

## 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:*

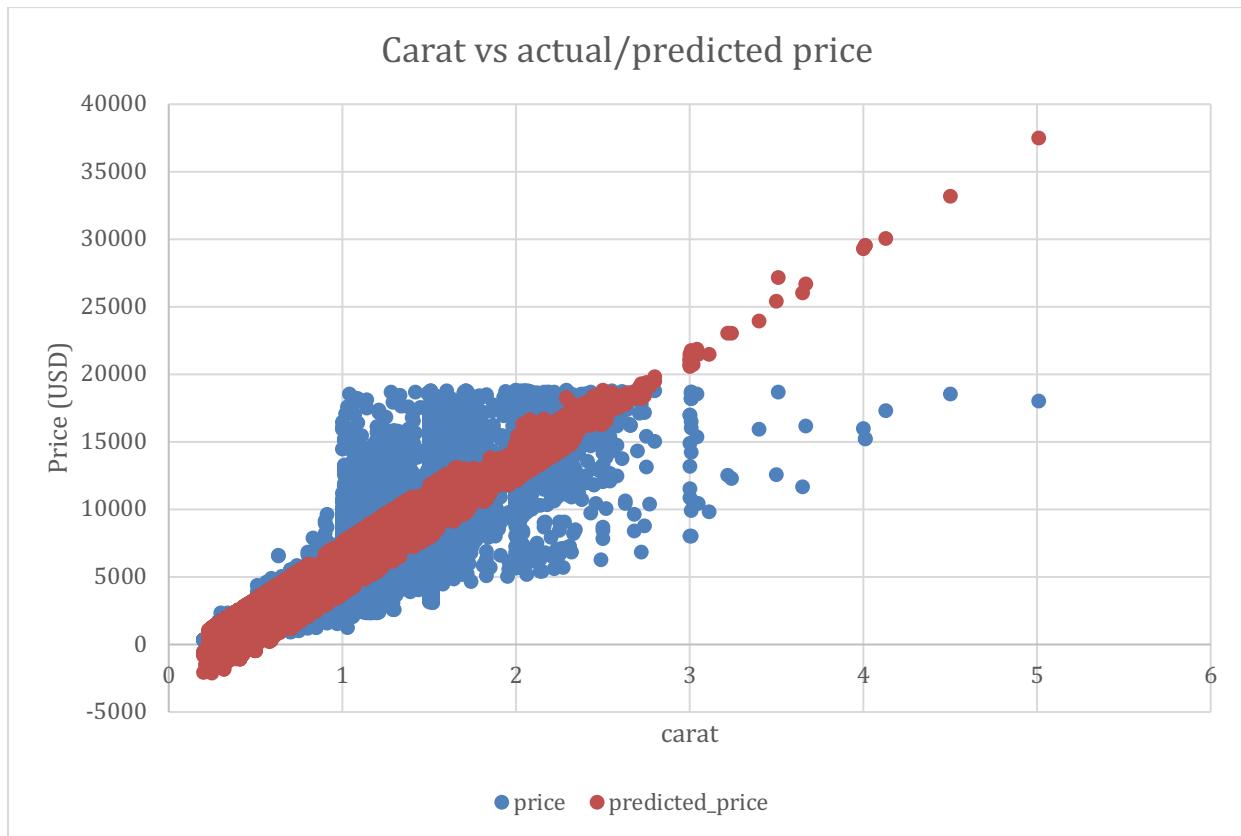
1. According to the model, if a diamond is 1 carat heavier than another with the same cut, how much more should I expect to pay? Why?  
According to the model, 1 carat heavier with the same cut will **cost 8,413 dollars more** as the coefficient for the carat in the multiple linear regression is positive 8,413 – 1 unit carat higher will lead to price increment of 8,413 dollars.
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?  
The formula is Price =  $-5,269 + 8413 \times \text{Carat} + 158.1 \times \text{Cut} + 454 \times \text{Clarity}$ .
  - With 1.5 carat, Very Good cut, and VS2 clarity rating,
  - Price =  $-5,269 + 8413 * 1.5 + 158.1 * 3 + 454 * 5$
  - I will have to pay for about **10094.80** dollars.

## 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?



The predicted prices are less widely spread compared to the actual prices. This may be because we have not account for all the possible variables that may affect the pricing, including the color of the diamond available in this dataset.

By looking at this plot, the model appears to average the prices, with many far off from actual prices. While the formula may not be accurate for all, it should do a decent job at predicting the price the jewelry company how much to bid since on average the predicted prices look 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.

Based on the model provided by the study designer, I recommend a bid of \$ **8,213,465.93**. I arrived at this number by using a formula provided by the regression model that was based on the previous diamond sales and applied the formula to the new diamonds that the jewelry company is interested. I then factored in the distribution price with 70% of the sale price, by multiplying the total predicted amount of \$ 11,733,522.76 by 0.70 to get the final predicted bid of \$ 8,213,465.93.