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: Before answering the question above, it is critical to clarify that the below equation is given:

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

   The above equation means that you would pay 8413 more for each increment in diamonds weight in carat. Since all other predictors are constant.

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: As per the equation below:

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

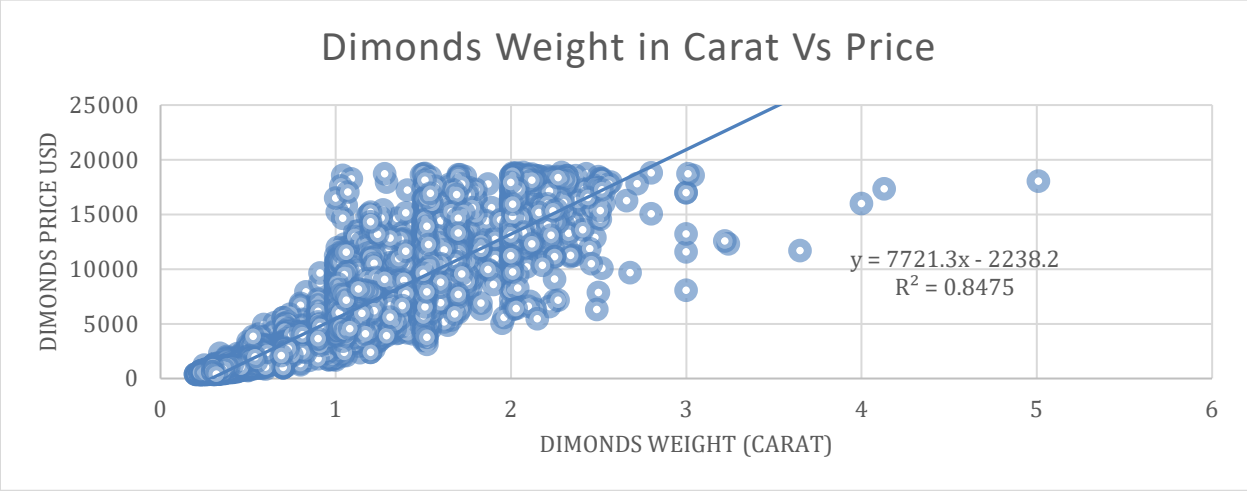
   Then, $\boldsymbol{Price = -5,269 + 8,413\, x\, 1.5 + 158.1\, x\, 3 + 454\, x\, 5}$

   the predicted diamond price would be: 10,094.8 USD

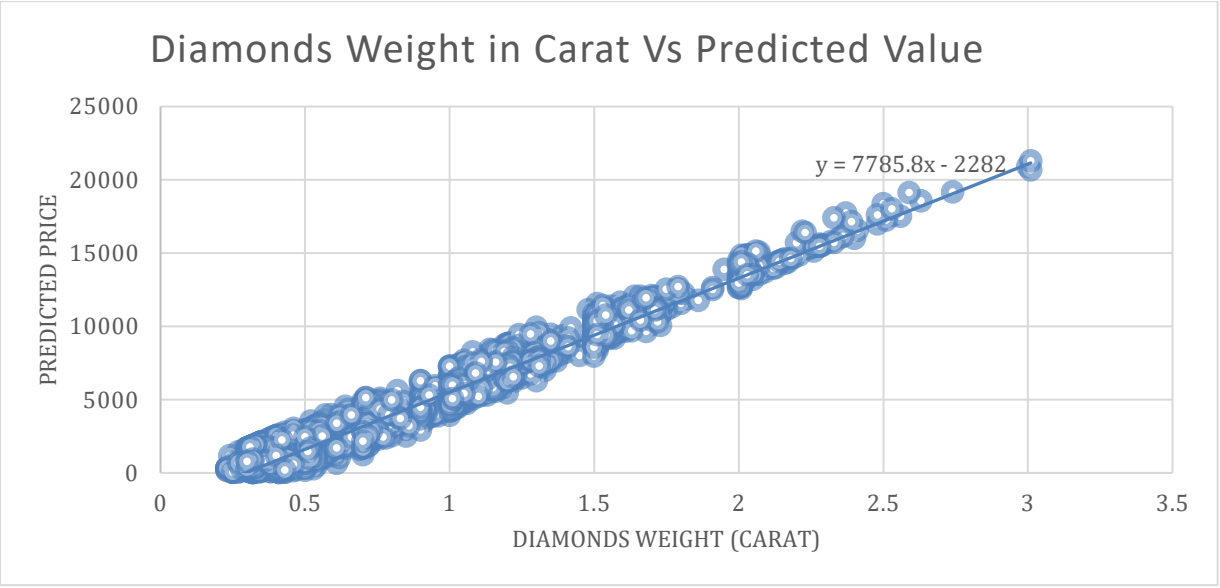
Step 2: Visualize the Data

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?

   Answer: The second plot shows more linear relationship than the first one. It also shows some data that would be classified as not realistic such as having diamonds with zero price.

   However, by looking at both equations, shown on the plots, it can be said that the built model is a good representative. Both equations are almost similar with minor differences. Moreover, the trendline of the first plot has a high R2 value (higher than 1). Adding this to the fact that both trend lines have similar equations gives even more confidence in the predictive model built.

   If R2 value of the first line was low, then even with having the same equations for both lines, the predictive model would be bad. But as mentioned earlier, the model has a good capability to predict the price.

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:

After applying the aforementioned equation over the 3000 diamonds, and summed all prices, the total value was found to 11,733,523 $. I recommend that the company shall not bid with a value higher than 8,213,466 $, which is 70% of the total value.

I arrived at the number by summing all predicted prices, for each price, I substituted the carat with its respective value under its column. Regarding clarity and cut, I substituted each of them with its respective order as shown in question 2 in step 1.