Project: Diamond Prices

Complete each section. When you are ready, save your file as a PDF document and submit it in your classroom.

Step 1: Understanding the Model

Answer the following questions:

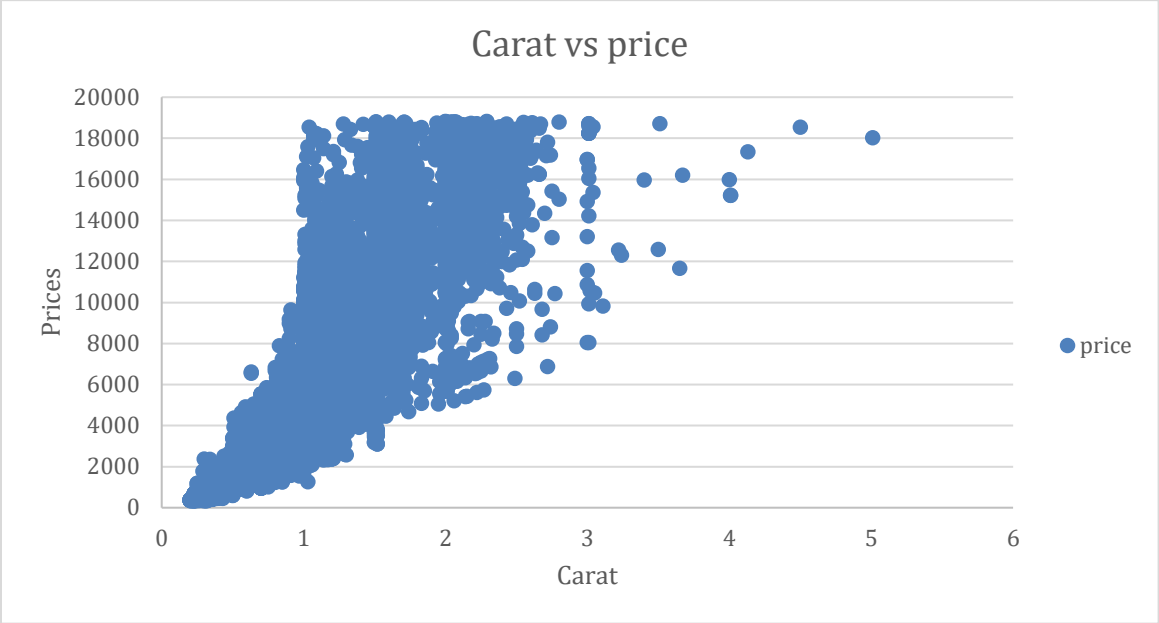
1. According to the model, if a diamond is 1 carat heavier than another with the same cut, how much more should I expect to pay? Why?
    - An extra carat would result in an addition of $8,413 in price. The formula created by regression determined that the coefficient for a carat is 8,413, so when ever the number of carats increase the price should be expected to increase by the amount of the coefficient.

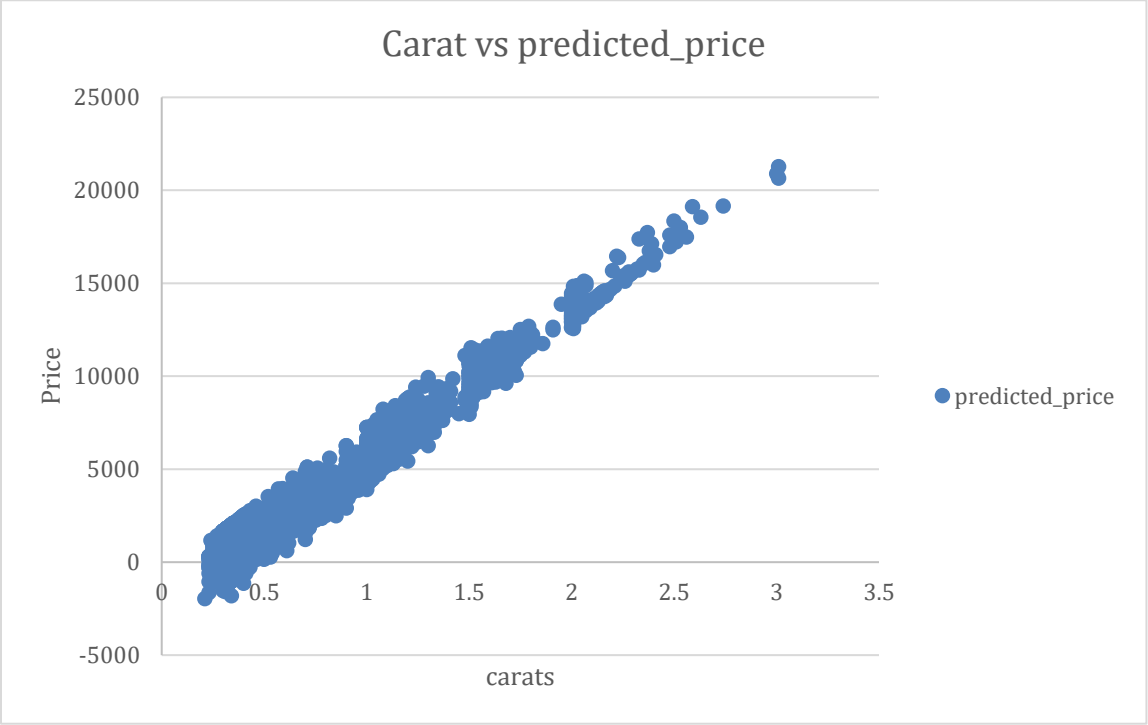
2. If you were interested in a 1.5 carat diamond with a **Very Good** cut (represented by a 3 in the model) and a **VS2** clarity rating (represented by a 5 in the model), how much would the model predict you should pay for it?
    - The formula is Price = -5,269 + 8,413 x Carat + 158.1 x Cut + 454 x Clarity
    - So now we will plug in the values of the different variables.
    - Price = -5,269 + 8,413 x 1.5 + 158.1 x 3 + 454 x 5
    - Price = $10094.8

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?





I have observed that in actual prices the plot is not linear when comparing to prices. In the beginning, the plot started as a liner but deteriorated further to be more scattered. Thus, the more the carats, the more outliers we find. This means that carats, clarity, and the cut are not enough to determine the price of actual diamonds.

In the predicted prices plot, the obvious issue is why are we having negative prices? This shows concern for inaccurate data. So maybe grading clarity and cut from alphabetical to numerical is not the best choice here.
From personal knowledge, the prices of diamonds should not increase when they are secondhand or being resold, but the plots are showing the opposite. Thus I believe that the data is inaccurate.

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 calculated predicted prices from the given formula then took the sum of all predicted prices
    - Sum of all predicted prices is 11733523
    - The company generally purchases diamonds from distributors at 70% of that price, I will take the sum and multiply it by 0.70 which equals to $8213466.1 which the company should bid for.