ECO 4421: Introduction to Econometric Methods
Department of Economics
Florida Atlantic University
Bonus Computer Project 1

You need do the project using R/Rstudio, compile the output into a (HTML/WORD/PDF) file. You may use "smart comments" to write your explanations. The output should contain both commands and results. Upload your output to Canvas.

The file `Earnings_and_Height.csv` contains data on earnings, height and other characteristics of a random sample of US workers. A detailed description is given in `Earnings_and_Height_Description.pdf`.

(a) Use the `read.csv` command to read the `Earnings_and_Height.csv` data set into R. Use the `attach` command to attach the data set into R.

(b) Print out an summary of the data set. In particular, find and report the sample average of the variables `earnings`, `height` and `sex`, respectively.

(c) Run a regression of `earnings` on `height`. In particular, find and use a sentence to interpret the meaning of the regression coefficient of the variables `height`.

(d) Plot a graph of `earnings` over `height`.

(e) On the graph, add a fitted line of the regression.

(f) Suppose Alex is 65 inches; Bob is 67 inches; Chris is 70 inches tall. Based on the regression, predict their corresponding earnings.

(g) Find the R^2 and SER from the regression in part (c). Use a sentence to interpret each of them.

(h) Based on the regression in part (c), find the p-value of the variables `height` and perform a t-test.

(i) Based on the regression in part (c), use the `confint` command to calculate the Confidence Interval (CI) of the variables `height`. Does your CI give you the same t-test conclusion?

(j) Run a regression of `earnings` on `sex`. For both the regression intercept and coefficient of the variables `sex`, use a sentence to interpret its meaning.