

Project: Diamond Prices

Complete each section. When you are ready, save your file as a PDF document and submit it here: <https://classroom.udacity.com/nanodegrees/nd008/parts/235a5408-0604-4871-8433-a6d670e37bbf/project#>

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?

$$\text{Price} = -5,269 + 8,413 \times \text{Carat} + 158.1 \times \text{Cut} + 454 \times \text{Clarity}$$

The one carat heavier for the same cut would result in an additional \$8,413 in price. The formula created by the regression determined that the coefficient for a carat is 8,413 so for every increase in the weight of diamond the price will 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?

$$\text{Price} = -5,269 + 8,413 \times \text{Carat} + 158.1 \times \text{Cut} + 454 \times \text{Clarity}$$

so now we will plug in the values for the different variables.

$$\text{Price} = -5,269 + 8,413 \times 1.5 + 158.1 \times 3 + 454 \times 5$$

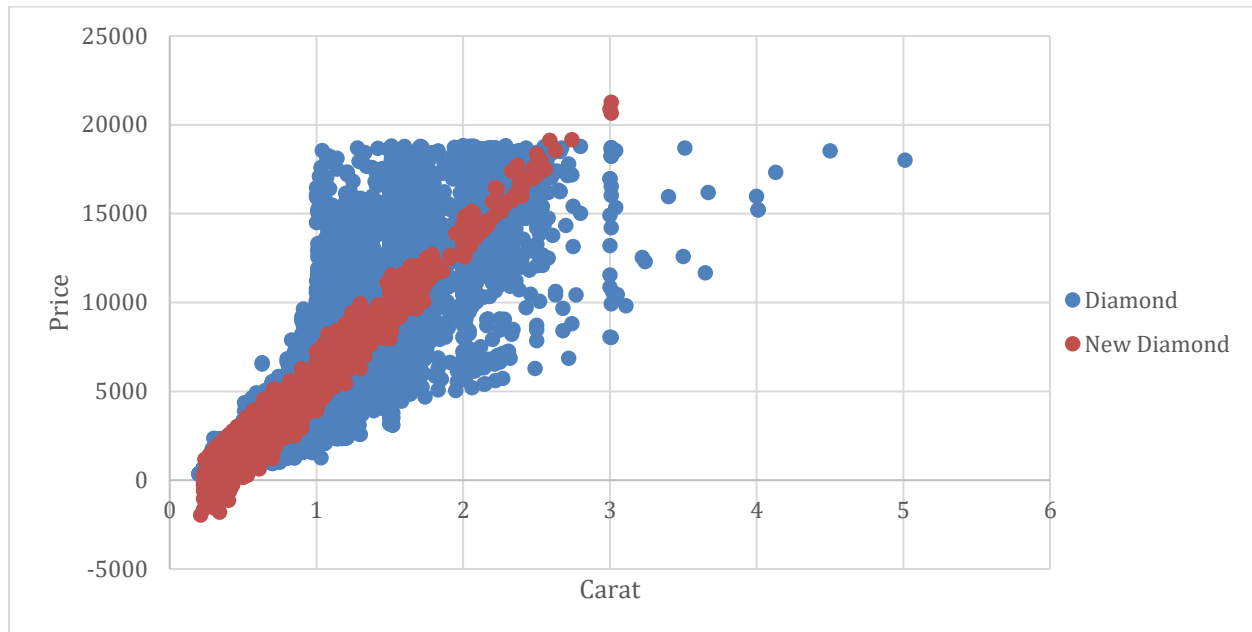
$$\text{Price} = \$10,095$$

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.
 - o **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?

Carat vs Price



1. For actual diamond data, we can conclude that Carat is not linearly distributed with prices. If we see the graph for carat 0 to 1, prices seem almost linear, however from 1 to 3 carats, prices are not linear with carats. Here the prices are divided into different ranges for each carat. There must be many other factors which are affecting the diamond prices.
2. For the diamond price we predicted based on the equation, the price of diamond is negative. It is like getting money while buying diamond.
3. Also Predicted prices are distributed linearly however old diamond prices are not linear.
4. As one of the assumptions of linear regression is that independent variables should be linearly distributed which did not meet in this diamond problem, linear regression is not the right approach 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.
 - Perform sum on all the predicted prices. (11733523)
 - Since this is the consumer price, take 70% of the sum and that is the amount a company should bid for diamonds. (8213465.932)

Name: Bipin Kumar Sultania