**ECO6763 - Group Project 1**

**Instruction:**

In this assignment, we will work with your teammate to practice an end-2-end regression modeling. You can choose any approaches to achieve the best prediction accuracy on a continuous variable.

Three sets of data from Kaggle website are provided for your choice. But you can choose other data source.

**Points:** 20 out of 100 of final grade

**Delivery Due Date:**

1) Project code and slides due on Nov 17 for my comments.

2) Oral presentation will happen on regular class time on Nov 18 during regular class session in person

3) The revised code and slides should be re-submitted to be graded by the end of Nov 19.

**Grading Scheme:**

1) submit your project code and presentation slides on Canvas (50% or 10 points)

2) prepare an oral presentation before the class by all teammates (25% or 5 points)

3) Q&A (25% or 5 points)

**Data Source in Kaggle:**

\*\*\* Attention\*\*\* (A Kaggle account is needed to access the data)

* [CalCOFI: Over 60 years of oceanographic data](https://www.kaggle.com/sohier/calcofi): Is there a relationship between water salinity & water temperature? Can you predict the water temperature based on salinity?
* [Weather in Szeged 2006-2016](https://www.kaggle.com/budincsevity/szeged-weather): Is there a relationship between humidity and temperature? What about between humidity and apparent temperature? Can you predict the apparent temperature given the humidity?
* [Weather Conditions in World War Two](https://www.kaggle.com/smid80/weatherww2/data): Is there a relationship between the daily minimum and maximum temperature? Can you predict the maximum temperature given the minimum temperature?