

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?

If a diamond is 1 carat heavier than another with the same cut, we should expect to pay an additional \$8,413. The linear regression model given shows that the coefficient for carat is 8,413, which means that for every increase in carat the price will increase 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?

The linear regression model to predict diamond prices for the set of 3,000 diamonds gives the following equation:

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

Substituting the value for each variable gives us:

$$\text{Price} = -5,269 + 8,413 * 1.5 + 158.1 * 3 + 454 * 5$$

$$\text{Price} = 10,094.8$$

Therefore, the model would predict that we should pay \$10,094.8 for it.

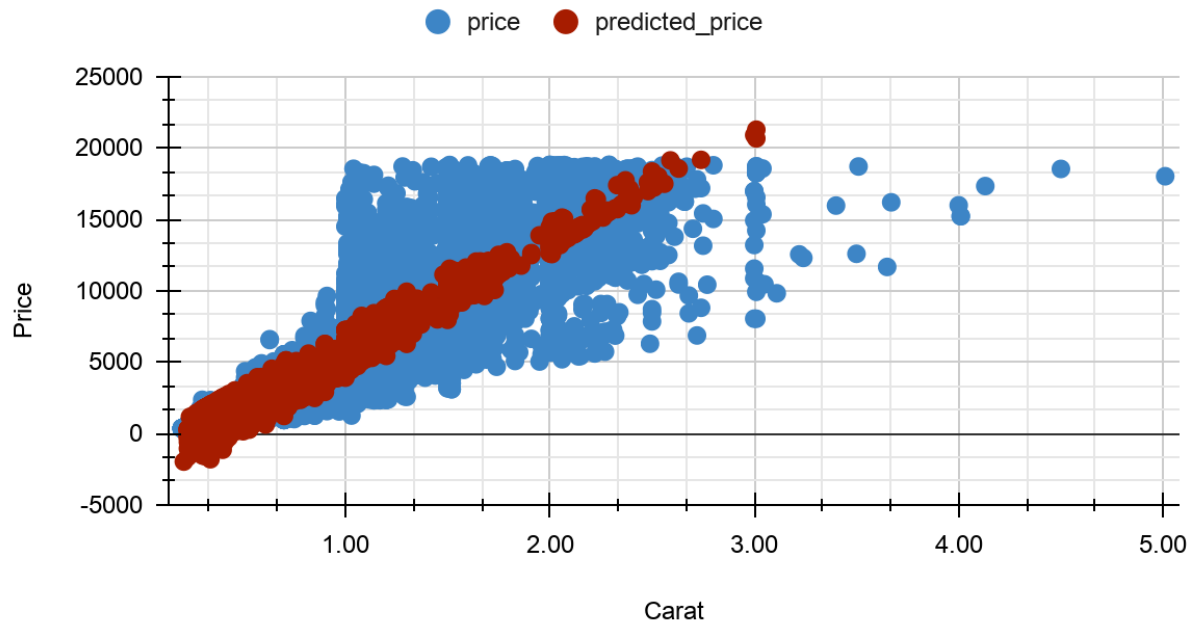
Step 2: Visualize the Data

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?

Carat vs Price



The predicted prices are more compact than the actual prices as we are not taking account of each and every factor which affects prices. Although cut, color and clarity are factored into our calculation, factors such as fluorescence and shape are also the major determinants of the price of diamonds.

On average, the model appears to predict the prices fairly well. However, it can be inaccurate for certain carats, for example, there is a diamond of 5.01 carats with an actual price of \$18,018 and another of 1.07 carats with an actual price of \$18,114. Additionally, the model predicts that some diamonds have negative prices. While the formula might not be accurate for the individual diamond, it can still predict the total price we should pay for a set of diamonds as the model is quite 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.

I recommend the jewelry company to bid \$8,213,465.93 for the entire set of 3,000 diamonds. I arrived at this number by totalling up the predicted prices from the linear regression model provided, which is based on a large database of diamond prices, and then factoring in the margin the company is looking for, which is 30% (as the company generally purchases diamonds from distributors at 70% of the final retail price a consumer will pay).

Bid price = total predicted price * 0.7

Bid price = 11,733,522.76 * 0.7

Bid price = 8,213,465.93