CSCI 183 Homework 2
Due Date: Feb 12, 2024

Regression for House Price Prediction

What do you need to do?

1. Load the dataset into a pandas dataframe and find the data types for each column in the dataset.
2. Find the names of the columns of this dataframe.
3. Find how many numerical features exist in the dataset.
4. Find the correlation matrix for this dataset. Report which features tend to have a high correlation with the target variable. (You can use the corr() function). Refer to supplementary slide 'Correlation'.
5. Create and compile as many graphs (feature vs target variable) as you can using the matplotlib library [https://matplotlib.org/gallery/index.html] for the given dataset. Select only numerical features.
6. Based on the graphs in step 5, identify features that have a linear relationship with the target variable.
7. Selecting different features from step 6, implement a linear regression algorithm and **find the slope, the intercept and the error of the regression model.**
8. Display the line of best fit from step 7.
9. Some options you can consider:
    a. linregress() from scipy.stats
    b. LinearRegression() from sklearn
    c. Manually code the gradient descent algorithm
10. Create a table similar to one given below for all the features selected in step 7.
11. Also attach images of the graphs to your report.

Observation Table:

Feature	Slope	Intercept
:		
:		

Submit your code as an .ipynb file and a document reporting your findings. You could also show them as output of your code.