

Project: Diamond Prices

Step 1: Understanding the Model

1. According to the model, if a diamond is 1 carat heavier than another with the same cut, how much more should I expect to pay? Why?

A diamond that is 1 carat heavier than another with the same cut will cost \$8413 more. From the linear regression model provided, the coefficient for the carat variable is 8413.

2. If you were interested in a 1.5 carat diamond with a **Very Good** cut (represented by a 3 in the model) and a **VS2** clarity rating (represented by a 5 in the model), how much would the model predict you should pay for it?

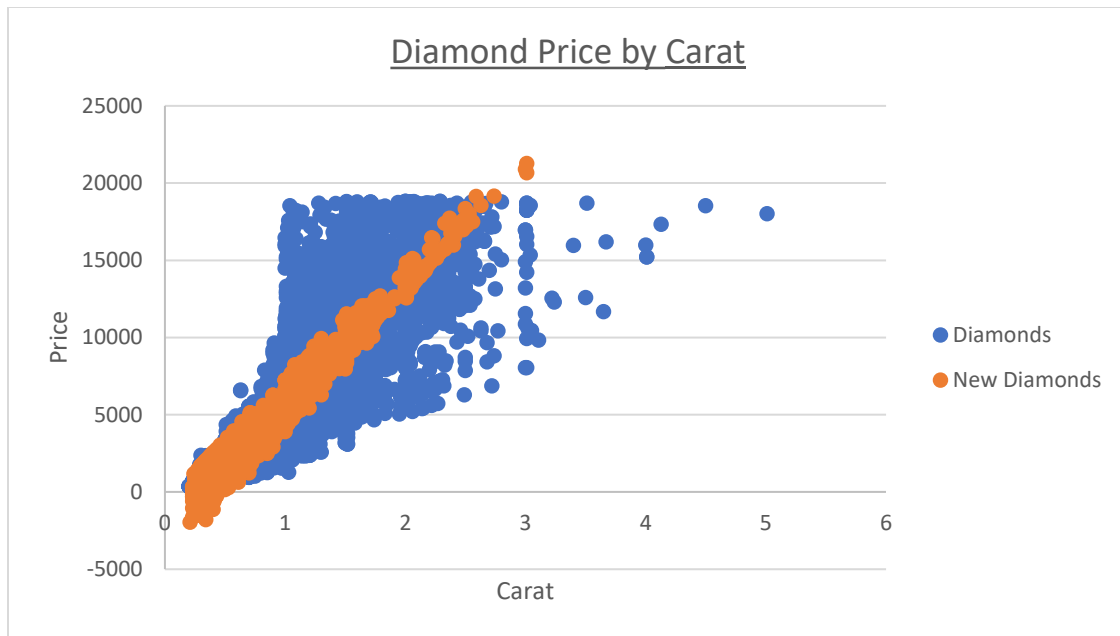
Using the linear regression model provided and adding the values given above results in the below equation.

$$= -5269 + 8413 * 1.5 + 158.1 * 3 + 454 * 5$$

This equation gives a result predicting \$10094.80 should be paid.

Step 2: Visualize the Data

1. Plot 1 - Plot the data for the diamonds in the database, with carat on the x-axis and price on the y-axis.
2. Plot 2 - Plot the data for the diamonds for which you are predicting prices with carat on the x-axis and predicted price on the y-axis.
 - **Note:** You can also plot both sets of data on the same chart in different colors.
3. What strikes you about this comparison? After seeing this plot, do you feel confident in the model's ability to predict prices?



The New Diamonds prices are more linear than the Diamonds prices which in turn show more outliers.

There are a wide variety of prices displayed by the Diamonds data which suggest that there are other variables to consider. Other variables may include the diamond market at certain times during the year and inflation.

From the graph there are also negative values which cannot be correct this could lead to the conclusion that the linear regression model does not provide the most suitable equation for this task. However the equation does result in values which display an average through the data points.

Step 3: Make a Recommendation

1. What price do you recommend the jewelry company to bid? Please explain how you arrived at that number.

The jewelry company should bid a price of \$8213465.93

I used the linear regression model equation provided to establish the predicted prices of the new diamonds and used the SUM function to provide a total of the predicted prices. This resulted in a price of \$11733522.76. The company purchases diamonds from distributors at 70% of this price, this results in a price of \$8213465.93.