

Project: Diamond Prices

Complete each section. When you are ready, save your file as a PDF document and submit it here: <https://classroom.udacity.com/nanodegrees/nd008/parts/235a5408-0604-4871-8433-a6d670e37bbf/project#>

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: The formula worked out from regression model is as follows:

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

If the second diamond is just 1carat heavier than the first diamond (provided Cut and clarity is same in both) then we expect the price of second diamond to be \$8,413 more. This is because it is the coefficient of carat. Let carat value of first diamond be represented by 'c'.

$$\text{Price of 1}^{\text{st}} \text{ diamond} = P_1 = \text{constant} + 8413 \times c$$

$$\text{Price of 2}^{\text{nd}} \text{ diamond} = P_2 = \text{constant} + 8413 \times (c+1)$$

$$\text{Hence, difference} = P_2 - P_1 = \$8,413$$

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?

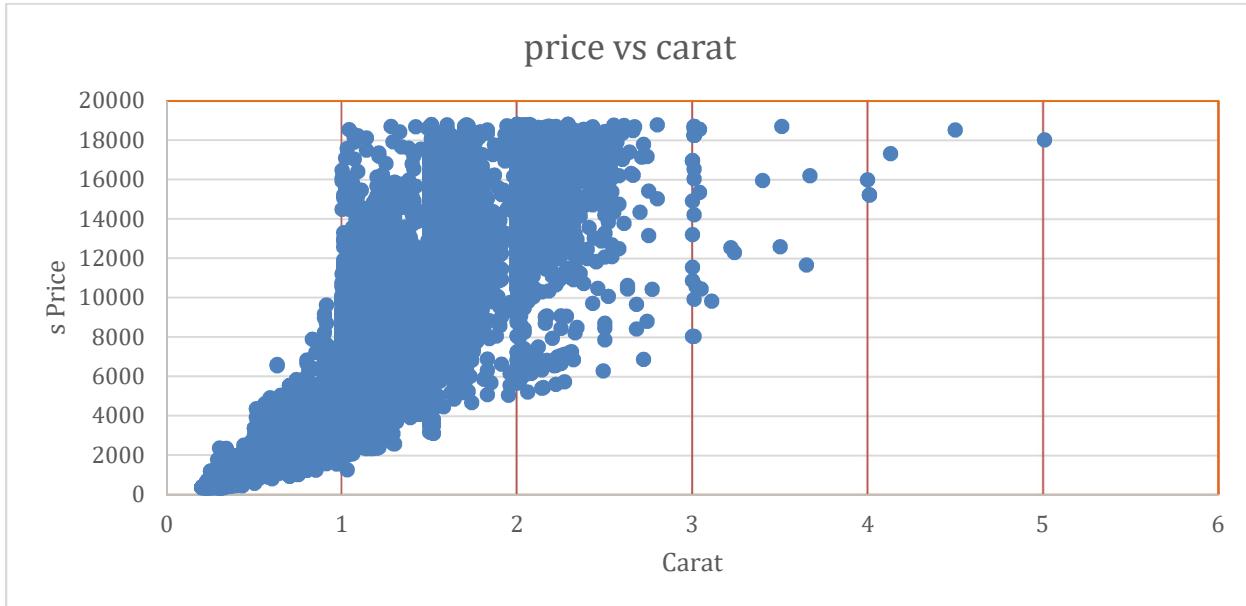
Answer: Substituting the values we have,

$$\begin{aligned}\text{Price} &= -5,269 + 8,413 \times 1.5 + 158.1 \times 3 + 454 \times 5 \\ &= \$10,094.80\end{aligned}$$

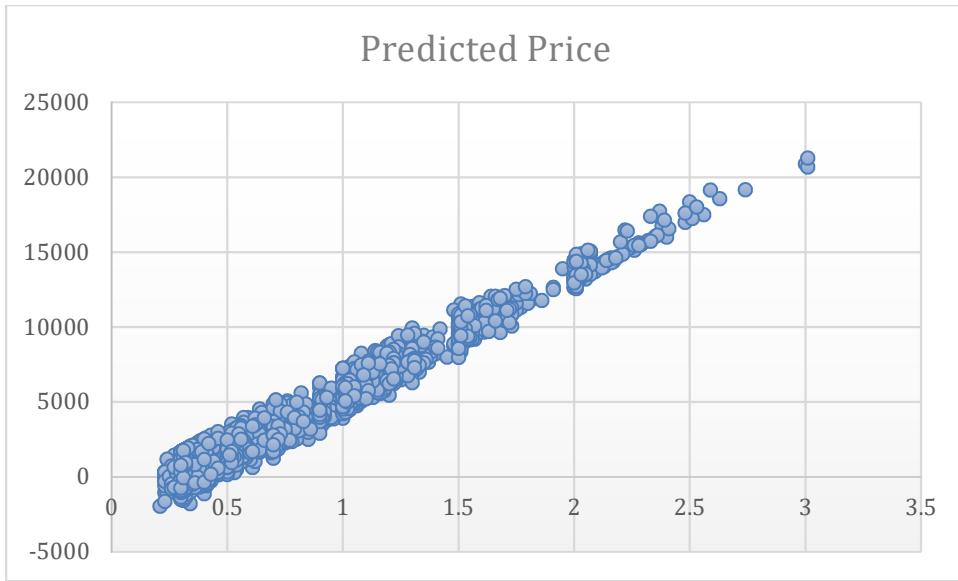
Step 2: Visualize the Data

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.
 - o **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?

Answer:

The predicted prices are more compact than the actual data is. This is because we are not accounting for everything that effects prices. There are many more things than carat that effect it. We had carat, clarity, and cut factored in to our formula but not even that will account for all the variation. For instance, this formula might look very different depending on the size of the diamond.

After looking at this plot the model appears on average to predict the prices ok, but it can be very off for certain diamonds. There appears to be an outlier price for some diamonds with negative value. The value of actual diamond cannot be negative. While the formula may not be accurate for an individual diamond, it should do a decent job at predicting the price we should pay for several diamond at once (the bid price) since it on average looks representative

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.

Answer:

I would recommend a bid price of \$8,213,466. This is 70% (required by our customer) of the total price that we expect retail customers to pay (\$11,733,523). We arrived at 11,733,523 by summing up the predicted price of each and every diamond using the formula.