Project: Diamond Prices

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 and same clarity, the one additional carat result in an additional $8413 in price. The formula created by the regression determined that the coefficient for a carat is 8,413, so for every increase in the number of carats 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

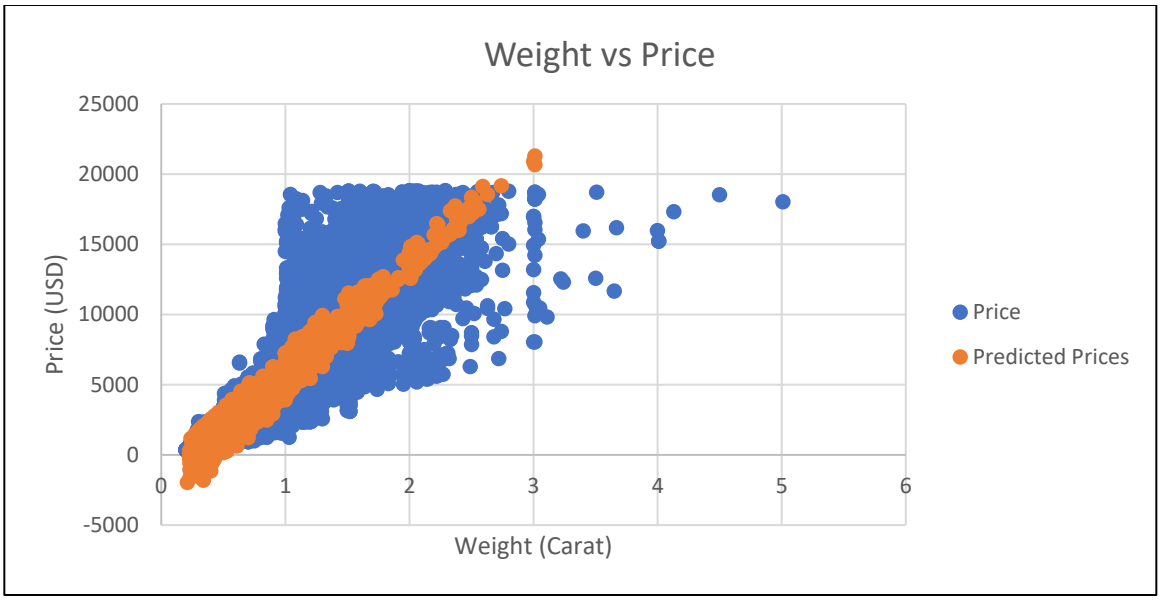
The Formula is Price = -5,269 + (8,413 x Carat) + (158.1 x Cut Ord) + (454 x Clarity Ord)

I. so now we can plug in the values for the different variables.
II. Price = -5,269 + (8,413*1.5) + (158.1*3) + (454*5)
III. Price = -5,269 + 12,619.5 + 474.3 + 2,270
IV. Price = 10,094.8

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?



- The predicted prices are more compact than the actual data is. This is because we are not accounting for everything that affects prices. There are many more things than Carat that affect it. We had Clarity and Cut factored into our formula but not even that will account for all the variation. For instance, this formula might look very different depending on the color you are training the model on.

- 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 Diamonds with only 1 Carat but sold for almost $18,000. We can see a negative value in the graph, which the model did not deal with these situations. 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 diamonds at once since 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.

    - As we refer to the question of how much the company should bid for the entire set of 3,000 diamonds, I recommend a bit of $8,213,465.93. I arrived at this number by using a formula from the regression model provided that was based on previous diamonds sales and applied it to the diamonds that were up for bid. I then factored in the margin of purchases diamonds form distributors at 70% of that price. so, I multiply the predicted amount $11,733,522.76 by 0.70 to get the final predicted bid of $8,213,465.93.