NOTE: ALL CLASS NOTES WILL BE POSTED ON BLACKBOARD IN "Course Content".

INTRODUCTION TO SPSS

CLASS NOTES

Prepared by R. Hassad, MPH, PhD

Reading assignment: Chapters 2 & 3 of the required text (*SPSS 16.0 Guide to Data Analysis*- Marija J. Norusis). You may also use the documents posted under "Learning Modules" to supplement your reading.

This is important background information that you need to know and have before we begin using SPSS. The important terms are highlighted in red. If you are encountering problems obtaining, installing or accessing SPSS, please contact me by telephone or email. You can also meet with me on campus (in the Bronx) on Monday, Friday and Saturday by appointment.

What is SPSS? - **S**tatistical **P**ackage for the **S**ocial Sciences

SPSS is a data management program, package or software. This program allows you to:

1. Perform data entry
2. Conduct statistical analysis; and
3. Generate tables and graphs (for presentation)

Data will be entered in the form of variables, which broadly speaking, can be either NUMERICAL (quantitative) or CATEGORICAL (qualitative). Note that categorical in SPSS is represented by the term "string".

What is a variable? A variable is any characteristic, quality or factor that can change. Knowing the specific type of variable/data that we are working with allows us to perform and generate appropriate statistical analyses and graphical formats.

For practical purposes there are two types of numerical variables/data:

1. Continuous – which we get from measuring, and the values can have a fraction of a unit, for example; a body temperature of 101.6 degrees Fahrenheit.

2. Discrete - which we get from counting, and the values are whole numbers only, for example; the number of students in this class is 22.

Also, for practical purposes, there are two types of categorical (or qualitative) variables/data:

1. Nominal – which refers to unordered categories, for example; gender, ethnicity, and marital status. The categories of these variables do not indicate or represent that one is higher or lower than the other. The category is simply a name.
2. Ordinal – which refers to ordered categories, for example; ranks in the military, severity of diseases (mild, moderate, severe), height (short, medium, tall), and stages of college life (freshman, sophomore, junior, senior). The categories of these variables have a particular rank or order, and hence one can be higher or lower than the other (unlike nominal variables/data).

Note that categorical variables/data are sometimes referred to as discrete, given that the categories are separate and generally different from the others, and this is what the word discrete means.

NOTE: There is a key difference between variables with respect to what they represent, and how they are measured.

Some characteristics are basic, objective, and easy to measure, for example; age, weight, height, ethnicity, and marital status. Each of these is UNIDIMESNIONAL (have one dimension or aspect), and if we want to ascertain any of these from our participants, a single (one) question or item will okay, for example: What was your age last birthday? What is your gender?

On the other hand, characteristics such as, IQ, self-esteem, depression, personality type, and anxiety are MULTIDIMENSIONAL (have more than one dimensions or aspects) and complex, and hence more challenging and difficult to measure. Such characteristics or factors are referred to as CONSTRUCTS, because they are built from, or are composed of many variables or items.

In order to measure a construct we must use a scale that is reliable and valid.

What is a scale: A scale is a set of inter-related items intended to tap or measure a particular concept or construct. See for example the Rosenberg self-esteem scale (http://www.yorku.ca/rokada/psyctest/rosenbrg.pdf).

Another commonly used scale in psychology and the behavior sciences is the BDI (Beck's Depression Inventory (http://www.ibogaine.desk.nl/graphics/3639b1c_23.pdf)

NOTE: When we begin to analyze data using SPSS, I will introduce you to statistical techniques/methods, which will be either descriptive (related to the sample) or inferential (related to the population).

Coming Attractions!!: The next handout will specifically address data entry, and we will begin by defining the variable(s). When we define a variable, we are basically assigning certain characteristics to it by designing a template, so that there is a standard format for entering the data collected.

To make this easy for you all, for each exercise, I will send a Microsoft Word file with the variable(s) and data, as well as the complete SPSS file for you to use a guide in coming up with your own - so at the end you can say: THAT WAS EASY!!