**Module Assignment**

**Module 1**

**QMB-6304 Analytical Methods for Business**

Map

Description automatically generatedWrite a simple R script to execute the following:

**Preprocessing**

1. Load the file “6304 Module 1 Assignment Data.xlsx” into R. This file contains information on crime in each of the 67 counties in Florida. This will be your master data frame.
2. Using the numerical portion of your U number as a random number seed, take a random sample of 30 counties using the method presented in class. This will be your primary data frame.

**Analysis**

Using R calculate and report the following using your primary data frame:

1. The structure of the data frame using the str() command.
2. Mean, Median, Standard Deviation, Skewness, and Kurtosis of the Population variable. Based on these descriptive measurements how closely do you think this variable conforms to a theoretical normal distribution?
3. A boxplot of the Total Crimes variable. Based on this boxplot what can you say about the symmetry/skewness of this variable?
4. Quartiles of the Aggravated Assault variable. Show your quartiles running from the minimum to maximum values for the variable, incrementing by .20.
5. A simple histogram of the Larceny variable. Color your histogram green and give it an appropriate main title. Make sure the bottom axis of your histogram covers a range from 0 to 80,000. Based on this graphical tool would you say from this histogram the distribution of Larceny follows a symmetric distribution, or a skewed distribution?
6. Three comparative boxplots for the Robbery, Aggravated Assault, and Burglary variables. Your boxplots should be colored red and shown side by side in a single graphic with an appropriate main title and labels for the crime categories on the bottom axis. Based on these boxplots what can you say about the similarity in the number of crimes in these categories?
7. Use R to determine and report the name of the county in primary data frame with the maximum number of Total Crimes.

Your deliverable will be a single MS-Word file showing 1) the R script which executes the above instructions and 2) the results of those instructions. The first line of your script file should be a “#” comment line showing your name as it appears in Canvas. Results should be presented in the order in which they are listed here. Deliverable due time will be announced in class and on Canvas. **No collaboration of any sort is allowed on this assignment.**