ME365D - Task 3

Dyab Asdi

The first thing I did was drop duplicate data in the data frame, and then for certain categories that had strings for data, I converted them to numbers for grouping using Label Encoding.

Next, I ran an optimization sweep for different parameters for the Decision Tree Regressor as well as the train and test split:

```
leaf_nodes_sweep = np.arange(650, 800, 5)
trainsplit_sweep = np.linspace(0.50, 0.95, 10)
model = optimization_sweep(df, leaf_nodes_sweep, trainsplit_sweep)
```

From this, whichever combination of these led to the smallest mean absolute error was printed:

```
Optimal Model:
   Mean Absolute Error = 20.91%
   Training Split = 80% Train, 20% Test
   Leaf Nodes = 795
```

I was only able to get as low as 20.91% for the MAE using this strategy.