DATA-445 IN-CLASS ASSIGNMENT SEPTEMBER 14, 2022

Consider the `CarPrice_Assignment.csv` data file (under the In-Class 6 assignment link). This data is public available on the Kaggle website, and has information on cars (characteristics related to car dimensions, engine and more). The goal is to use car information to predict the price of the car. **In Python**, answer the following:

1. (5 points) Load the data file to you S3 bucket. Using the pandas library, read the csv data file and create a data-frame called `car_price`.

2. (7 points) Using `wheelbase`, `horsepower`, `peakrpm`, `citympg`, and `highwaympg` as the predictor variables, and `price` is the target variable build a linear regression model. Perform a leave-one-out cross validation. Report the MSE of this model.

3. (7 points) Using `carwidth`, `carheigh`, `horsepower`, `citympg`, and `highwaympg` as the predictor variables, and `price` is the target variable build a linear regression model. Perform a leave-one-out cross validation. Report the MSE of this model.

4. (3 points) Using the results from parts (2) and (3), what model would you use to predict car prices? Explain.