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?

According to the linear regression equation:

Price = -5,269 + 8,413 x Carat (+1) + 158.1 x Cut + 454 x Clarity

The non- constant variable is the carat and it's multiplied by 8413, hence, by adding 1 carat the price is expected to increase by 8413.

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?

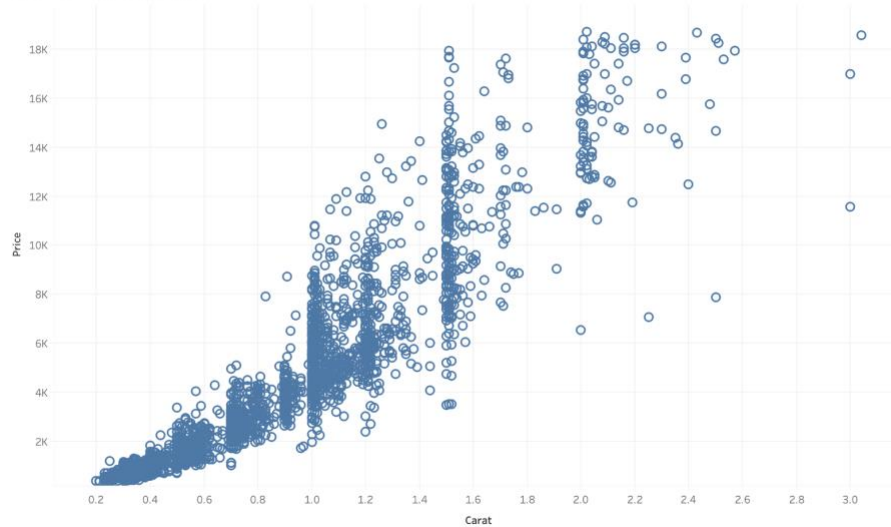Price = -5,269 + 1.5 + 158.1 x 3 + 454 x 5

Predicted amount: 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.
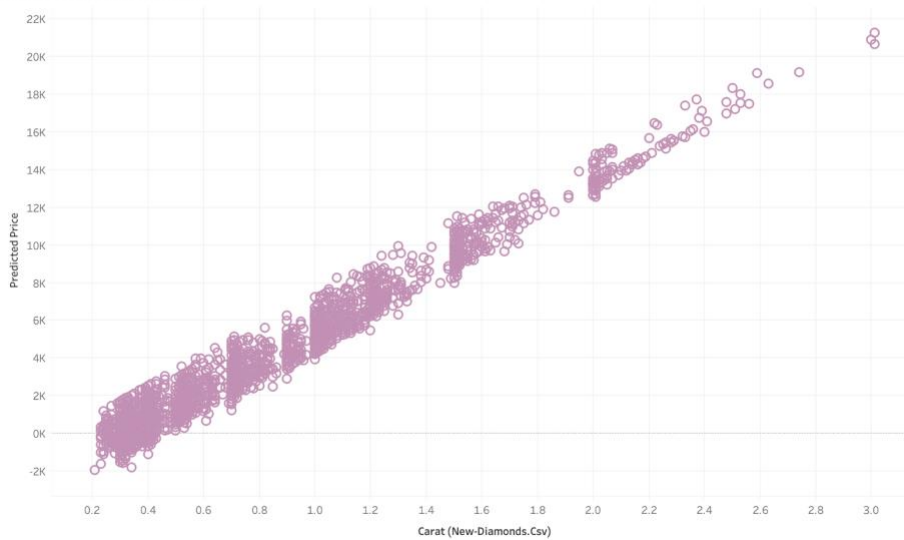
Training Data Graph [1]

Carat vs. Price.

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.

Predicted Data Graph [2]

Carat (New-Diamonds.Csv) vs. Predicted Price.

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

The regression model is not accurate enough. Because, while comparing the training data to the predicted data the distribution is not similar. The predicted data has negative values, in spite prices cannot be in negative.

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.

The sum of the predicted prices: 11733522.76
The recommended price: 11733522.76 x 0.70 = 8,213,466.89