Dear Students,

We learned how ordinary least squares (OLS) method models the linear relationship, which is known as linear regression, between dependent variable and predictor variables. This assignment helps you remember the key steps in OLS implementation using Jupyter Notebook and interpretation of the outcomes. We will use Real estate valuation data set, which is accessible through [1], in this assignment. We intended to examine the role of the following features in the house price of unit area in this data set:

X2=the house age (unit: year)

X3=the distance to the nearest MRT station (unit: meter)

X4=the number of convenience stores in the living circle on foot (integer)

X5=the geographic coordinate, latitude. (unit: degree)

X6=the geographic coordinate, longitude. (unit: degree)

In this assignment, please create a Jupyter Notebook with the name of "Assignment01_YourStudentID" and follow the steps below:

1. Create a cell of code and paste the code provided in "Example 2: Applying OLS regression on a dataset from UCI repository" available in OLS_Example.ipynb, which you have access OLS folder at Lectures Learning module in Blackboard learning management system.
2. Use pd.read_excel() instead of pd.read_csv() to read the excel file of "Real estate valuation data set.xlsx" accessible through link [2]. If you faced with the requirements for installation of a library, please remember you can install the requirements by pip instruction.
3. Since we no longer work in heart data set, please change the name of the pandas.core.frame.DataFrame from Heart to House. Apply this change in all the other instructions, which you use data frame to work on.
4. Select 'Y house price of unit area' as your dependent variable via
    Y = house['Y house price of unit area']
5. Select 'X2 house age' as your independent variable via
    X = house['X2 house age']
6. Run the code and get coefficients, mean squared error and coefficient of determination.
7. Add a Markdown cell and describe what you learn from these outcomes.

References:

[1]. https://archive.ics.uci.edu/ml/datasets/Real+estate+valuation+data+set

[2]. https://archive.ics.uci.edu/ml/machine-learning-databases/00477/