Exercise 1: Exporting Data from Online Platforms

Author: Leilani Madrigal

Department of Kinesiology

California State University, Long Beach

Long Beach, CA 90840

Email: Leilani.madrigal@csulb.edu

**Note to the Instructor**: The data set used in this exercise (**Sport Injury and Anxiety**) is data that was collected and put together for the purpose of this exercise module. The exercise explains how to export online data from platforms such as qualtrics and the steps needed to check data prior to placing in an SPSS file. You have permission to use this exercise and to revise it to fit your needs.  Please send a copy of any revision to the author. Included with this exercise (as separate files) are the excel file for students to practice (Sport Injury and Anxiety raw data), the cleaned excel file for students to compare answers (Cleaned Excel Fil**e**). Please contact the author for additional information.

Attached are files for this exercise:

-date set (excel format) –for students to practice (Sport Injury and Anxiety\_raw data)

-data set (excel format) –answers (cleaned excel file)

**Goals of Exercise**

The goal of this exercise is to learn how to export an excel file from Qualtrics and steps needed to “clean” the file so it can be easily copied to an SPSS file.

**Exporting Data from Qualtrics**

Although not everyone may use qualtrics to collect data using their online platform, the steps of this exercise are universal in how data should be checked and managed when exporting from an online site in preparation to inserting in an SPSS file.

For those using qualtrics, once you have logged in to export your data, you would click on the existing project (e.g., “Sport Injury Study”) and then click on “**Data & Analysis**” then on “**Export & Import**”. This would give you a drop down list, which you would then click on “**Export Data**…” You are then prompted to select which method you want to download the data. The pre-set selection is a CSV file (similar to Excel), I recommend using this format. All you would need to do at this point is click on “Download”. Once your file is downloaded, I recommend saving it to your Desktop or a location you are able to find on your computer.

Once you export the data file, into a CVS or excel format, the file may appear similar to **Excel Raw,** which is attached to this exercise.

**Reducing Exported Data ~ Cleaning the file**

I refer to this step as “cleaning” the file because often when data is exported there is information that has been added that is unnecessary. For example, the first few columns are the IP addresses and length of time the survey took each individual. While some people may find this added information useful, it often just clutters the screen and makes the data set appear intimidating. Additional problems may be the use of written text instead of numerical values, so checking the data and cleaning at this stage is important. Follow the steps below to begin cleaning the exported data file:

 **Deleting Unnecessary Columns of Data.** Notice on the **Excel Raw File**, columns that should be deleted. For example, columns A through C, should be removed as they were not part of the data collected. Can you identify where the first question of the survey begins? It is on *column S*, where participants are asked to report their current age. So everything prior can be removed.

Exercise 1a: Remove all unnecessary columns from the data. Add a column and label it ID, then number off each row (e.g., 1, 2, 3, etc). Remember each row represents an individual’s response so this is a good way to organize your data (assign IDs).

 **Creating Codebook.** This steps is done to help you remember what variables are when you remove the text and replace with a numerical value. For instance, gender may have originally been listed as a series of females and males on the dataset, but in order to transfer the file to SPSS, assigning a numerical value to each would be needed. Open a new tab on the excel file and label it “Codebook”, beginning with gender you would indicate 1 = male, 2= female. You would want to do this for all variables on your remaining excel list.

Exercise 2a: The variables on your codebook tab should include: Year in School, geographic location, gender, ethnicity, sport, seasons competed in college, NCAA level, injury label, went to athletic trainer, season injury occurred, length unable to train/compete, surgery required, rehab required, currently injured. Create your codebook on another tab of your excel file.

Exercise 2b: Now go through the data and change all written values for their numerical coding. At the end of this step, there should be no written text on your excel file expect for the first two rows.

 **Missing data.** Often times in survey data, there are places in which a person skipped or did not answer an item. In order for SPSS to appropriately calculate for the missing information, you will need to assign a value so the system will know when it sees that number it means missing value. What would be a good number to assign as a missing data code? 0? What may be problematic with using 0? Remember whatever code you give it, you would want to make sure it wouldn’t get mixed up with a value in which you were interested in collecting. I generally use 999 as my missing variable code because the data I collect generally does not go up to or beyond that number.

Exercise 3: Go through the data set and insert 999 wherever you see a blank cell.

**Check Your Work**

Go back through the steps and see if your answer match. For deleting unnecessary columns of data, how many columns did you delete? Do your columns match with that on Excel Clean. You can also check this by looking at the Codebook tab on the Excel Clean, does your codebook match? Lastly, did you do a visual scan and catch any 999s that should be inserted? Always make sure to save your work. Save the excel data file as Exercise 1\_dataset.