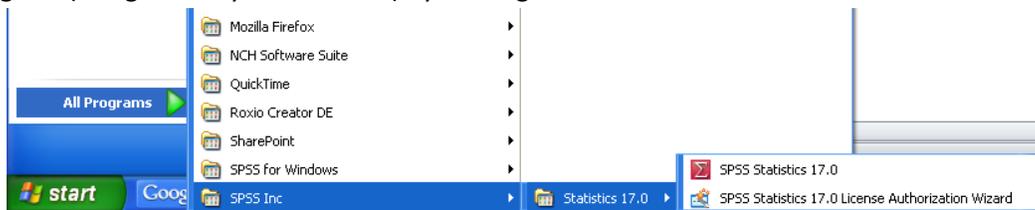


## Chapter 2 Dealing with Data SPSS Tutorial

1. **Visit** the student website at for this textbook at [www.clowjames.net/students](http://www.clowjames.net/students).
2. **Download** the following files:  
*"Chapter 02 Dealing with Data"* (SPSS data file)  
*"Chapter 02 Dealing with Data Survey"* (Word file)

The first file is the actual data file that contains the results obtained from a student research project. The second file is the questionnaire composed by the student, with response coding instructions included in red type. These codes correspond to the value labels, discussed later.

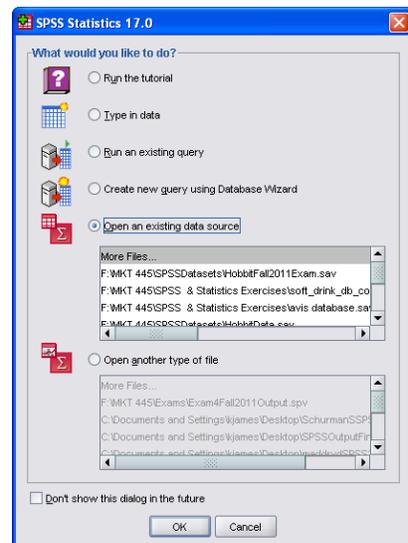
3. Open the questionnaire file. It may be helpful to print this one page questionnaire before continuing.
4. Open SPSS 17.0.  
 The program may appear as an icon on your computer desktop, or you may need to use the START > ALL PROGRAMS command in Windows to find the "SPSS Inc." folder. Once you have located this folder, open it, then open the "Statistics 17.0" folder. Finally, launch the SPSS Statistics 17.0 program (designated by the red icon) by clicking on that item.



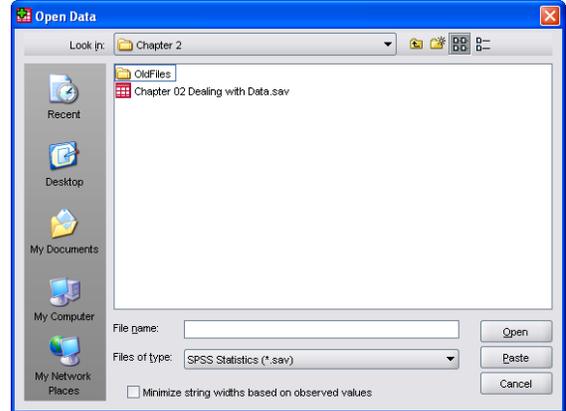
When first launched, the window at right appears and SPSS asks users what they would like to do.

Options include running the SPSS tutorial, creating a new data file (which includes defining variables and typing in data), running either an existing query (analysis) or a query using the database wizard, or opening either an existing data source file, or another type of SPSS file, such as an output file.

File type identifiers include .sav for data files, or .spv for output/viewer files. Most of the time you will not see these file type identifiers listed as part of the file name, except through this window. To acquaint yourself with SPSS, continue following these instructions.



- For this exercise, you will be selecting the fifth option, “Open an existing data source” to open the data file named, “Chapter 02 Dealing with Data”. If the file name does not appear in the window below this option, select the “More Files” option, then click the OK button.
- The OPEN DATA Window will appear (see right).
- Browse to the location where you saved the SPSS data file. Once you have located the file, click once on the file name to highlight it, then click the OPEN button.



The opened data file will look similar to that shown below.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	IDNO	Numeric	8	2		None	None	8	Right	Scale
2	Q1Often	Numeric	8	2	How often do y...	(1.00, Less ...	None	8	Right	Scale
3	Q2Money	Numeric	8	2	How much do y...	(1.00, Less ...	None	8	Right	Scale
4	Q3RankBuc...	Numeric	8	2	Rank Buckle	(1.00, Most ...	None	8	Right	Scale
5	Q3RankDil...	Numeric	8	2	Rank Dillard's	(1.00, Most ...	None	8	Right	Scale
6	Q3RankJCP...	Numeric	8	2	Rank JCPenney's	(1.00, Most ...	None	8	Right	Scale
7	Q3RankTarg...	Numeric	8	2	Rank Target	(1.00, Most ...	None	8	Right	Scale
8	Q3RankWal...	Numeric	8	2	Rank Wal Mart	(1.00, Most ...	None	8	Right	Scale
9	Q4ShopBuc...	Numeric	8	2	Shop at Buckle	None	None	9	Right	Scale
10	Q4ShopDil...	Numeric	8	2	Shop at Dillard's	None	None	9	Right	Scale
11	Q4ShopJCP...	Numeric	8	2	Shop at JCPen...	None	None	8	Right	Scale
12	Q4ShopTarg...	Numeric	8	2	Shop at Target	None	None	8	Right	Scale
13	Q4ShopWal...	Numeric	8	2	Shop at Wal-Mart	None	None	8	Right	Scale
14	Q4ShopOther	Numeric	8	2	Shop at Other ...	None	None	8	Right	Scale
15	Q5Gender	Numeric	8	2	Gender	(1.00, Fema...	None	8	Right	Scale
16	Q6Class	Numeric	8	2	Class status	(1.00, Fresh...	None	8	Right	Scale
17	Q7Age	Numeric	8	2	Age	(1.00, Youn...	None	7	Right	Scale
18	Q8Income	Numeric	8	2	Annual income	(1.00, Less ...	None	8	Right	Scale
19										
20										
21										
22										
23										
24										
25										
26										
27										
28										
29										
30										
31										

Notice the “Data View” and “Variable View” tabs in the lower left-hand corner. SPSS defaults to the variable view, or second tab, that shows information about each variable. The “Data View” contains the raw numbers that are used to represent information in the data file.

### Variable View Window

The “Variable View” window displays a listing of the properties of each variable contained in the file. In the example shown here, the SPSS data file contains a number of variables related to a survey. The first four variables are named “IDNO”, “Q1Often”, “Q2Money”, and “Q3aPleasure”. In this example, the “IDNO” variable represents a unique identification number assigned to each individual participant’s completed survey, and the other three variables correspond to the survey question answers provided by the study participants. SPSS allows users to create **labels** for variables – an important advantage – as these labels will appear on all output, charts, or graphs resulting from data analysis procedures. Labels may replicate the entire question exactly as asked on a survey, as

shown the second through fourth variables in the screenshot. Alternately, labels can be descriptive of a particular variable or survey question, as is the case for the remaining variables, beginning with variable Q3bBrand.

Because data analysis uses numerical (counts, percentages, means and more) and statistical processes (chi-square, t-tests, regression calculations, etc.), most information must be represented numerically in SPSS data files. The **values** portion of the variable description allows users to define what the numbers associated with a particular variable represent. For example, data represented as a “1” in response to the question 3a, corresponds to an answer of “Strongly Disagree”. While researchers have control over additional settings, such as the number of decimals in the data, and how missing data is treated, most researchers leave the default settings in place for these remaining columns. When creating a new SPSS data file, the variable view information is always entered first.

8. Click the “Data View” tab. Your screen will look similar to that shown below.

	IDNO	Q1Often	Q2Money	Q3RankBuckle	Q3RankDillards	Q3RankJCPe	Q3RankTarget	Q3RankWalMart	Q4ShopBuckle	Q4ShopDillards	Q4ShopJCPe	Q4ShopTarget	Q4ShopM
1	1.00	3.00	3.00	1.00	3.00	2.00	4.00	5.00	5.00	5.00	10.00	20.00	
2	2.00	1.00	5.00	1.00	3.00	2.00	4.00	5.00	25.00	10.00	0.00	10.00	
3	3.00	1.00	4.00	3.00	1.00	2.00	4.00	5.00	0.00	0.00	15.00	30.00	
4	4.00	1.00	5.00	5.00	4.00	2.00	1.00	3.00	0.00	5.00	20.00	20.00	
5	5.00	1.00	3.00	5.00	2.00	1.00	3.00	4.00	0.00	0.00	10.00	40.00	
6	6.00	3.00	3.00	1.00	3.00	2.00	4.00	5.00	0.00	10.00	15.00	25.00	
7	7.00	1.00	3.00	3.00	4.00	2.00	4.00	5.00	0.00	0.00	20.00	30.00	
8	8.00	4.00	4.00	1.00	2.00	3.00	4.00	5.00	25.00	20.00	5.00	0.00	
9	9.00	3.00	2.00	3.00	2.00	1.00	4.00	5.00	5.00	20.00	5.00	5.00	
10	10.00	4.00	2.00	5.00	3.00	4.00	1.00	2.00	10.00	10.00	20.00	10.00	
11	11.00	3.00	3.00	5.00	3.00	2.00	1.00	4.00	10.00	0.00	20.00	10.00	
12	12.00	3.00	5.00	1.00	2.00	3.00	4.00	5.00	0.00	0.00	0.00	40.00	
13	13.00	1.00	2.00	1.00	5.00	4.00	3.00	2.00	0.00	0.00	15.00	50.00	
14	14.00	2.00	4.00	3.00	2.00	1.00	5.00	4.00	0.00	5.00	10.00	20.00	
15	15.00	2.00		2.00	1.00	4.00	3.00	5.00	0.00	5.00	5.00	40.00	
16	16.00	1.00	3.00	1.00	4.00	3.00	2.00	5.00	5.00	10.00	30.00	15.00	
17	17.00	2.00	2.00	3.00	2.00	4.00	1.00	5.00	10.00	20.00	20.00	25.00	
18	18.00	3.00	3.00	1.00	3.00	2.00	4.00	5.00	0.00	0.00	10.00	20.00	
19	19.00	1.00	5.00	1.00	3.00	2.00	4.00	5.00	20.00	20.00	20.00	10.00	
20	20.00	1.00	4.00	3.00	1.00	2.00	4.00	5.00	0.00	0.00	10.00	40.00	
21	21.00	1.00	5.00	5.00	4.00	2.00	1.00	3.00	10.00	10.00	10.00	20.00	
22	22.00	1.00	3.00	5.00	2.00	1.00	3.00	4.00	0.00	10.00	5.00	25.00	
23	23.00	4.00	4.00	5.00	1.00	2.00	3.00	4.00	5.00	5.00	20.00	20.00	
24	24.00	1.00	3.00	3.00	4.00	2.00	4.00	5.00	0.00	0.00	10.00	50.00	
25	25.00	4.00	4.00	1.00	2.00	3.00	4.00	5.00	40.00	20.00	5.00	0.00	
26	26.00	3.00	2.00	3.00	2.00	1.00	4.00	5.00	0.00	0.00	0.00	25.00	
27	27.00	4.00	2.00	5.00	3.00	4.00	1.00	2.00	0.00	0.00	5.00	25.00	
28	29.00	3.00	3.00	5.00	3.00	2.00	1.00	4.00	0.00	5.00	5.00	10.00	

### Data View Window

As the name suggests, this view contains data, which is usually in the form of numbers that represent information. The type of data represented by a particular variable is defined via the type column, which defaults to “numeric”. While several types of data can be represented in SPSS, the most commonly used types of data include:

**Numeric.** A variable whose values are numbers. Values are displayed in standard numeric format. The Data Editor accepts numeric values in standard format or in scientific notation.

**Scientific notation.** A numeric variable whose values are displayed with an embedded E and a signed power-of-10 exponent. The Data Editor accepts numeric values for such variables with

or without an exponent. The exponent can be preceded by E or D with an optional sign or by the sign alone--for example, 123, 1.23E2, 1.23D2, 1.23E+2, and 1.23+2

**Date.** A numeric variable whose values are displayed in one of several calendar-date or clock-time formats. Select a format from the list. You can enter dates with slashes, hyphens, periods, commas, or blank spaces as delimiters. The century range for two-digit year values is determined by your Options settings (from the Edit menu, choose Options, and then click the Data tab)

**Dollar.** A numeric variable displayed with a leading dollar sign (\$), commas delimiting every three places, and a period as the decimal delimiter. You can enter data values with or without the leading dollar sign

**Custom currency.** A numeric variable whose values are displayed in one of the custom currency formats that you have defined on the Currency tab of the Options dialog box. Defined custom currency characters cannot be used in data entry but are displayed in the Data Editor

**String.** A variable whose values are not numeric and therefore are not used in calculations. The values can contain any characters up to the defined length. Uppercase and lowercase letters are considered distinct. This type is also known as an alphanumeric variable.

It is important to understand that data in each *row* represents the responses pertaining to a particular survey, interview, or observation, while columns represent the nature of the variable data – in this instance, different respondents' answers to particular questions.

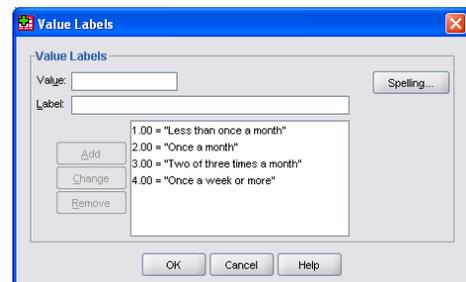
For example, usable data was collected for one-hundred and ninety-five survey respondents. However, the unique identifier numbers (column 1) range between 1 and 200, indicating that some respondents were eliminated from the data set, probably due to incomplete information. The data associated with IDNO 1 indicates that this particular respondent shopped at least two to three times per month, spending \$50-\$74 dollars on average per trip. Furthermore, the respondent ranked Buckle as the most desirable store, while Walmart was ranked as the least desirable store.

- From the top menu, select VIEW > VALUE LABELS. Notice how the numbers in the "Data View" change to show the information that these numbers represent. Repeat this process to turn off the value labels.

So, where did this information come from?

- Return to the "Variable View".

The screenshot at right shows the "Value Labels" pop-up window associated with *Q10ften*, "How often do you shop for clothes?"



- To access this information, click the little grey box with

three dots next to the under the “Values” column in the row associated with the *Q1Often* variable.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	IDNO	Numeric	8	2		None	None	8	Right	Scale
2	Q1Often	Numeric	8	2	How often do y...	month)...	None	8	Right	Scale
3	Q2Money	Numeric	8	2	How much do y...	{1.00, Less ...	None	8	Right	Scale

- Review question 1 of the survey. Now look at the “Label” column for *Q1Often*. If you have trouble viewing the entire label, position your cursor on the line between the “Label” and “Values” column. Once your cursor turns into a double-headed arrow, hold the mouse button down and pull to the right to expand the width of this column.

### 13. Your Assignment:

#### Part A:

Write a short description of each column shown in the “Variable View” tab. For instance, the first column is a short abbreviated name of the variable. Most of these columns should be fairly self-explanatory and easy to figure out. Use the grey button when available to explore different options in each column, and if you still don’t understand what a particular column represents, select HELP from the main menu, then TUTORIAL. Go through the “Using the Data Editor” section until you find a bit more information about the column in question.

#### Part B:

Two columns that are especially important for you to examine closely are the columns entitled, “Labels” and “Values”. Why do you think these are important?

**NOTE:** This simplistic data file is offered merely for the purpose of providing a preliminary overview of SPSS. In “real-life”, data files may contain a hundred or more variables, gathered from hundreds or even thousands of participants.