ECO240 R Homework 2
[Due Date: May 18th (Friday) 2018 at 13:00]

> Task1: Hypothesis Testing
> Task2: Two Variable Regression Analysis

Requirement: This is a group homework (max. four students). You must use R program to complete the required tasks.

Task 1: Hypothesis Testing
<u>NOTE: For all the questions, include your interpretation of results.</u>

Refer to: https://rstudio-pubs-static.s3.amazonaws.com/23230_b2e9b87251a2488da0fba51325e26040.html

(Refer also to: http://htmlpreview.github.io/?https://github.com/andrewpbray/oiLabs-base-R/blob/master/inf_for_numerical_data/inf_for_numerical_data.html)

Run
`load(url('http://s3.amazonaws.com/assets.datacamp.com/course/dasi/nc.Rdata'))`
to download "inference" function. (You may need to download other necessary packages to make this function work. Read the error message and download accordingly, if it doesn't work.)

I. Gifted [data = gifted.csv] (Pay attention to the sample type).
   a) Plot mother IQ and father IQ and observe the difference using BoxPlot.
   b) Derive the 95% confidence interval for the difference between mother IQ and father IQ.
   c) Test if population mean mother IQ and population mean father IQ of gifted children are equal against that they are different at 5% significance level.

II: HSB [data = hsb2.csv]
   a) Plot the exam scores using boxplot (reading, writing, math, science, social study).
   b) Generate one *Hypothesis Test of Single Population Mean* and conduct the test.
   c) Generate one *Hypothesis Test of the Difference between Two Population Means using Paired Observation* and conduct the test. (Hint: Define new variable for the difference.)
   d) Generate one *Hypothesis Test of the Difference between Two Population Means using Case 3: Independent Sample without any assumption on population variance* and conduct the test. (Hint: Select one numerical variable (y) and one categorical variable (x), then define two populations as one from A category and the other from B category.)
   e) Report the results which are unexpected and/or interesting. (If there is none, your research is boring!)

Task 2: Regression Analysis
Murder [data=murder.csv]
Refer to: http://htmlpreview.github.io/?https://github.com/andrewpbray/oiLabs-base-R/blob/master/simple_regression/simple_regression.html

a) Select one variable which can be a dependent variable of the model, and another variable which becomes an independent variable explaining the variability of the dependent variable.
b) Plot the selected variables using scatter plot.
c) Calculate the correlation coefficient.
d) Run the simple linear regression model.
e) Interpret the intercept.
f) Interpret the slope.
g) Interpret R2.
h) Add the regression line to the plot from part b.
i) Check if the residuals are nearly normal.
j) Include multiple independent variables to the linear regression model and run. Interpret the slope coefficients for all the variabl

Enjoy!