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?

   According to the linear regression model, the formula to predict the diamond prices is: **Price** = -5,269 + 8,413 x **Carat** + 158.1 x **Cut** + 454 x **Clarity**

   We can see that for each extra Carat the price will be 8,413 more as in the formula we have ***"8,413 x Carat"***

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?
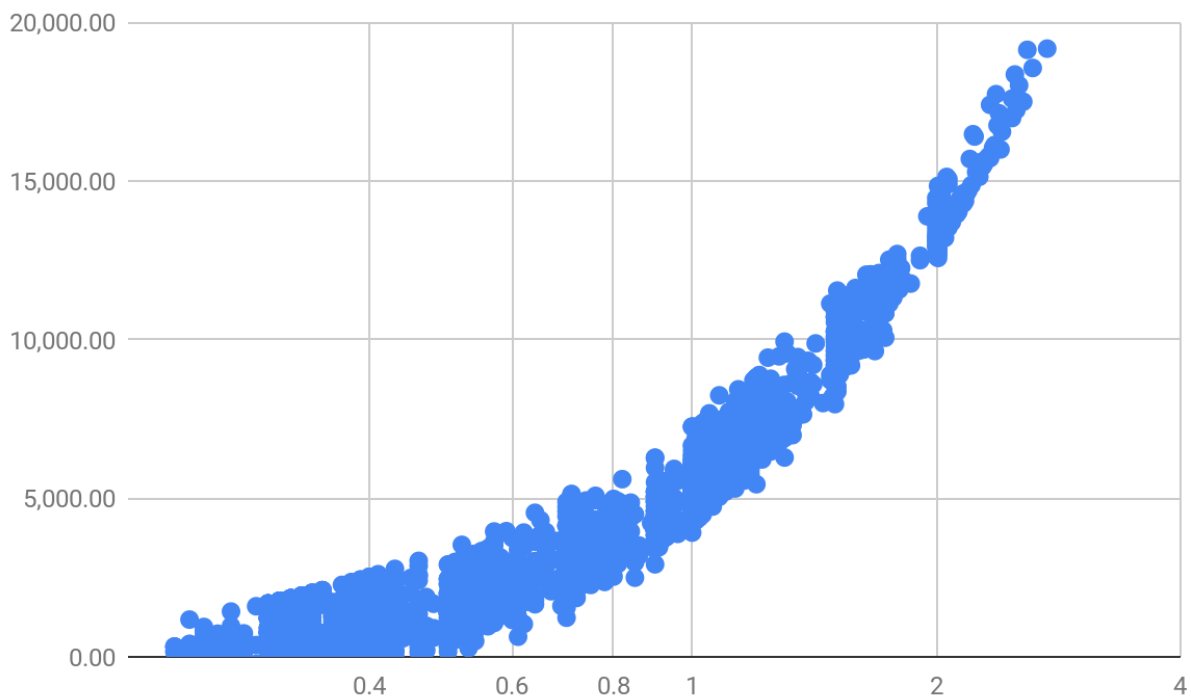
   The price to pay would be: -5269 + 8413 * 1.5 + 158.1 * 3 + 454 * 5 = **10094.8**

Step 2: Visualize the Data

1. Plot 1 - Plot the data for the diamonds in the database, with carat on the x-axis and price on the y-axis.
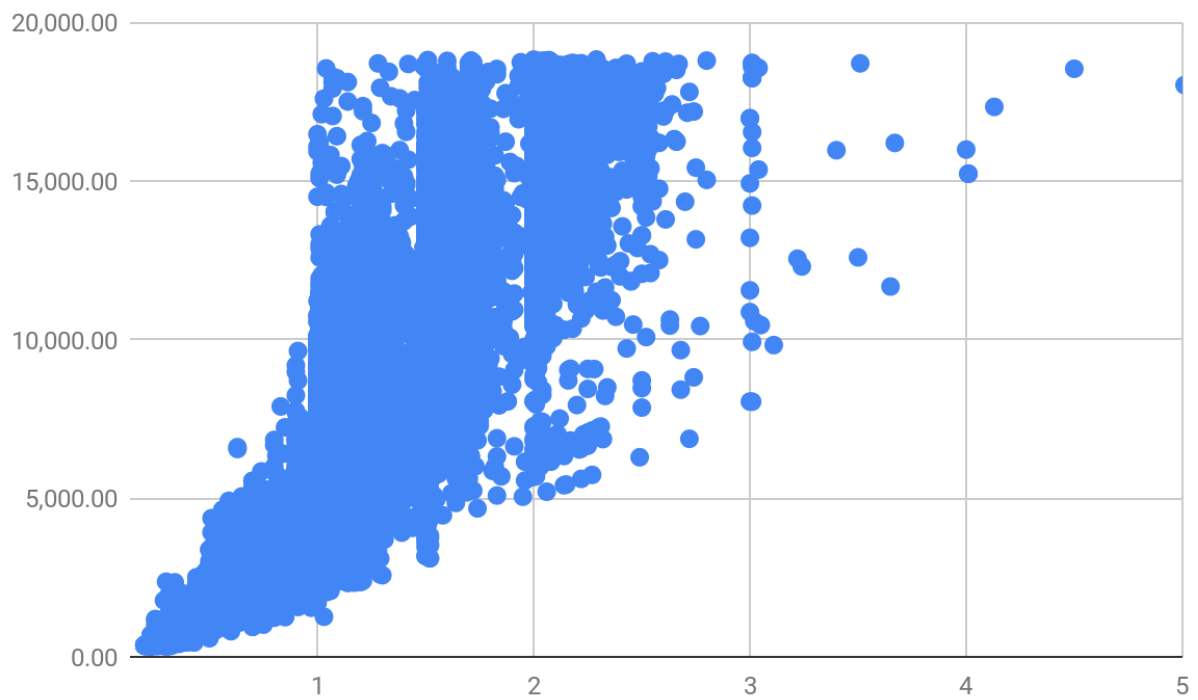
2. Plot 2 - Plot the data for the diamonds for which you are predicting prices with carat on the x-axis and predicted price on the y-axis.

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

The model seems to capture the correlation between carats and price through the increasing tendency that it displays. What is different is the prices for diamonds between 1 and 3 carats in the diamonds dataset. Even though it follows the tendency in the low end, the high end has not a very clear structure.

Despite the fact that the model does a decent job for a simple linear regression model, I would definitely like to try a better model to improve the prediction ability.

Step 3: Make a Recommendation

1. What price do you recommend the jewelry company to bid? Please explain how you arrived at that number.

Since the company wants a 20% margin to make the profit they want, and since they generally purchase diamonds from distributors at 70% I would recommend a bid at

the 20-30% lower than the prediction. Even lower as the model doesn't seem too accurate. I would say 65% of the total worth of the diamonds which is: 0.65 * 11,733,522.76 = **7,626,789.79**