MGT 6203 Group Project Final Report

The Fair Price of Diamonds

Team 14
William Scheib
Jae Wook Yoo
Cindy Hoang
Suen Kwon
Michael Salvatore DeMatteo

Contents

Total gradable page-count: 10 (2-11)

Background

There are many reasons for purchasing or owning diamonds, but the ultimate reasoning is the perceived value of diamonds being rare, expensive, and meaningful. Afterall, the slogan “A diamond is forever” was created and popularized. In today’s diamond market, there are two mining companies with a combination of over 55% market share: Alrosa and DeBeers. Alrosa, Russia’s largest diamond miner, currently leads the world in diamond output with a production of 32.4 million carats in 2021. DeBeers, the world’s largest supplier of rough stones, once controlled 90% of the world’s diamond production and had a production of 32 million carats in 2021. With both companies in control of the diamond market, an oligopoly market was constructed, creating a diamond market with economical disadvantages for consumers and investors.

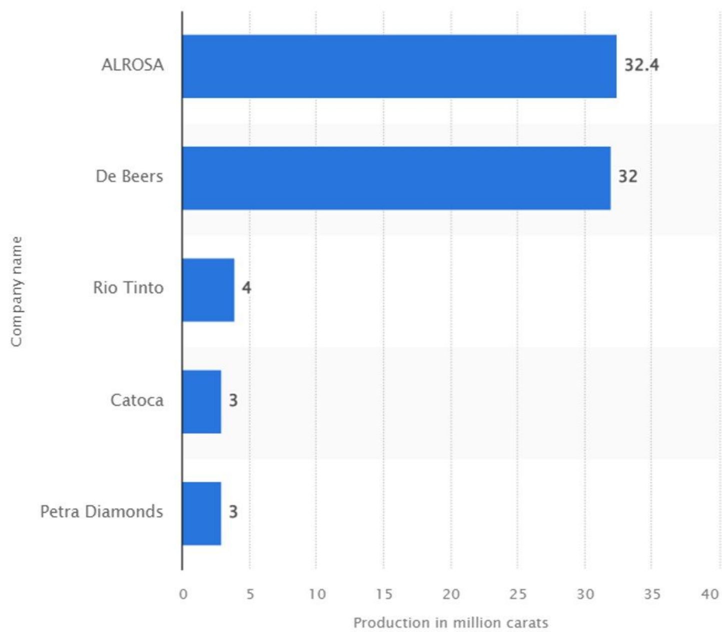


Figure 1. Diamond Mine Production of the Leading Companies Worldwide in 2021

Generally, there are several factors that contribute to the varying prices of diamonds. This includes rarity, cut, color, clarity, market demand, and marketing efforts. Carat is a standard unit of measurement for weight in gemstones. The larger the carat, the more valuable the diamond. Cut refers to the way diamonds are shaped, which is directly related to the amount of sparkle a diamond has or how well the diamond reflects light. Color is based on the diamond grading scale which ranges from D (colorless) to Z (light yellow or brown), with D being the highest grade and the rarest color of diamonds. Clarity is determined by the presence or absence of imperfections and inclusions, with the highest perceived value being little to no blemishes or inclusions.

Although it is important to note that the value of diamond is subjective and can be influenced by a range of factors, even among diamonds with similar characteristics, there may be other factors influencing the prices, such as the control of diamond supply. Due to other factors impacting the diamond market, there seems to be a lack of transparency and standardization in diamond pricing. Analyzing the diamond market can provide important insights into the industry, support ethical and sustainable practices, and help ensure fair pricing and transparency for consumers and investors. It can also help us determine which attributes impact the prices of diamonds the most and whether the diamond industry being an oligopoly would affect available consumer information.

Literature Review

Within the "*Markets: Continuity and Change in the International Diamond Market*", explanations pertaining to the past diamond market patterns and the effects of supply and demand were shown. At the time, the top diamond-producing organizations all adapted to an explicit set of rules set by the dominant company, DeBeers, the parent company of the Diamond Trading Company (previously known as the Central Selling Organization) who created a diamond monopoly. The rules stated that these production organizations must "manage their production in line with expected demand, stockpile excess stones, and sell the bulk of their rough diamonds to the Diamond Trading Company" (Spar 196). These efforts continued for well over a century, to control the diamond market and maximize its long-term prospects. As seen in Figure 2, this worked in favor of DeBeers and the diamond-producing organizations.

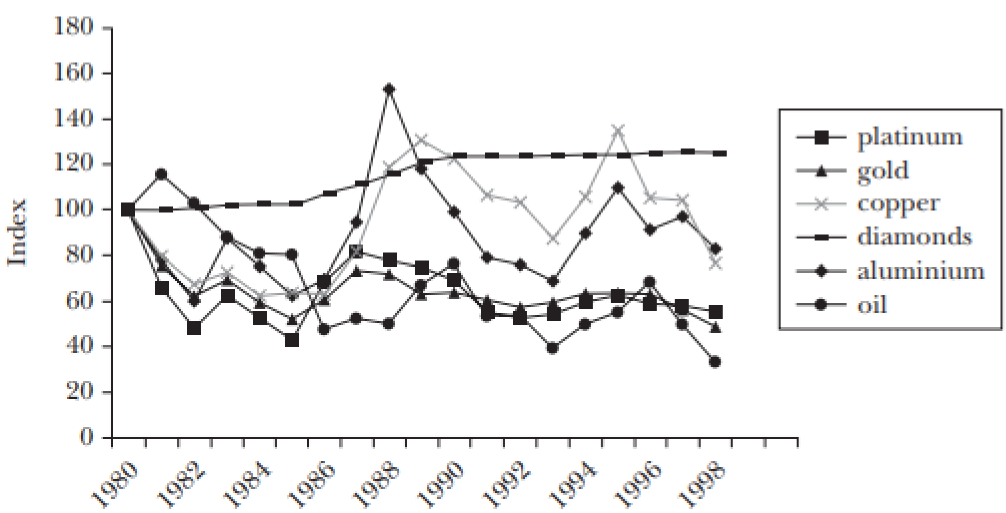


Figure 2. Commodity Prices between 1980-1998

Although these rules were widely understood, there were a few occurrences when the supply and demand of diamonds were threatened, and DeBeers responded to it by controlling the market. DeBeers knew that if diamonds were common, their demand and selling point as being luxurious would decline. In the 1970s, diamond prices began to rise due to diamond dealers converting financial assets into diamonds. Because DeBeers didn't want to risk the price decrease that tends to follow a speculative increase, DeBeers placed a drastic price hike on diamonds sold by the Central Selling Organization. This stopped dealers from pursuing this venture because of its riskiness. Later, when diamond sales began to drop, DeBeers spent hundreds of millions of dollars to support diamond prices. Another occurrence was when the global diamond industry was hit by a huge demand, and diamond sales rose by 6 percent. This was perceived as a threat to the diamond market, thus the CSO raised its rough diamond prices by 14 percent over the course of the year, forcing other major producers to follow suit.

All of these efforts were made for the stability the diamond market provides to its investors. It comes from DeBeers' efforts of convincing its consumers that diamonds are valuable, sentimental, and scarce. If either supply or demand were to change rapidly, scarcity and value would diminish, shattering the allurement of diamonds to consumers. Through all of DeBeers efforts to control the diamond market, the general rule became that diamonds never declined in value, and the diamond industry remained consistently profitable.

To collect more information about the factors that may affect diamond prices and whether there is a need to pursue an analysis of the diamond market, a research study *"Subjectivity of Diamond Prices in Online Retail: Insights from a Data Mining Study*" was used. This study explored the relationship between the physical properties, in particular, the various attributes such as carat weight, cut, color, and clarity, to determine which factors affected diamond prices the most and to what extent. It revealed that the various attributes did have an effect on the diamond market; however, those attributes were not the only factors influencing the prices of diamonds. The study suggests that a high degree of subjectivity in diamond pricing exists largely due to different obscure strategies being employed by diamond retailers, including the retailer's pricing strategy, market demand, and the availability of similar diamonds. It also suggests that the subjectivity of diamond prices may be due to the lack of transparency and information asymmetry in the market.

In conclusion, an analysis of the diamond market is deemed necessary for both investors and consumers. There is a need for greater transparency and information disclosure in the diamond market. Stakeholders, including retailers and industry organizations, may need to consider strategies to improve transparency and provide consumers with more information about diamond prices and quality.

Objective and Problem Statement

With the increase in diamond prices and the imbalance between supply and demand, there seems to be a lack of transparency and standardization in diamond pricing that has led to concerns that consumers may be paying more for diamonds than they are actually worth. Therefore, the aim of this study is to assess whether there is a correlation between the quality of diamonds and their market value, and to identify any factors that may be contributing to price discrepancies. Additionally, this study can help us determine which attributes impact the prices of diamonds the most, explain why certain attributes affect the prices of diamonds more than others qualitatively and empirically, and resolve whether the diamond industry being an oligopoly would affect available consumer information.

Our team believes an analysis of the diamond market is critical because the prices for diamonds are constantly growing at a rapid rate in comparison to other commodities. The global market generated an average of 77 billion dollars within the last 13 years. Additionally, the global diamond market is expected to have a growth rate of 7.5% through 2027, with an expected annual increase in market size of 3.41 billion dollars. This rate is almost double the average inflation rate (3.8% per year) between 1960-2021 in the United States.

Furthermore, the demand for diamonds continues to grow. Diamonds and similar keywords (Kays Jewelry, engagement rings, diamond ring, diamond earrings, and engagement rings for women) have an average monthly search and interest of 100,000-1,000,000 over the last two years, which confirms the demand, sentiment, and importance of diamonds. Additionally, as seen in Figure 3, the demand for diamonds is predicted to grow over the next few years.

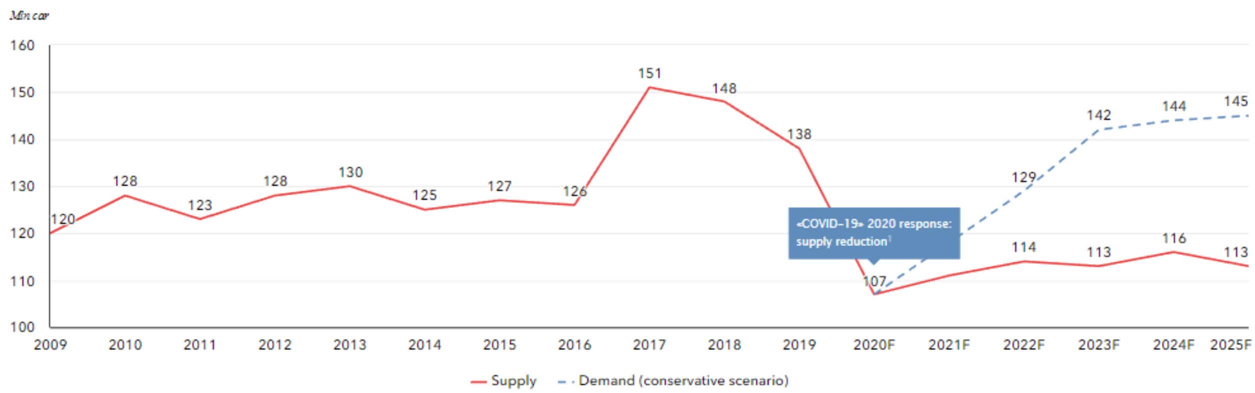


Figure 3. Projected Demands for Diamonds (Conservative Scenario)

By analyzing the diamond market, our team can help increase the transparency and standardization in diamond pricing, ensuring that consumers are getting a fair price for their diamonds. Additionally, the diamond market is an important investment sector, and analyzing the market can help investors make informed decisions about trading diamonds by understanding the factors that contribute to the prices of diamonds and identifying opportunities for economic growth and development.

Initial Hypotheses

Our team hypothesizes that the prices of diamonds are largely profit-driven and that the correlation between price and quality is not strong. Quality refers to the characteristics of the factors (cut, clarity, color, and carat weight) used in our analysis. Since there are many external factors that fluctuate over time (market demand, market supply, inflation, and social changes), we believe that as time progresses, the disparity between quality and price has grown larger in magnitude due to the lack of transparency and standardization in the diamond market.

Data (Available Data, Data Wrangling, and Transformation)

Expanding on our initial research and theory, we looked to open sources to find consumable data that would allow us to tackle this problem from an empirical approach. Our discovery concluded with finding two open-source data sets that provide a wealth of insight. An initial overview of these data sources is as follows:

Our first dataset–diamonds.csv–was pulled from an open-source data sharing platform, Kaggle. This dataset includes various information and factors of 54,000 diamonds. Various key attributes within this dataset include diamond prices, depth, table, carat, cut, color, clarity, and diamond measurements in millimeters.

Our second dataset—diamond_index_transformed.csv—was sourced in a more complex manner. Our group leveraged Python to web scrape information from a large diamond price index website with diamond prices between January 2009 through January 2023. The nature of this data set is time-series, where the key variables are index.price and month/year. We also provided additional feature engineering which can be seen in more details below.

Our data cleaning process was quite involved. Due to the vast amount of data, we spent considerable time loading the data into R and performing initial exploration of the data. Initially, we performed numerous data validations such as checking for null values, ensuring data types were correct for analysis, and summarizing our data sets. Additionally, we also checked for structural errors within the datasets. We found several entry errors such as faulty data types and string inconsistencies. We were able to correct for this by leveraging the as.factor() R function to convert all data types to be usable in our analysis. We leveraged regular expressions in combination with gsub() to identify common character patterns and convert cut/color/carat/clarity into integer values. Lastly, we identified outliers from our dataset. Using the Z-score approach, we classified data points that fell outside the 3rd standard deviation as an outlier. Following this identification, we leveraged the mean/median imputation. In practice, this involved replacing the outlier price and quality integer values with the median value.

Additionally, we leveraged minor feature engineering successfully to convert raw character data into usable data for our analysis. In our second dataset, after performing the necessary Python operations, we were able to populate a .csv file with the web scraped time-series information. We leveraged basic text splitting and feature engineering to remove the index-integer column, which was essentially a row count, as well as split the Month column, which was in the format 'MM-YYYY' as type character, to two columns "Month" and "Year" which both were as type data. This ensured we would be able to aggregate values and compare and contrast indice performance over time, in the traditional sense month-over-month / year-over-year, as well as determine if there was implicit seasonality.

Approach, Models, and Methodology (Analysis Included)

Stated simply, we were most interested in leveraging our available data to understand if diamond prices could be predicted well, given the aforementioned key variables. Intuitively, our approach was to begin with better understanding our observations and isolating various variables to examine their relationship to diamond prices. As we thought of diamond quality, carat, color, cut, and clarity seemed the most important to investigate.

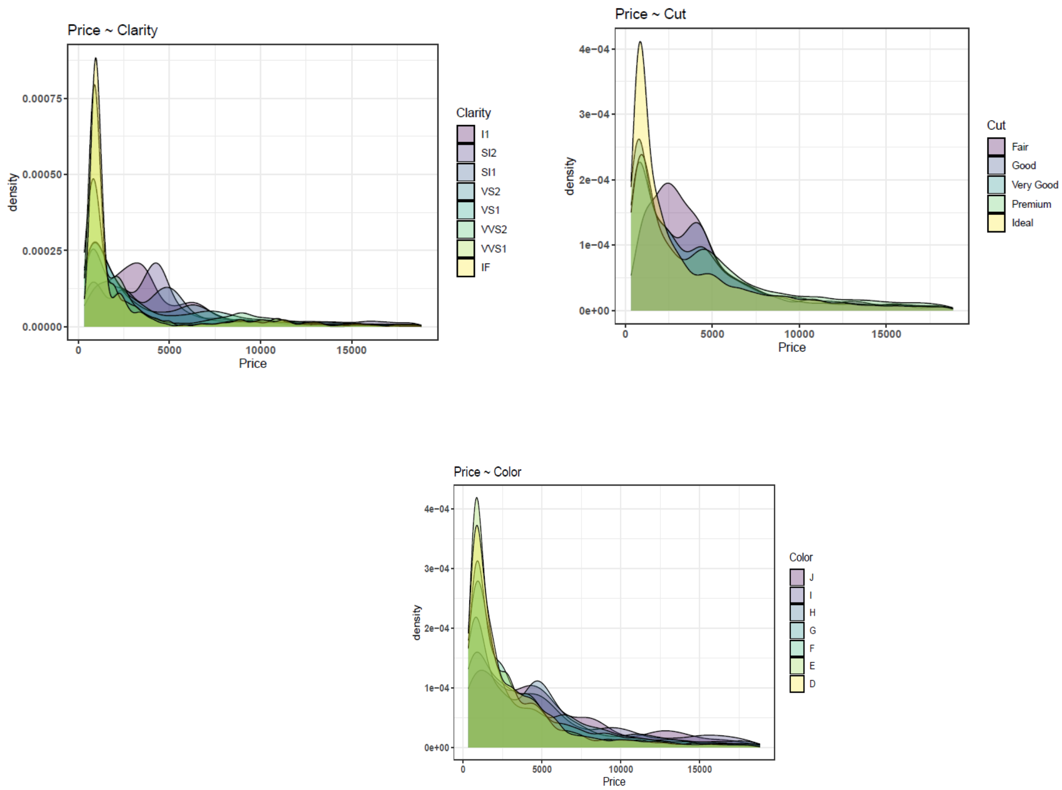


Figure 4. Density vs. Variables (Cut, Clarity, and Color)

Figure 4 was crucial to our modeling considerations. Contrary to our initial hypothesis, we saw several quality variables that had densities which outlined greater quality observations based on characteristics were tending to be associated with a higher price and noticeably so. As a result, we wanted to understand the relationship between our remaining variables deeper and performed correlation ggpairs as shown below.

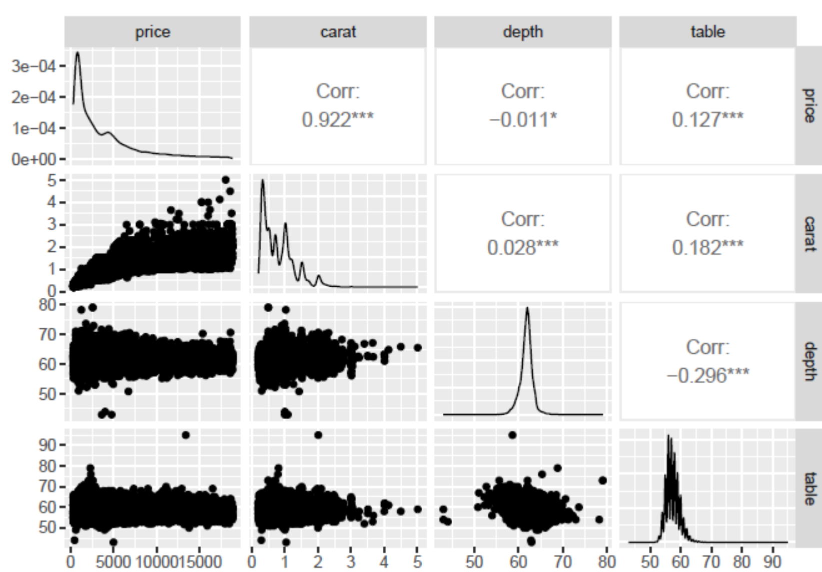


Figure 5. Correlations

The correlation between price and table as well as the correlation between price and depth were particularly underwhelming for us. However, we also saw a strong and statistically significant correlation between price and carat. Therefore, we decided to proceed with building our prediction models with the key variables: color, cut, clarity, and carat.

Our goal in modeling was to try and find the best model to predict diamond price. Our comparison criteria was 1) goodness-of-fit, 2) R^2 $both\ standard\ and\ adjusted$, and 3) qualitative intuition of which model interpretation would best address our problem statement.

We created four models based on our first data set which leveraged training and hyperparameter optimization. The models that we created were a linear-linear model, logarithmic-linear model, logarithmic(root)-linear model, and a root-linear model. Initially, we started with just a linear-linear model, however, we noticed heteroskedasticity and continued to try new model constructs/designs until we found the root-linear model to be the best. Additionally, across all of these models we leveraged 70% of the observations for training, 20% for testing, and 10% for validation.

Generally speaking, our models performed well. From the goodness-of-fit perspective, we were encouraged that our sample sets were fitting the observation sets well. Additionally, the R^2 $(both\ standard\ and\ adjusted)$ was extremely high across all of our models so we knew that our models were providing very strong approximations on the data. Additionally, our model selection and choice to leverage the square root-linear model was motivated by analytical intuition as we saw our dependent variable, price, had a skewed distribution and wanted to reduce heteroskedasticity for the sake of having a better model fit and R^2 value. The outputs of these four models can be examined below and will be explained in detail in the results section:

Linear-Linear

```
Call:
lm(formula = price ~ carat + cut + color + clarity, data = train)

Residuals:
     Min       1Q   Median       3Q      Max 
-12573.6   -679.9   -194.0    464.0  10362.4 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  -7479.31      62.77 -119.15   <2e-16 ***
carat         8890.94      14.25  624.00   <2e-16 ***
cutGood        684.43      39.84   17.18   <2e-16 ***
cutIdeal      1004.02      36.34   27.63   <2e-16 ***
cutPremium     866.68      36.67   23.63   <2e-16 ***
cutVery Good   855.58      37.10   23.06   <2e-16 ***
colorE        -225.75      21.65  -10.43   <2e-16 ***
colorF        -305.45      21.89  -13.96   <2e-16 ***
colorG        -504.13      21.42  -23.54   <2e-16 ***
colorH        -963.45      22.72  -42.41   <2e-16 ***
colorI       -1457.61      25.56  -57.03   <2e-16 ***
colorJ       -2331.38      31.59  -73.79   <2e-16 ***
clarityIF     5556.78      63.15   87.99   <2e-16 ***
claritySI1    3682.45      54.53   67.53   <2e-16 ***
claritySI2    2730.34      54.75   49.87   <2e-16 ***
clarityVS1    4634.08      55.61   83.34   <2e-16 ***
clarityVS2    4324.45      54.82   78.88   <2e-16 ***
clarityVVS1   5184.74      58.64   88.41   <2e-16 ***
clarityVVS2   5086.88      57.20   88.93   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1144 on 37739 degrees of freedom
Multiple R-squared:  0.9175,	Adjusted R-squared:  0.9175 
F-statistic: 2.332e+04 on 18 and 37739 DF,  p-value: < 2.2e-16
```

Log-Linear

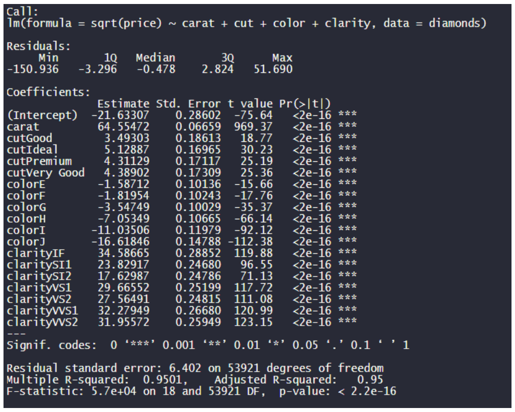
```
Call:
lm(formula = log(price) ~ carat + cut + color + clarity, data = train)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.1625 -0.2181  0.0567  0.2488  1.0903 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)   5.321071   0.018482 287.906  < 2e-16 ***
carat         2.205617   0.004195 525.769  < 2e-16 ***
cutGood       0.047640   0.011731   4.061 4.90e-05 ***
cutIdeal      0.080751   0.010700   7.547 4.56e-14 ***
cutPremium    0.053645   0.010797   4.968 6.78e-07 ***
cutVery Good  0.055710   0.010922   5.101 3.40e-07 ***
colorE       -0.055256   0.006374  -8.669  < 2e-16 ***
colorF       -0.058964   0.006444  -9.150  < 2e-16 ***
colorG       -0.131356   0.006306 -20.829  < 2e-16 ***
colorH       -0.269113   0.006688 -40.236  < 2e-16 ***
colorI       -0.426171   0.007524 -56.639  < 2e-16 ***
colorJ       -0.580320   0.009302 -62.388  < 2e-16 ***
clarityIF     1.066824   0.018593  57.378  < 2e-16 ***
claritySI1    0.754953   0.016054  47.025  < 2e-16 ***
claritySI2    0.575869   0.016121  35.722  < 2e-16 ***
clarityVS1    0.913574   0.016372  55.801  < 2e-16 ***
clarityVS2    0.852341   0.016142  52.804  < 2e-16 ***
clarityVVS1   0.978093   0.017266  56.650  < 2e-16 ***
clarityVVS2   0.964590   0.016841  57.277  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3368 on 37739 degrees of freedom
Multiple R-squared:  0.8901,	Adjusted R-squared:   0.89 
F-statistic: 1.698e+04 on 18 and 37739 DF,  p-value: < 2.2e-16
```

Root-linear



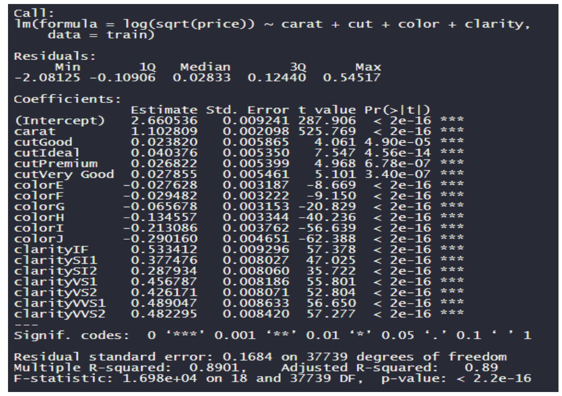
```
Call:
lm(formula = sqrt(price) ~ carat + cut + color + clarity, data = diamonds)

Residuals:
     Min       1Q   Median       3Q      Max
-150.936   -3.296   -0.478    2.824   51.690

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -21.63307    0.28602  -75.64   <2e-16 ***
carat         64.55472    0.06659  969.37   <2e-16 ***
cutGood        3.49303    0.18613   18.77   <2e-16 ***
cutIdeal       5.12887    0.16965   30.23   <2e-16 ***
cutPremium     4.31129    0.17117   25.19   <2e-16 ***
cutVery Good   4.38902    0.17309   25.36   <2e-16 ***
colorE        -1.58712    0.10136  -15.66   <2e-16 ***
colorF        -1.81954    0.10243  -17.76   <2e-16 ***
colorG        -3.54749    0.10029  -35.37   <2e-16 ***
colorH        -7.05349    0.10665  -66.14   <2e-16 ***
colorI       -11.03506    0.11979  -92.12   <2e-16 ***
colorJ       -16.61846    0.14788 -112.38   <2e-16 ***
clarityIF     34.58665    0.28852  119.88   <2e-16 ***
claritySI1    23.82917    0.24680   96.55   <2e-16 ***
claritySI2    17.62987    0.24786   71.13   <2e-16 ***
clarityVS1    29.66552    0.25199  117.72   <2e-16 ***
clarityVS2    27.56491    0.24815  111.08   <2e-16 ***
clarityVVS1   32.27949    0.26680  120.99   <2e-16 ***
clarityVVS2   31.95572    0.25949  123.15   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.402 on 53921 degrees of freedom
Multiple R-squared:  0.9501,    Adjusted R-squared:   0.95
F-statistic: 5.7e+04 on 18 and 53921 DF,  p-value: < 2.2e-16
```

Log(root)-linear



```
Call:
lm(formula = log(sqrt(price)) ~ carat + cut + color + clarity,
    data = train)

Residuals:
     Min       1Q   Median       3Q      Max
-2.08125 -0.10906  0.02833  0.12440  0.54517

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.660536   0.009241 287.906  < 2e-16 ***
carat         1.102809   0.002098 525.769  < 2e-16 ***
cutGood       0.023820   0.005865   4.061 4.90e-05 ***
cutIdeal      0.040376   0.005350   7.547 4.56e-14 ***
cutPremium    0.026822   0.005399   4.968 6.78e-07 ***
cutVery Good  0.027855   0.005461   5.101 3.40e-07 ***
colorE       -0.027628   0.003187  -8.669  < 2e-16 ***
colorF       -0.029482   0.003222  -9.150  < 2e-16 ***
colorG       -0.065678   0.003153 -20.829  < 2e-16 ***
colorH       -0.134557   0.003344 -40.236  < 2e-16 ***
colorI       -0.213086   0.003762 -56.639  < 2e-16 ***
colorJ       -0.290160   0.004651 -62.388  < 2e-16 ***
clarityIF     0.533412   0.009296  57.378  < 2e-16 ***
claritySI1    0.377476   0.008027  47.025  < 2e-16 ***
claritySI2    0.287934   0.008060  35.722  < 2e-16 ***
clarityVS1    0.456787   0.008186  55.801  < 2e-16 ***
clarityVS2    0.426171   0.008071  52.804  < 2e-16 ***
clarityVVS1   0.489047   0.008633  56.650  < 2e-16 ***
clarityVVS2   0.482295   0.008420  57.277  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1684 on 37739 degrees of freedom
Multiple R-squared:  0.8901,    Adjusted R-squared:   0.89
F-statistic: 1.698e+04 on 18 and 37739 DF,  p-value: < 2.2e-16
```

Figure 6. Models

In the context of our problem statement and supporting research questions, our models have real-life value. Across all of these prediction models, we can see that the key quality variables are determined to be statistically significant and mirror the correlation and density exploration. The highly desired characteristics, such as premium cut and better clarity, tend to relate to price. From a consumer's perspective, these models provide clarity and transparency. A customer could request all of the diamond characteristics and leverage one of our models to see what they should pay. Additionally, the parameter estimates provide great value to the investment sector in understanding what characteristics of diamonds have the best prediction of a diamond's price.

Results and Analysis cont:

As touched on at length during the modeling section, our models greatly show that diamond pricing and diamond quality were related. We determined from our model outputs that diamond price was most largely driven by the diamond carat value. Within Figure 7, across all of our models, the diamond quality characteristics were each statistically significant. Within the context of our chosen model, the interpretation is as follows:

$Y_{1/2} = \beta_0 + \beta_1 + X_1 + \beta_2 X_2 + \ldots\ldots + \varepsilon$

Our model predicts that a one unit change in

X_i *will be associated with a change in Diamond Price* (Y) *of* $2\beta_i * Y^{1/2} = 2\beta_i (0 + \beta 1 + X1 + \beta 2X2 + \ldots$

For instance, $2\beta_{carat} \rightarrow$ 2(66.4) = 132.8 indicates that a unit change in diamond carat is associated with a change in diamond price 132.8 times the square of the current diamond price

---

While this prediction can be mathematically rigorous, this model allows us to examine how incremental changes in diamond quality variables impact diamond price. However, we were still interested in understanding how diamond pricing changes have affected consumers.

We decided to perform additional analysis by seeing how diamond price has changed in recent years, leveraging our second dataset.

Figure 7. Diamond Index Price between 2006-2023

As seen in Figure 7, diamond index price (which accounts for inflation and provides the diamond price in real US dollars) has increased by ~30% since 2006. However, diamond quality has remained largely unchanged. This indicates that while diamond pricing is significantly impacted by diamond quality, there still exists a large premium that DeBeers and other industry leaders force consumers to pay. If we were to have more time throughout the semester, we would want to explore the time-series relationship and implications in more detail. Additionally, our group was generally curious if we could estimate or approximate the delta between the predicted value and actual value if this would be helpful in computing the premium value.

In a different vein, despite our initial hypothesis and intuition being that diamond pricing would not be correlated with diamond quality, our methodology and approach throughout the project led us to conduct numerous analyses that provide some evidence that key variables or characteristics associated with diamond quality do impact price.

As a result, we believe that our research, literature understanding, as well as modeling work further the conversation on diamond pricing. Our results are not exceptionally useful by themselves, however, coupled with the business justification, they greatly provide consumer confidence that there are standards of quality with respect to price within this market. Additionally, it also begs the question that despite deliverance of uniform quality, how are diamond premiums and non-measurable factors driving diamond prices.

Future Research

If given more time and available data, we would like to spend additional time and consideration around consumer research and diamond price premiums. Our team originally hypothesized that the prices of diamonds were largely profit-driven and that the correlation between prices and the quality of diamonds were not strong. Although, as seen in this analysis, diamond quality and prices are highly correlated, we would want to challenge this further and continue to analyze different factors (market demand, market supply, and social changes) that could also be affecting diamond pricing.

Conclusion

Throughout the semester, our group has been largely motivated to provide consumer value through the pursuits of research, understanding of existing literature, and conducting statistical analysis and modeling to uncover if diamond pricing is consistent with diamond quality. Despite our initial belief that economic structures and incentives would lead to a disparity between these two things, we have found concrete data-driven evidence that this is not the case.

Moreover, our analysis and investigation has validated claims and provided additional evidence that the characteristics of diamonds that consumers are principally concerned with as they think of quality do have statistically significant impacts on price.

Our supporting research questions were also addressed. Our model was able to highlight that throughout recent data observations diamond carat was the largest predictor of price and other diamond quality metrics such as color, cut, clarity were all statistically significant and important. Secondly, we also uncovered throughout our research that available information about this market is extremely sparse, implying that consumers have a need for additional literature and guidance which we hope our paper and findings will provide.

Additionally, we were able to leverage existing open source data sets to create various prediction models that allow us to directly tackle our problem statement. Namely, we wanted to identify if consumers are paying more for diamonds than they are actually worth and if diamond prices are correlated with their quality. From our work throughout the semester, we have strong evidence that diamond prices are consistent with their perceived quality. Additionally, we have proposed an idea that can potentially explain the rise in diamond prices since 2006 being an increase in diamond price premiums posed by industry leaders. As a result, it appears consumers are likely paying prices over what our model predicts and therefore likely overpaying. From an investors perspective, we feel accomplished in identifying important factors to consider when investing in diamonds.

Data Sources

Diamond Prices Index. Diamond Prices Index™ - Monthly Pricing Chart for Loose Diamonds. (n.d.). Retrieved April 16, 2023, from https://www.diamondse.info/diamonds-price-index.asp

Pedersen, U. T. (2023, February 1). *Diamonds*. Kaggle. Retrieved April 16, 2023, from https://www.kaggle.com/datasets/ulrikthygepedersen/diamonds

References

Diamond Market Size: Trends and Global Forecast to 2032. The Business Research Company. (n.d.). Retrieved April 15, 2023, from https://www.thebusinessresearchcompany.com/report/diamond-global-market-report

Inflation Rates in the United States of America. Worlddata.info. (n.d.). Retrieved April 15, 2023, from https://www.worlddata.info/america/usa/inflation-rates.php#:~:text=During%20the%20observation%20period%20from,year%20inflation%20rate%20was%206.0%25.

The 4Cs of Diamond Quality. The Diamond Center: Where Wisconsin Gets Engaged. (2019, January 23). Retrieved March 29, 2023, from https://www.thediamondcenter.com/about-us/the-4-cs/

Alrosa. (n.d.). *Outlook*. ALROSA. Retrieved April 15, 2023, from https://www.alrosa.ru/en/investors/results-reports/forecast/

Arora, G. (2021, July 18). *The Marketing Strategy that Made De Beers a Monopoly*. LinkedIn. Retrieved March 29, 2023, from https://www.linkedin.com/pulse/marketing-strategy-made-de-beers-monopoly-gaurav-arora-/

Garside, M. (2022, October 5). *Diamond Production by Top Companies Worldwide 2021*. Statista. Retrieved March 29, 2023, from https://www.statista.com/statistics/585369/diamond-production-of-top-diamond-companies

Garside, M. (2023, February 14). *Diamond Jewelry Global Market Value 2021*. Statista. Retrieved April 15, 2023, from https://www.statista.com/statistics/585267/diamond-jewelry-market-value-worldwide/

Goryainov, S. (2021, September 6). *Polished Diamonds and Inflation*. Rough & Polished. Retrieved March 29, 2023, from https://www.rough-polished.com/en/analytics/122625.html

International Gemological Institute. (n.d.). *Diamond 4cs*. IGI. Retrieved April 15, 2023, from https://www.igi.org/education/diamond-4cs

Mamonov, S., & Triantoro, T. (2018, May 1). *Subjectivity of Diamond Prices in Online Retail: Insights from a Data Mining Study*. MDPI. Retrieved March 29, 2023, from https://www.mdpi.com/0718-1876/13/2/10

Rice, A. (2022, July 9). *Different Diamond Shapes: Ultimate Guide with Size & Price Chart*. International Gem Society. Retrieved March 29, 2023, from https://www.gemsociety.org/article/diamond-shapes-price-size/

Spar, D. L. (2006). *Markets: Continuity and Change in the International Diamond Market*. Journal of Economic Perspectives. Retrieved March 29, 2023, from https://www.aeaweb.org/articles?id=10.1257%2Fjep.20.3.195