**Module Assignment**

**Module 9**

**QMB-6304 Analytical Methods for Business**



Write a simple R script to execute the following data preprocessing and statistical analysis. Where required show analytical output and interpretations.

**Preprocessing**

1. Load the file “Module 10 Assignment Data.xlsx” into an R data frame. This file contains information on 437 counties in 5 Midwestern states in the US. Relevant variables for this work will be state, area, poptotal, and inmetro.
2. Using the numerical portion of your U number as a random number seed, create a new data frame create a new data frame for your work. The new data frame should have the following characteristics:
   1. It should include a new variable showing population density of each county, calculated as total population divided by county area.
   2. It should include only counties flagged as being rural, i.e., the inmetro variable of all included cases should be 0.
   3. The state variable should be a factor variable.
   4. The data frame should have 100 cases in it, these cases being a random sample of n=10 taken from each state. You will have twenty counties from Illinois, twenty from Indiana, etc., merged into a single data frame.

**Analysis**

1. Conduct a Levene test on the density variable when categorized by the state variable. Report the results of this test and give a written interpretation of the results.
2. Conduct a One Way Analysis of Variance on the density variable using the state variable as the factor. Report and interpret the results of the F test.
3. Conduct a Tukey HSD test and show the results as well as a plot of the results. Give a correct interpretation of these results, indicating any significant (or close-to-significant) factor contrasts.

Your deliverable will be a single MS-Word file created using R Markdown. Your file will show 1) the R script which executes the above instructions and 2) the results of those instructions. The first two lines of your deliverable will state this is “Assignment 10” of our course and your name as it appears in Canvas. Your code chunks and analysis results should be presented in the order in which they are listed here. Deliverable due time will be announced in class and on Canvas. This is an individual assignment to be completed before you leave the classroom. **No collaboration of any sort is allowed on this assignment.**