DATA-445 IN-CLASS ASSIGNMENT AUGUST 30, 2021

Consider the `Real_Estate.csv` data file posted on Blackboard (under the In-Class 3f assignment link). This file contains information related to 414 houses. The goal is to predict the house price of unit area. In **Python**, answer the following:

1. (3 points) Using the pandas library, read the csv data file and create a data-frame called `house_price`.
2. (3 points) Drop the `No` and `transaction_date` columns.
3. (4 points) Split the data-frame into two data-frames: train (80%) and test (20%).
4. (5 points) Using the train data-frame, build a linear regression model in which: `house_age`, `distance_to_the_nearest_MRT_ station`, `number_of_convenience_stores`, `latitude`, and `longitude` are the input variables, and `house_price_of_unit_area` is the target variable.
5. (5 points) Using the model from part 4, predict the house price of unit area of a the houses in the test data-frame. Report the MSE.