# Stakeholder - Diamond Outlet

Price Predictions

By Kahuma Walter

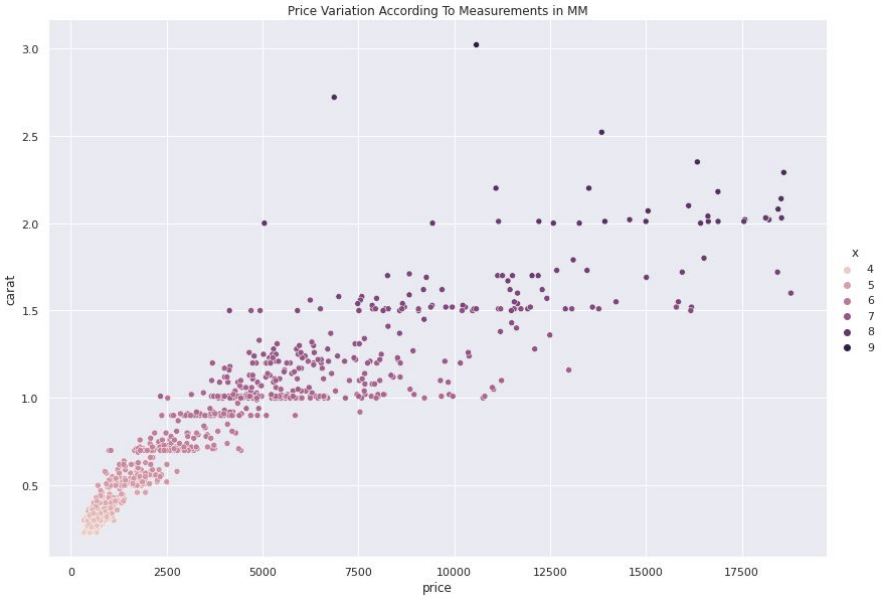
# Introduction

Business Problem - Price Predictions for a Diamond Store

There are 53,940 diamonds in the dataset with 10 features (carat, cut, color, clarity, depth, table, price, x, y, and z). Most variables are numeric in nature, but the variables cut, color, and clarity are categorical.

The dataset was visualized to give insights about the data and price predictions were made.

# Visualizations - Scattered Plot showing Price Variations




According to the scatter plot which sampled 1000 data points, the more the carats the higher the price of the diamonds. However there are some outliers where the carats were roughly high as 2.0 but were valued at 5000. There were also times when the carat was relatively low at 1.3 but the diamond was valued at 12,500USD.

The scatter plot also shows us the length(x) of the diamond which varies in color. As the length of the diamond grows the more expensive the diamond turned out to be.

However there are some outliers where length was 8mm but the diamond price was roughly low at 7,000 USD and when the length was 9mm (the highest) but the price was relatively low at about 11,000USD. While in some other situations the diamond length was 6mm but the price was relatively high at 13,000 USD.

# Visualization - Density Graph


According to the density graph above, the majority of the diamonds in our sample data have 0.2 to 0.8 carats as shown by the dark contours and these range between 300 and 2,500 USD. The graph also shows that the majority of the diamonds are less than 5,000 USD.

# Brief Description of The Model

It’s a regression problem meaning that it gives a continuous value of the target in this case being price.  I modelled the target against the other set predictor column features.

After using different models like Random Forest Regressor, Linear Regression, Decision Tree Regressor,, the Random Forest Regressor performed the best and  was able to predict 98% of the price data.

# Final Recommendations

- The diamond carat and length should be streamlined so that with every increase in carats or length, the more the diamond should cost.
- There are external factors which can affect the predictive model like insecurity, inflation and scarcity among others.

# Thank You