

Diamonds Price Predict

Fawzia Alsubaie
Deem Alshammri

Instructor
Dr. Mejdal Alqahtani

Outlines

INTRODUCTION

EXPLORATORY
DATA ANALYSIS

CONCLUSION

DATASET

REGRESSION
MODELS





Introduction

Our goal from this project is to perform a regression model to predict the  price.

Dataset

This dataset contains of 10 variables such as Carat, Price, Cut, Color, Clarity, depth and Table with 53940 rows.

Clean

- Drop some columns we don't need in our model training.
- Convert Price feature type from int to float.

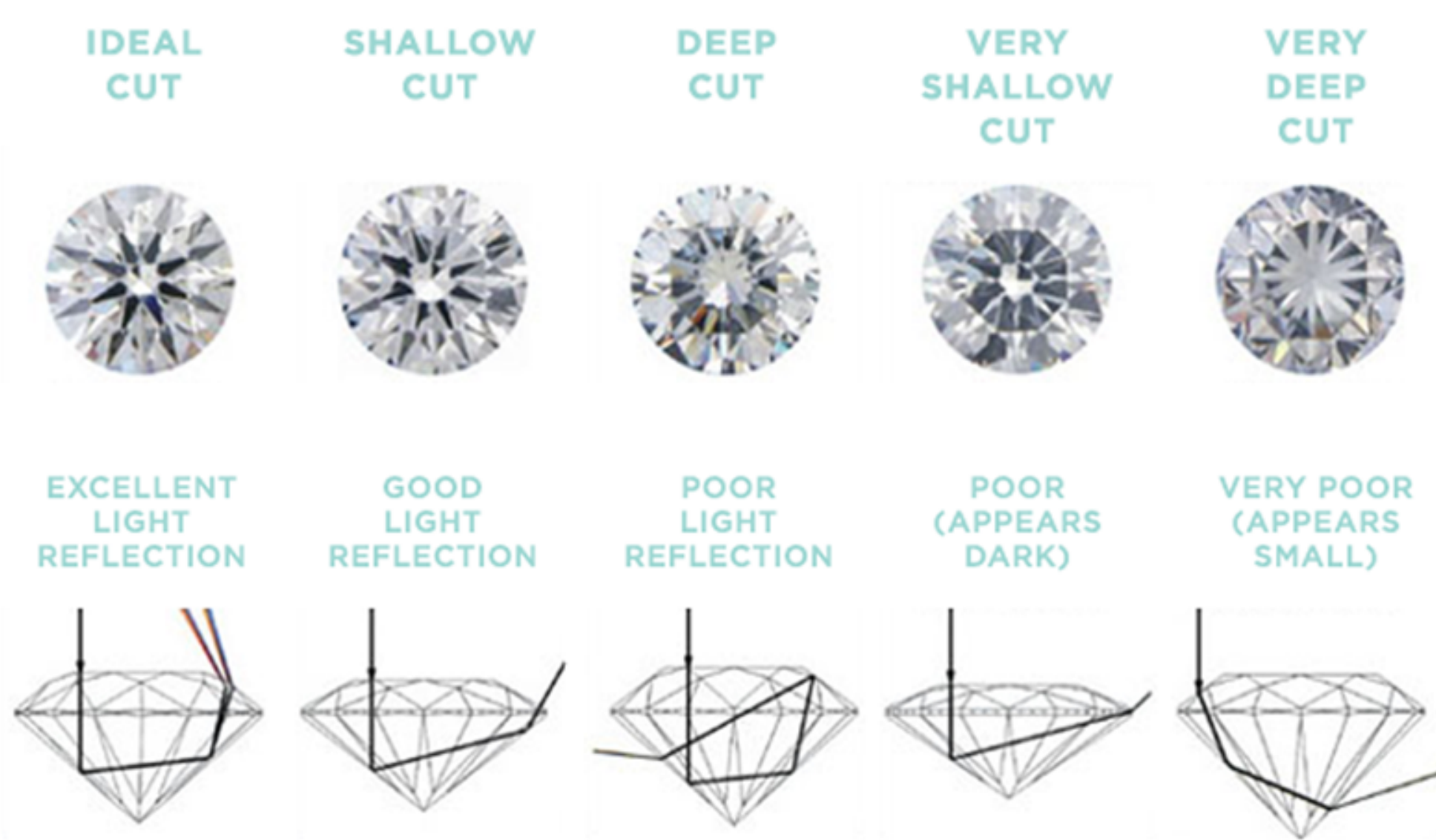
DataSet

	Unnamed: 0	carat	cut	color	clarity	depth	table	price	x	y	z
0	1	0.23	Ideal	E	SI2	61.5	55.0	326	3.95	3.98	2.43
1	2	0.21	Premium	E	SI1	59.8	61.0	326	3.89	3.84	2.31

After Clean :

	carat	cut	color	clarity	price	cut_lable	clarity_lable
0	0.23	Ideal	2	SI2	326.0	2	3
1	0.21	Premium	2	SI1	326.0	3	2

Diamonds Cut



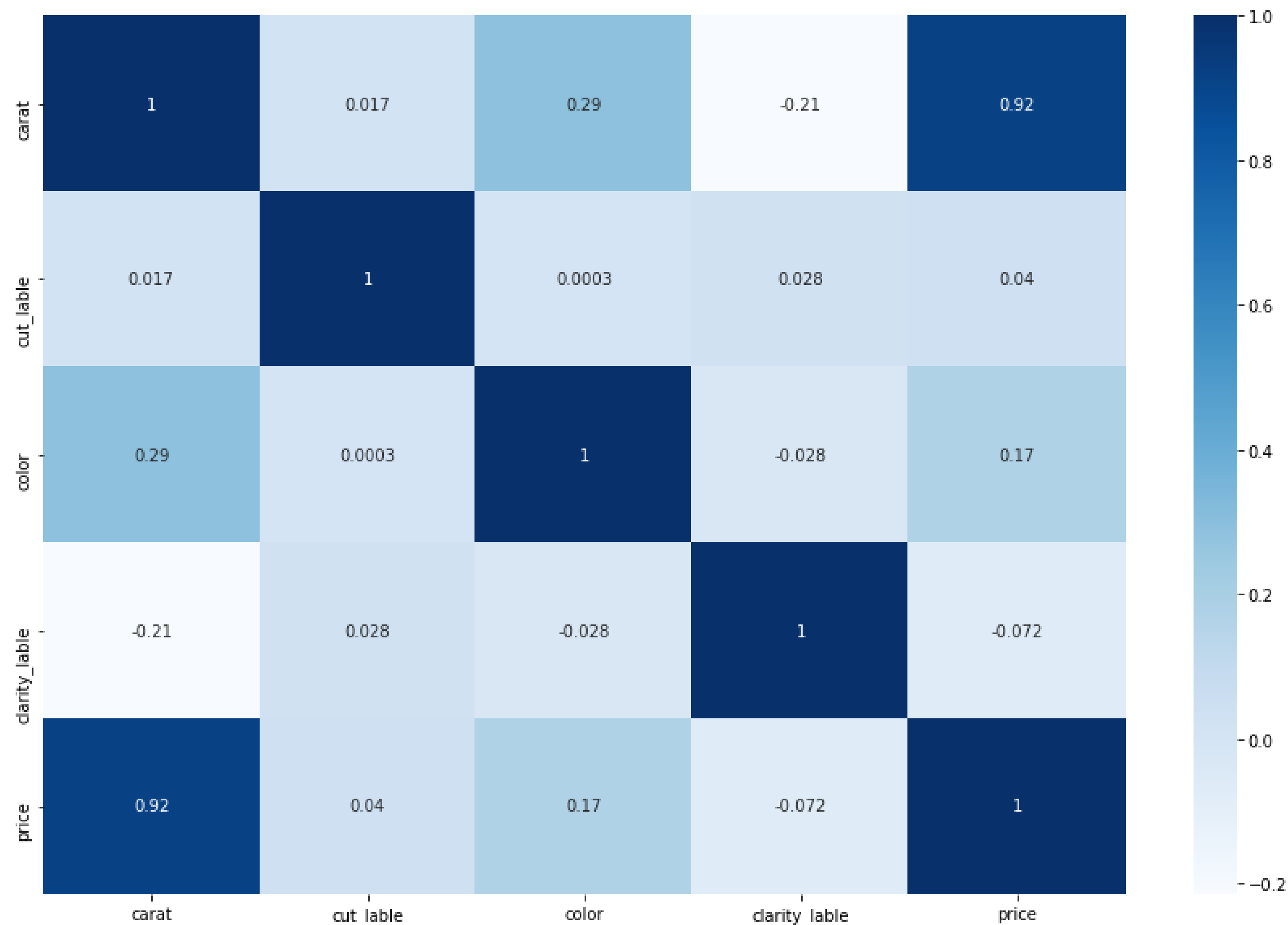
Diamonds Color



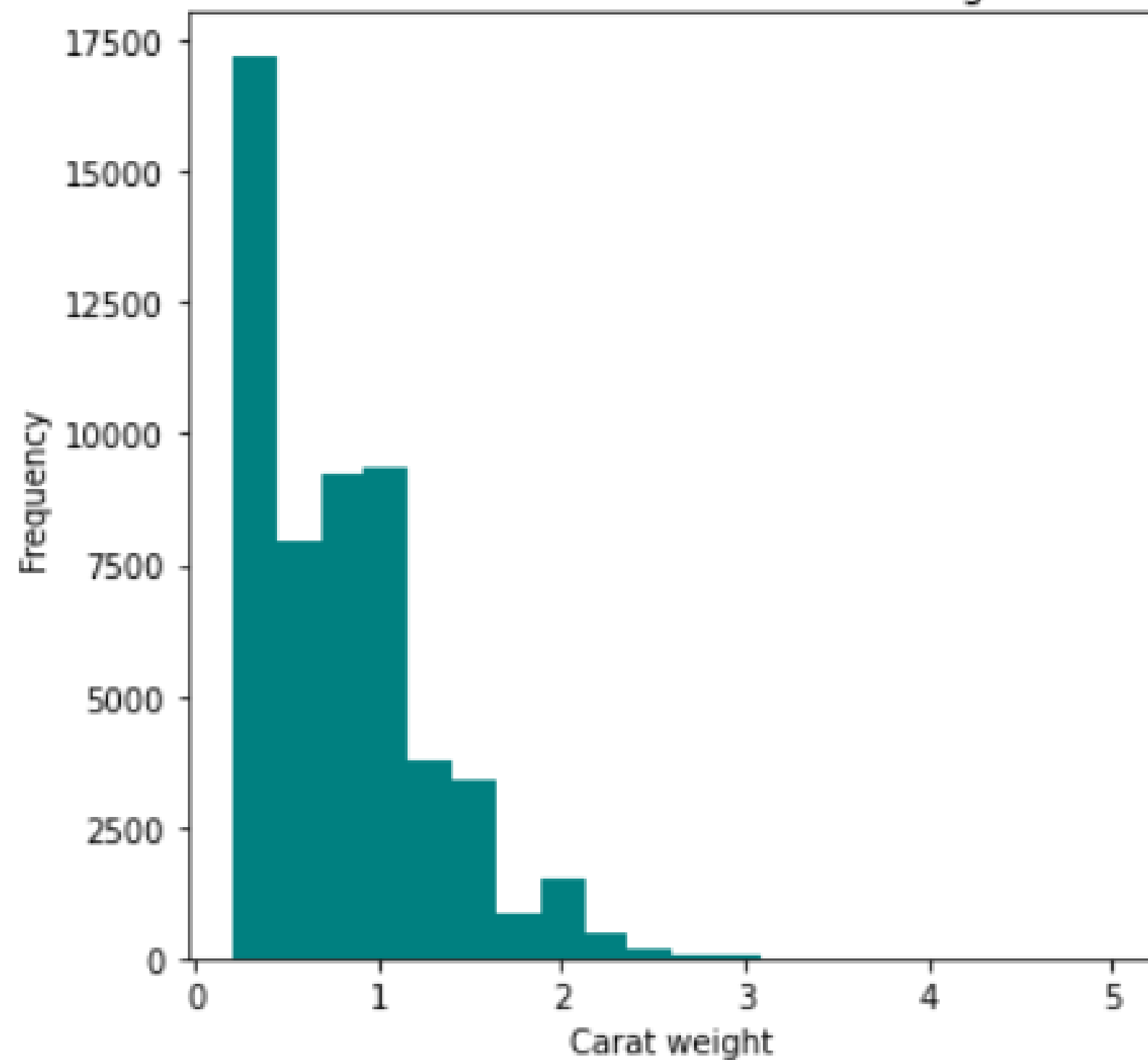
Exploratory Data Analysis

Heatmap

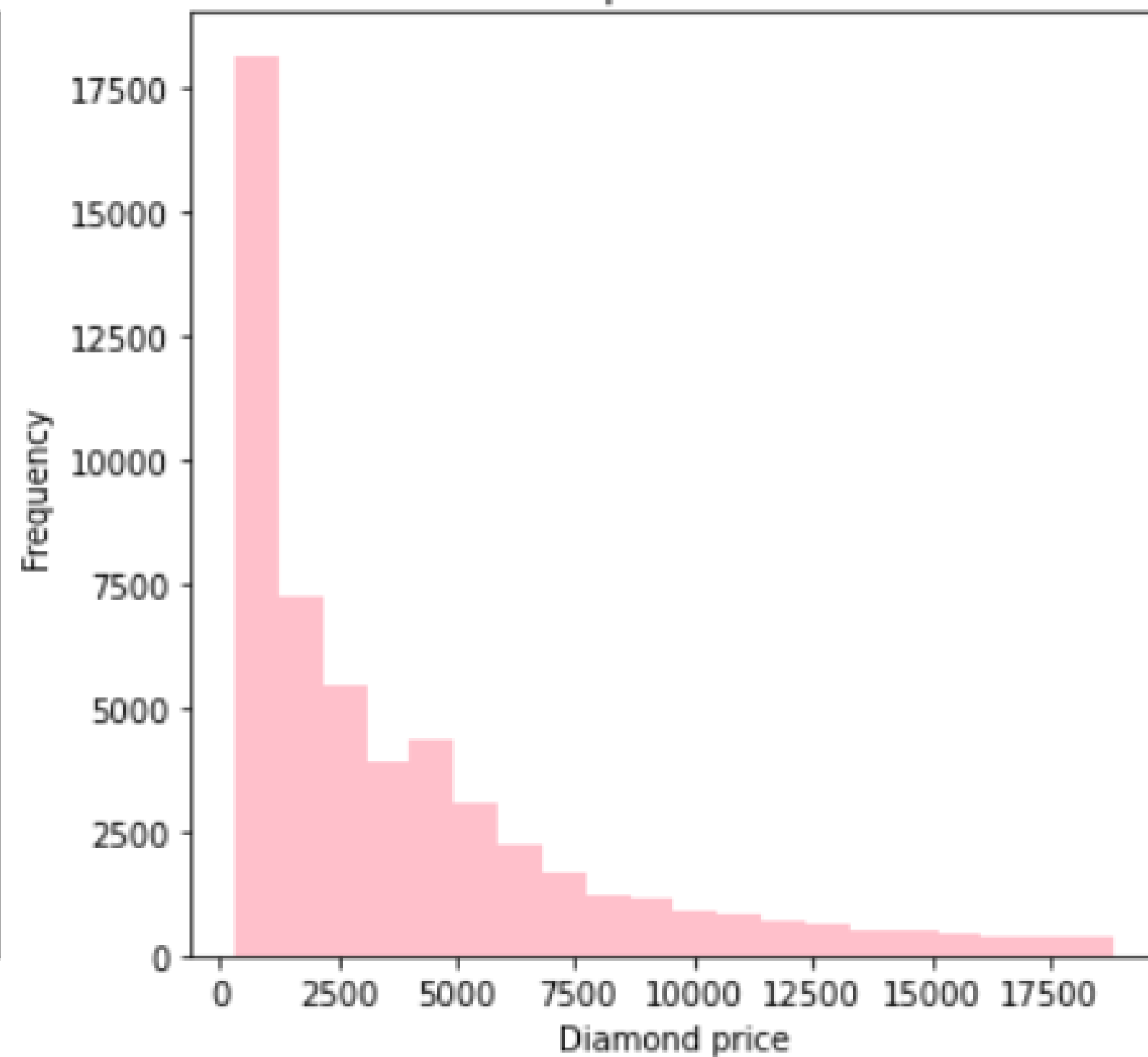
To shows the relations between the features. This result shows that Carat has the strong correlation with Price.



Distribution of diamond carat weight



Diamond price distribution

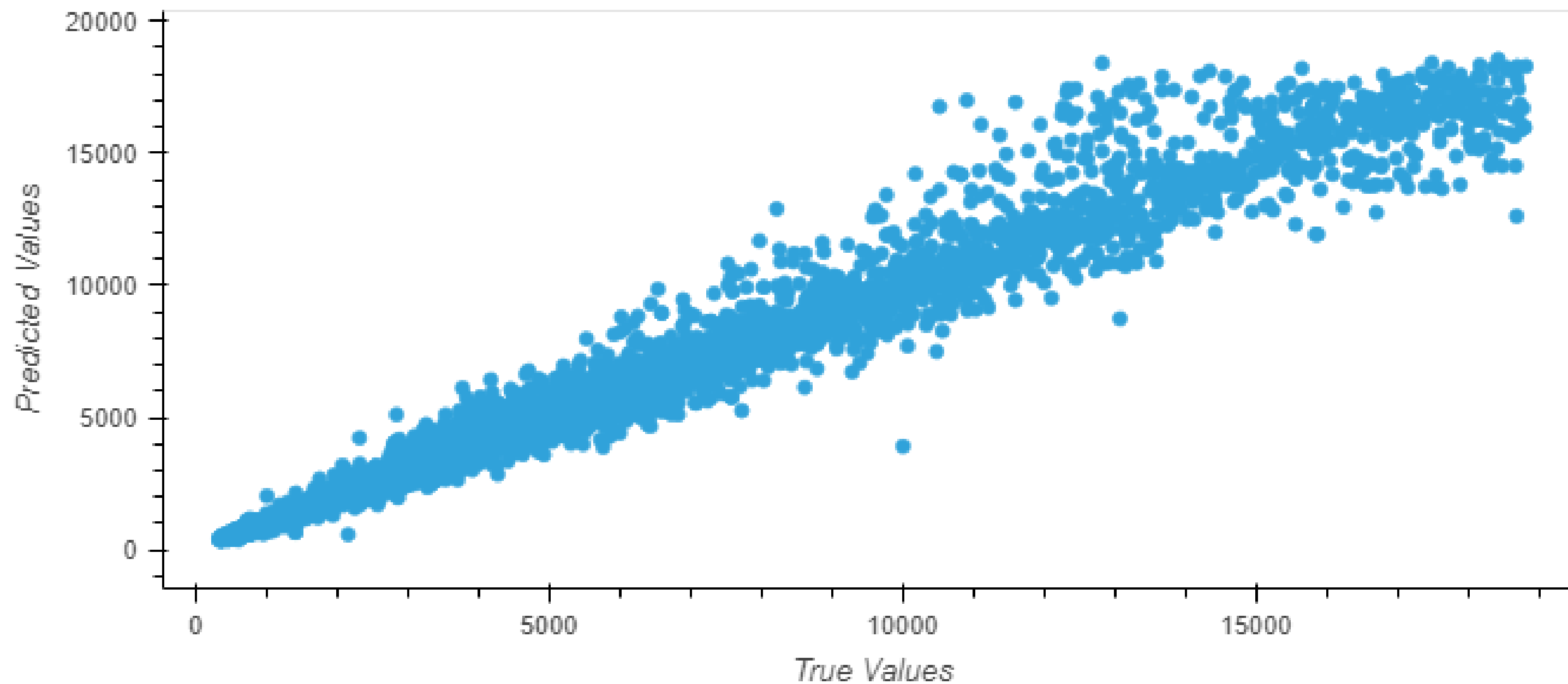


REGRESSION MODELS

Models	R2 Square
Linear Regression	0.877637
Polynomial Regression	0.843160
Random Forest	0.977686
Ridge	0.764608

The best Model to predict Diamonds price is Random Forest Regression

Prediction result





Conclusion

- Random forest model able to predict the diamond price.
- its depends strongly on the carat also clarity and the cut.

Thank you!

Shine bright like a Diamond!



Guess the Diamond!

