**ASSIGNMENT 1**

This assignment will make use of the RAI\_short dataset, included in the “Assignments” folder. Description of the variables for the RAI and LSI instruments can also be found in the corresponding documents. Please complete each question in the assignment and include any relevant output as part of your answer. You can complete the assignment alone or in pairs (no more than 2 people per team). Please submit through Brightspace by the due date; only ONE copy per team. Note that students are entirely responsible for submitting their assignments on time and in a format that can be read by the faculty (e.g., no Google Apps “.pages” format). The acceptable formats for submissions are MSWord “.doc” or “.docx” or “rtf” if you are using other word processors.

**Question 1**

The RAI-MH dataset includes a sample of over 500 inmates recruited in the province of Ontario, Canada. The study was intended to collect a wide array of data pertaining to inmates’ background, mental and physical health, life events, criminogenic risk factors and psychosocial rehabilitative activities taken while in the institution. The inmates were recruited in 14 correctional centers throughout the province. A purposeful non-probability sampling design was used for efficiency and to insure adequate representation of specific subgroups of offenders (e.g., women and Aboriginal inmates).

All inmates in the participating institutions were considered for the study. Research assistants visited the various units within the institutions to describe the requirements associated with taking part in the study, and to answer any questions. Inmates who volunteered to participate became part of the study sample and subsequently completed a consent form. These inmates consented to an interview as well as a review of their correctional and health care files. All interviews were conducted by trained research assistants, who were all undergraduate university students at the time of the data collection.

*Based on this description, provide a brief description of three methodological problems or issues that may impact the validity of the research findings. If the information outlines above is found to be lacking, identify specific details that would need to be provided in order to make the description of the methodology more sound. You do not have to prove that these problems exist, just to point to potential areas of weakness either in the study or the description of the method. (3 points)*

**Question 2**

Looking at the variable “race” located at the beginning of the dataset, identify the number and proportion of Aboriginal inmates included in the dataset (2 points).

**Question 3**

For the same variable (race), identify the prevalence of missing data (in frequency and proportion) (4 points).

**Question 4**

For the same variable (race), determine whether the missing data is missing at random by verifying associations with sex (“Sex”, for which 1=Female and 2=Male) and the relative severity of the sentence (“remsever”). (6 points)

**Question 5**

Looking at the inmates for whom a risk assessment was completed (n=398), determine whether the variables “totscore” and “totrisk” are normally distributed and provide the most relevant descriptive statistics for each variable. (5 points)