Project: Diamond Prices

Step 1: Understanding the Model

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?

Price = -5,269 + 8,413 x **Carat** + 158.1 x **Cut** + 454 x **Clarity**

= **-5,269 + 8,413 x** 7 **+ 158.1 x** 5 **+ 454 x 5**
= $56,682.5

= **-5,269 + 8,413 x** 6 **+ 158.1 x** 5 **+ 454 x** 5
= $48,269.5

=$56,682.5 - $48,269.5= $8,413

It can be concluded that the coefficient value for carat is 8,413 which means that each carat is heavier than the other one that is with the same cut and clarity which leads to increasing in the price that is equal to $8,413.

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?

Price = -5,269 + 8,413 x **Carat** + 158.1 x **Cut** + 454 x **Clarity**

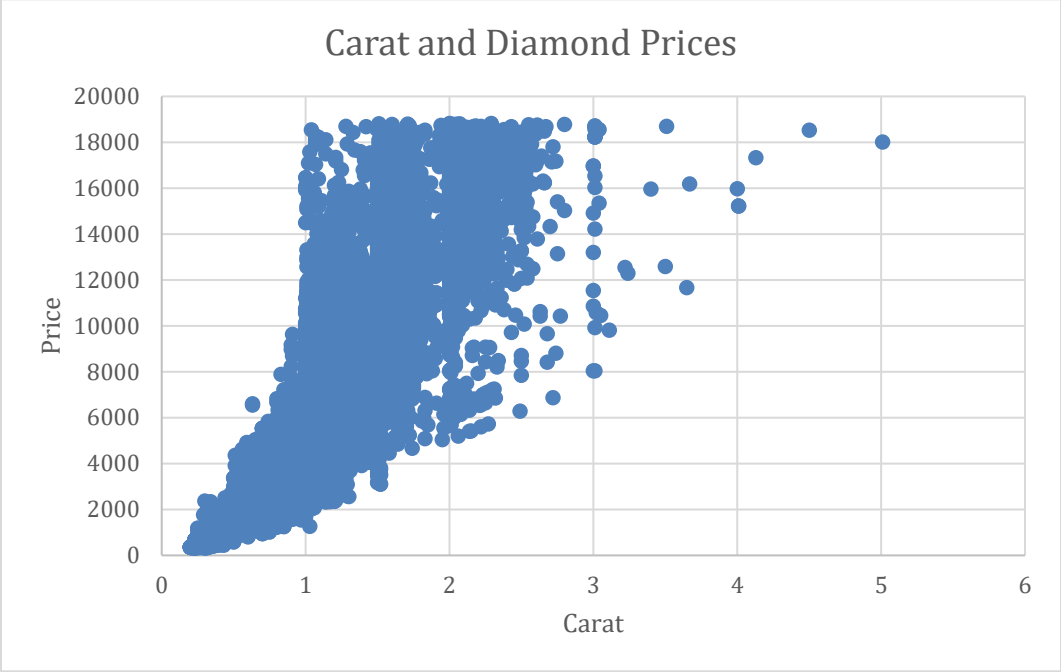
=**-5,269 + (8,413 * 1.5**) **+ (158.1 * 3) + (454 * 5)**

= $10,094.8 they have to pay in order to gain 1.5 carat, cut and clarity of diamond.

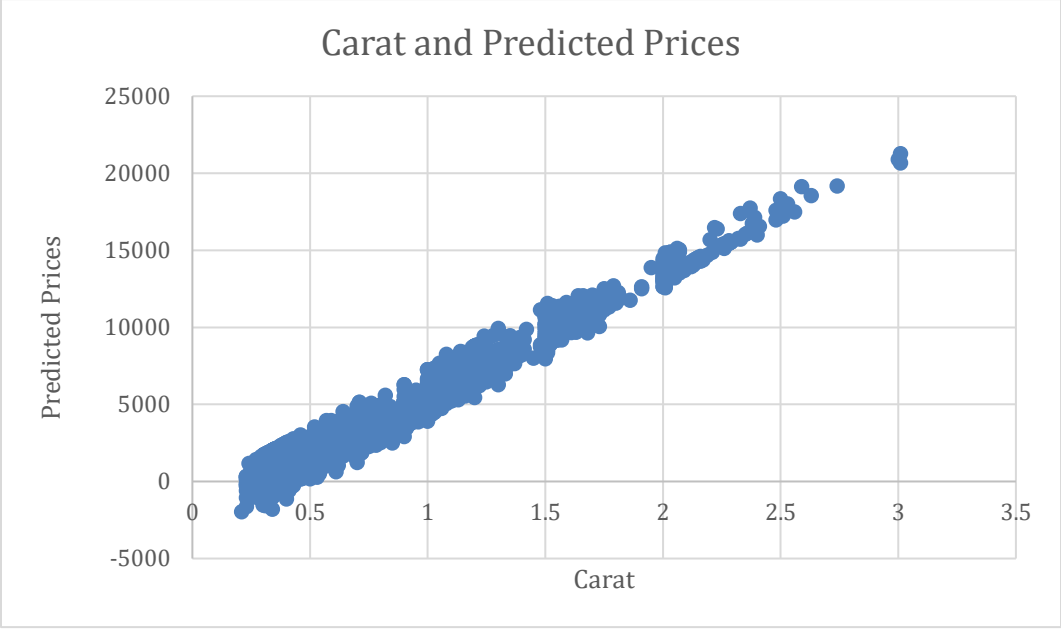
Step 2: Visualize the Data

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.



3. What strikes you about this comparison? After seeing this plot, do you feel confident in the model's ability to predict prices?

It can be seen from the first scatter plot that from 0 to 1 the carat and diamond prices are linear, but from 1 till 4 they are not because we cannot draw one line at all while the correlation is not strong also. The range for prices is distributed in the plot which can be predicted that other reasons affect the diamond prices to be changed like color and clarity. In addition, there is a strong positive correlation between carat and predicted prices, on the other hand there are some negative prices in the plot which can be indicated that this is not the correct method to solve the issue, so we should look for another new appropriate method.

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 think that the linear regression is not the correct method to be used to solve with because there are some negative values in the column. There is another method that can be used by calculating the sum of predicted prices which is equal to $11,733,522.76 and multiply the sum by 0.7 to get $8,213,465.93