Predict Diamond Prices
Step 1 - Understanding the Model:

1. According to the linear model provided, if a diamond is 1 carat heavier than another with the same cut and clarity, how much more would the retail price of the heavier diamond be? Why?

   Linear Regression Equation:

   Price = -5,269 + 8413 * **Carat** + 158.1 * **Cut** + 454 * **Clarity**

   - If a diamond is 1 carat heavier than another with the same cut and clarity, the additional carat would be an extra $8,413 in the retail price. The formula created by the regression determined that the coefficient for a carat is $8413. Therefore, for each additional carat, 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), what retail price would the model predict for the diamond?

   - The formula for retail price is:
     Price = -5,269 + 8413 * **Carat** + 158.1 * **Cut** + 454 * **Clarity**

   - Plug the values into the formula:
     Price = -5269 + 8413 * 1.5 + 158.1 * 3 + 454 * 5

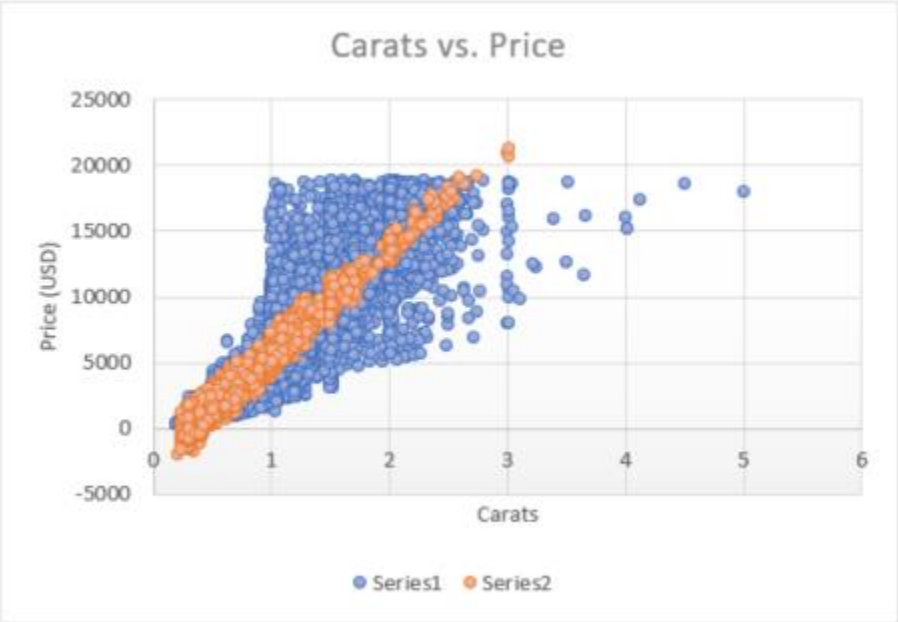
   - **Price** = **10094.8**

Step 2 - Visualize the Data:

Create two scatter plots. If you're not sure what a scatter plot is, see **here**.

- 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.
- Note: You can also plot both sets of data on the same chart in different colors.

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



The predicted diamond prices are more compact than the actual prices. In this chart, only carats and its effect on price is be accounted for. Cut and clarity is also factored into the formula and these factors along with other factors can cause variations in price. For example, color also has a huge impact on price, but this was not factored into the formula. If the diamond has less color, the price of the diamond increases per carat.

After looking at this plot the model appears to do an average job in predicting prices. However, variations in price for certain diamonds do exists. The formula does not consider all factors that can determine price including color.

Step 3 - The Recommendation:

What bid do you recommend for the jewelry company? Please explain how you arrived at that number. HINT: The number should be 7 digits.

The bid price that I recommend for the whole set of diamonds is $ 8,213,466. I arrived at this number by using the linear regression model to predict the price of diamonds based on its attributes. I then summed the predicted prices for the set of 3000 diamonds and arrived at the amount of 11,733,522.76. I multiplied 11733522.76 by .70 to get the final predicted amount of **$8,213,466,** which was rounded to the nearest dollar.