ECO476 Lab

Group Project/presentation Guidelines

The final project and group presentation for this course is a short empirical presentation and the preparation of an R-script file/R-markdown paper that answers any research questions related to a data-set you have chosen or related to any economic/financial theories you have studied in your class. You are encouraged to utilize the techniques we've covered in ECO 476 class such as data cleaning, recoding, creating new variables, the simple and multivariate regression models, creating publishable summary/regression tables using stargazer package, interpreting coefficients, visualization using ggplot2 package, checking t and F-tests, multicollinearity, and heteroscedasticity test, when addressing your question. Any other innovative techniques, output, and presentation are also welcome!

Presentation date: Aug 23, 2022.

The report submission is due at 11.59 pm on Aug 22, 2022

The Objectives

-To apply basic econometrics techniques using R discussed and covered in the ECO 476LAB class

-To show how a basic data mining, visualizations, and regression technique (in addition to simple/multivariate OLS) using a medium scale data set

-To produce a complete (a draft version at this stage) data analysis project/portfolio you may use later in your job interview.

The Group

Each group consists of at least four and at most ten students. Students must take the initiative to form their own group. The instructor will be happy to serve as a resource if you need assistance. In case a student is unable to join a group, I will assign him/her to a particular group.

How to start

Once a group is formed, the group members should come together and choose a data-set either from the GitHub page (https://github.com/masud-alam/ECO476LAB) or from any of the following open sources:

https://www.altexsoft.com/blog/datascience/best-public-machine-learning-datasets/

https://vincentarelbundock.github.io/Rdatasets/datasets.html?fbclid=IwAR3uZY3unxNlJLYjhfvv8GcwYTlL98DI0kvOZjBR2LaqA_A-sBdAt9Xko4w

https://mc.ai/the-50-best-public-datasets-for-machine-learning/

Before you start analyzing with the data set, do some reading. See how other people have addressed a particular problem in the past. Please read the related chapters and chapter 19 of the Wooldridge textbook. It has a number of useful suggestions about how to write an empirical paper. Put some thought into the model. Think about what variables should be used as outcomes (y's), what should be used as control variables (x's). Think about how the variables should be measured? Should they be logged or linear or non-linear? Just because a variable is in the data set does not mean it should be used. Think about what an appropriate control variable is and what is not.

Submission

Each group must submit an R-script/R-markdown file. **A bonus point will be awarded for R-markdown or Quarto (https://quarto.org/) submission/presentation.**

The Presentation

Each group is required to create a max.20 presentation slide for their presentation.

Presentations should be about 20-25 minutes long

The presentation slides must contain the following items:

- A list of the names of the group members and the presentation title (1 slide)
- The data set, topic or issue you have selected for your presentation (1 slide)
- A description of your presentation, i.e., what data mining and econometrics techniques you used and why did you choose these techniques.
- A potential or exactly theoretical area (microeconomics/macroeconomics/health economics/environmental economics/financial economics) where this topic would fit in. (1 slide)
- Visualizations
- Regression equation and result (must be stargazer output)
- Diagnostic tests (multicollinearity, and heteroscedasticity test)

The Grade

The project report will be counted **as 50% of your total grade.**

The Presentation will be counted **as 40% of your total grade** and will be evaluated using the following criteria:

Selection of data-set (10%) + Application of Econometric technique (30%) +Analysis and presentation of result (30%) + Overall Presentation, group agreement and Q&A (20%) + Q&A (10%).

Each member will be awarded an individual grade on the project; there is no "group grade." The group must submit **a credits page** with the brief contribution that lists the specific tasks, roles, responsibilities, and work products contributed by each group member. The entire group should be involved in writing the brief contribution/credits page. The assigned grade can be different for each individual depending on the quantity and quality of their contribution to the group effort and performance during the presentation (i.e., Q&A session). **Good Luck!**