Complete each section. When you are ready, save your file as a PDF document and submit it in your classroom.

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?

    Ans: Given Regression equation showss the coefficient for the predictor variable carat is 8413. This clearly shows that, whenever there is an increase by 1 carat, the extra amount to be paid 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?

    Ans: The given Linear Regression Equation is,

    Price = -5269 + 8413 * Carat + 158.1 * Cut + 454 * Clarity

   Asked to estimate Price for 1.5 carat diamond with Very Good cut & VS2 clarity,

    Price = -5269 + 8413 * 1.5 + 158.1 * 3 + 454 * 5

    = 10,944.8

    So, the model predicts that I have to pay **$ 10,944.8** for the given specifications.

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.

Scatterplot of carat versus price
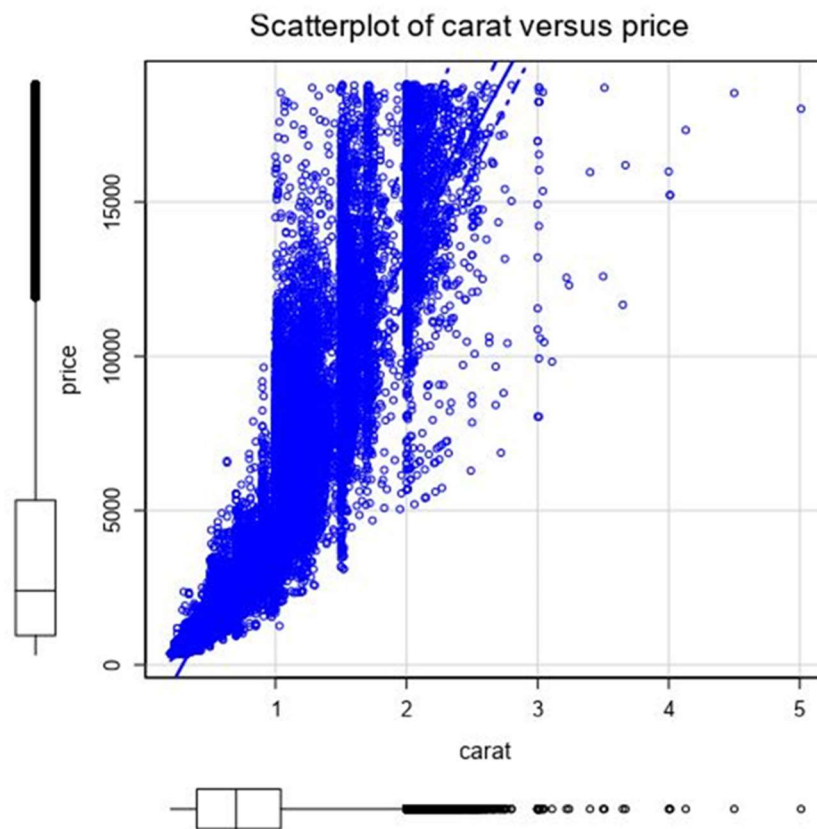
2. Plot 2 - Plot the data for the diamonds for which you are predicting prices with carat on the x-axis and predicted price on the y-axis.
   - o **Note**: You can also plot both sets of data on the same chart in different colors.

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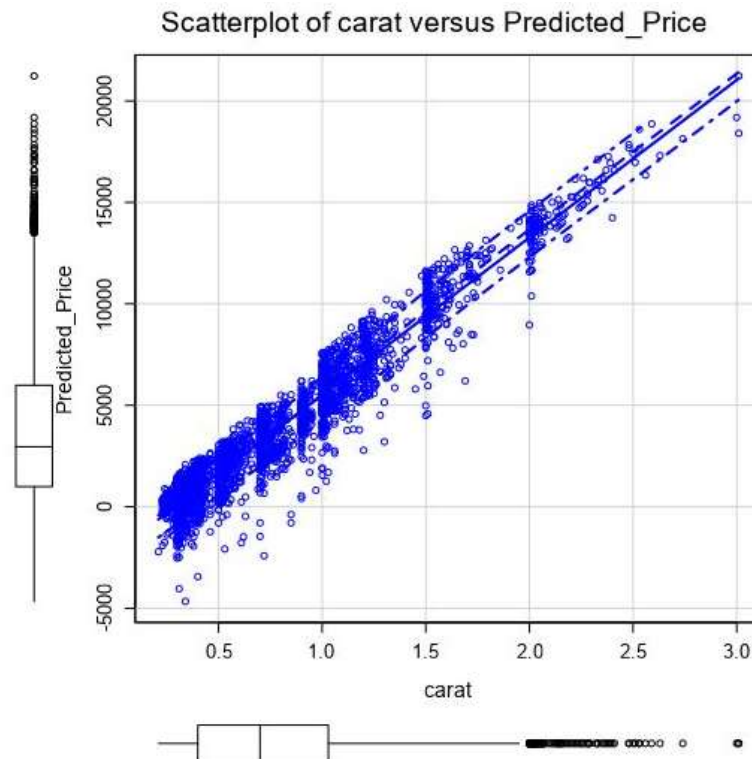
3. What strikes you about this comparison? After seeing this plot, do you feel confident in the model's ability to predict prices?

Ans: The above two scatter plots shows that there is a linear relationship exists between the independent (carat) variable and the dependent (price) variable.

However the first plot, for the diamonds dataset is not 100% linear. When the carat is higher than 1.5, it's not strongly positive.

But, the second plot shows the linearity between carat and predicted_prices.

Hence, I think the model is pretty good in predicting prices for diamonds up to 1.5 carat. Beyond that (for carat > 1.5), it's not so good. But other factors like cut and clarity also have greater impact in predicting 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.

    Ans: From my model predictions, the total predicted_price is: $11,724,983.055. But, the jewelry company purchases diamonds from distributors at 70% of that price, which is 11,724,983.055 * 0.7 = 8,207,488.14. Hence, 70% of 11,724,983 is 8,207,488, which is the amount the jewelry company should pay to the diamond distributors.
    So, my recommended price that the jewelry company has to bid is: $ **8,207,488**.