

I. Getting to Know the General Social Survey (GSS)

1. Go to the [National Opinion Research Center at the University of Chicago](http://www.norc.uchicago.edu/) at <http://www.norc.uchicago.edu/> and read a little about NORC and some of its reports.
2. Go to the [GSS 1998 Codebook](http://webapp.icpsr.umich.edu/GSS/) at <http://webapp.icpsr.umich.edu/GSS/>.
 - a. Click on and read [About GSSDirs](#)
 - b. Click on and read **Overview of GSSDirs**
 - c. Close window.
 - d. Click on and read **About the National Opinion Research Center**.
3. Go to back to the [GSS 1998 Codebook](http://webapp.icpsr.umich.edu/GSS/) at <http://webapp.icpsr.umich.edu/GSS/>.
4. Go to [Codebook Indexes by Subject](http://webapp.icpsr.umich.edu/GSS/rnd1998/merged/indx-sub/indexa.htm) at <http://webapp.icpsr.umich.edu/GSS/rnd1998/merged/indx-sub/indexa.htm>. Note that you are taken to a list of subjects beginning with the letter “A” – on the first page, you see **Abortion to Automobiles**.
5. Note that “**Abortion**” is a hot link. Click on it. This will take you to the variables in the 1998 dataset related to abortion, specifically opinions regarding abortion.
6. Click on a variable in that list, such as **ABRAPE**, which is an opinion question about whether or not the subject respondent thinks a woman should be able to obtain a legal abortion if she became pregnant as a result of rape. This is what you will see:

RESPONSE	PUNCH	YEAR									COL: 784
		1972-82	1982B	1983-87	1987B	1988-91	1993	1994	1996	1998	ALL
Yes	1	10922	234	4703	246	3,131	850	1616	1550	1439	24,691
No	2	2123	97	1080	84	602	175	318	289	357	5,125
Don't know	8	533	20	236	20	167	46	60	77	83	1,242
No answer	9	48	3	53	3	20	4	2	7	3	143
Not applicable	BK	0	0	1470	0	1,987	531	996	981	950	6,915

The primary usefulness of this table is that it provides information regarding the *level of measurement* for the variable, ABRAPE. Here, it is categorical: specifically, dichotomous (yes/no), with three additional choices of “missing variable” (don’t know, no answer, not applicable). Note that responses are shown for other years during which this same question was asked.

Be sure to review Utts, Chapter 3, regarding levels of measurement, if you need assistance.

For another example, click on **B** on the **tan alphabetical guide** and go to the subject “**Birth Order.**”

Subject Index: **B**

					Help																				
Previous	Pick Page				Next																				
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Click on it. This will take you to the primary variable in the 1998 dataset related to birth order, SIBORDER. **Click on SIBORDER.** This is a response to a question about how many of the respondent subject’s brothers and sisters were born earlier than the respondent. See the resulting table.

RESPONSE	PUNCH	YEAR										COL: 4054- 4055
		1972-82	1982B	1983-87	1987B	1988-91	1993	1994	1996	1998	ALL	
0	0	0	0	0	0	231	0	0	0	0	231	
1	1	0	0	0	0	244	0	0	0	0	244	
2	2	0	0	0	0	148	0	0	0	0	148	
3	3	0	0	0	0	86	0	0	0	0	86	
4	4	0	0	0	0	58	0	0	0	0	58	
5	5	0	0	0	0	36	0	0	0	0	36	
6	6	0	0	0	0	18	0	0	0	0	18	
7	7	0	0	0	0	18	0	0	0	0	18	
8	8	0	0	0	0	11	0	0	0	0	11	
9	9	0	0	0	0	6	0	0	0	0	6	
10	10	0	0	0	0	5	0	0	0	0	5	
11	11	0	0	0	0	2	0	0	0	0	2	
16	16	0	0	0	0	1	0	0	0	0	1	
Only child	95	0	0	0	0	45	0	0	0	0	45	
Don't know	98	0	0	0	0	1	0	0	0	0	1	
No answer	99	0	0	0	0	18	0	0	0	0	18	
Not applicable	BK	13626	354	7542	353	4,979	1606	2992	2904	2832	37,188	

Here we can see that the data is ratio-level data (has an absolute zero and can be counted). Respondents are giving a ratio-level number in response to the question.

The level of measurement is important information for you because it determines what kind of analyses you will be able to run – that is, what kinds of questions you will be able to ask. For example, we can’t ask what the “average” response about attitudes toward whether women should be permitted to have a legal abortion if they become pregnant as a result of rape – because “yes” and “no” are categorical values and can’t be “averaged.” On the other hand, we can ask about the average number of younger siblings respondents have. *Notice how the last four response categories are given special numbers (95, 98, 99, BK). That way they will not be included in any analysis that attempts to calculate the average (we know that “98” means the respondent is an only child, not that she or he has 98 younger siblings).*

Before you run analyses – that is, ask questions that can be answered through data analysis – you need to know the nature of your data, particularly the level of measurement.

The first type of analyses we will examine are appropriate for categorical (nominal), ordinal, or short-range ratio-level data, such as number of vehicles per household.

II. Frequencies and Bar Charts

Frequencies tables and bar charts are useful for categorical, ordinal, and short-range ratio data. The first exercise will take you step-by-step through running both a frequencies table and a bar chart for a variable in the 1998 GSS dataset called **WRKWAYUP**. The specific question asked of respondents was this:

Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with the following statement?

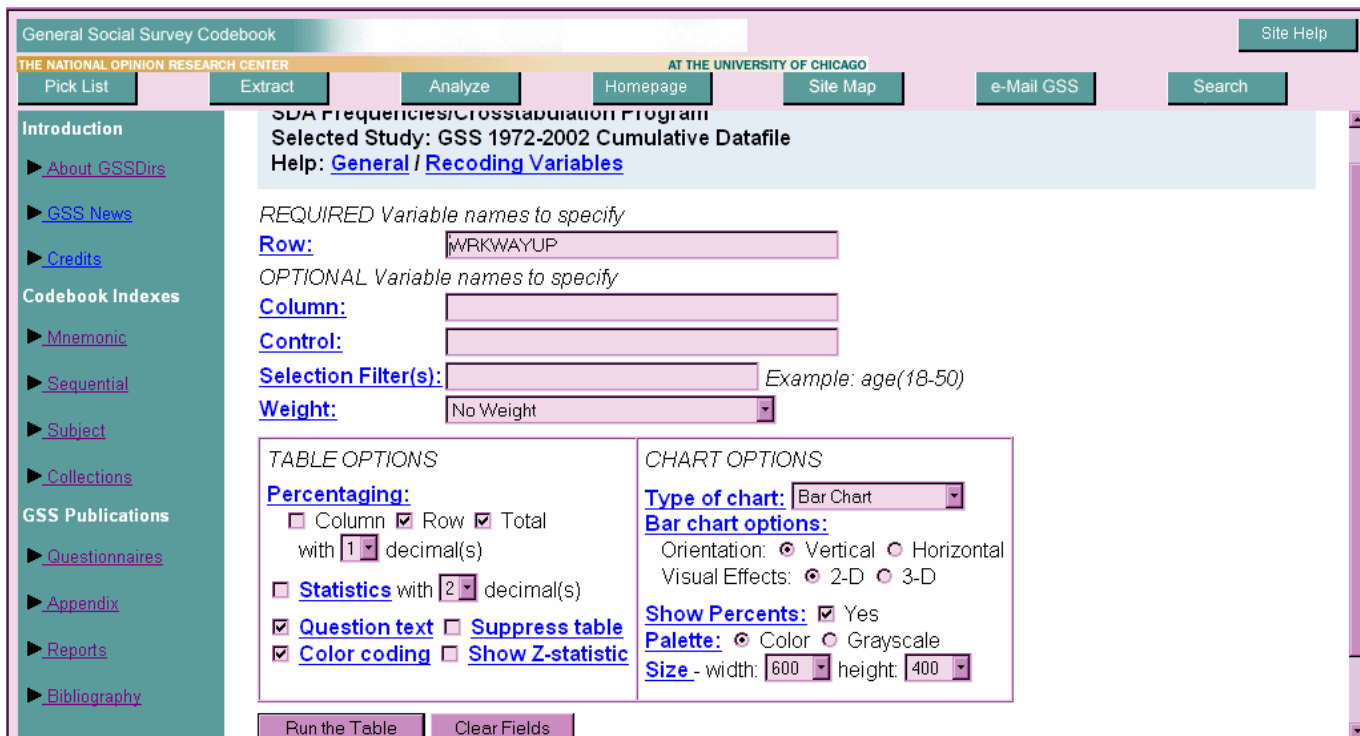
Irish, Italians, Jewish and many other minorities overcame prejudice and worked their way up. Blacks should do the same without special favors.

A scale that ranges from strongly agree to disagree strongly is called a *Likert Scale*. This kind of scale is sometimes considered to be ordinal, but really, it is simply categorical (nominal), because the responses go in two opposite directions from neutral.

Make sure you see the menu bar shown below near the top of your browser. If at any time you don't see it, go to <http://webapp.icpsr.umich.edu/GSS/> which takes you to the GSS home page.



1. Click on **Analyze**.
2. Select **Frequencies or Cross-Tabulations with Charts**
3. Click on the **Start** button.
4. In the **new window** that comes up, type in as follows and as shown in the illustration below:
 - a. In the **Row** field, type in the variable name (case doesn't matter): **WRKWAYUP**
 - b. Under **Table Options**, select **Row** and **Total**, with **1 decimal**; **Question text**; and **Color coding**.
 - c. Under **Chart Options**, select **Bar Chart Orientation, Vertical**; **Visual Effects, 2-D**; **Show Percents, Yes**; **Palette, Color**.
 - d. Click **Run the Table**.



You will get the results shown below. I have highlighted information of interest and used red text to explain the results.

SDA 1.4: Tables

GSS 1972-2002 Cumulative Datafile

Aug 07, 2005 (Sun 03:42 PM PDT)

Variables					
Role	Name	Label	Range	MD	Dataset
Row	wrkwayup	BLACKS OVERCOME PREJUDICE WITHOUT FAVORS	1-5	0,8,9	1

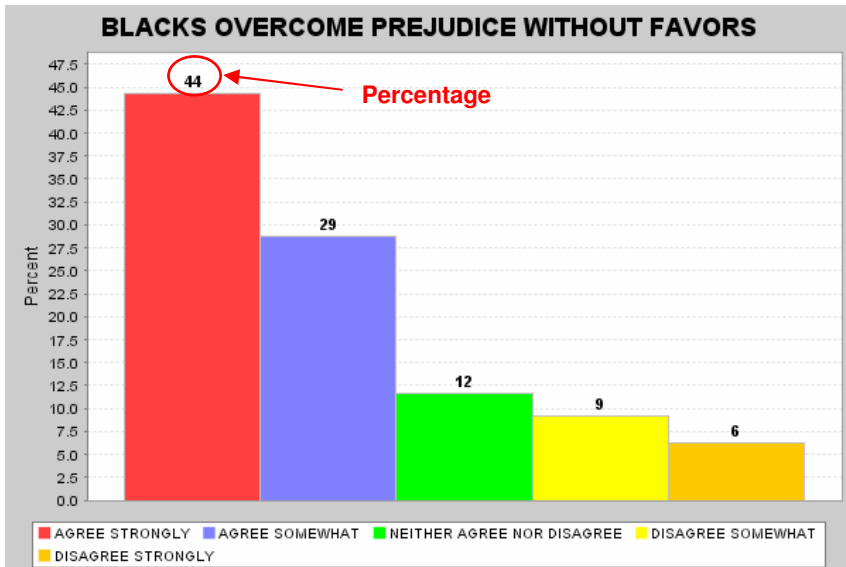
Frequency Distribution

Cells contain: **N = number of respondents**

		Distribution
wrkwayup	1: AGREE STRONGLY	44.3 3,499
	2: AGREE SOMEWHAT	28.7 2,270
	3: NEITHER AGREE NOR DISAGREE	11.6 915
	4: DISAGREE SOMEWHAT	9.2 724
	5: DISAGREE STRONGLY	6.2 489
	COL TOTAL	100.0 7,897

44.3 + 28.7 = 73 percent, or a majority

total sample size = 7,897, which is more than adequate



instructions to survey administrator

Text for 'wrkwayup'

B. Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with the following statement (HAND CARD TO RESPONDENT): Irish, Italians, Jewish and many other minorities overcame prejudice and worked their way up. Blacks should do the same without special favors.

What are some ways to interpret these data? Well, if we combine the “Agree Strongly” and “Agree Somewhat” categories, to yield a general “Agree” category, we can see that 44% + 29% (or about 73%) of Americans surveyed in 1998 agree with the statement that blacks should overcome prejudice and work their way up without special favors. In terms of percentages, 73 percent is a fairly strong “majority.” Still, despite the care with which the GSS Survey is conducted, these data don’t tell us anything within, say, a ± 5 -percent margin of error. At this point, we are dealing with descriptive statistics that are purely suggestive.

Exercise 1:

Run FREQUENCIES and BAR CHARTS, following the procedure above, for three or four additional variables from the 1998b dataset, having to do with any subject matter that interests you. Simply skim through the variables available for 1998 (a or b) by clicking on [Collections by Year](http://webapp.icpsr.umich.edu/GSS/rnd1998/merged/indx-mod/years.htm) at <http://webapp.icpsr.umich.edu/GSS/rnd1998/merged/indx-mod/years.htm>. Remember that if you can’t find the ANALYZE button, just go back to the [GSS homepage](http://webapp.icpsr.umich.edu/GSS/), at <http://webapp.icpsr.umich.edu/GSS/>. Look for variables that seem related to your area of interest and then examine them to identify the level of measurement. Once you find three to four variables that are **categorical through small-range ratio**, run FREQUENCIES and BAR CHARTS for each.

Copy and paste the results into a Word document.

Resize the graph as necessary. Select and edit the font size in the table. Move the table margins. Try to make the results for each procedure fit on one page.

Note that the bar chart will not copy and paste along with the tables. You will need to go back and **select the bar chart, right click**, and choose **Copy**. Then, **position your cursor in the table in Word in the row immediately below the frequencies**. Select **Edit/Paste Special** and click on **Device Independent Bitmap**. That will insert the graphical image.

Directly beneath your results, type up a paragraph interpreting the results – one paragraph per variable. Please note whether there were any surprises. And remember that at this point, your analyses are only suggestive. You aren’t “proving” anything – or even getting close to doing so. Also remember to try to keep the procedure output and your interpretive paragraph on one page (one page per variable).