

Project: Diamond Prices

Project Overview

A jewelry company wants to put in a bid to purchase a large set of diamonds, but is unsure how much it should bid. In this project, you will use the results from a predictive model to make a recommendation on how much the jewelry company should bid for the diamonds.

Step 1: Understanding the Model

Answer the following questions:

1. According to the linear model provided, if a diamond is 1 carat heavier than another with the same cut and clarity, how much more should we expect to pay? Why?

According to the coefficient of the explanatory variable “carat”, a diamond with 1 carat heavier than another with same cut and clarity would have \$8,413 higher value. For every increase in “carat”, the price will increase by the amount of the coefficient which is 8413.

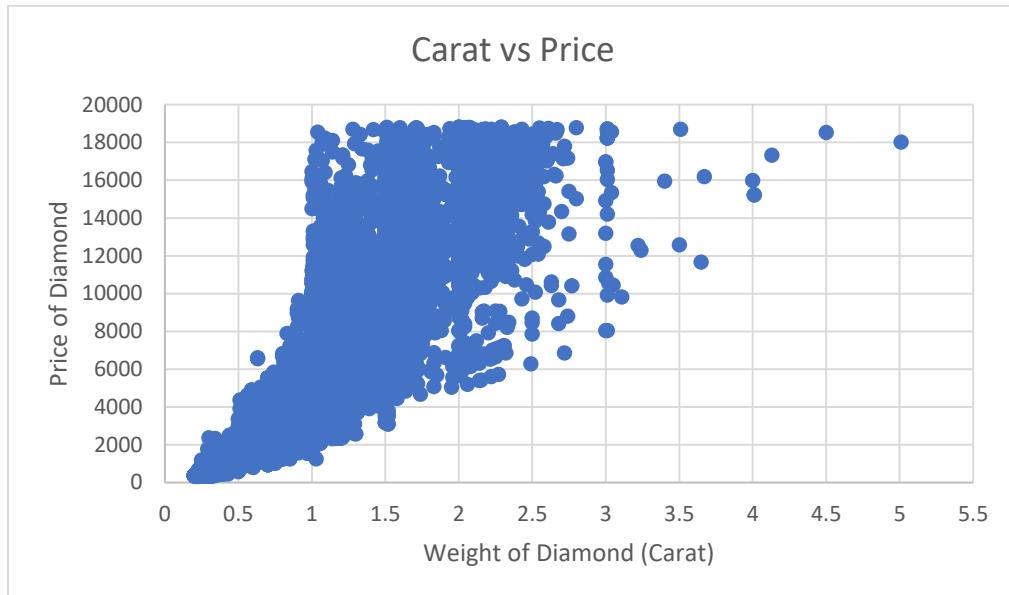
2. If you were interested in a 1.5 carat diamond with a **Very Good** cut (represented by a 3 in the model) and a **VS2** clarity rating (represented by a 5 in the model), how much would the model predict you should pay for it?

*Plugging the values into the provided formula, $\text{Price} = -5,269 + 8,413 \times \text{Carat} + 158.1 \times \text{Cut} + 454 \times \text{Clarity}$, we would expect the price of the diamond to be around \$10,094.8. This expected price is calculated by $-5269+(8413*1.5)+158.1*(3)+454*(5)$*

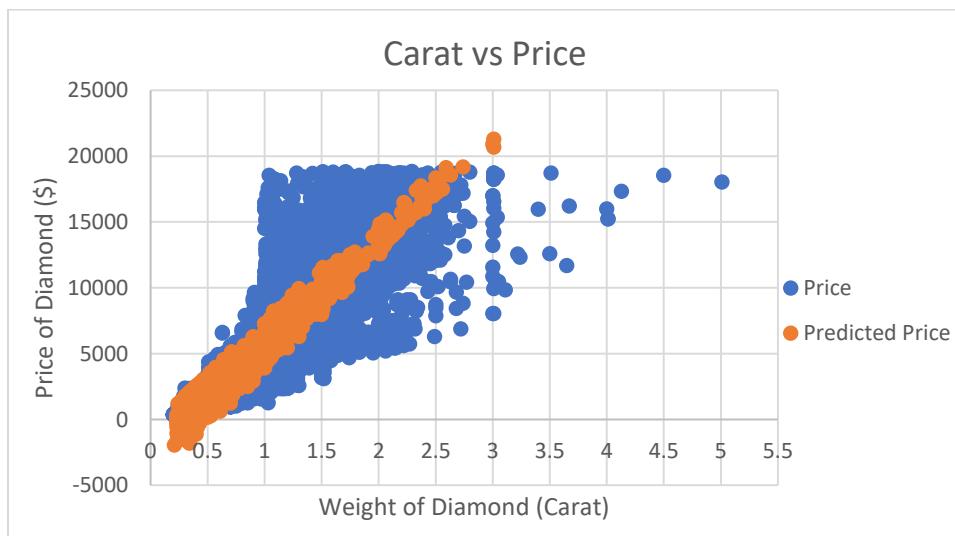
Step 2: Visualize the Data

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?

One of the noticeable point is that some diamonds have negative expected price, which is usually impossible. This happened because variable "Carat" has the strongest power in explaining prices of diamonds (coefficient value highest). This means that any diamonds with size of a Carat 0.62 (=5269/8413) or below have high possibilities to have negative expected prices.

Also, the predicted prices of diamonds are more compact than the actual data is. This is so because there are other factors besides the weight, Quality, and purity of diamond that contributes to the price of a diamond. For example, the prices would look much different from now if other explanatory variables such as "shape" and "Fluorescence" were added to the model.

Besides the fact that some diamonds have negative expected price, the model in general does a satisfactory job in predicting the prices of diamonds.

Step 3: Make a Recommendation

Answer the following questions:

What price do you recommend the jewelry company to bid? Please explain how you arrived at that number. HINT: The number should be 7 digits.

Assuming the model is valid for predicting prices of diamonds, I would suggest a bid of \$8,213,465.93. This bid was calculated by using a formula from the regression model provided that was based on previous diamond prices and applied it to the diamonds that were up for the auction. The total sum of the prices of the diamonds came out to be \$11,733,522.76 (including negative expected diamond prices). Then I factored in the circumstance that the company would like to purchase diamonds at 70% of the expected price, so I multiplied the predicted total amount of 11733522.76 by .70 to get the final predicted bid of \$8,213,465.93. If we assume that all the negative expected prices of diamonds as \$0 value, the final bid would be \$8,309,267.72, using the same logic explained above.