

A series of white, overlapping geometric lines and polygons in the top-left corner of the slide, creating a complex, abstract pattern.

GEMSTONE PRICE PREDICTION

Gemstone Price Prediction: Enhancing Profitability for Gem Stones Co Pvt. Ltd

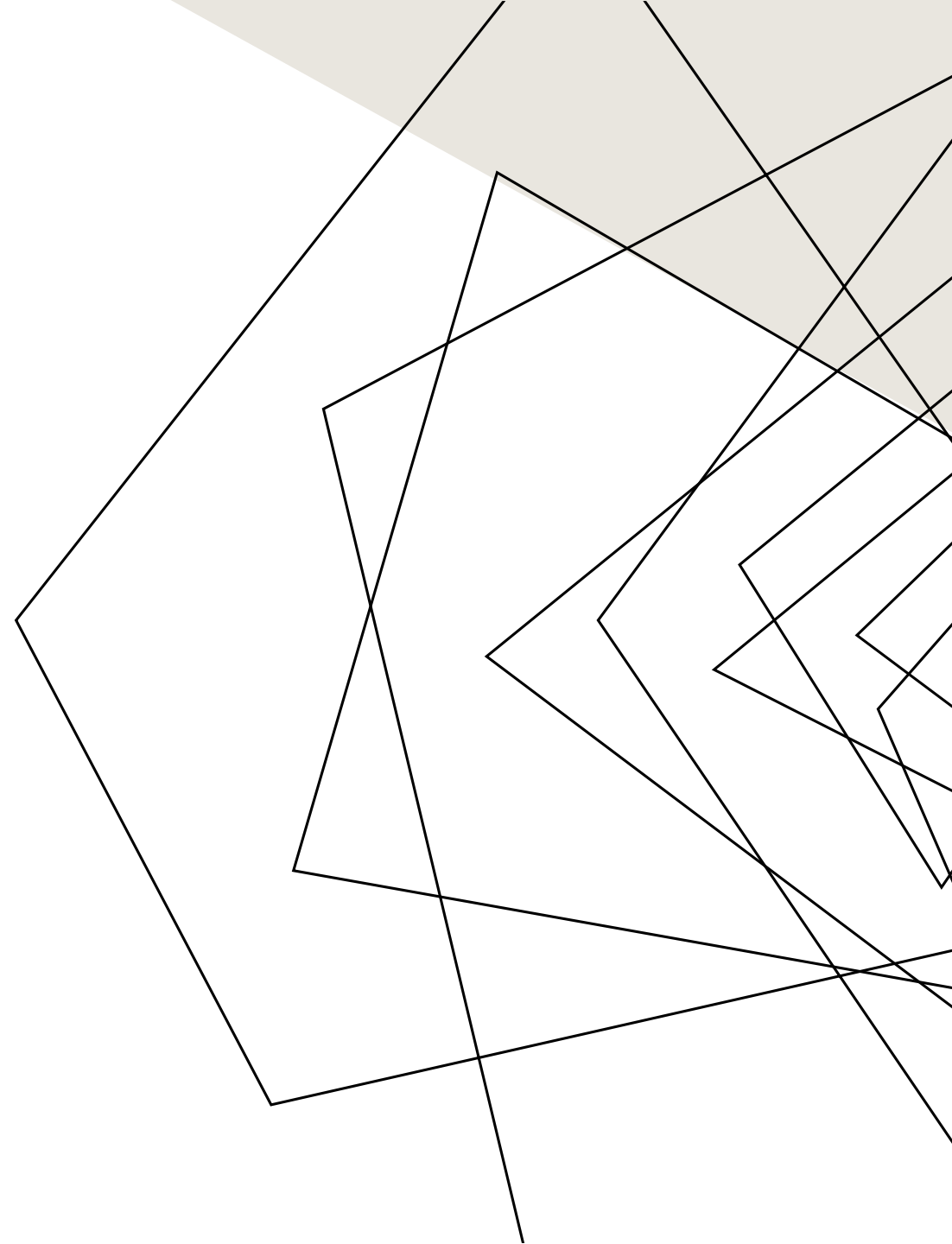
Presented by:

Team Data Vortex

(Harsh Shah | Pranali Sawant | Parag Khankari | Vidish Kale)

INTRODUCTION

Gem Stones Co Pvt. Ltd is a leading manufacturer of cubic zirconia, an affordable diamond alternative with similar qualities. The company aims to optimize its profits by accurately predicting gemstone prices based on various attributes provided in the dataset. This project focuses on building predictive models to distinguish between higher and lower profitable stones, ultimately improving the company's profit margins.

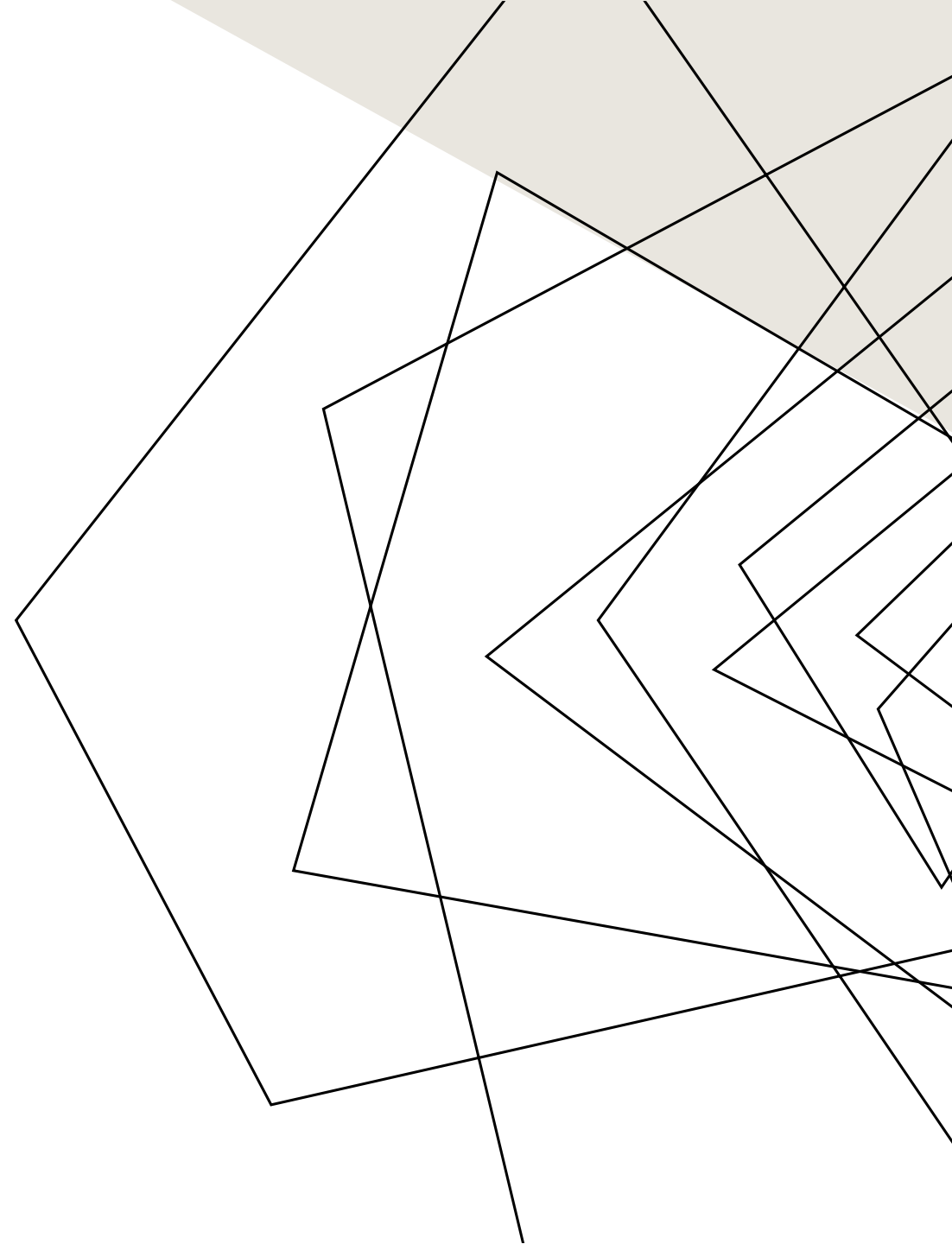


Introduction to Gem Stones Co Pvt. Ltd:

- Gem Stones Co Pvt. Ltd is a leading manufacturer of cubic zirconia, providing an affordable alternative to natural diamonds.
- With a focus on quality and affordability, the company has gained a significant market share.

Project Purpose:

- The aim of this project is to predict gemstone prices accurately using machine learning techniques.
- By doing so, Gem Stones Co Pvt. Ltd can optimize its pricing strategy and maximize profitability.



PROJECT GOALS

Price Optimization:

- Develop predictive models to accurately forecast gemstone prices, enabling the company to set competitive yet profitable prices in the market.

Profit Maximization:

- Identify key factors influencing gemstone prices to guide strategic decisions aimed at maximizing profits.

Quality Control:

- Gain insights into the relationship between gemstone attributes (e.g., carat weight, cut, color, clarity) and prices to improve quality control measures and enhance product value.

DATASET AND FEATURES

Dataset Overview:

- The dataset contains records of almost 27,000 cubic zirconia gemstones, each with various attributes.

Key Features:

- Carat, Cut, Color, Clarity, Depth, Table, Dimensions (X, Y, Z), Price

Importance of Features:

- These attributes play a crucial role in determining the value of a gemstone in the market.

	Unnamed: 0	carat	cut	color	clarity	depth	table	x	y	\
0	1	0.30	Ideal	E	SI1	62.1	58.0	4.27	4.29	
1	2	0.33	Premium	G	IF	60.8	58.0	4.42	4.46	
2	3	0.90	Very Good	E	VVS2	62.2	60.0	6.04	6.12	
3	4	0.42	Ideal	F	VS1	61.6	56.0	4.82	4.80	
4	5	0.31	Ideal	F	VVS1	60.4	59.0	4.35	4.43	
...	
26962	26963	1.11	Premium	G	SI1	62.3	58.0	6.61	6.52	
26963	26964	0.33	Ideal	H	IF	61.9	55.0	4.44	4.42	
26964	26965	0.51	Premium	E	VS2	61.7	58.0	5.12	5.15	
26965	26966	0.27	Very Good	F	VVS2	61.8	56.0	4.19	4.20	
26966	26967	1.25	Premium	J	SI1	62.0	58.0	6.90	6.88	

METHODOLOGY

Data Preprocessing:

- Handle missing values: Imputed using median for 'depth' and mean for 'price'.
- Remove duplicates: Ensured data integrity.
- Outlier detection and removal: Enhanced model robustness using the Interquartile Range (IQR) method.

Feature engineering:

Performed one-hot encoding for categorical variables and normalize numerical features using Min-Max scaling.

Exploratory Data Analysis (EDA):

- Analyze categorical columns: Visualized distribution using pie and count plots.
- Analyze numerical columns: Examined histograms to understand distributions.
- Check correlations: Investigated relationships and potential multicollinearity using a heatmap.
- Q-Q Plot: Assessed the normality of numerical features.

RESULTS AND INSIGHTS

Model Development:

- Built pipelines for various regression models including Linear Regression, XGBoost, Random Forest, KNeighbors, and Decision Tree regressors.
- Evaluated model performance using negative root mean squared error (RMSE) during cross-validation.

Model Interpretability and Evaluation:

- Assessed the significance of predictor variables using Ordinary Least Squares (OLS) regression based on p-values.
- Calculated R-squared (R^2) values to measure the proportion of variance explained by the models on the test data.
- The highest accuracy we received is for linear regression model.

Insights and Recommendations:

- Identified important features influencing gemstone prices and provide actionable insights for pricing strategies and quality control measures.

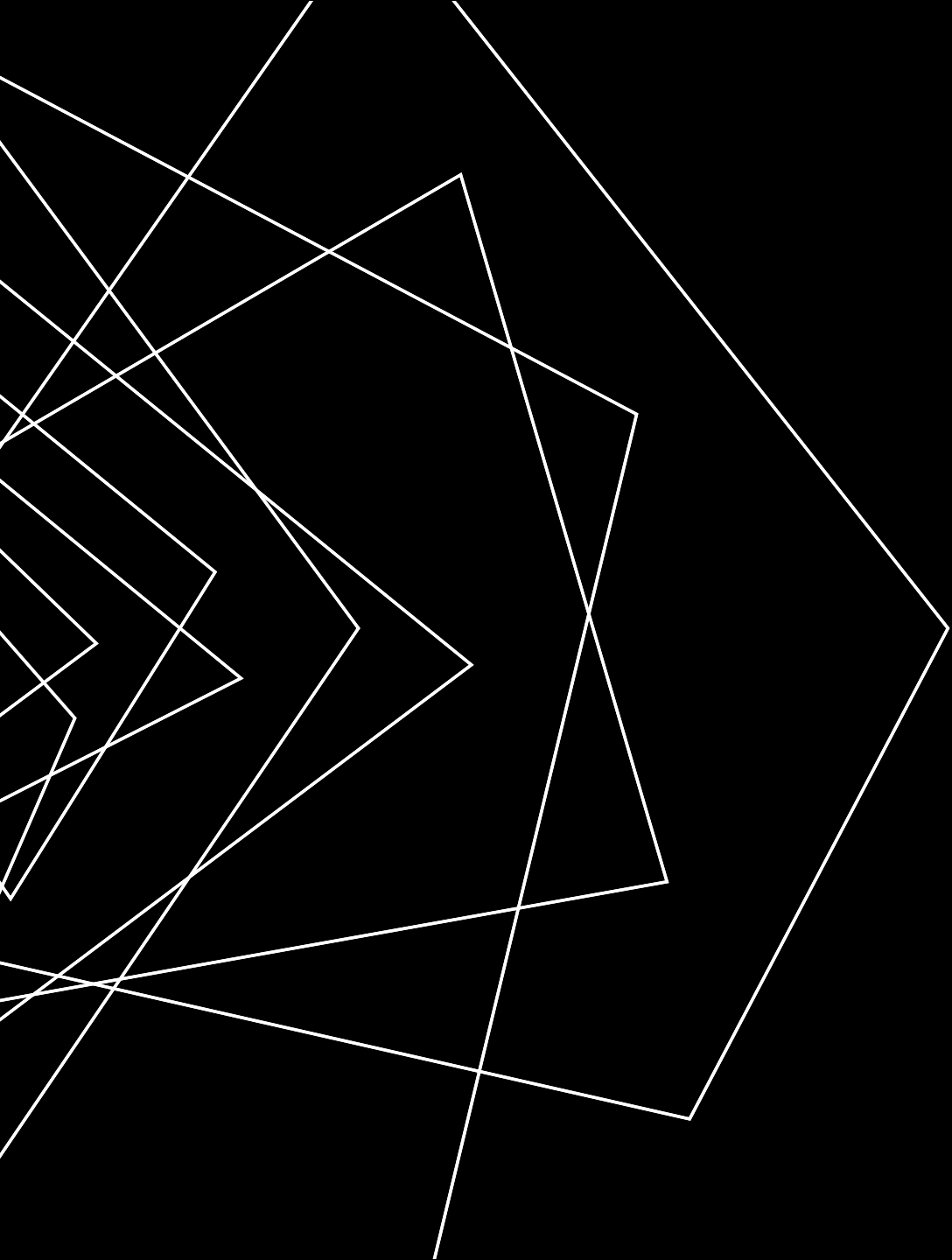
CONCLUSION

Project Summary:

- Through thorough data preprocessing, exploratory analysis, and model development, this project provided valuable insights and predictive models to support Gem Stones Co Pvt. Ltd.

Future Directions:

- Continuous refinement and validation of models are essential to ensure reliability and accuracy in real-world applications.
- Further enhancements and validations can lead to more accurate predictions and better decision-making processes for Gem Stones Co Pvt. Ltd.



THANK YOU