<u>**PROJECT: PREDICTING DIAMOND PRICES**</u>

<u>**PROJECT SUBMISSION WRITTEN BY CHIDIEBERE STEPHEN NWOSU**</u>

Step1 – 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?

   Answer:
   From the Model equation below;
   Price   = -5269 + (8413*Carat) + (158.1*Cut) + (454*Clarity)

Therefore, let's take Carat = 1, Cut = 3 and Clarity = 4

   Price   = -5269 + (8143*1) + (158.1*3) + (454*4)
   Price   = 5,164.3

   Now, if a diamond is 1 Carat heavier;

Then, Carat = 2, Cut = 3, and Clarity = 4.

   Price per-carat increase   = - 5269 + (8143*2) + (158.1*3) + (454*4)

   Price per-carat increase   = 13,307.3

   Retail price = $(Price per-carat increase – Price)
   Retail price = $(13,307.3 – 5,164.3)
   Retail price = $8,143

   The retail price of the heavier diamond is 7,873(USD), this is so because the increase in carat will increase the price of the diamond. A 2.00carat diamond will have a much higher per-carat price than a 1.00carat diamond.

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?

   Answer:
   Predicted retail price = -5269 + (8413*Carat) + (158.1*Cut) + (454*Clarity)
   Predicted retail price = -5269 + (8413*1.5) + (158.1*3) + (454*5)
   Predicted retail price = $10,094.8

Step2 – Visualize the Date

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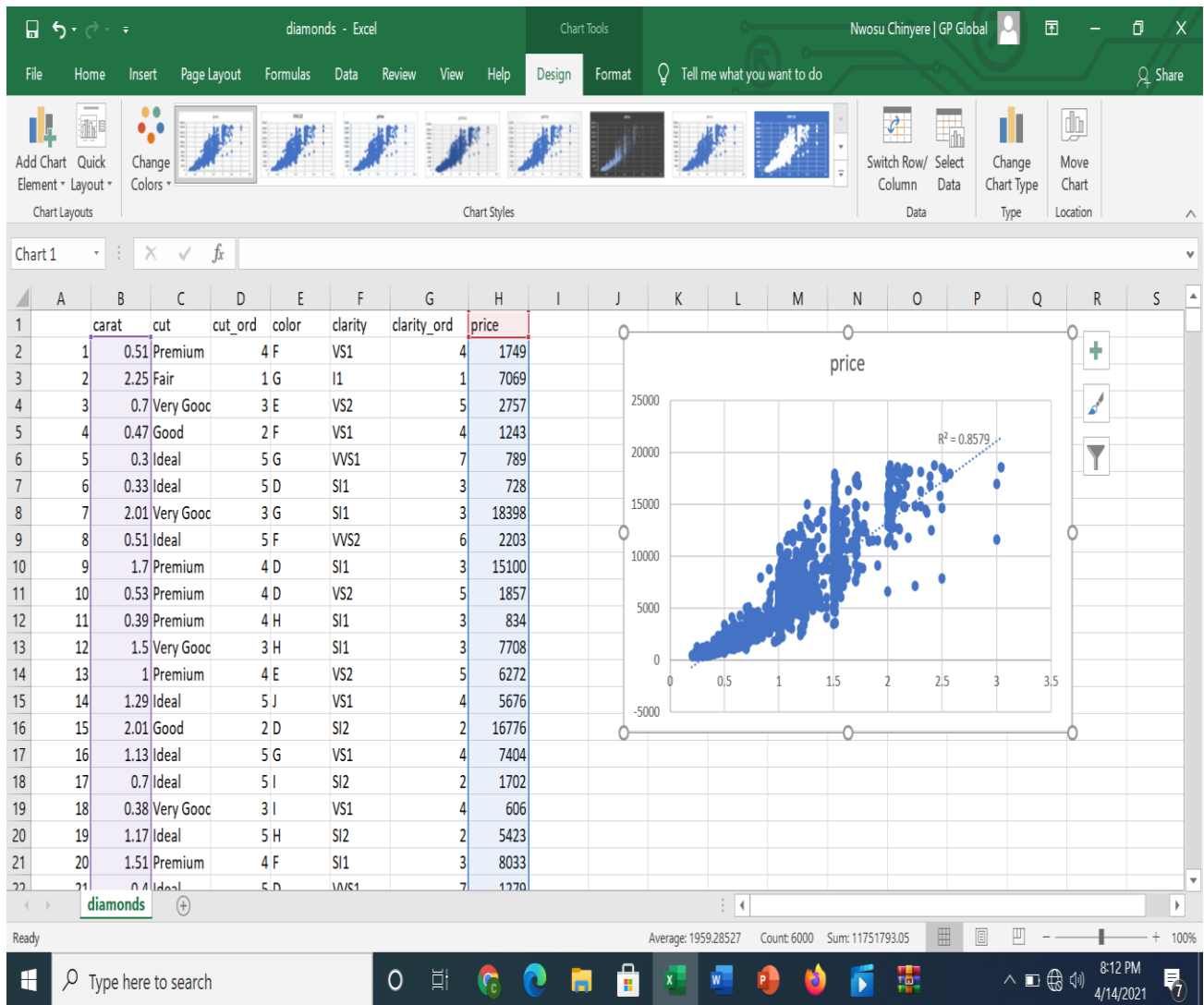

Figure1. Graph of Price vs carat (Plot of price on the Y-axis and carat on the x-axis).

- Plot 1 – 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.
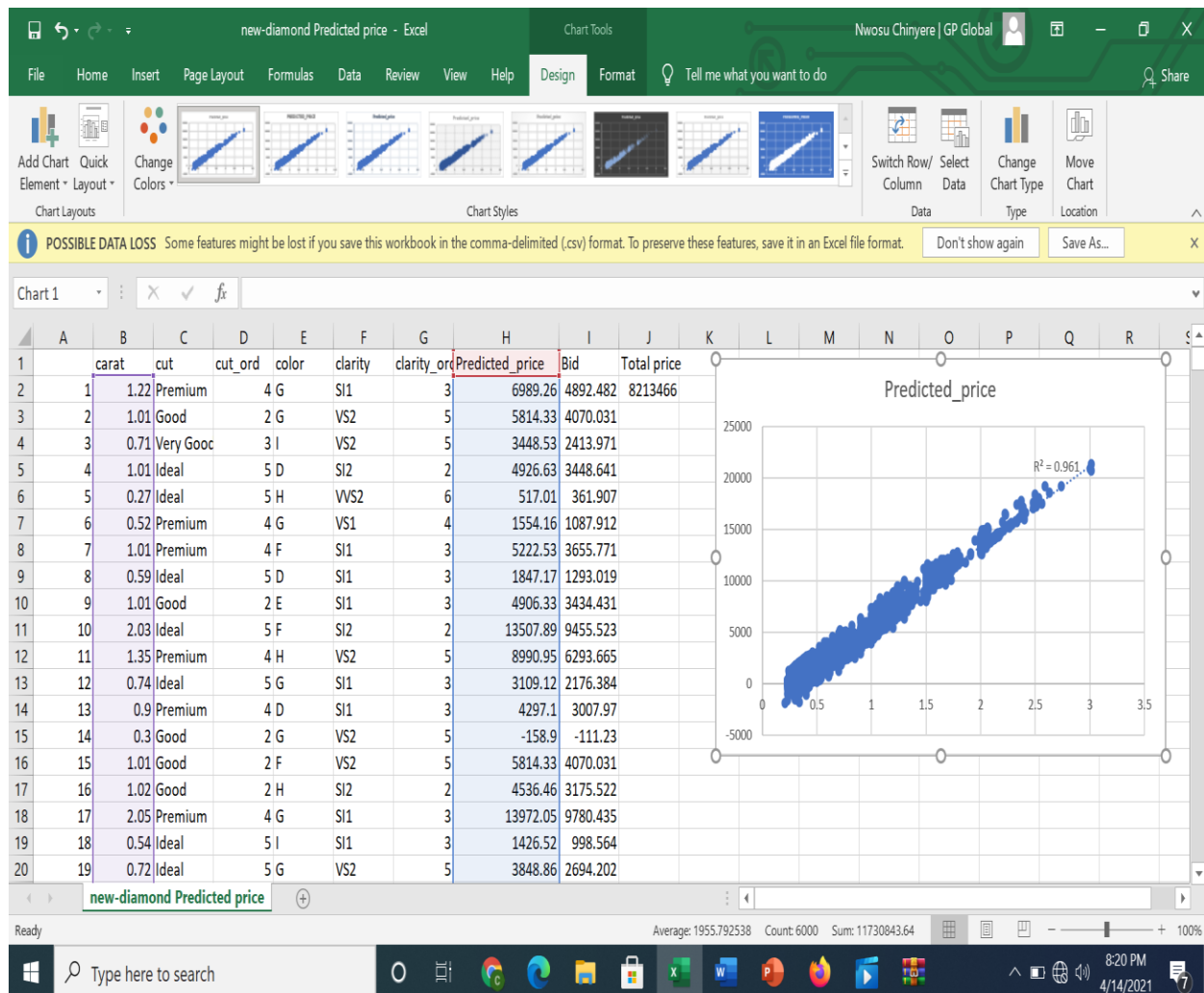
Figure2. Graph of Predicted_price vs Carat (Plot of Predicted_price on the Y-axis and Carat on the X-axis).

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

  From the scatter plot above, both graphs show a strong correlation. In Figure1 above, plot of price vs carat shows a strong linear regression along the straight line and some offset can be seen. The offset is due to the fact that it has lager data set than the data set in Figure2. Most importantly, it has an R square value of 0.8579
  Similarly, in Figure2 above, plot of Predicted_price vs carat, shows a stronger correlation because there is little or no offset. This shows that there is stronger linear relationship between the predicted price and the variables: carat, cut, and clarity along the straight line. The R square value is 0.961, which is closer to 1, would mean that nearly all variance in the target variable is explained by the model.
  I absolutely feel confident in the model's ability to predict price.

Step 3 – The Recommendation: What bid do you recommend for the jewelry company? Please explain how you arrive at that number.
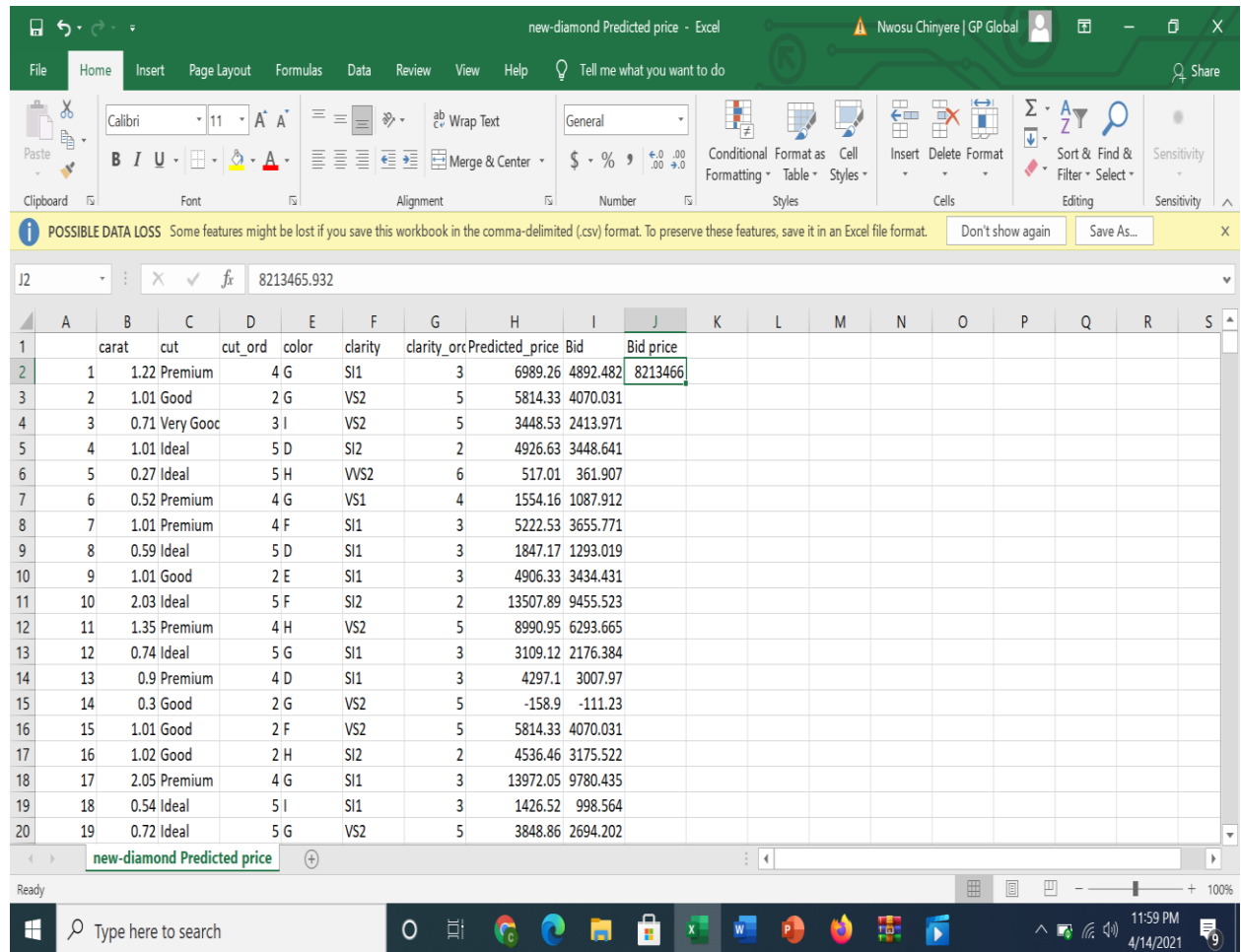

Figure3. Bid price

I recommend the Jewelry company to bid at $8213466.

Considering the fact that the company purchases diamond at 70% from the distributor. From the 3000 data set of the predicted_price, each cell was multiplied by 0.7 to get the Bid, then the 3000 data set of the Bid was summed up to get the Bid price of $8213466.