

Project: Diamond Prices

Step 1: Understanding the Model

Answer the following questions:

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?
 - Answer: A diamond that is one carat heavier than another with same cut would result in \$8,413 in price. The formula created by the regression, determines that the coefficient for a carat is 8,413. So, for every increase in carat, the price also increases by the amount of the coefficient.
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?

Answer: The regression formula is:

$$\text{Price} = -5,269 + 8,413 * \text{Carat} + 158.1 * \text{Cut} + 454 * \text{Clarity}$$

- To calculate how much the model predicts plug in the values of the variables.
- $\text{Price} = -5269 + 8,413 * 1.5 + 158.1 * 3 + 454 * 5$
- $\text{Price} = 10,094.8$

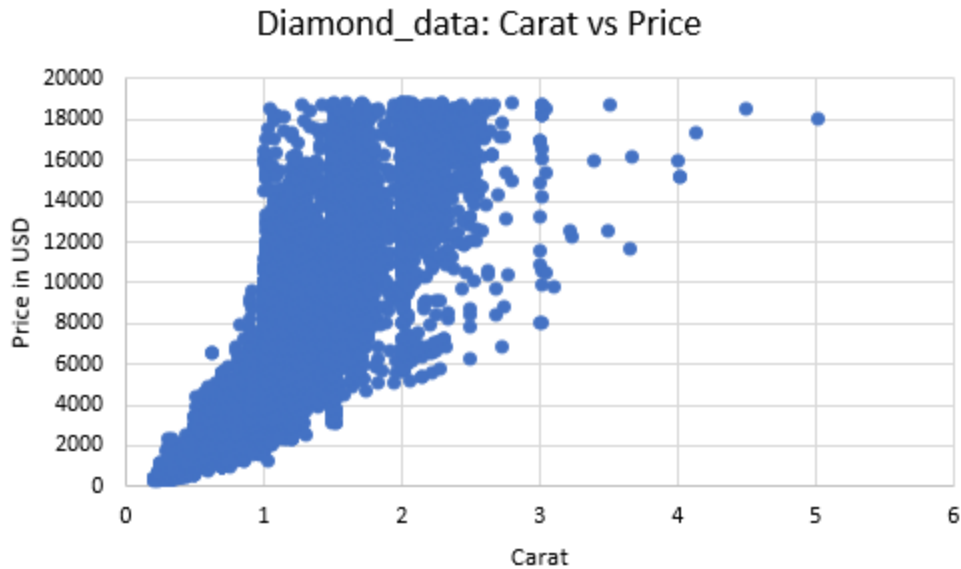
The model will predict a price of \$10,094.8

Step 2: Visualize the Data

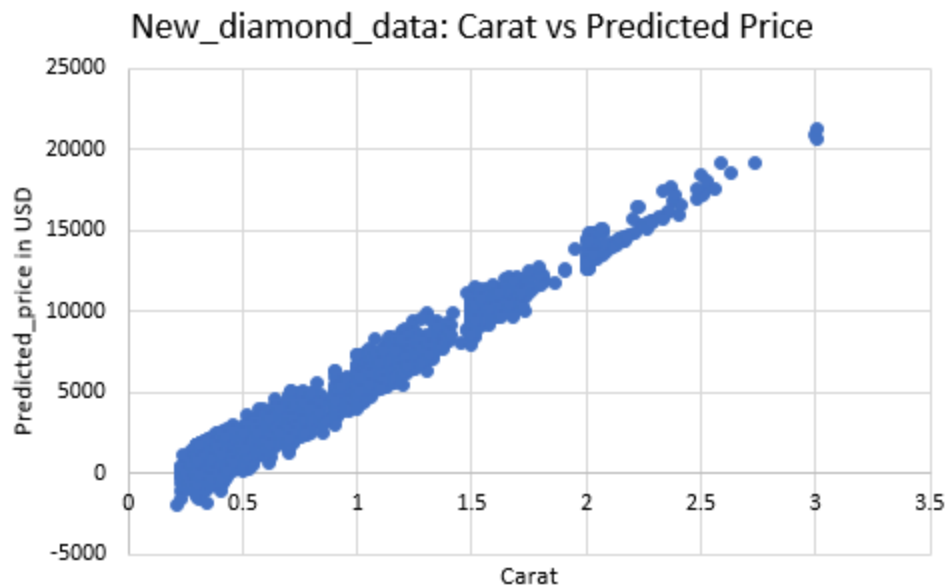
Make sure to plot and include the visualizations in this report. For example, you can create graphs in Excel and copy and paste the graphs into this Word document.

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?

Plot 1:



Plot 2:



From the two plots above we can see that the predicted prices of the new diamond data seem to have a strong positive linear relationship with the carat of each diamond as compared to the prices of the diamond data with each diamond carat. Though the regression formula wasn't based only on the carat, the coefficients of the cut_order and clarity_order was used to generate the model formula, so we cannot clearly prove that only carat of the diamond predicts the price.

The model seems to predict the prices of the new diamonds quite well, as we can see a clear relationship in the plots above. On average, majority of the new diamond predicted prices compact

in the same area having similar carat sizes while there are three diamonds with high carat and high predicted price. With this, the formula should be able to offer the company a reasonable bidding price for the new diamonds.

Step 3: Make a Recommendation

Answer the following questions:

1. What price do you recommend the jewelry company to bid? Please explain how you arrived at that number.
 - I recommend a bid of \$8,213,465.932. I arrived at this number by using a formula from the regression model provided that was based on previous diamond data prices and applied it to the new diamonds that were up for bidding. I then factored in that company purchases diamonds from distributors at 70% of the price so I multiplied the predicted price amount 11733522.76 by .70 to get the final predicted bid of \$8,213,465.932.