# Data to Story Project: SPSS Cheat Sheet for Analyzing *General Social Survey* Data

This guide is intended to help you explore and analyze the variables you have selected for your group project.


Conducting Descriptive Statistics for Qualitative Variables

To compute frequencies for the gender, race, and occupational variables, [SEX, RACE, INDUS80] follow these steps:

1. Click **Analyze**, click **Descriptive Statistics**, and then click **Frequencie**s. When you do this, you will see the **Frequencies** dialog box as shown below.
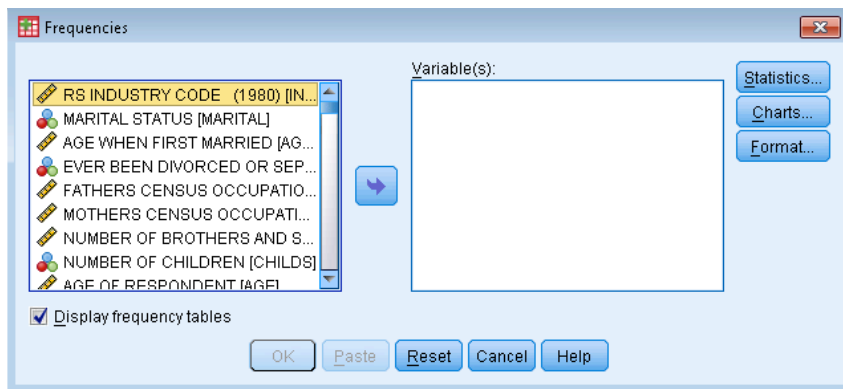
2. Holding down the Ctrl key, click **INDUS80, SEX**, and **RACE** to select them, then click ►to place them in the Variable(s) box.

3. Click **Statistics**. You will see the Frequencies: Statistics dialog box as shown below
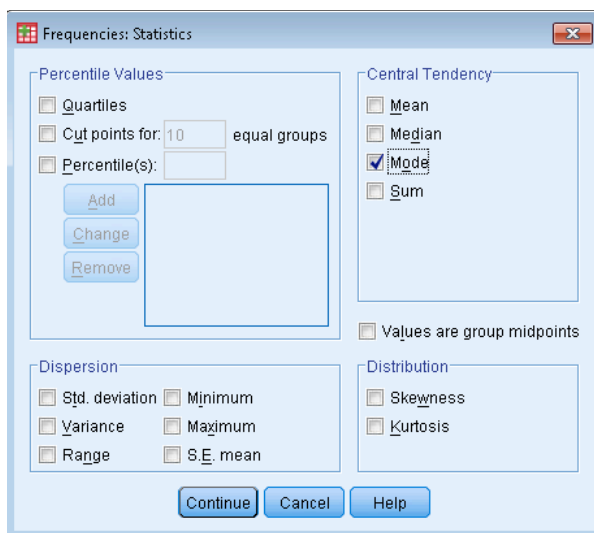
4. Under the Central Tendency area, click **Mode**.

5. Click **Continue.**

6. Click **OK**.

The output for the analysis is shown in part below. The output initially presents the mode for each of the three variables. The tables labeled **RS INDUSTRY CODE, RESPONDENTS SEX, RACE OF RESPONDENT** show the frequencies (counts) for these variables. In addition, the tables show the percentage of participants in each category. Because the variables are qualitative, the column labeled Cumulative Percent is not meaningful.

Frequency Table

RS INDUSTRY CODE (1980)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Agricultural production, crops	20	.4	.5	.5
	Agricultural production, livestock	25	.6	.6	1.1
	Agricultural services, except horticultural	14	.3	.3	1.4
	Horticultural services	30	.7	.7	2.1
	Forestry	4	.1	.1	2.2
	Fishing, hunting, and trapping, etc	5	.1	.1	2.3
	Metal mining	1	.0	.0	2.4
	Coal mining	4	.1	.1	2.5
	Crude petroleum and natural gas extraction	5	.1	.1	2.6
	Nonmetallic mining and quarrying, except fuel	3	.1	.1	2.6
	Construction	272	6.0	6.5	9.1
	Meat products	17	.4	.4	9.5
	Dairy products	5	.1	.1	9.7
	Canned and preserved fruits and vegetables	6	.1	.1	9.8
	Bakery products	6	.1	.1	9.9
	Beverage industries	9	.2	.2	10.2

RESPONDENTS SEX

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	MALE	2003	44.4	44.4	44.4
	FEMALE	2507	55.6	55.6	100.0
	Total	4510	100.0	100.0	

RACE OF RESPONDENT

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	WHITE	3284	72.8	72.8	72.8
	BLACK	634	14.1	14.1	86.9
	OTHER	592	13.1	13.1	100.0
	Total	4510	100.0	100.0	

**Using SPSS Graphs to Display the Results**

Bar charts are often used to display the results of categorical or qualitative variables. Bar charts allow for the visual representation of the frequencies in each category and can be used to display results of variables with a relatively large number of categories. An alternative to the bar chart is the pie chart. Pie charts also present the frequencies or percentages in each category but in a form that can be edited to highlight specific categories. Pie charts are more useful for displaying results of variables that have relatively few categories, if there are more than a few, pie charts become cluttered and difficult to read if variables have many categories. For example the **RS INDUSTRY CODE** has over 100 categories – way too many for a pie chart.

To create a bar chart, follow these steps:

1. Click **Analyze**, click **Descriptive Statistics**, then click **Frequencies**.
2. Click **Reset** to clear the dialog box.
3. Click **RACE**, then click ▶to place it in the Variable(s) box.
4. Click **Charts**. You will see the Frequencies: Charts dialog box as shown below
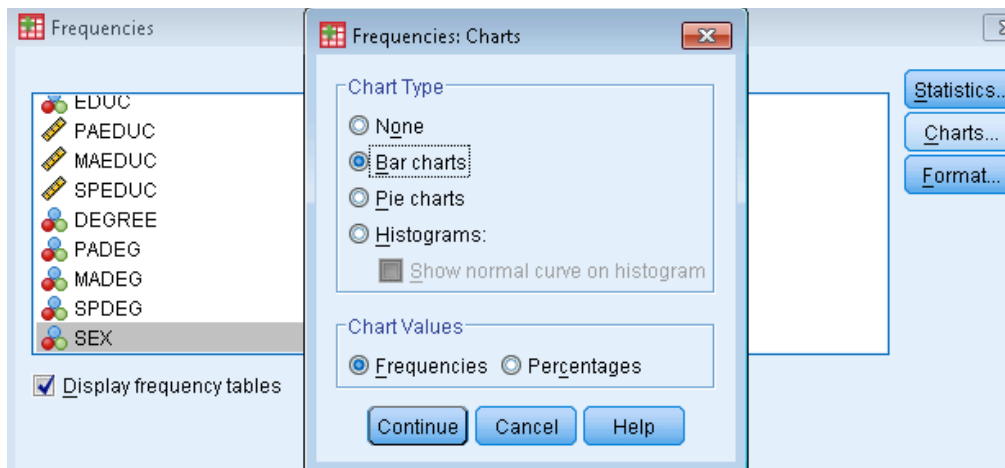
5. Click **Bar charts.**
6. Click **Continue.**
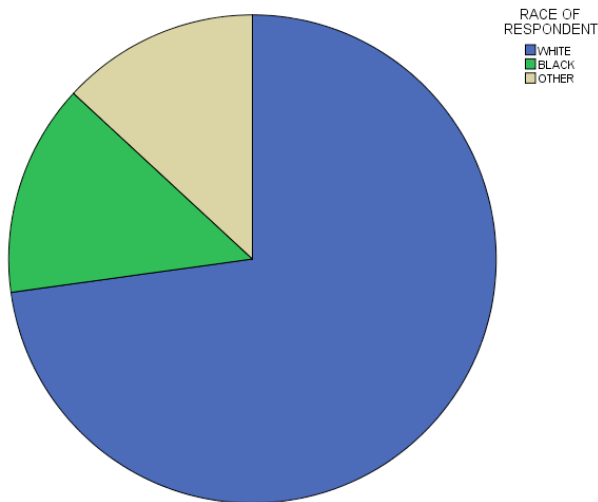7. Click **OK**
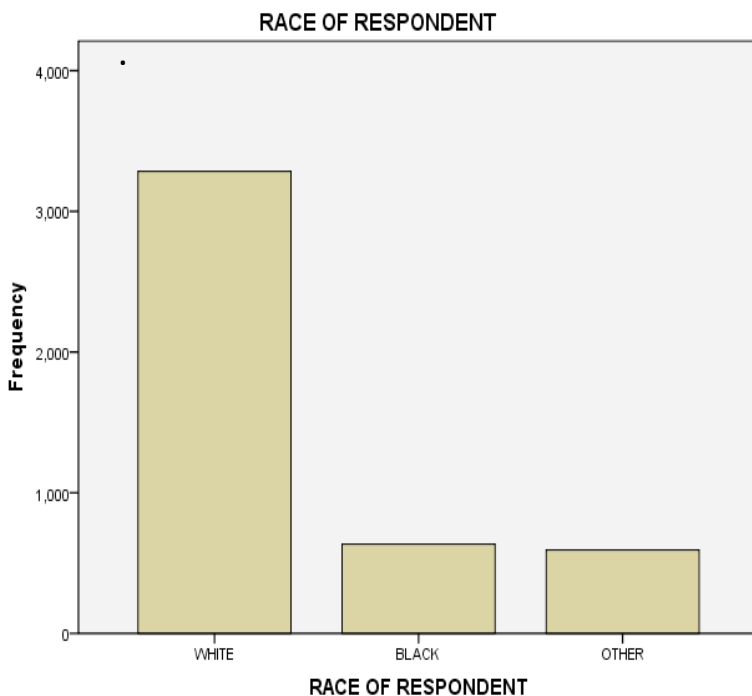
To edit the graph:
1. Double-click on the chart to select it for editing
2. Click inside the axes of the graph but not inside the bars.
3. Click on **Edit** and click on **Properties**
4. Click on the **Fill and Border** tab of the Properties box if not currently selected
5. Click on box labeled **Fill** and then click on the white rectangle in the color chart on the right
6. Click on box labeled **Border** and then click on the white rectangle with the red diagonal line in the color chart on the right.
7. Click **Apply** in the Properties dialog box.
8. Click once on the label of the –axis. The graph should now appear like the chart on the left.

To create a pie chart for the RACE variable, follow these steps:

1. Click **Graphs**, click on **Legacy Dialogs**, and click **Pie**.

2. Click **Summaries for groups of cases**.

3. Click **Define**. You will see the Define Pie: Summaries for Groups of Cases dialog box as on the right:

4. Highlight **RACE** and click ► to move it to the Define Slices by box.

5. Click **% of cases** in the Slice Represent area.
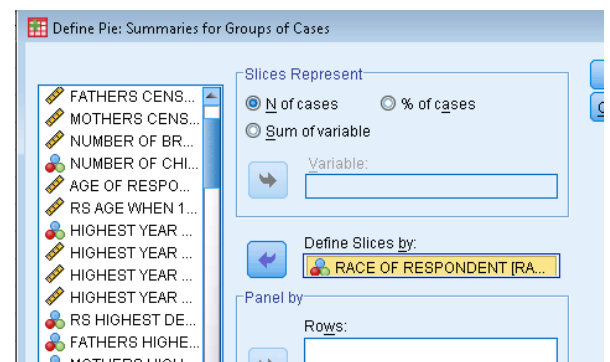
6. Click **OK**

To edit the pie chart:
1. Double-click on the chart to select it for editing
2. Click on **Elements** and click on **Show Data Labels**. You should now see the Properties dialog box.
3. Click on the **Data Value Labels** tab of the Properties box if not currently selected.
4. Click on **RACE** in the **Labels Not Displayed** area and then click on the **green upward pointing arrow** to move RACE Status to the Labels Displayed area.
5. Click **Apply** in the Properties dialog box. You should now see appropriate labels in the slices of the pie chart.
6. Click on the **Text Style** tab of the Properties dialog box.
7. In the front area, click on the **Size** drop-down menu and choose a font size (e.g. 10).
8. Click **Apply** in the Properties dialog box. The labels should now have a larger font size.
9. Click **Close** in the Properties dialog box.
10. Click on **Options** from the main menu. Click on **Hide Legend**. The graph should look like this:
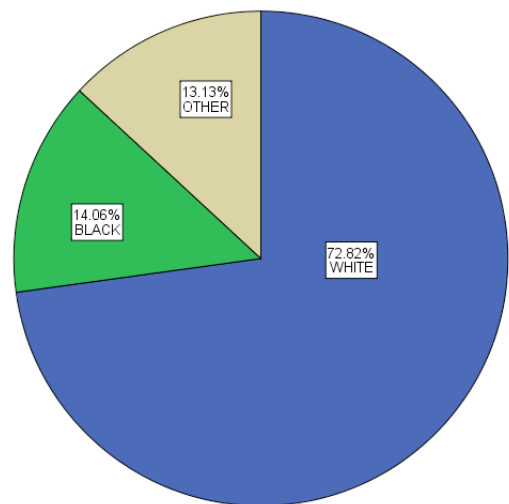

You may want to select only some of the "cases" in the data set. For example you may want to look only at females, or only respondents who live in a certain region of the country or only respondents who identify themselves as 'extremely conservative'. Before you begin to analyze the data, you will want to **select cases** that meet your criteria.

To select cases which meet a certain criteria:

1. First you need to look at the response options for the variable you are interested in limiting. For example, look at the Variable view of the data set and find a particular variable – **HAPPY** [General Happiness]. In the **Value** column you may click on the possible values/options that were provided to the respondent. In this case each respondent would chose *Not Applicable, Very Happy, Pretty Happy, Not too Happy, Don't Know,* or *Didn't Answer*. Each of these choices is assigned a numeric value [0, 1, 2, 3, 8, or 9].

Select which answer(s) you want to limit your cases to – for example if you only want to look at respondents who indicated "Very Happy" – note that the corresponding value is "1". You will need this information in a minute.
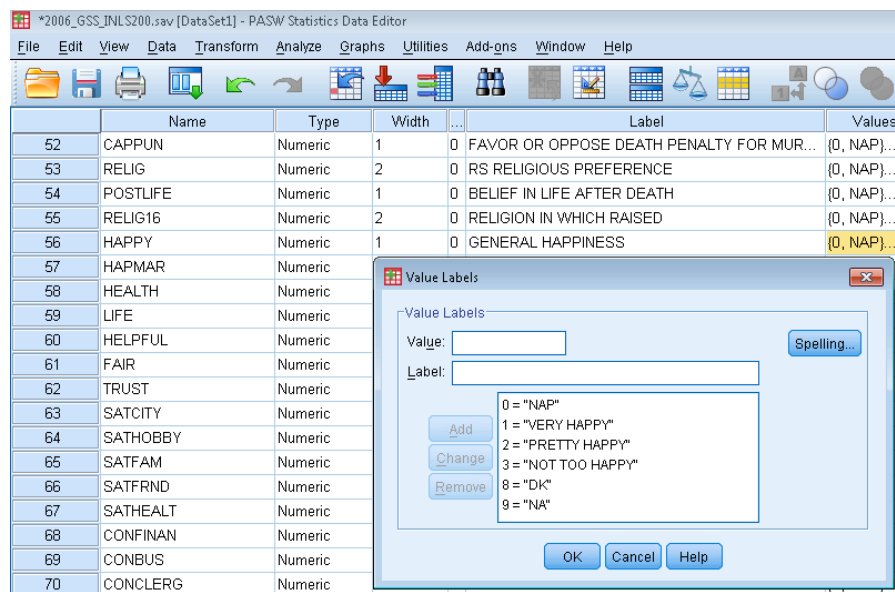

2. Click **Data** -- **Select Cases**
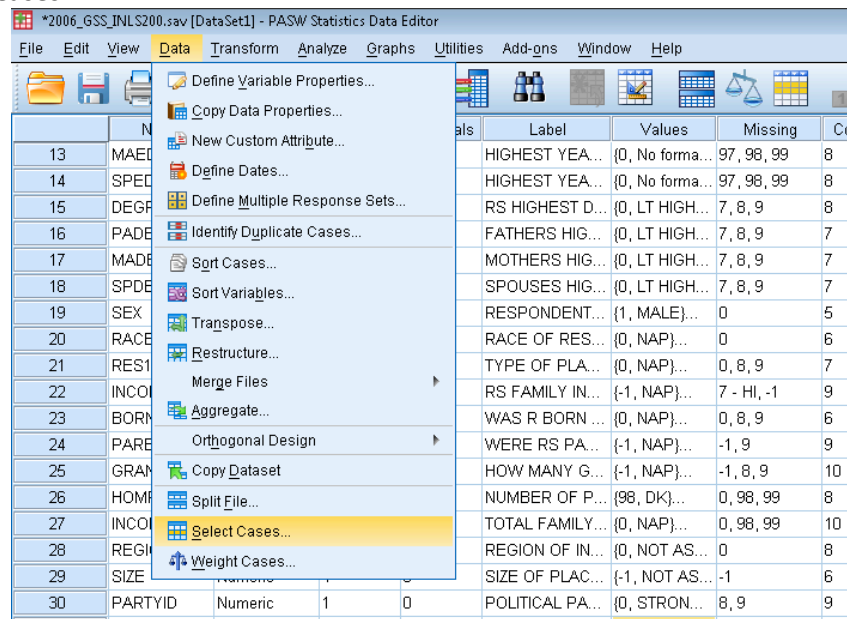
3. There are a number of choices. A random sample of cases, a sample based on a range, a filter variable, or a condition can be used to select cases. For this example, "If condition is satisfied" will be used. Click on **"If condition is satisfied"** Then click "**If…"**
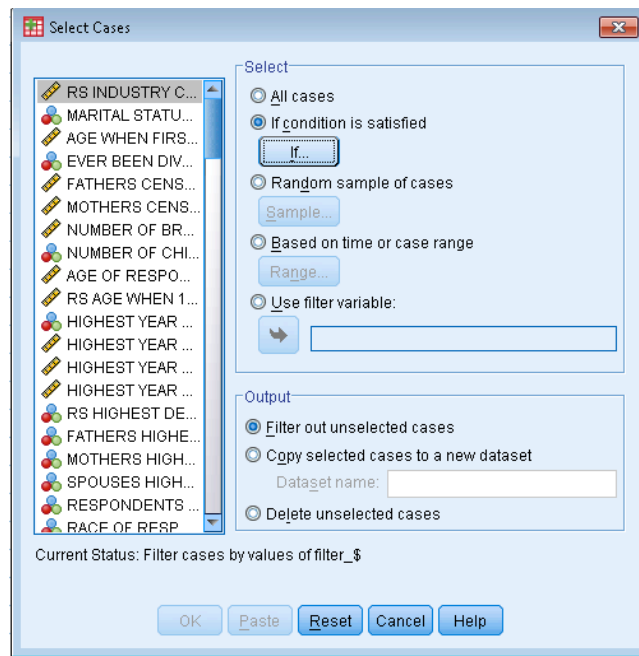
4. Here is where you need the value of the answer choice for the variable of interest. Highlight **GENERAL HAPPINESS** and click ► to move it to empty box. Once there type an "=" sign and the value you want. In this case type **"=1"** (no quotation marks though).
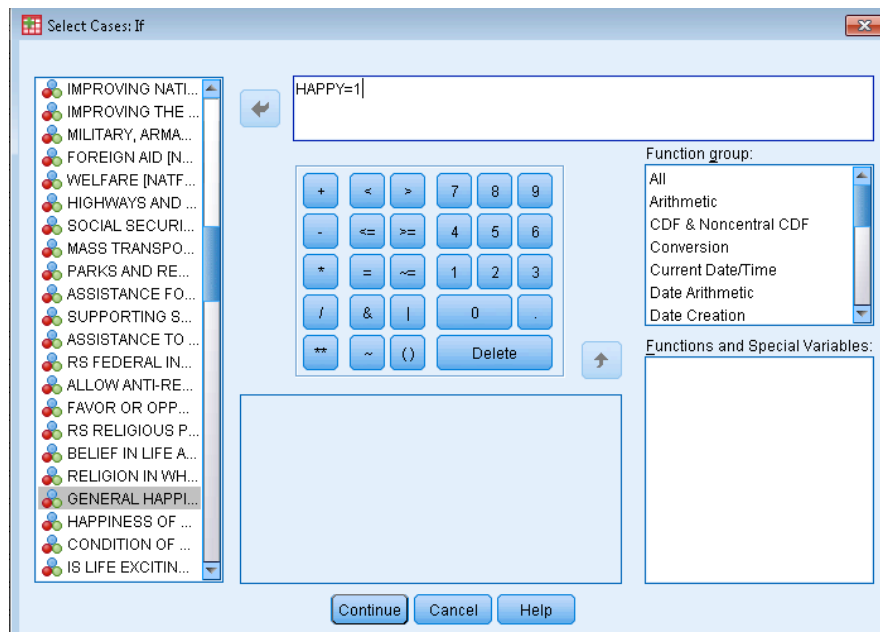
5. Click **Continue.**

6. In terms of Output, there are several options. The selected cases could be copied into a new dataset, deleted, or filtered out. In this instance, the unselected cases will be filtered out. Click **OK**.

7. In the **DATA VIEW**, note the slashes through many of the cases. These cases were respondents who selected any other choice than *Very Happy*. Cases 2, 4, and 10 have a general happiness coding of 1 for *Very Happy* and remain in the sample. All analyses done with the cases selected will only include *Very Happy* respondents.

★ In order to include all cases again or to select a different set of respondents, click
Data -- Select Cases -- All cases – OK ★