Short Assignment Two
Political Science 210, Empirical Political Theory
Paul Gronke

Lab Sessions:

Wednesday, September 7th, 5:00-5:50, 6:00-6:50
Thursday, September 8th, 5:00-5:50, 6:00-6:50

Assignment Handed Out: September 7-8, 2005
Assignment Due Back In: September 23, 2005

In this short assignment, you must report on the results of a short quantitative analysis of survey data.

The paper must contain, on one page, a table or graphic displaying the different scores (values) for one variable among one or more subgroups (the second variable).

You must also pen a short paragraph saying what the table / graphic shows.

All of this needs to be in a word processed document—i.e. you must learn to cut and paste your results into the word processing program and not physically cut the output and glue it onto a piece of paper.

Specific research question (example—you CANNOT use this for your assignment):

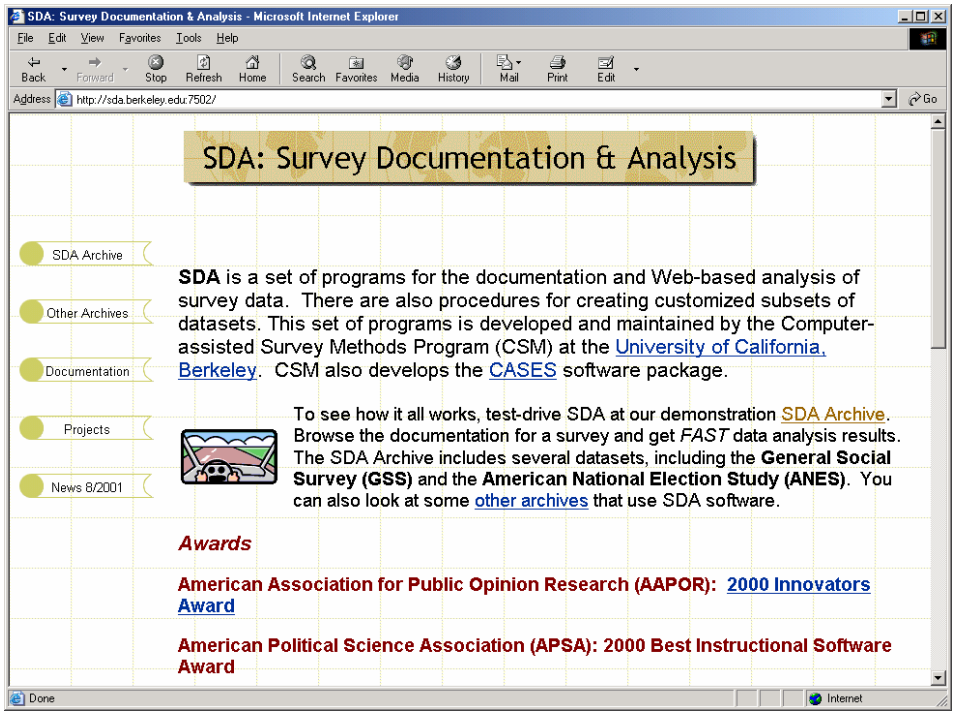
> **Did more African Americans vote for George Bush in 2000 than voted for Albert Gore?**

I. Getting Started: http://sda.berkeley.edu

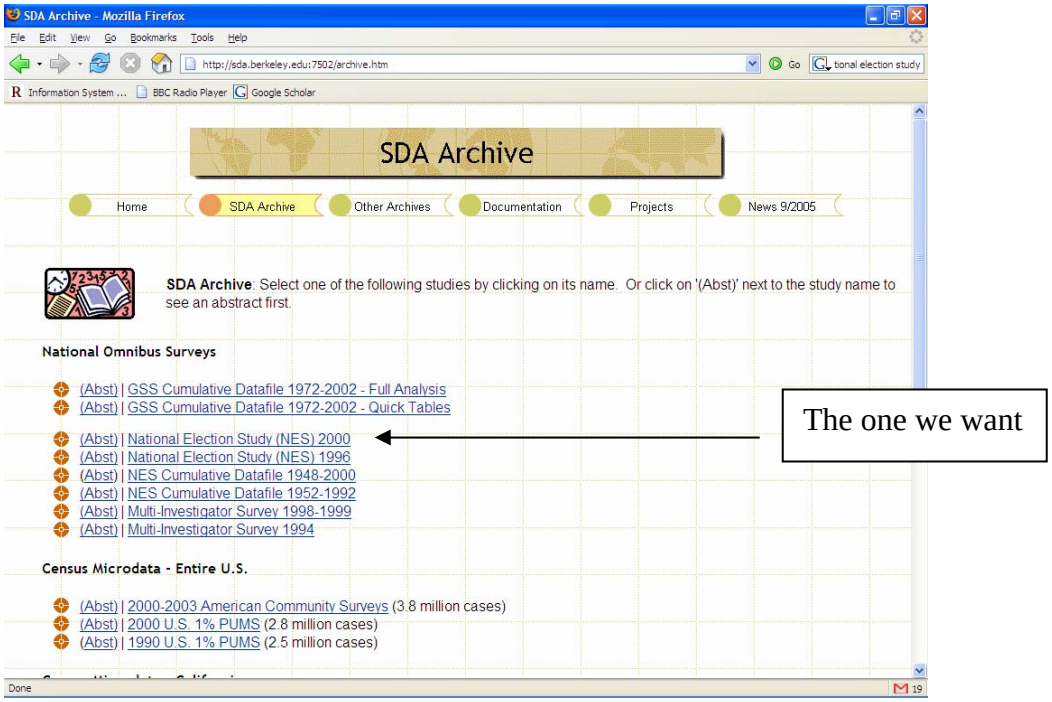
The web site you will be using is at Berkeley that has a variety of datasets, an index of questions and variables, and tools for conducting statistical analyses. You may also be interested in the Interuniversity Consortium for Political and Social Research (www.icpsr.umich.edu) which also has many datasets that you can analyze online.

1. Run an internet browser (Internet Explorer or Mozilla are fine)

2. **Getting to the website.** Type: http://sda.berkeley.edu
   In the "website" window at the top of the window and then hit return.



Note the items along left side. You should click on **SDA Archive**.



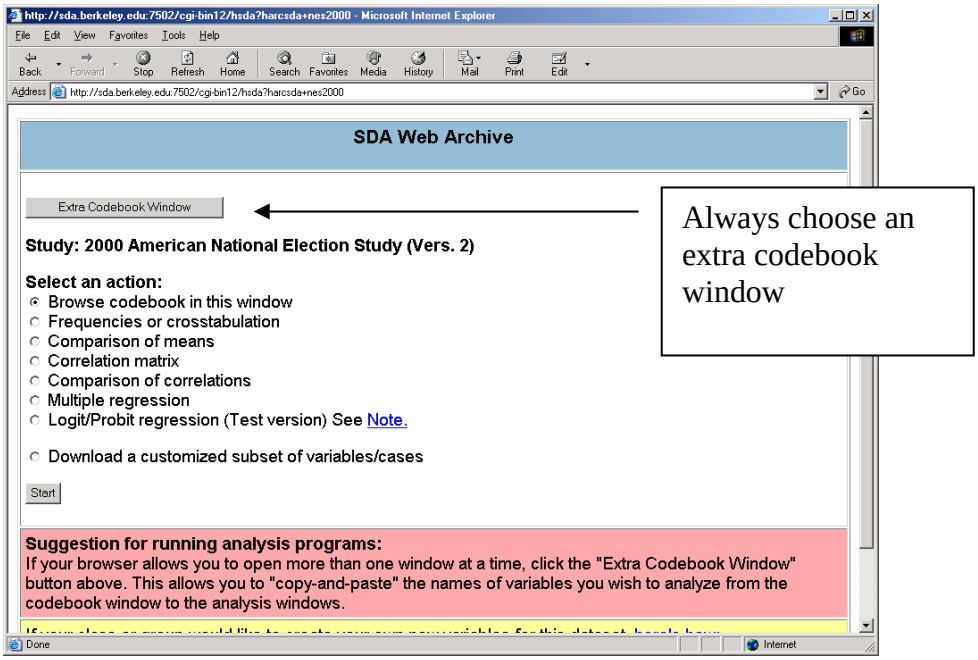
II. For our example, we will choose the "**National Elections Studies (NES) 2000**". This survey is conducted biennially during the presidential and congressional elections. It is the most common data set used by scholars in American politics

The other survey that is prominently listed is the General Social Survey (GSS). This is a survey with more sociological and public opinion questions.

You can browse the codebooks for both of the surveys to see what kinds of subjects they cover. They are **very** large, so you should find variables that strike your fancy.

III. Finding Variables

1. When you choose the survey, a page will pop up asking you what sort of analysis you wish to conduct

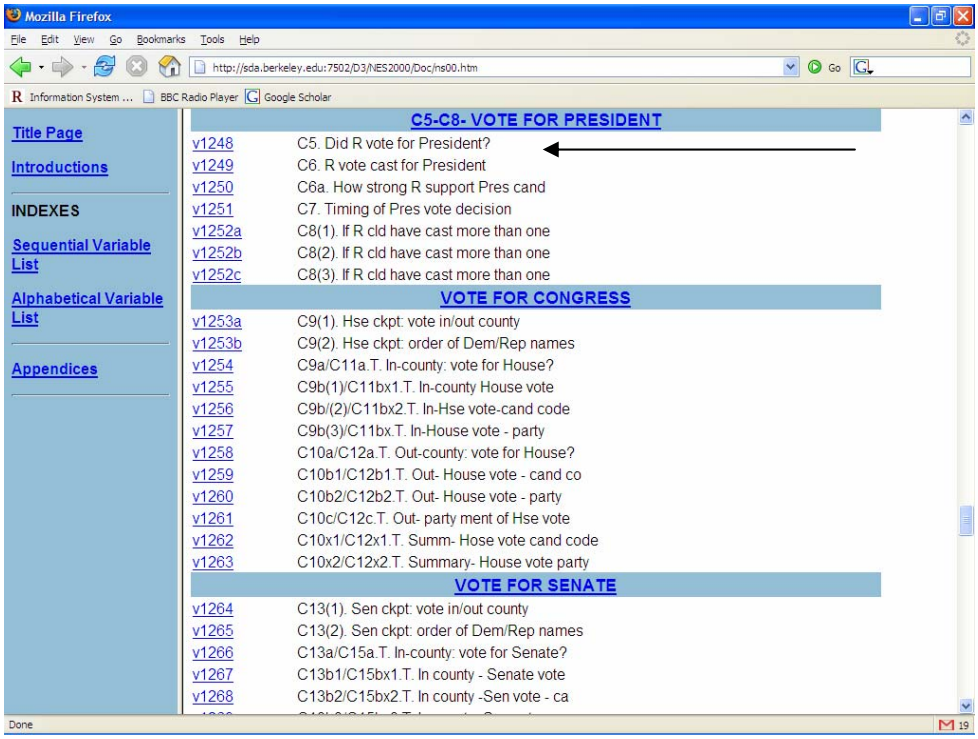


:

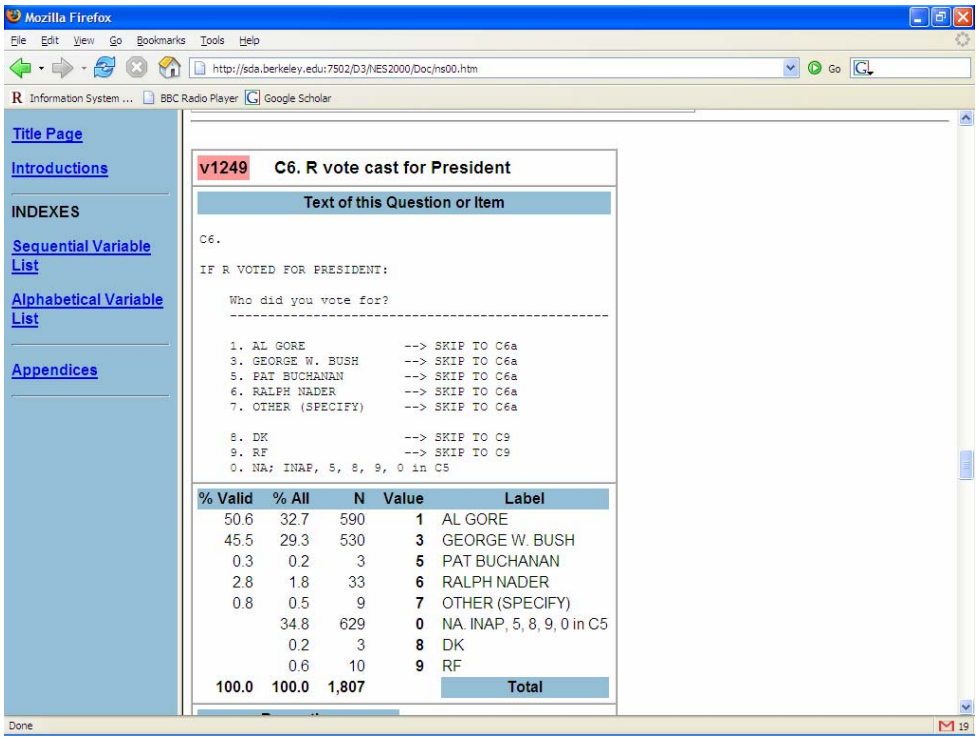
2. You need to click "**Extra Codebook Window**" so you can examine the contents of the survey.

   Here you will find an index of subjects included on the survey, by letter of the alphabet, by group, and by category.. Take a look at the items to get a sense of what is available. You will be returning here again.

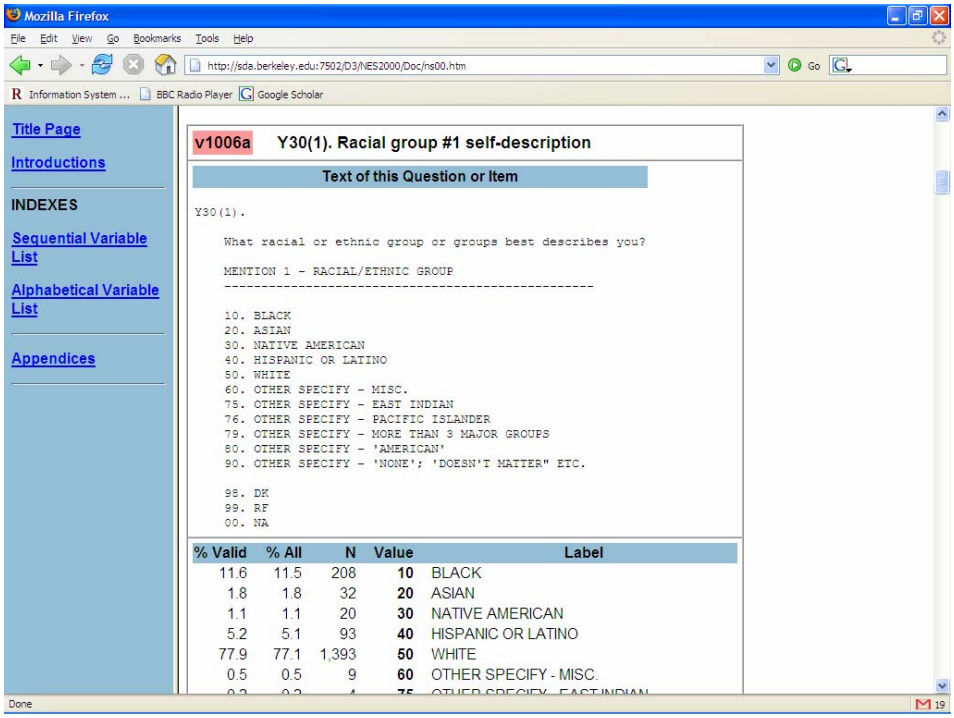
   I clicked on "**sequential variable list**" and scrolled down until I saw "C5-C8: Vote for President:" These are the variables that I want.



I clicked on this set and discovered that George Bush lost the election!



3. Now I need to go find race. I will scroll down the previous list until I find it. Here it is, but see if you can find it:

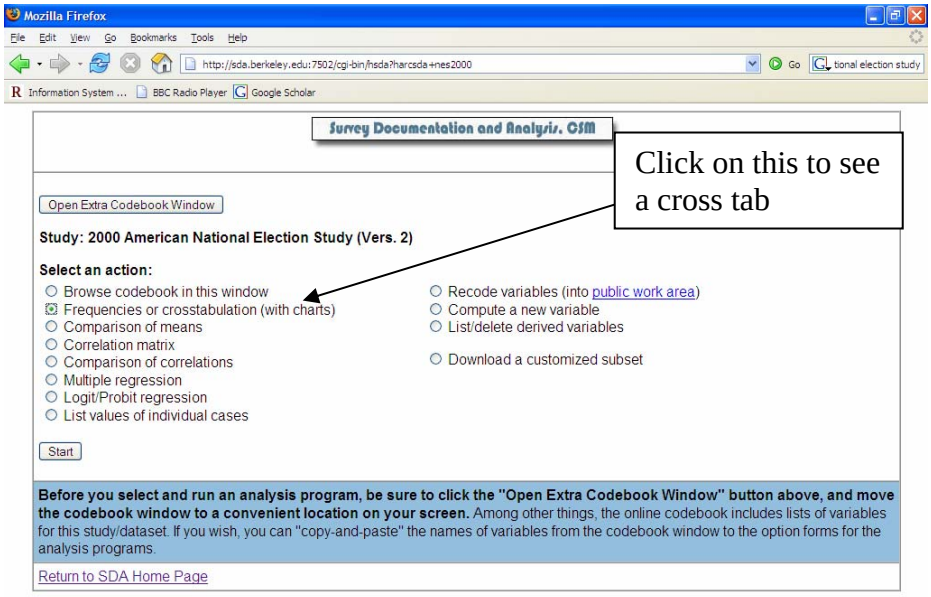


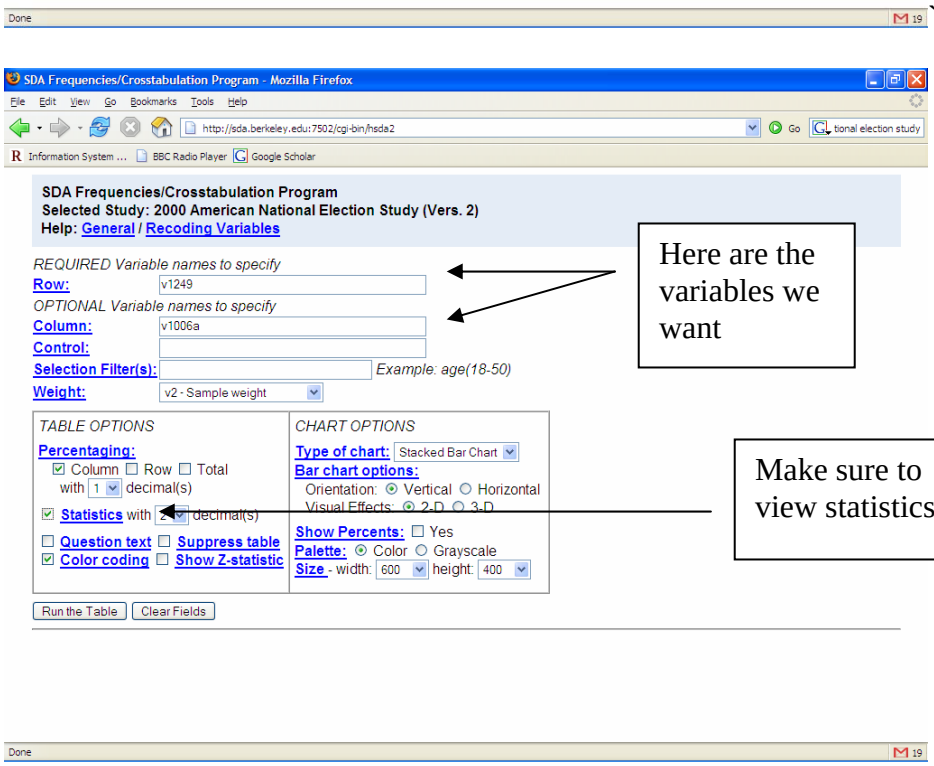
IV. Getting A Crosstab

Did African Americans and Whites differ in the level of support they gave to Bush and Gore?

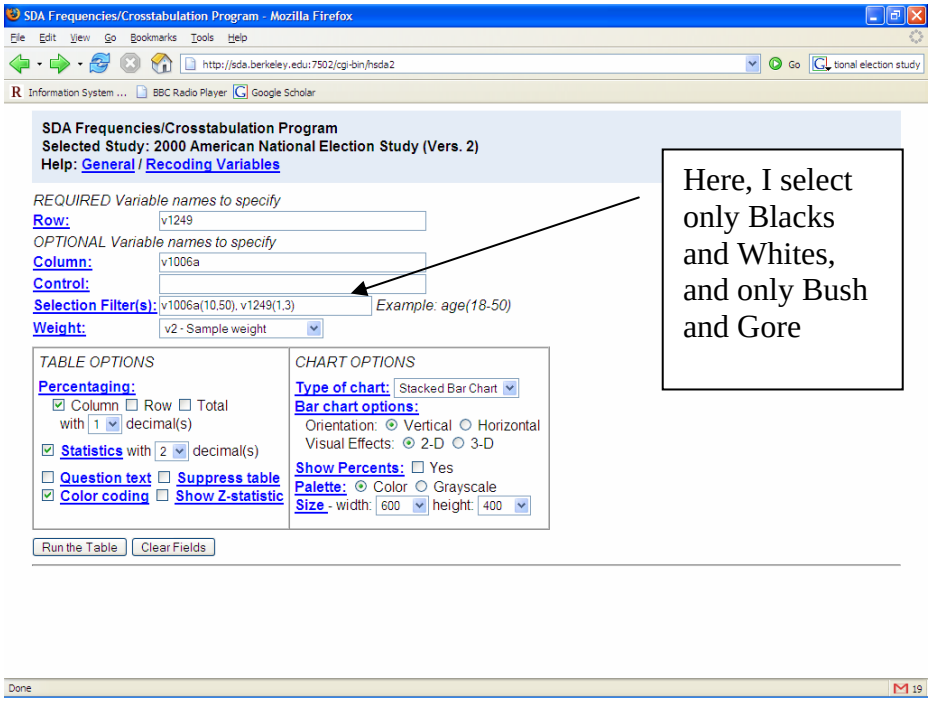
What this means in practical terms is: did the percentage of African Americans who said they voted for Bush differ from the percentage of Whites who said they voted for Bush?

We get to this by a "crosstab" which is just two variables "crossed" against one another. Go back to the first screen and choose "**Frequencies or crosstabs"**

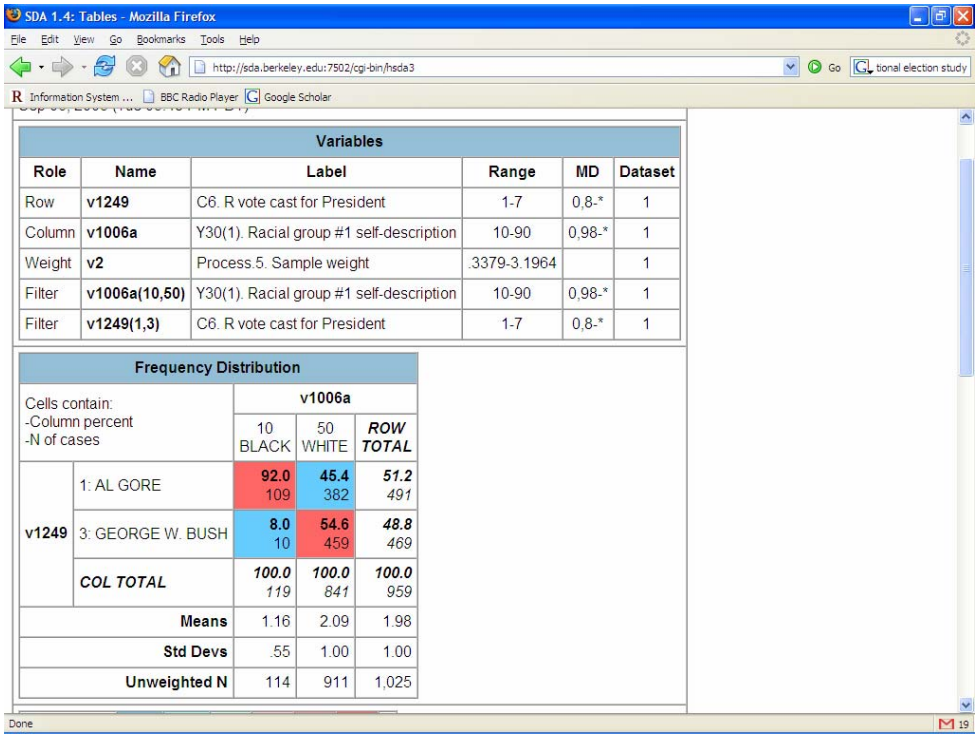




… and the output is incredibly ugly. So I know that I need to clean this up, I figure out how to **select only the cases that I am interested in** (by reading the online manual), and I run this instead:



And this looks good. 92% of Blacks voted for Al Gore vs. only 8% voting for George Bush, and this difference is statistically significant. 45% of Whites voted for Gore vs. 55% of Blacks. Nader is not considered in the analysis.I'll show you how to cut/paste this into Word in Class.



Variables

Role	Name	Label	Range	MD	Dataset
Row	**v1249**	C6. R vote cast for President	1-7	0,8-*	1
Column	**v1006a**	Y30(1). Racial group #1 self-description	10-90	0,98-*	1
Weight	**v2**	Process.5. Sample weight	.3379-3.1964		1
Filter	**v1006a(10,50)**	Y30(1). Racial group #1 self-description	10-90	0,98-*	1
Filter	**v1249(1,3)**	C6. R vote cast for President	1-7	0,8-*	1

Frequency Distribution

		v1006a		
Cells contain: -Column percent -N of cases		10 BLACK	50 WHITE	***ROW TOTAL***
v1249	1: AL GORE	**92.0** 109	**45.4** 382	***51.2*** *491*
	3: GEORGE W. BUSH	**8.0** 10	**54.6** 459	***48.8*** *469*
	COL TOTAL	***100.0*** *119*	***100.0*** *841*	***100.0*** *959*
Means		1.16	2.09	1.98
Std Devs		.55	1.00	1.00
Unweighted N		114	911	1,025