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?

   From the linear regression equation given, the coefficient for carat is 8413. For a diamond to be 1 carat heavier than another with same cut, the price should be $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?

   From the linear regression equation,
   Price = -5,269 + 8,413 x Carat + 158.1 x Cut + 454 x Clarity,
   we put the values in the equation.

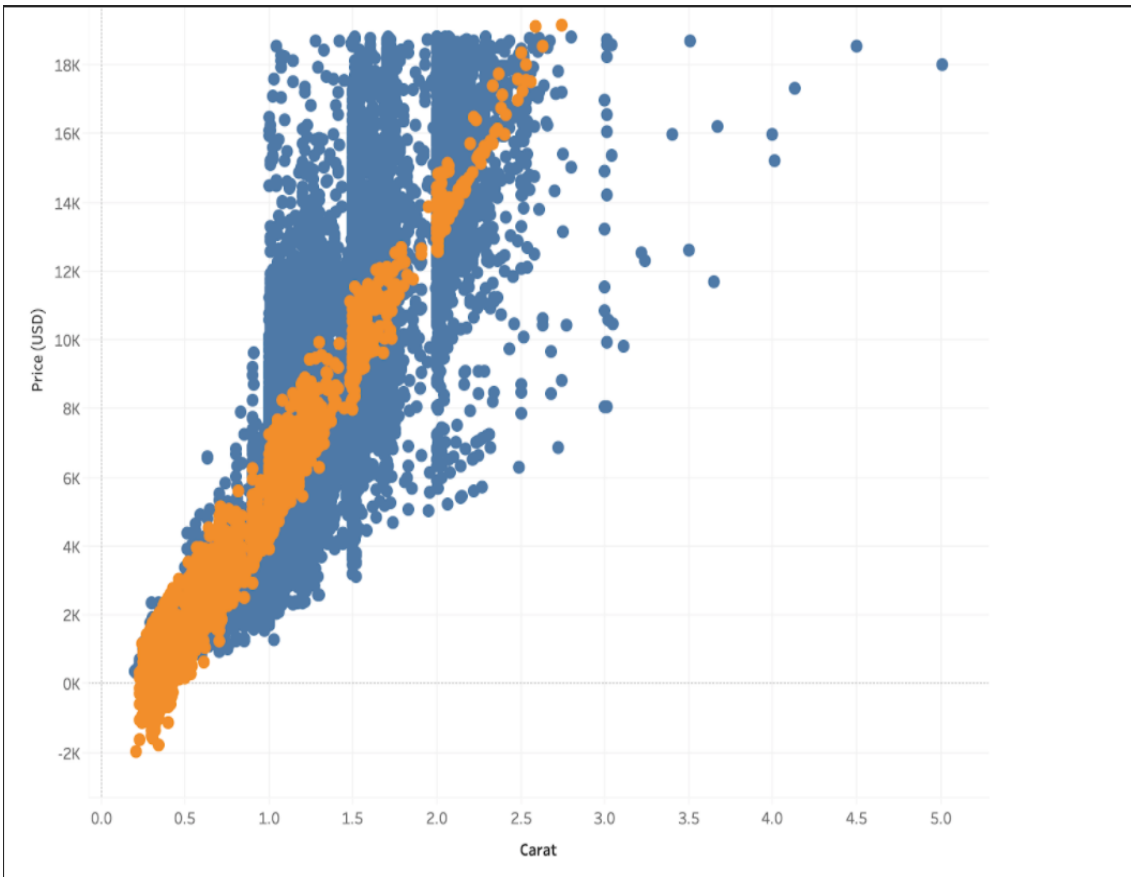
   Price = -5269 + 8413 * 1.5 + 158.1 * 3 + 454 * 5
   Price = $10,094.8
   Therefore $10,094.8 should be paid.

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?

In the scatter plot above, I noticed that the relationship in the previous Diamond graph, between prices and carat are nonlinear but in the new Diamonds graph, the Predicted prices are distributed linearly. Also there are negative values in the graph which the model does not take care of the situation. So I don't feel confident that linear regression model is the best to predict the prices.

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.

For the entire set of 3,000 diamonds, I will recommend the company to bid $8,213,465.93.

I arrived at this number by using the linear regression equation model which was derived from the previous sales, applied it on the new diamond data to get the predicted prices. These amount was the summed together to be $11,733,522.76. Since the company purchases diamonds from distributors at 70% of that price, I multiplied the predicted amount by 0.70 to get the final predicted bid of $8,213,465.93.