Project: Diamond Prices

Dung Ngoc Nguyen

Step 1: Understanding the Model

Answer the following questions:

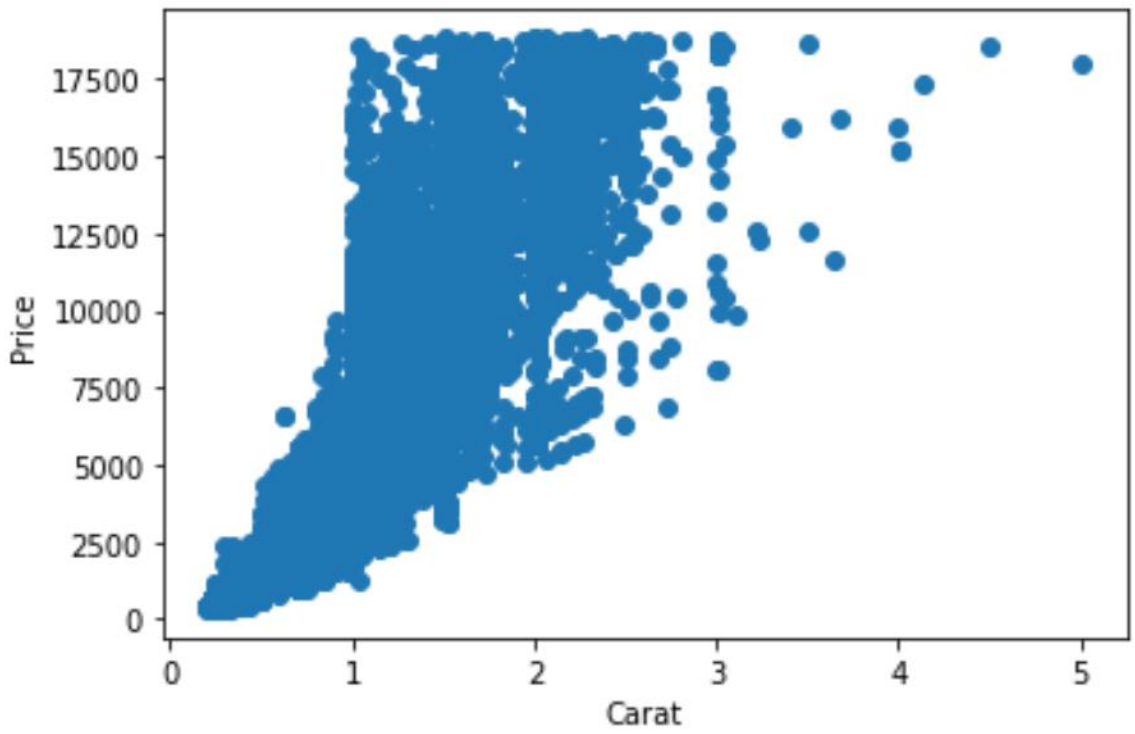
1. *According to the model, if a diamond is 1 carat heavier than another with the same cut, how much more should I expect to pay? Why?*
   8,413 unit of price (I guess it is dollar) I should expect to pay. Because the coefficient for carat in the model is 8,413.

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?*
   The price I should pay = -5,269 + 8,413*1.5 + 158.1*3 + 454*5 = 10094.8

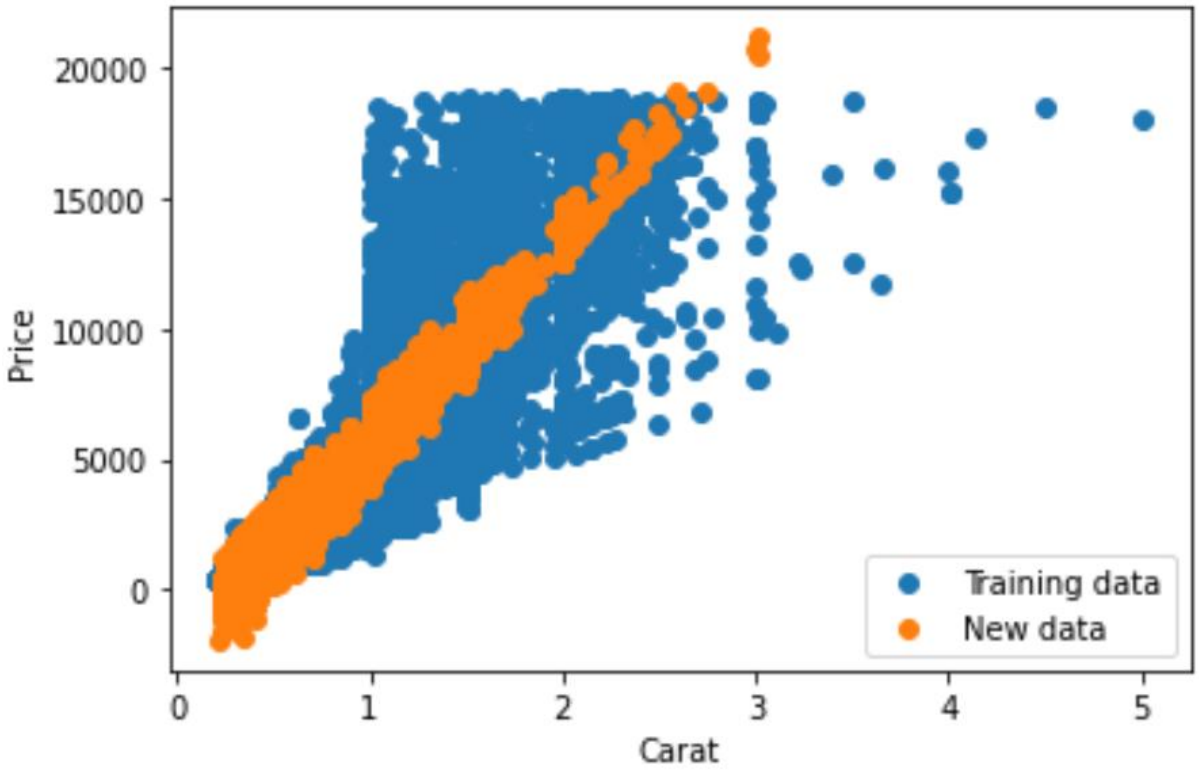
Step 2: Visualize the Data

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.*
   As my understanding, I have to plot the data of diamonds.csv



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*?

There is still certain extent of variability in price that the linear regression model of carat cannot explain. I am not very confident in the model's ability to predict 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.*
   Using this formula: Price = -5,269 + 8,413 x Carat + 158.1 x Cut + 454 x Clarity, I summed up all the predicted price in the new data and come up with 11730233.07 as the price that the company should bid.