

## Lab Two Directions

Before starting, read the introduction to the data set found on the sheet with the data and lab questions. Once you know about the data you will be using you can start the program as you did last week by going to the **start** button, clicking on **programs** and then selecting SPSS from the side menu. When the software comes up you can use the blank data screen.

### Defining variables:

Last week we labeled two variables by giving each variable a name. There were also several other characteristics about the variable that we could have set. You will do some of this in this lab. You will need to define three variables. Each variable will occupy one column. The first variable will be a string variable, in other words, text. It will designate each town. (This variable is only there to help you keep track of the data.) The second and third variables will be numeric, holding the dollar amount and the geographic area that each town belongs to. Remember that to begin defining the variables you will need to click on the **variable view** tab at the bottom left hand corner of your screen. This produces a new spread sheet that is oriented in the opposite direction from the data sheet. Each column now contains information concerning each variable e.g., the variable name, the width of the column, etc.)

*Naming the variable:* Begin by defining the variable for the name of the town. It doesn't matter what you use as a variable name. Just remember that you are naming the general variable for town names, not a particular town. Next, go to the "type" column. You will notice that the default setting is *numeric*, in other words, data that are in the form of a number. If you place the cursor on the spreadsheet box under this column you will see a small gray area. Click there and a dialogue box will open that allows you to select the type of data that you want. You should click on "**string**".

Move the cursor to the second row so that you can define the variable that will hold the data concerning the average amount spent per child. Create a variable name for that as well. You can leave this variable as "numeric" since you will want to use the data in this column as a set of numbers.

*Number of decimal places:* Now move over to the "decimals" column. As you place the cursor on the box you will see arrow keys that will allow you to change the number of decimals that you can have. Since we will use whole numbers change this value to 0.

Now go to name the third variable, the one that will identify each town's geographic area. First, give this variable a name. Again, the variable will be numeric. You may wonder why you don't define it as a "string" variable so that you could enter the words "western",

“south shore” and “metro west”. The problem with this is that SPSS won’t analyze string data in ways that I want you to use these data. Again, you can define this variable as having 0 decimals.

*Labeling the values that you are using:* You are going to indicate each town’s geographic location with a number. It doesn’t really matter what the number is as long as you are consistent throughout the data. I happened to use “1” for “metro west”, “2” for “south shore” and “3” for “western”. I can have these numbers identified in any printout by using the “labels” column in the defining sheet. Go to that once you have given the geographical area variable a name. Go to the column labeled “values”. By clicking on a box under this column a gray spot will appear. Click on the gray spot and you will get a dialogue box.

This dialogue box lets you label the values that you are using. Type the value for metro west (in my case a 1) in the upper box. Then, type the word or phrase that value represents (e.g. metro west) in the box below. Once this is done you can click on the **add** button. The label will be added to a list of labels in the larger box below. Now you can return to the first box and type the next value, the one for south shore. Repeat the labeling process. Once you have labeled your three possible values and added them to the list you can click on the **OK** button. Labeling or not labeling will not affect your data, but providing labels during the defining process will make reading your output easier.

Now you are ready to return to the **data view** to enter data.

### **Entering data:**

You are going to enter the data in almost the same way as you did last week. The only difference will be that each row this week will represent one town. Make sure to enter the name of the town, money spent and the number that designates the geographical area.

### **Analyzing the data:**

I want you to run two analyses on these data. First, analyze the data from all of the towns together. Then, you will find descriptive statistics grouping by geographical location.

To begin, use the **descriptive statistics** from the **analyze** pull down menu like last week. Select **frequencies** from the side menu. Highlight the variable name for money spent and move it over to the box for variables to be analyzed by clicking on the arrow. Using the **statistics** button, select the following descriptive statistics:

mean, median, standard deviation, minimum and maximum.

Click on the **continue** button to return to the main dialogue box. Now click on the **charts** button. Ask for a histogram but also check the box about the normal distribution so that

you can see how your data compares to the normal distribution. Make sure to click on the **ok** button to get the output.

Once you have looked at the output return to the **Analyze** command. Now you will analyze the data, grouping the towns by geographical region. To do this, choose **compare means** from the pull down menu under **Analyze**. Select **means**, the first option, from the side menu. This will open a dialogue box that will again list the variable names in a box on the left hand side. There will be two boxes on the right into which you can place variable names. The top box is for the **dependent list**. In this case that is the variable name that you gave to the average money spent. Below this box is a second one labeled **Independent list**. Put the variable name that you gave to the geographical area variable in this box. Click on the **options** button and you will find a list of descriptive statistics. Choose mean and standard deviation and move these to the “cell statistics” box using the arrow. Click on the **continue** button which will return you to the dialogue box. Then click on **OK**.

Notice that the second output gets added to the bottom of the first output that you requested that gave information about all of the states together.