

Step1: Understanding the Model

- 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?
 - I expect to pay 8,413 for one carat heavier.
 Because each 1 carat with the same cut will got increase the price With the rest of the variables held constant
- 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?

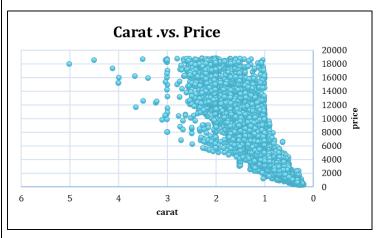
```
The formula is:
```

```
Price = -5,269 + (8,413 x Carat) + (158.1 x Cut) + (454 x Clarity)
= -5,269 + (8,413 x 1.5) + (158.1 x 3) + (454 x 5)
= -5,269 + 12,619.5 + 474.3 + 2,270
Price = $ 10.094.8
```

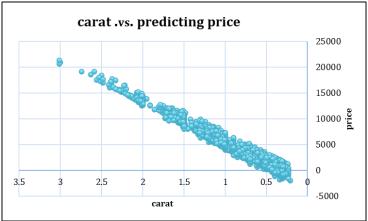
We will pay 10,094.8 to get 1.5 carat diamond.

Step2: Visualize the Data.

Plot 1 - Plot the data for the diamonds in the database, with carat on the x-axis and price on the y-axis.



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.



- 1. What strikes you about this comparison? After seeing this plot, do you feel confident in the model's ability to predict prices?
 - After looking plot 1 appears great dispersion, it can be caused by other factors such as clarity and color.

in **plot 2** There is a strong linear regression equation, and there is no dispersion.

Step3: Make a Recommendation.

- 1. What price do you recommend the jewelry company to bid? Please explain how you arrived at that number.
 - I recommend a bid of \$8,213,465.93 I used decision tree approach and I arrived at this number by using the random forest model, so I multiply the predicted amount by 0.7

I don't used liner regression because it gave some negative prices.

sum predicting prices	07% of the sum of predicting prices	\$ 8213465.93
\$ 11,733,522.76	=J5*0.7	

Illustrative image