Diamond Price Analysis

Thomas Dolan, Kemalcan Jimmerson, Curtis Escobar Rood, Sean Kenny
An Analysis by LOTUS (League of the Undying Statisticians)

Importance and Context

The struggle to quantify the relationship between the physical properties of diamonds and their market price dates back centuries [2]. Once only available to the wealthiest of society, these precious gems are now purchased by all as symbols of love, status, and more. Until the 1940s, consumers, and industry lacked a standardized system of evaluation for grading and purchasing diamonds, mainly relying on binary distinctions of clarity and cut [3]. However, by the early 1950s, the measurement of the 4Cs: Cut, Color, Clarity, and Carat in conjunction with new rating scales, allowed for the objective grading of a diamond's quality [5]. Despite the establishment of the 4Cs and the Gemological Institute of America's International Diamond Grading System, the diamond industry keeps pricing information secret making it difficult for those who shape the diamonds for jewelers to maximize profits [4].
Before these gemstones are sold as jewelry, some critical steps will be taken to determine how much a diamond will be worth. Once extracted from the ground, rough diamonds will undergo sorting based on characteristics such as size, shape, quality, and color [1]. These diamonds are then sold to professional diamond cutters, such as our client Diamonds in the Rough Cutters (DRC), where these diamonds are shaped and polished into their finest and most profitable form. These skilled artisans are subject to thin profit margins and must optimize for the best attributes possible for each diamond [4]. While maximizing the four Cs is important, diamond cutters often must compromise between saving weight (Carat) and the other three Cs [6]. The work for our client DRC focuses on quantifying the worth of a cut diamond based on its Carat weight. The client exclusively conducts business with luxury goods dealers looking for the finest diamonds on the market. With adequate knowledge of how their diamonds will be priced by luxury dealers based on the diamond's Carat, DRC will be able to make better decisions regarding the frequent problem cutters encounter between saving weight and the other factors determining a diamond's worth. The basis of our study is a snapshot of renowned Tiffany and Co.'s 2017 price list for natural unmounted diamonds. We will use this data to produce an explanatory model aimed at addressing the following:

> *How does the Carat weight of a natural cut diamond determine its price when selling to luxury diamond dealers?*

We propose an initial causal path where increases in Carat (treatment variable) will drive positive change in Price (response variable). We must also account for several confounding variables that might affect the end price of a diamond. These confounding variables include Color, Cut, and Clarity with an error term epsilon capturing the information not explained by our Carat or our confounding variables.

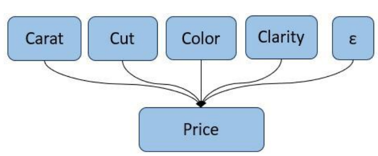

Figure 1: Causal Path Diagram

Data and Methodology

The dataset leveraged for our model was retrieved from Kaggle after having been compiled by Sivam Agrawal. Each row in the data represents a single natural diamond from a 2017 Tiffany and Co's snapshot price list. This will ensure that our model is suited to our clients target market– luxury diamond retailers. The entire dataset contains 53,940 observations. 20 Records included a 0 value in either the X, Y, or Z axis and were subsequently removed since it would imply a non three dimensional diamond. The data was otherwise clean with no additional data transformation required. There are 10 features in our dataset. Price (in USD$) and carat are both metric variables while cut, color and clarity are ordinal variables that we mapped to a likert scale. Finally, there were the dimensional variables x, y, z, depth and table. These dimensional features were not used in the regression due to the obvious reverse causal pathways present. For example, the formula for diamond depth is comprised of x, y, and z attributes. In order to build and train the model, we randomly sampled 600 of these price observations and split that sampling into a training set and test set (420 observations 180 observations). In order to operationalize our model, it was necessary to transform both carat and price which were both positively skewed. The performed transforms (Fig. 2) removed the tails and thus created a more suitable dataset against which to build our model. It is worth noting that the resulting transformations revealed a multimodal distribution grouped around half carat intervals. We interpreted this as an artifact of current diamond cutting procedures which try to maximize the perceived symmetry of a diamond to consumers. At this point, we were prepared to start building our model. Our regression form took the following shape after all ordinal variables were one hot encoded to allow for a more accurate model.

$$\begin{aligned} log(\widehat{price}) = \beta_0 + \beta_1 \cdot log(carat) + \beta_2 \cdot cut(Ideal) + \beta_3 \cdot cut(Premium) + \beta_4 \cdot cut(VeryGood) + \\ \beta_5 \cdot cut(Good) + \beta_6 \cdot color(D) + \beta_7 \cdot color(E) + \beta_8 \cdot color(F) + \beta_9 \cdot color(G) + \\ \beta_{10} \cdot color(H) + \beta_{11} \cdot color(I) + \beta_{12} \cdot clarity(IF) + \beta_{13} \cdot clarity(VVS1) + \beta_{14} \cdot \\ clarity(VVS2) + \beta_{15} \cdot clarity(VS1) + \beta_{16} \cdot clarity(VS2) + \beta_{17} \cdot clarity(SI2) \end{aligned}$$

Results

We defined 4 different models and compared their adjusted R-squared, and F-statistics. In the Stargazer table (Table 1) below, the model 1 refers to the model where we estimated price using carat only. In this model the R-squared value is 0.932. So, 93% of the changes in the price can be explained by the changes in the carat. This is a significant relationship. We started to add other variables to the model and created different models. Model 2 estimates the price using carat and clarity. This model increased the adjusted R-Squared more than 3%. So, carat and clarity together explains the variance in the price almost 97%. In model 3, we used carat, clarity, and color to estimate the price. In model 4, we added cut into the linear model. Model 4 has the highest adjusted R-squared score. We defined that model 4 is our best model to make predictions for the price. The F-Statistic score is significant for all models, so we reject the null hypothesis that the coefficients are zero.

In Fig. 2, we are comparing the predictions of Model 1 and Model 4. Y-axis shows the predicted price (logged), and X-axis shows the actual price (logged). The blue line represents where the perfect prediction would be (y=x).The red data points represent model 1's predictions and all black data points represent model 4's predictions on the test data. This plot shows that the black data points are closer to the line than the red data points. In other words, there is higher variance in model 1's predictions. That means that the total magnitude of errors are higher in model 1. Because of that, model 4 is a better predictor. On the test data the model 4's RMSE score is 535.

Since our best-performing model is model 4, we will interpret the coefficients of the model. As it is seen in the Stargazer table (Table 1), in model 4, all coefficients are statistically significant at a 1% level except for the "Good Cut". Every 1% increase in carat is associated with about 1.9% increase in the price, ceteris paribus. For diamond clarity the base case is 'I1' category. Changing clarity from 'I1' to 'IF' category increases the price by a factor of about 3.19, or 219%, ceteris paribus. For the Color category, the base case is J-Colorless and the best color is D. Changing color from 'J' category to 'D' increases the price by a factor of 1.66 or 66%, ceteris paribus. Similarly, for the cut category, the base is "Fair Cut". Changing the cut from 'Fair Cut' to 'Very Good' category increases the price by a factor of 1.16 or 16% ceteris paribus.

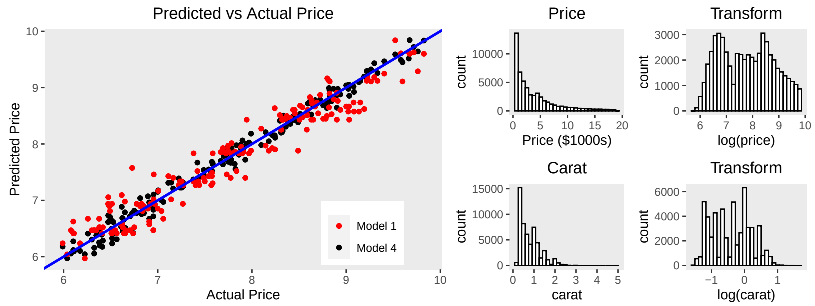

Figure 2: Model Comparisons and Data Transformations

Table 1: Model Summary

	Dependent variable:			
	Log Transform of Natural Cut Diamond Selling Price			
	(1)	(2)	(3)	(4)
Log Transform of Carat Weight	1.675*** (0.024)	1.827*** (0.020)	1.890*** (0.016)	1.894*** (0.016)
Clarity: IF (Best Clarity)				0.075 (0.050)
Clarity: SI1				0.153*** (0.049)
Clarity: SI2				0.159*** (0.048)
Clarity: VS1				0.181*** (0.048)
Clarity: VS2		0.485*** (0.094)	0.465*** (0.107)	0.426*** (0.080)
Clarity: VVS1		0.656*** (0.094)	0.648*** (0.105)	0.616*** (0.080)
Clarity: VVS2		0.821*** (0.095)	0.820*** (0.106)	0.780*** (0.081)
Color: I		0.845*** (0.096)	0.878*** (0.107)	0.839*** (0.082)
Color: H		0.983*** (0.099)	0.999*** (0.108)	0.966*** (0.084)
Color: G		1.148*** (0.101)	1.140*** (0.111)	1.086*** (0.088)
Color: F		1.139*** (0.158)	1.219*** (0.141)	1.161*** (0.129)
Color: E			0.158*** (0.038)	0.131*** (0.040)
Color: D (Best Color)			0.298*** (0.036)	0.275*** (0.038)
Cut: Good			0.387*** (0.035)	0.363*** (0.038)
Cut: Ideal (Best Cut)			0.420*** (0.037)	0.404*** (0.039)
Cut: Premium			0.454*** (0.037)	0.436*** (0.039)
Cut: Very Good			0.530*** (0.043)	0.508*** (0.045)
Constant	8.431*** (0.017)	7.747*** (0.092)	7.397*** (0.110)	7.305*** (0.103)
Observations	420	420	420	420
R^2	0.932	0.969	0.982	0.984
Adjusted R^2	0.932	0.968	0.982	0.983
Residual Std. Error	0.264 (df = 418)	0.180 (df = 411)	0.137 (df = 405)	0.131 (df = 401)
F Statistic	5,753.353*** (df = 1; 418)	1,602.059*** (df = 8; 411)	1,613.595*** (df = 14; 405)	1,372.801*** (df = 18; 401)
Note:				*p<0.1; **p<0.05; ***p<0.01

Limitations

In order to genuinely assess the practical significance of our model and its results, we must contend with the assumptions required to support the model as well as the model's limitations. Given the size of our dataset, there are only two assumptions that must be vetted: whether or not our data is independent and identically distributed (IID) and does a unique best linear predictor (BLP) exist. Common violations of IID include clustering in geographic areas as well as autocorrelation from factors carrying across time. While our data does not have these issues, it isn't immune to IID concerns. For example, a larger stone could have been broken down to produce 2-3 smaller diamonds. In such a case, the size of one diamond would without exception necessarily have an influence on the size of the other diamonds. However, for all the factors that we could make knowledgeable claims about, the data set was ideal. The second assumption is that there is a Unique BLP. This entails both that the BLP exists and that it is unique. In order to assess that the BLP exists, we evaluated our training dataset with a variance covariance test (table 2) to ensure that the covariances within the dataset were finite.

	price_log	carat_log	cut_num	clarity_num	color_num
price_log	1.03	0.57	-0.12	-0.30	-0.24
carat_log	0.57	0.34	-0.11	-0.33	-0.23
cut_num	-0.12	-0.11	1.27	0.32	0.01
clarity_num	-0.30	-0.33	0.32	2.53	-0.13
color_num	-0.24	-0.23	0.01	-0.13	2.79

Table 2: Variance-Covariance Matrix

In order for the BLP to be unique, we need to ensure that no perfect collinearity exists. You can see in table 3 below that this is not a concern for our model.

	GVIF	Df	GVIF^(1/(2*Df))
log(carat)	1.37	1.00	1.17
cut_factor	1.22	4.00	1.03
clar_factor	1.62	7.00	1.03
col_factor	1.34	6.00	1.02

Table 3: GVIF Test Result for Model 4

Now that we have satisfied the assumptions for a large sample linear model, it is worthwhile to discuss the limitations of the resulting model. The model's limitations center around incompleteness either in the training data or in the model. As our data comes from Tiffany & Co's price list, not all colors and clarities of diamonds are covered. This limits the type of diamonds against which our model can reliably predict price. Moreover, the values in this list utilize the GIA diamond grading system. Yet, there are numerous other grading systems recognized around the world [7]. Again, this limits the pool of diamonds against which our model can be leveraged. Perhaps more importantly, our model is likely subject to a missing variable bias. Our dataset and consequently our model does not incorporate diamond shape as a reported category. However, there is a seemingly strong argument to be made that diamond shape does have a pertinent impact on the pricing of a diamond [8]. All in all, though, given the results of our model, its predictive capacity would likely still provide meaningful practical insights for our client.

Discussion

Our analysis clearly indicates the significance of carat weight and how it determines the diamond's selling price. Carat alone explains most variance in our model with an adjusted R2 of 0.93, underpinning its importance in the model. With all other confounding variables included in our final model we were able to explain an additional 5% of the variance and found that a 1% increase in carat weight could increase the diamond's selling price by approximately 1.89%. Given the importance of carat weight in determining price, we advise the DRC to heavily prioritize carat weight when cutting diamonds for luxury retailers.

As the model currently stands it is a good start to understanding the luxury diamond market for diamond cutters. We recommend using this model to inform your cutters on the general trends of cut diamond physical characteristics and their selling price in the luxury market. Additionally, when utilizing the information from this model it is pertinent to keep in mind the limitations of the model and how it could benefit from several improvements. Improvements such as the recency of data, the inclusion of additional cut diamond characteristics, and the inclusion of price data from additional luxury diamond retailers would all aid in expanding our results in the luxury retail market and the accuracy of our model. With DRC's cutters equipped with this knowledge they should be much better prepared to balance their cut diamond's properties and get a better selling price in the luxury diamond market.

References

1. ASHI. (n.d.). Journey of a Diamond. ASHI Diamonds. Retrieved [July 28, 2023] from https://www.ashidiamonds.com/Education/DiamondGuide/Journey_of_Diamond.aspx
2. GIA. (n.d.). Diamond History and Lore. Gemological Institute of America Inc. Retrieved [July 26, 2023], from https://www.gia.edu/diamond-history-lore
3. GIA. (n.d.). Diamond History and Lore. Gemological Institute of America Inc. Retrieved [July 26, 2023], from https://4cs.gia.edu/en-us/4cs-diamond-quality/
4. Clark Donald. (n.d.). Diamond Rough and Equipment for the Occasional Diamond Cutter. International Gem Society. Retrieved [July 28, 2023] from https://www.gemsociety.org/article/diamond-rough-equipment-occasional-diamond-cutter/
5. Luke Amanda J. (January 14, 2016). Diamond Cut: The Wow Factor. Gemological Institute of America Inc. Retrieved [July 28, 2023], from https://www.gia.edu/diamond-cut/diamond-cut-basic-overview
6. Shang Phoebe. (n.d.). Diamond Buying and the Four Cs, Part 3: Evaluating Diamond Cuts. Gemological Institute of America Inc. Retrieved [July 27, 2023], from https://www.gemsociety.org/article/diamond-cuts/
7. https://www.brides.com/diamond-certification-5112923
8. https://blog.jamesallen.com/diamonds/how-does-diamond-shape-affect-its-price/?gclid=CjwKCAjw8symBhAqEiwAa
_H-fUlWMJezjG9LhCThm-qgMsHRGOfdMFi5562YVyjHxDqc08d01txoC8PIQAvD_BwE