UCLA Extension Data Science Summer Course

Instructor: William Yu

Term Project

- The goal of the term project is that you will use the skills and models you learned in this course and apply them onto a real-world dataset. Do the data visualization and analytics and tell a coherent and good story from your results.
- In the project, it will be the best that you can demonstrate your **problem-solving** skills or strategies. Problem-solving could be in the following fields:
  - Finding and collecting a good-quality, meaningful, relevant, or interesting dataset that are related to China's businesses, economy, or society.
  - National Bureau of Statistics of China:
  - http://www.stats.gov.cn/english/
  - Or you can analyze the international dataset, including China:
  - Your dataset might be good but contain a lot of missing values, noises, etc. You can demonstrate your effort and skills on how to manage a messy dataset and turn it into a well-organized one.
  - Demonstrate your knowledge and skills on machine-learning models: such as a linear model and a logistic model.
  - Find a problem that has yet to be discussed or analyzed.
  - Find a practical solution to the problem you are facing in your current/past work projects.
- Prepare your term paper around **10 to 15** pages, including all the charts.
- In addition to submit your term paper, you also need to submit your R script/markdown and dataset.
- Due Day: August 9 at 11:59pm.