Multiple Linear Regression in Python

In Python, the .fit function requires a **list** of all independent variables. By adding multiple variables to the list, we now are doing multiple linear regression!

When we used the carat weight of the diamond to predict the price, our list of independent variables was simply `["carat"]`:

Python:	`model = model.fit( df[ ["carat"] ], df["price"] )`
Description:	Fits the linear regression model with the data in the DataFrame df.

Suppose we want to use **both** the weight and the table (the large, flat surface you see when looking at a diamond inset into a ring) to predict the price. The list of our independent variable column names is now: `["carat", "table"]`.

In our LinearRegression model:

Python:	`model = model.fit( df[ ['carat', 'table'] ], df['price'] )`
Description:	Fits the linear regression model with the independent variables "carat" and "table", used to predict the variable "price".

In this new model, we now must always have both a carat weight and a table size for a diamond to use this model for prediction. However, we expect a better prediction! *...let's discover if the addition of the table resulted in a smaller error when applying our dataset to the original dataset.*

Puzzle #1: First, we need to find the prediction using **simple linear regression**:

Python:	
Description:	Computes and adds a new column to the DataFrame for the error in a linear regression model predicting the price of a diamond **based only on the carat weight**.

Puzzle #2: Next, we need to find the prediction using **multiple linear regression**:

Python:	
Description:	Computes and adds a new column to the DataFrame for the error in a linear regression model predicting the price of a diamond based on **both the carat weight and the table size**.

Puzzle #3: Compute the **absolute value of the errors** for both predictions:

Python:	
Description:	Computes and adds a new column to the DataFrame for the error in a linear regression model predicting the price of a diamond based on **both the carat weight and the table size**.

Puzzle #4: Finally, find the average absolute error for each model:

Python:		
Results:	Simple Linear Regression: *(Using only carat weight)*	Multiple Linear Regression: *(Using carat weight and table size)*