

*Basic SPSS Instructions for Displaying Data
(January 30, 2006)*

Locations of data files

Data page on the course website <http://ilt.ilstu.edu/kbarbou/eco138.htm> . These files can be opened from the webpage (SPSS must be installed on the computer for the file to open) and saved to the desktop or removable disk.

To Open a data file in SPSS

- Open SPSS by clicking the shortcut icon on the desktop or navigating through the program files on the start menu.
- Once SPSS has been launched, a dialog box will appear prompting you to select a particular objective for your session. In this class, our objective will be to open an existing data set for analysis; however, in the future (outside of this class) you may find yourself in the position where you are required to enter your own data (requires data entry similar to that done in a spreadsheet program like Excel).
 - Click the “Open an existing data source option” and select the specific data file you would like to work with in the text box below this option. If the file name does not appear, click on “more files” and navigate to the folder in which the file is located. You will likely select either the desktop, removable desk or My Network Places (to access my public file) depending on where you have stored the data. If you are trying to access the file from my public directory, first select My Network Places, then select econ on Samba 3.0.2a (stvcas) and then select the following folders in sequential order: kbarbou, public, data and then select the specific data set you would like to work with.
 - If you are attempting to open an excel file you must first change the file type in the “Files of type:” dialog box near the bottom center of the “Open File” dialog box. The default option is SPSS(*.sav), to open an excel file click the down arrow and select Excel (*.xls). In the “File name” box select the file you would like to open. In the next dialog box that appears, make sure that the appropriate excel worksheet (the one that contains the actual data) has been selected.

Two main Windows in SPSS

1. “Data View” shows the raw data where individual observations are reported in the rows and the variables containing information on each of the observations are reported in the columns under a given variable name.
2. “Variable View” shows specific details related to each of the variables including a descriptive label and value definitions for those categorical variables that have been coded with numbers. (Example of labels for the categorical variable gender 1= Male, 2 = Female). I recommend that anytime you start work on a new data set that you first visit the variable view window to learn more specific details on what the data set has to offer. Note: If you have opened an excel data file you will need to manually supply your own labels and value definitions. (Therefore, for the purpose of the data assignments for this class it is best to open a data file that I have converted to an SPSS format for you.)

To navigate back and forth between the windows selected the “Data View” and “Variable View” tabs in the lower right-hand corner of your screen.

To create Frequency Distributions

- Click Analyze/DescriptiveStatistics/Frequencies . . .
- This command creates a frequency table for the variable that appears in the *target variable window* (i.e. the window on the right-hand-side of the box). To place a variable in the target window, first clicking on the variable name from the list on the left and then click the right arrow key in the center of the box to move the variable into the target window.
- Click OK to complete the task. The table will appear in *output window*.
- Here is an example for a categorical variable

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	61	43.0	43.3	43.3
	Female	80	56.3	56.7	100.0
	Total	141	99.3	100.0	
Missing	System	1	.7		
Total		142	100.0		

Note: The results above indicate that one student did not respond to the gender question, thus he or she is recorded as a missing observation. Notice the Percent column shows the percentage in each category including the missing observation (i.e. the denominator for these calculations represents all students (142)). However, the valid percent column shows percentages based on only those who responded to the question (i.e. the denominator for these calculations is 141). You need to keep this distinction in mind when formulating your conclusions.

- Unfortunately, to my knowledge SPSS can not create frequency tables for quantitative variables similar to those we created in class (see notes from lecture on August 29). Instead, SPSS generates a table where each value that appears in the data set appears as an individual row. This is certainly not helpful if the

variable takes on many different values across observations. Therefore, to describe the distribution of a quantitative variable, I recommend that you go directly to a graphical analysis such as a histogram (which is perhaps the most frequently used), stem-and-leaf display or dotplot.

To create Bivariate Frequency Tables (i.e. frequency tables that include two variables)

- Click Analyze/Tables/Tables of Frequencies
- Variable Selection
 - Identify the variable you would like to count and move to Frequencies for box.
 - Identify the variable you would like to split the sample by and move to the Subgroups in each table box.
 - To show both counts and relative frequencies (i.e. percentages) Click Statistics and check both the count and percents display boxes and then hit continue.
 - Click OK to complete the task

Example: Frequency of Gender by Baseball team preference

To create the appropriate table, the gender variable should be in the frequencies for box and the baseball team variable should be in the Subgroups in each table box.

The table in the output will appear as

	Gender			
	Male		Female	
	Favorite NL central baseball team		Favorite NL central baseball team	
	Count	%	Count	%
Cardinals	14	23.3%	15	19.2%
Cubs	46	76.7%	63	80.8%

To create Contingency Tables (i.e. Crosstabs)

- Click Analyze/DescriptiveStatistics/Crosstabs
 - Select your variables of interest
 - Move one to the row box and the other to the column box. (Note: It really does not matter which one is which so long as you understand how to interpret the results of the table.)
- Click OK to complete the task

Example:

Gender * Favorite NL central baseball team Crosstabulation

Count

		Favorite NL central baseball team		Total
		Cardinals	Cubs	
Gender	Male	14	46	60
	Female	15	63	78
Total		29	109	138

To generate a more detailed table that includes (% of row, % of column, % of total), repeat the steps above, but before you click OK, select the “Cells” tab (bottom center of Crosstab dialog box). In the options box that appears under percentages check the type of percentage you would like to see reported in the table (row, column, total or all three).

Gender * Favorite NL central baseball team Crosstabulation

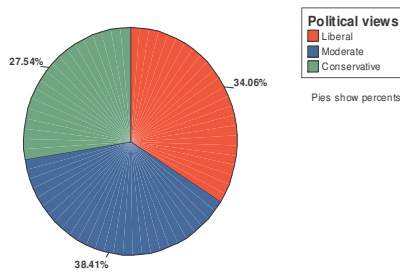
			Favorite NL central baseball team		Total
			Cardinals	Cubs	
Gender	Male	Count	14	46	60
		% within Gender	23.3%	76.7%	100.0%
		% within Favorite NL central baseball team	48.3%	42.2%	43.5%
		% of Total	10.1%	33.3%	43.5%
	Female	Count	15	63	78
		% within Gender	19.2%	80.8%	100.0%
		% within Favorite NL central baseball team	51.7%	57.8%	56.5%
		% of Total	10.9%	45.7%	56.5%
Total		Count	29	109	138
		% within Gender	21.0%	79.0%	100.0%
		% within Favorite NL central baseball team	100.0%	100.0%	100.0%
		% of Total	21.0%	79.0%	100.0%

To create Pie Charts

- Click Graphs/Interactive/Pie/Simple . . .
 - Click on the “Assign Variables” tab
 - Select the variable of interest from the list on the left and move it to the “Slice by” text box located on the right of your screen.
 - To report the results as percentages, move the percent variable from the list on the left to the “Slice summary” text box on the right
 - To add percentage labels for each slice of the pie as shown below, click on the “Pies” tab at the top of the Pie Graph dialog box and check the box next to percent under the “slice labels” option.
 - Select your variable of interest and move it to the define slices by box

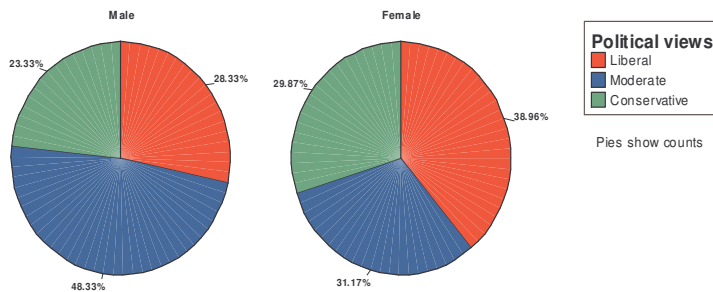
- Click OK to finish the task

Example:



To create pie charts for two different groups, say for example you want to compare political views by gender, repeat the above steps but before you click okay move the second variable (the variable you would like to split the sample on – in this case gender) to the “panel variables” text box

Example: Comparison of political views across genders

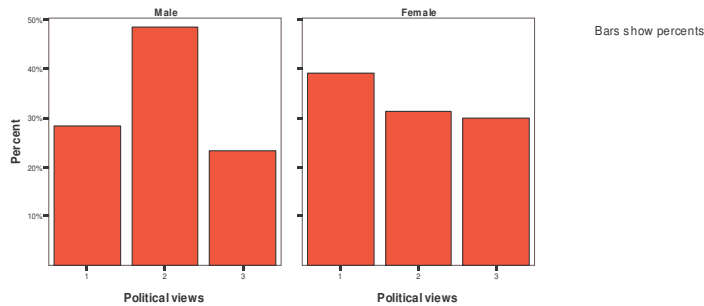
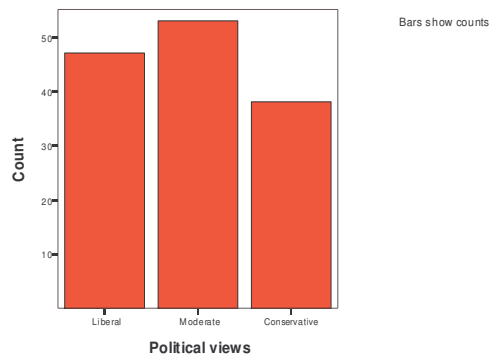


To create Bar Charts

The best method for comparing bar charts across groups is to use the interactive bar chart option.

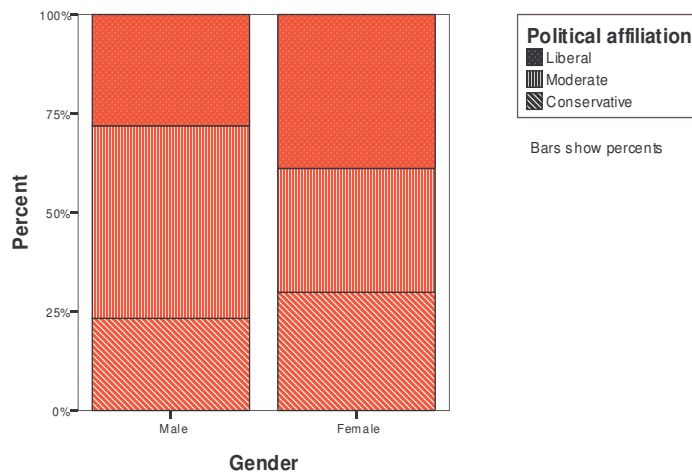
- Select Graph/Interactive/bar
- Move the SPSS generated variable “percent” or “count” (depending on whether you would like to report the actual frequencies or relative frequencies (%)) to the text box on the y axis of the chart sketch on the right of your screen.
- Select the variable of interest from the list on the left of the screen and move it to the text box on the x axis in the sketch (i.e. category axis) to define the bars.
- To compare results for two different groups, move the variable name for which you would like to split the sample on (gender in this example) to the Panel variables textbox in the lower half of the screen. Note: This variable must also be categorical.

Examples:



To Create a Stacked Bar Chart

- Recall that stacked bar charts are a great method for exploring whether or not an association exists between two variables. However, for this method to be effective we must follow the appropriate procedures for creating the bar charts we desire. Specifically we must be sure that the height of each bar reaches 100%. For example if we wanted to show the distribution of political affiliation by gender, we would want the bars for male and female to total 100% as shown below.
- To create a stacked bar chart, click on graph/interactive/bar move the first variable (the one that defines the categories for the bars – in this example gender) to the horizontal axis in the graph sketch on the right of the screen. Move the second variable (the variable that defines the bar segments) to the text box labeled style. Final make sure the box labeled 100% stacked (located top center of this submenu) is checked, then click okay.



•

To Create A Clustered Bar Chart

- SPSS has some troubles generating the types of cluster bar charts similar to those presented in lecture and in the textbook. I would recommend that you do not use this option unless you become familiar with an understand SPSS' deficiencies.

To create a Histogram

- Click Graphs/Interactive/Histogram
- Move the count or percent variable to the text box on the y axis in the chart sketch depending on which type of histogram you would like to create. [Using the count variable creates a frequency histogram while using the percent variable creates a relative frequency histogram].
- Move variable of interest (i.e. the variable whose distribution you are trying to explore) to box on the x-axis in the chart sketch.
- Click the histogram tab to set the number of intervals (i.e. classes) by selecting "number of intervals" and then use the up and down arrow keys to select the appropriate number. A second option would be to use the default option (i.e. let SPSS choose the number of class for you automatically). For data analysis assignments, I will usually specify the number of classes you should use in your histogram.
- Click OK to complete the task
- If you want to compare histograms for two different groups, in the "Assign variables" window drag the variable that defines the groups into the "Panel variables" window.

Examples:

