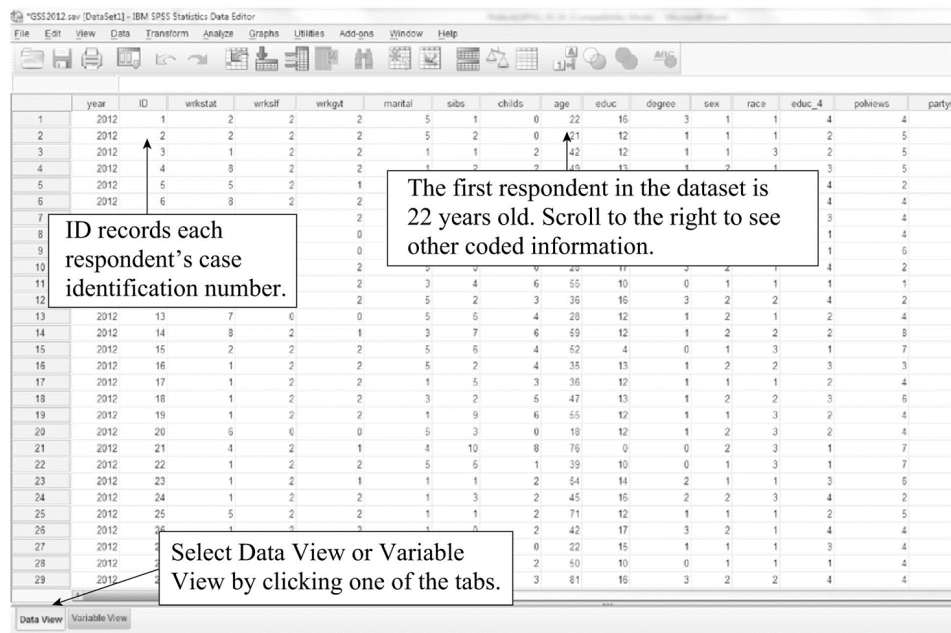
Week 1: Analysis

Part I: Data Editor & Describing Respondents

Data structure: We are going to look at how the data are structured in the GSS dataset and you are going to describe several GSS respondents using these data.

1. Open the 2022 GSS dataset to see how the data look.
 - a. Notice that there are two tabs at the bottom (“variable view” and “data view”). Click on the “Data View” tab.
 - b. “Data View”
 - i. How the data are organized
 1. Rows: each row of data in the GSS file represents an individual respondent who was a participant in the 2022 survey.
 2. Columns: each column represents one of the survey questions a respondent was asked and the cells give their numeric answers. The names across the top row are the “code names” designated for each survey question on the GSS. Some datasets will contain an “ID” variable, which uniquely identifies each person in our extracted file. In our class data, the row number can be used to identify each individual respondent and act as their ID number.
 3. Scrolling down you can see that there are over 3500 respondents to the GSS survey and 595 variables.

Figure 1-1 SPSS Data Editor: Data View



- c. “Variable View”:
 - i. To see more information about the variables/questions, click on the “Variable View” tab on the bottom of the SPSS window.
 1. This view shows you information about each variable. For example, in “variable view” you can obtain information about the name of

each variable, get a short description under “Label” of the variable question, a list of the numeric codes that correspond with the survey question’s response options under “Values”, and shows you the “Missing” codes for respondents who did not answer the question.

Describing individual respondents:

1. We are going to explore six variables and describe 4 respondents’ characteristics based on these variables (respondent numbers **636, 1794, 1865, and 2152**).
 - a. The variables you will use are: 1. age, 2. education level, 3. health status, 4. participant’s opinion on whether the government should be able to require masks, and 5. political party identification 6. subjective class identification
 - b. The variable code names are: **1. age, 2. educ, 3. health, 4. reqmasks, 5. partyid, 6. class**
2. To help make sense of the numeric values assigned to each answer of a survey question, you must use the GSS Codebook. Here’s how to read the codebook:

Variable: HAPMAR Type: Numeric				
Label: (IF CURRENTLY MARRIED, ASK HAPMAR) Taking things all together, how would you describe your marriage? Would you say that your marriage is very happy, pretty happy, or not too happy?				
Notes: See HAPPY				
LABEL	VALUE	COUNT	PCT	PCT Excl. Reserve Codes
VERY HAPPY	1	1211	30.0%	61.0%
PRETTY HAPPY	2	701	17.4%	35.3%
NOT TOO HAPPY	3	74	1.8%	3.7%
SUBTOTALS:		1986	49.3%	100.0%
RESERVED CODES:				
NO ANSWER	N	3	0.1%	n/a
NOT APPLICABLE	I	2031	50.4%	n/a
SKIPPED ON WEB	S	12	0.3%	n/a
TOTALS:		4032	100.0%	100.0%

KEY:

At the top you can see the variable code name “HAPMAR” and under “Label” the exact wording of the survey question.

Value is the numeric code assigned to each response option

LABEL is the qualitative meaning of each numeric code (i.e. the response option wording)

Count is how many respondents selected each response option

3. Click on “Data View” so that you can find these respondents answers to the survey questions.
 - a. Remember, we can identify respondents by their row number. That will uniquely identify each individual respondent.
 - i. Note that respondent row/ID number is the first column in data view.
 - b. Scroll down to ID numbers **636, 1794, 1865, and 2152**. Look up the numeric responses for each respondent’s answers to the six variables listed on the previous page of the assignment.
 - c. Sometimes respondents won’t answer all the questions in a survey or won’t be asked a particular question. If the answer is missing, it may be represented by a period “.” in the “Data View” window.

- d. To make it easier, we will copy the information about these respondents into a table in Microsoft Word (or Google Docs). The instructions to build the table are found in the next step of the assignment.
4. Open Microsoft Word.
- a. Put the below identifying information at the top of your document:
 - i. Name
 - ii. Course and your lab section (Tuesday, Wednesday or Thursday and the starting time for the lab).
 - iii. Assignment #1.
 - b. Create a table where we can store the information about these four respondents. (You can also copy and paste the below table into your document)
 - c. Click Insert → Table → Insert Table (7 columns x 5 rows)
 1. Label the Columns “ID” “Age” “Education”, “Health status”, “Gov’t should be able to require masks”, “Party identification”
 2. Label your table. Table 1:[Descriptive title] and give your table a descriptive title in the brackets.
 3. Label the rows with the numbers of the respondents.
 4. Your table should look like the one below.

Table 1: [Descriptive title]

ID	Age	Education	Health status	Gov’t should be able to require masks	Party identification	Subjective Class Identification
636						
1794						
1865						
2152						

- d. Save your Word document as “**Assignment1_YourLastName**” on your One Drive, your thumb drive or GoogleDocs.
 - iv. After naming and saving the file, you can update it by simply using **Save**. However, if you wish to save a copy of the file under a new name, use the **Save As...** [enter file name] and repeat the process.

Analysis – Be complete and specific in your answers! Keep in mind it may take up to one paragraph to answer each part of the question.

PART I Analysis write up:

1. **Describe the respondents:** In your Word document, write in paragraph form (i.e. not just bullet points) about **each** of these four respondents in Table 1 using information from the variables in the table and the codebook. Please write these up in paragraph form that shows that you understand what the variables mean. At this point, you have information about four

respondents in your dataset in Table 1, and to make sense of it, you are going to need to reference the codebook.

For example: respondent 159 has the value of “66” for the variable **age**, and the value “14” for the variable **educ**. This person is 66 years old and has completed 14 years of school (i.e., this person has two years of college).

- Note that not all the variables are that easy to interpret. For instance, what is the variable “HAPPY” measuring and what does a value of “3” mean? To find out, you need to match the variable names to the survey questions and the variable values to their response option. To look up the variable names for a variable in the codebook, use the search function in Adobe and search for the abbreviation of the variable name. For instance, looking in the codebook on page 203, you can find that a value of “1” for the variable **sex** represents a male and the value of “2” for that variable represents a female. (You may find it helpful to highlight the variables that we’re using in the pdf version of your codebook and write down the pages where you can find these variables.)

PART 2 Analyses: Thinking about how the GSS captures data

Please put on your social-scientist-in-training hat and answer the questions below about the concepts behind the survey questions (Note: please do not answer this question about yourself).

1. The following question about racial identity (variable: RACE) has been asked on the GSS since 1972.

What race do you consider yourself?

1. White 2. Black 3. Other

- a. Survey questions may measure concepts such as beliefs, attitudes (i.e. what you think about a person or object), identifications (i.e. how you see yourself), or knowledge. What concept is this question intending to measure?
- b. Are there any problems that you can identify with the question wording or the response options (i.e. Did you have any problems answering? Are there response options missing? Are all the response options different enough to be able to just choose one?) Be specific.
- c. How else could you measure the concept this question intends to measure? Write an alternative survey question including the response options. Explain why this is a better alternative and what it captures that the original question does not.
- a. In 2021, the GSS moved to a different way to ask about race. Instead of one question, the GSS allow people to pick multiple categories for their race along with ethnicity categories. What is the benefit of this new set of questions? What is one way that it complicates analysis?

2. Until 2020, the GSS was conducted as an in-person interview. Survey takers come to a respondent's home and ask them questions about a variety of political and social topics, including sensitive ones such as health behaviors. Below is an example of a health behavior question. (Note: Please do not answer this question.).

Have you ever, even once, taken drugs by injection with a needle (like heroin, cocaine, amphetamines, or steroids)...Do not include anything you took under a doctor's orders.

Yes, No, Don't answer, No answer

- a. How might an in-person interview affect the answers that a respondent would give to this question? Are there other types of ways to ask these types of questions that may elicit different answers? Explain.
- b. Respondents have the option of not giving an answer or saying "don't know". What could an answer of "don't know" mean in this context of asking about drug use.
- c. In 2020, because of the COVID 19 pandemic, the GSS was conducted either as a web survey or a phone survey (with no in person contact). What could the effect be of moving away from in person interviews to a web or phone survey?

WHAT TO HAND IN:

Word document with name, lab session, Assignment #1 on top.

1. Table 1 (information about respondents)
2. Analysis with answers to Part 1, Question 1 in paragraph form.
3. Analysis with answers to Part 2, Questions 1 and 2 (answering all the parts) in paragraph form.

PLEASE NOTE: We will accept late assignments up until 11:59PM on Thursday each week for a 5% penalty per day late. No assignments will be accepted after that time unless alternative arrangements have been made in advance.