

## Introduction to SPSS

### Module 1: Preparing for Analysis

#### SLIDE 1,2,3

The module introduces you to the basic data entry steps. Attachments: Instructions, a sample data set, and a sample survey. We will use it to get started with preparing data for analysis.

#### SLIDE 4

##### Step 1 - Check the surveys

A sample survey is attached. You are looking for data that is obviously wrong or information that leads you to believe that the survey needs to be removed from the study. This is a good time to number your surveys.

#### SLIDE 5

##### Step 2 - Locate the variables on the survey

Using the sample survey, identify the items which are measuring the variables in the study.

#### SLIDE 6

##### Step 3 - Assign a layout for SPSS

Open SPSS: **Start > Programs > SPSS for Windows > SPSS for Windows**. This first screen is your SPSS Data Editor (enter data here - one participant per row, one variable per column).

#### SLIDE 7

##### Step 4 - Develop your codefile

The codefile is where you define and code your variables. Switch to the Variable View: at the bottom left of the SPSS Data Editor screen there are two tabs (Data View and Variable View).

Click on **Variable View**.

#### Define

The Name column is the first column (format is one variable per row). Starting with your first variable, enter the variable name. For example, in the first column of the first row type in *case* (up to 8 characters with no spaces).

Example Variables and Coding:

- \*case (participant number)

- \*gender: 1 = male, 2 = female

- \*age

- \*ethnicity

- \*school: 1 = community college, 2 = 4-year university 3 = other

- \*college GPA

- \*high school GPA

- \*RSR rank

- \*RSR math

- \*RSR writing

- \*confidence: "I knew I would be successful in college." 1 = strongly disagree, 2 = somewhat disagree, 3 somewhat agree, 4 = strongly agree

- \*Desire Control (4 questions entered as separate variables)

The Type column's default setting is Numeric data. If this variable is non-numeric data (words), click on the cell and then on the 3 dots shaded in grey. You can then select the new setting. You can also change the column width and number of decimal places (these are the next two columns on the spreadsheet). Click **OK**.

The Label column is an area where you can write in the full question or any other information to help you identify the variable. Later, you may want to (or not to) have the labels on your printouts. To make this choice, select **Edit > Options**. Select the **Output Labels** tab and make your selection under *Labeling*. Click **OK**.

The Measure column is where you indicate the level of measurement of the variable. Choices are Nominal, Ordinal, or Scale (interval or ratio). Click on the cell, then on the down arrow, and make your selection.

#### Code

In the first column of the second row type in *gender*, your second variable. Since this is a nominal variable (participant's responses will be words), you will need to give each possible response a value (a number/code) such as 1 = male and 2 = female. (Another option is to change the Type to String but transfers with Excel are easier if it is all coded.) In the Values column, click on the cell and then on the 3 dots shaded in grey. A Value Labels window will appear. Enter "1" in the Value box and "male" in the Value Label box. Select **Add**. Then enter "2" and "female." Select **Add > OK**.

#### SLIDE 8

##### Step 5 - Enter Data

Go to the Data View spreadsheet by clicking on the Data View tab in the bottom left corner of the screen. Start with the first participant by entering the case number in the first row of the *case* column. Leave the cells blank when there is missing data (i.e. a participant does not respond to an item). Each analysis will indicate if there is missing data and how many cases were used in each analysis. For variables that you have coded, you can enter them as numbers (i.e. 1 for male) or select from a menu (clicking on the cell will produce a down arrow in the box that presents you with the choice of male or female). Select **View > Value Labels** (check or uncheck as you choose). [Remember to save your file frequently.]

#### SLIDE 9

##### Step 6 - Check your entries

Once you have entered your data, it is important to check carefully for errors. You can enter the data twice and have the computer compare the two datasets. You could run descriptive statistics on all your variables and verify that the minimum and maximum values for each variable do not exceed the possible value range for that variable (e.g., if you find your maximum value for gender is 3, and you have 2 levels, you know there is a problem). Also, examine the standard deviation and sample size for all variables. You may want to make a scatter plot or a histogram to check for outliers or normality.

#### SLIDE 10,11

##### Recap and Next Modules