



**PRIN. L. N. WELINGKAR INSTITUTE OF MANAGEMENT
DEVELOPMENT & RESEARCH (PGDM)**

**TRIMESTER VI
PROJECT REPORT**

ON
Market Analysis of The Indian Diamond Industry

BY
Arham Khan
PGDM
Research and Business Analytics
2022-24

ROLL NO.- 96

PROJECT FACULTY GUIDE

Dr. P.V. Chandrika

APPENDIX ‘C’

CAPSTONE PROJECT COMPLETION CERTIFICATE

This is to certify that project titled “Market Analysis of The Indian Diamond Industry” is successfully done by Mr. Arham Khan in partial fulfillment of his two years full time course ‘Post Graduation Diploma in Research & Business Analytics Management’ recognized by AICTE through the S. P. Manali’s Prin. L. N. Welingkar Institute of Management Development & Research (PGDM), Matunga, Mumbai.

This project in general is done under my guidance.



(Signature of Faculty Guide)

Name: Arham Khan

Date:

Acknowledgement

I would like to express my sincere gratitude to Dr. P.V. Chandrika for her invaluable guidance, support, and encouragement throughout the duration of this project, "Market Analysis of the Indian Diamond Industry." Her expertise in Business Analytics has been instrumental in shaping my research and providing valuable insights.

I extend my heartfelt thanks to my peers, classmates, and survey takers for their active participation and contribution to this project. Their feedback and involvement have enriched my study and enhanced its quality.

I also extend my gratitude to the people of internet for their assistance and cooperation in providing data, resources, and feedback, which significantly contributed to the completion of this project.

I am also grateful to my family and friends for their unwavering support and understanding during the research process.

Thankyou
Arham Khan

APPENDIX 'B'

SYNOPSIS

APPENDIX 'B'

Student Copy

SYNOPSIS

(Copy No 1: To be handed over to Faculty Guide after completion of synopsis and obtaining Faculty Guides Signature)

1. Name of the Student : Arham Khan
2. Program & Year : PGDM - RBA 2022-2024
3. Area of Project Research : Business Analytics
4. Name of the Faculty Guide: Dr. P V Chandrika
5. Title of the Project : Develop a Diamond Price Prediction model for the Indian market using machine learning and market research insights for compliance and consumer understanding, refining algorithms for precise predictions

6. Project Details

- (A) Objective of study :
- Build an accurate diamond price prediction model for India.
 - Refine the model with Indian market research (consumer preferences & regulations).
 - Empower Indian diamond industry with market insights and precise predictions.

- (B) Research Methodology:
(With bibliography, if available) may be as follows:

Step I : Gather insights - Analyze market research and collect data on Indian diamond market trends, consumer preferences, and regulations.

Step II : Prepare data - Organize, clean, and potentially refine data for machine learning model training.

Step III : Develop the model - Select, train, and refine a machine learning algorithm for accurate diamond price prediction in India.

Step IV: Deploy and evaluate - Evaluate and refine the model for optimal performance, and deploy it to provide reliable price predictions.

- (C) Expected Results of the study (hypothesis): A market-specific diamond price prediction model, leveraging both machine learning and comprehensive market research, will empower the Indian industry with improved pricing insights and informed decision-making.

Faculty Guide:

Name: Dr. P V Chandrika

Signature: P.V.Chandrika

Date: 01 March 2024

Student:

Name: Arham Khan

Contact No: 7219658497

Signature: Arham Khan

Date: 5 March 2024

TABLE OF CONTENTS

Chapters	Topic	Page No.
I	List of figures and tables	5
II	List of Appendices	5
III	Executive Summary	8
1	Introduction- - Background of the problem statement - Objectives	9
2	Literature Review	11
3	Research Methodology	13
4	Indian Diamond Market Analysis - Consumer Behavior Analysis - Strategic Support and Indian Regulations	15
5	Data Analysis	21
6	Data Visualization - Observations - Modelling	22
7	Model Prediction	27
8	Results and Discussions	29
9	Recommendations & Future Scope of Study	30
10	Survey Questionnaire	32
11	References	33
12	Plagiarism Report	

LIST OF FIGURES AND TABLES

Figure/Table	Page No.
Table 1,2	6
Table 3	16
Figure1,2	21
Figure 3	22
Figure 4	18
Output 1	25
Output 2,3,4	26

LIST OF APPENDICES AND DEFINITIONS

1. Carat: Carat denotes the weight of a diamond, measured in metric carats where one carat equals 0.20 grams and is further divided into 100 points.
2. Cut: Refers to the quality of the diamond's cut, with higher precision leading to enhanced visual appeal and grading.
3. Colour: Denotes the hue of gem-quality diamonds, ranging from colourless to light yellow or brown, with colourless diamonds being the rarest.
4. Clarity: Describes the presence of internal (inclusions) or external (blemishes) characteristics in diamonds, typically visible only under magnification.
5. Depth: The total depth percentage of a diamond, calculated as the height from the culet to the table relative to the mean of its length and width.
6. Table: Represents the width of the diamond's top relative to its widest point, contributing to its brilliance and luster by reflecting light in various directions.
7. Price: Indicates the monetary value of the diamond in US dollars, serving as the primary target column in the dataset.
8. x: Refers to the length of the diamond in millimetres.
9. y: Denotes the width of the diamond in millimetres.
10. z: Represents the depth of the diamond in millimetres

Types of Cuts: Excellent, Very Good, Good, Fair Poor

GIA Diamond Colour Grading Chart:

Table

Color Grade	Description
D	Colorless
E	Colorless
F	Colorless
G	Near Colorless
H	Near Colorless
I	Near Colorless
J	Faint Yellow
K	Faint Yellow
L	Very Light Yellow
M	Very Light Yellow
N	Very Light Yellow
O-P-Q-R	Very Light Yellow to Light Yellow (Not Shown)
S-T-U-V-W-X-Y-Z	Light Yellow to Yellow (Not Shown)

Clarity Index:

Table

Shortform	Full Form
I	Internally Flawless
IF	Internally Flawless
VVS1	Very Very Slightly Included Grade 1
VVS2	Very Very Slightly Included Grade 2
VS1	Very Slightly Included Grade 1
VS2	Very Slightly Included Grade 2
SI1	Slightly Included Grade 1
SI2	Slightly Included Grade 2
I1	Included Grade 1
I2	Included Grade 2
I3	Included Grade 3

EXECUTIVE SUMMARY

This project undertakes a detailed investigation of the Indian diamond market, fusing advanced machine learning techniques with in-depth market research to build a robust analytical model. This model is specifically designed for the Indian context, offering a comprehensive view of diamond pricing that accounts for the unique preferences of Indian consumers and the specific regulatory environment. The aim is to provide decision-makers in the diamond industry with clear, data-driven insights by refining algorithms and integrating tailored market data.

Our research highlights the competitive landscape of the Indian diamond industry, dominated by Titan Company Ltd., with an annual turnover of INR 3,333 crore. Other key players, including Kalyan Jewellers, Senco Gold, and Vaibhav Global, also demonstrate substantial turnovers, signalling their strong market positions. The study further explores the determinants of diamond pricing, revealing that in 2021, the cost for a 1-carat diamond spanned from INR 50,000 to INR 300,000. The analysis shows that pricing is significantly influenced by the four Cs—carat, color, clarity, and cut—with the cut being the most crucial factor for price determination.

In conclusion, the development of a customized Diamond Price Analysis model, leveraging sophisticated business analytics and insightful market intelligence, offers significant advantages to the Indian diamond industry. The model is poised to transform pricing strategies and enhance the decision-making process by integrating advanced analytics that consider both consumer behaviour patterns and regulatory guidelines. The findings stress the importance of developing specialized analytical tools capable of navigating the complexities of the Indian market, equipping industry stakeholders with strategic insights and supporting the sustainable growth of the diamond industry in India.

CHAPTER 1

INTRODUCTION

Background of the Problem Statement-

The Indian diamond industry, globally renowned for its expertise in cutting and polishing, encounters various challenges affecting its growth and stability. Here's a brief overview of these issues:

1. **Declining Exports:** India's exports of cut and polished diamonds are expected to reach a five-year low, raising concerns within the industry. This downturn highlights the need for strategic measures to revive exports and enhance the industry's economic contribution.
2. **Ethical Considerations:** Ethical issues, such as responsible sourcing and labour practices, present hurdles for the Indian diamond market. Addressing these concerns is crucial to maintain industry credibility and promote sustainable practices.
3. **Financial Constraints:** The banking sector expresses apprehensions about the repayment capacity of players in the Indian diamond market, indicating financial stress. Improving financial stability and access to capital is vital for fostering resilience and growth.
4. **Impact of Global Crisis:** The global crisis has significantly impacted India's diamond cutting and polishing sector, revealing vulnerabilities. Adapting to changing market conditions and mitigating external shocks are essential for navigating uncertainties.

In response to these challenges, targeted interventions and innovative solutions are necessary to rejuvenate the Indian diamond industry, promote competitiveness, and ensure long-term viability. Initiatives like the Diamond Price Analysis project aim to provide industry stakeholders with valuable insights to address existing challenges and drive positive transformation.

Problem Statement (elaborated): Conducting a comprehensive market analysis of the Indian diamond industry and developing a machine learning model to forecast diamond prices based on key features like cut, color, clarity and carat.

Objectives:

1. To conduct a comprehensive analysis of the Indian diamond industry, capturing market trends, dominant forces, and sector growth.
2. To gain a nuanced understanding of diamond valuation factors—color, carat, clarity, and cut—by employing data analysis and visualization techniques.
3. To determine the primary attributes influencing diamond pricing and enable a strategic pricing approach through the use of machine learning algorithms for building predictive models.
4. To delve into the purchasing patterns and market perceptions of Indian consumers regarding diamonds.
5. To examine the regulatory landscape and forecast emerging trends within the Indian diamond market.

CHAPTER 2

LITERATURE REVIEW

In the domain of gemology, significant shifts in the diamond industry have been meticulously documented by scholars who have delved into its political and economic underpinnings. Shor (2005) provides a critical analysis of the industry's evolution, charting the decline of De Beers' near-monopoly over the diamond market. This shift was precipitated by geopolitical changes, such as the dissolution of the Soviet Union and the onset of civil wars in Africa, which, in turn, paved the way for new diamond discoveries in Canada and Australia. Shor further discusses the impact of technological advancements in diamond cutting and polishing that emerged in the 1990s. These innovations enabled more efficient production and led to a demand for higher quality, premium-cut diamonds, shifting the industry's focus towards excellence in craft and quality.

Building on this narrative of transformation, Rao (2009) casts a lens on the Indian diamond industry, often perceived as informal yet operating as a well-structured and potent economic network. This detailed exploration into the industry's complex operational dynamics and networking strategies unveils its significant role in India's foreign exchange earnings, with an impressive annual export value. Rao's study offers a comprehensive examination of the growth trajectory and unique characteristics of the industry, delving into the operational nuances of a prominent 40-year-old Cutting and Polishing (CPD) unit. The paper's depth is further enhanced by its analysis of the industry's fortitude in the face of the 2008 global financial crisis. This resilience is depicted through the adaptive strategies employed, demonstrating the sector's capacity to withstand, and rebound from economic shocks.

Rose (2022) contributes to this literature by highlighting the Indian diamond industry's remarkable ability to navigate through adversity. The recession of 2008, the challenges posed by demonetization, and the unprecedented COVID-19 pandemic represent significant hurdles that the industry has successfully overcome. This resilience is attributed to technological advancements and the expertise of a skilled workforce. Rose emphasizes the critical role of migrant labor from regions like Saurashtra in forming the backbone of the industry's operations.

However, the industry's susceptibility to global events is also noted, with recent geopolitical tensions, such as the Russia-Ukraine conflict, impacting diamond supplies and leading to workforce disruptions in diamond hubs like Surat.

Complementing these macro-level insights, Jain (2016) sheds light on a microcosmic yet pivotal aspect of the diamond market—the burgeoning jewelry demand in non-urban Indian settings. This surge is propelled by an increase in disposable income and changing preferences in rural and semi-urban locales. Jain's research is incisive in its analysis of the cultural, social, and investment factors that influence jewelry purchasing behaviors. Moreover, the study offers strategic marketing insights, advocating for tailored approaches to cater to the distinct tastes and purchasing patterns of non-urban consumers.

In the sphere of technical precision, Cowing (2014) tackles the subjective nature of diamond clarity grading. The introduction of an objective grading system based on quantifiable factors such as the size, number, contrast, position, and nature of inclusions marks a paradigm shift towards consistency. Cowing's methodological rigor, evidenced by the examination of high-quality diamond photographs, aligns closely with established standards of esteemed gemological laboratories like the GIA and AGSL. The findings underscore a substantial increase in inclusion area across various clarity grades, advocating for the system's potential to harmonize clarity grading practices in the diamond industry.

Collectively, these scholarly contributions offer a mosaic of perspectives on the diamond industry, encompassing its economic impact, operational agility, technological progress, consumer dynamics, and the pursuit of standardization in valuation. The interplay of these dimensions encapsulates the multifaceted nature of the diamond trade and sets the stage for future research directions that could propel the industry towards greater heights of transparency, efficiency, and customer satisfaction.

This comprehensive body of work underscores the necessity of continual analysis and the application of innovative strategies to maintain the diamond industry's vibrancy and relevance in a rapidly evolving global economic landscape. As each study builds upon the last, they create a rich tapestry of knowledge, crucial for those who seek to navigate the complexities of the gem market with acumen and foresight.

CHAPTER 3

RESEARCH METHODOLOGY

Research Design:

The study will adopt a mixed-methods approach, combining qualitative and quantitative research designs. Qualitative methods will explore the intricacies of consumer behavior and regulatory environments through surveys and case questionnaires. Quantitative analysis will involve statistical and predictive modeling to examine market trends, pricing attributes, and sector growth. This convergent design will cross-validate findings for robust conclusions.

Data Collection Tools: Data will be gathered through several avenues -

- Surveys and structured questionnaires aimed at consumers to assess purchasing patterns and market perceptions for primary research.
- Market reports and industry publications for secondary data on trends and growth.
- Public and proprietary databases for quantitative data on pricing, valuation factors, and economic indicators relevant to the diamond industry.

Data Description: The dataset describes a collection of diamonds, with each entry providing details on the diamond's physical and quality characteristics. The attributes captured include:

- **Weight Measurement:** The dataset records the weight of each diamond in carats, which is a key indicator of size and has a direct impact on value.
- **Craftsmanship Quality:** It includes a qualitative assessment of each diamond's cut, affecting its sparkle and overall appeal.
- **Color Rating:** Each diamond is graded on its colour, with ratings indicating how close it is to being colourless, a trait associated with higher-quality diamonds.
- **Purity Scale:** Clarity ratings are included to denote the presence of any internal imperfections, with higher clarity contributing to greater worth.

- **Physical Dimensions and Shape:** The dataset captures the detailed measurements of each diamond's physical dimensions, which help to determine its shape and how it interacts with light.

Statistical Tests Used:

- **Descriptive Statistics:** Summarizing the central tendency and dispersion of continuous variables.
- **Correlation Analysis:** Quantifying the strength and direction of the relationship between continuous variables.
- **Regression Analysis:** Exploring the relationship between independent variables and a continuous dependent variable.
- **Pair Plots and Violin Plots:** Visualizing the distributions and relationships between variables, aiding in data exploration and analysis.
- **Regression Analysis with XGBoost (XGB):** Utilizing the XGBoost algorithm to predict a continuous dependent variable based on independent variables.
- **Regression Analysis with Random Forest (RF):** Employing the Random Forest algorithm to predict a continuous dependent variable based on independent variables.

Limitations of the Study:

1. **Data Constraints:** The study may face limitations due to data availability and quality, potentially affecting the accuracy and generalizability of findings.
2. **Scope Limitations:** Constraints such as time, budget, and resource availability may restrict the depth and breadth of the study, impacting the comprehensiveness of analysis.
3. **Assumption Risks:** Assumptions made during the research, particularly in predictive modelling, may introduce bias or inaccuracies in results, warranting cautious interpretation.
4. **External Influences:** External factors like economic fluctuations, regulatory changes, and geopolitical events can exert unforeseen influences on the diamond market, challenging the relevance of study conclusions over time.

Chapter 4

Indian Diamond Market Analysis

The Indian diamond market stands as a formidable force on the global stage, boasting a staggering valuation of USD 65.8 billion. Renowned for its expertise in diamond cutting and polishing, India leads the world as the largest manufacturer of cut and polished diamonds, with a remarkable 93% of its production destined for international markets. This dominance underscores India's pivotal role in shaping the dynamics of the global diamond trade.

Driven by relentless growth, the Indian diamond cutting and polishing sector witnessed substantial expansion, reaching a valuation of USD 18.69 billion in 2021. This impressive figure, coupled with a projected compound annual growth rate (CAGR) of 3.4%, underscores the market's resilience and potential for further advancement. Moreover, the steady increase in the production value of diamonds in India over the years reflects the industry's robust trajectory and unwavering upward momentum.

The historical evolution of the Indian diamond industry paints a compelling narrative of transformation. Once perceived as an informal sector, the industry has undergone a remarkable metamorphosis, emerging as a well-organized community renowned for its precision and craftsmanship. This evolution not only signifies the industry's maturation but also highlights its pivotal contribution to India's foreign exchange earnings and economic growth.

In essence, the Indian diamond market stands as a beacon of excellence and innovation, epitomizing the nation's prowess in the global trade landscape. With its remarkable valuations, dominant manufacturing capabilities, and steady growth trajectory, the Indian diamond industry continues to shape the course of the global diamond market, solidifying its position as a formidable leader in the realm of luxury goods.

As the Indian diamond industry continues to thrive, it presents a plethora of opportunities for both domestic players and international stakeholders. With its rich heritage of craftsmanship and cutting-edge technologies, India serves as a hub for innovation and collaboration in the diamond trade.

Ranking Top 10 Diamond Manufacturing Companies in India (Based on their Annual Turnover)

Table

Ranking	Diamond Company	Yearly Turnover (INR Crores)
1	Titan Company	3,333
2	Kalyan Jeweller	389.89
3	Senco Gold	160.91
4	Vaibhav Global	99.90
5	Thangamayil	79.74
6	Asian Star	57.31
7	Goldiam Inter	56.88
8	Tribhovandas	39.67
9	Rajesh Exports Ltd	30.38
10	Radhika Jewel	29.69

In 2021, the price range for a 1-carat diamond in India typically varied from INR 50,000 to INR 300,000, contingent upon its quality attributes. The valuation of a diamond is determined by four key characteristics collectively known as the "4C's": carat weight, color grade, clarity grade, and cut quality, with the cut being considered the most influential factor.

Below are examples of diamond prices sourced from India MART:

- 1 ct (0.2 gm) IJ Color Natural Diamond, suitable for jewelry purposes, with a size of -2 Star Melee, priced at ₹18,000 per carat.
- 1 ct (0.2 gm) Solitaire White Real Natural Diamond, specifically crafted for jewelry, available at ₹1,00,000 per piece.
- 1 Ct (0.2 gm) White Pear Shape Pie Cut Diamonds, intended for fancy jewelry applications, featuring a size of 6 mm, priced at ₹95,000 per carat.
- 1 ct (0.2 gm) White Natural Diamond, weighing 0.05 carats with a size of 1 mm, suitable for jewelry applications, priced at ₹10,500 per carat.

These examples underscore the diversity in diamond pricing based on various attributes such as color, cut, and shape, reflecting the intricate nature of the diamond market.

- Consumer Behavior Analysis –

In the dynamic realm of diamond jewelry, consumer awareness and preferences around the 4Cs—Color, Cut, Clarity, and Carat—exhibit a wide spectrum. While a basic awareness of these fundamental grading scales is relatively common, the depth of understanding varies (Fig.1). A substantial number of consumers, armed with varying degrees of information, approach the purchase of these precious stones with individualized priorities. For many, the carat weight of a diamond often takes precedence due to its direct influence on the stone's size and value. In contrast, others may find the cut's influence on a diamond's brilliance, or the subtleties of color grading, more significant in their quest for aesthetic perfection.

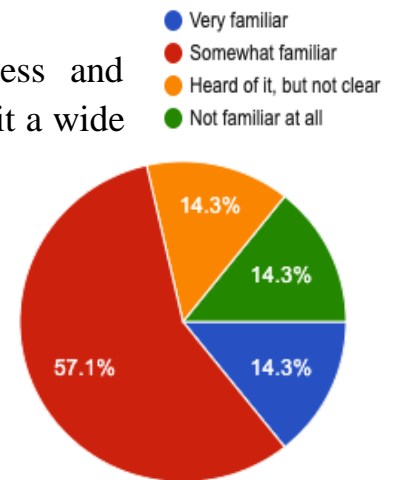


Fig.1

The evaluation of a diamond's clarity frequently factors into the purchasing decision; however, the emphasis placed on this attribute can differ among buyers. While some meticulously scrutinize a diamond's purity, others may place less weight on this aspect. The sufficiency of information available to consumers also spans a broad range, with some feeling well-equipped to make informed decisions, while others express a need for further guidance. Trusted sources such as jewelers, online resources, and personal networks play pivotal roles in shaping consumer knowledge and choices (Fig.2).

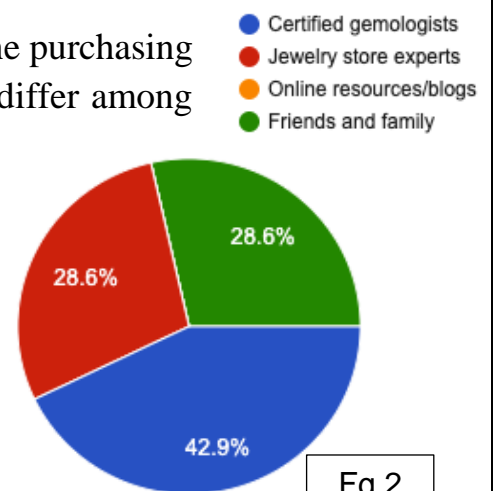


Fig.2

The burgeoning awareness of lab-grown diamonds reflects a shift in consumer sentiment (Fig.3), showcasing an openness to alternatives that promise ethical sourcing and a more attractive price-to-quality ratio. This openness, coupled with an increasing interest in investment security through buyback schemes (Fig.4), underscores the evolving landscape

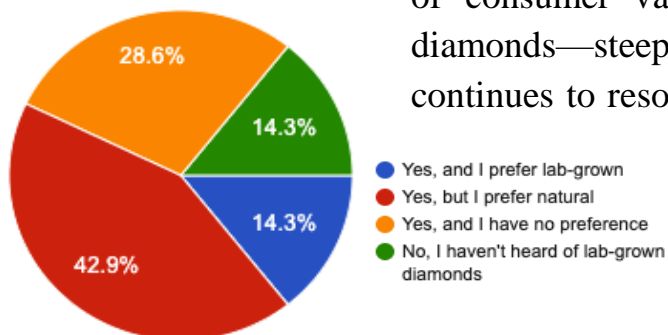
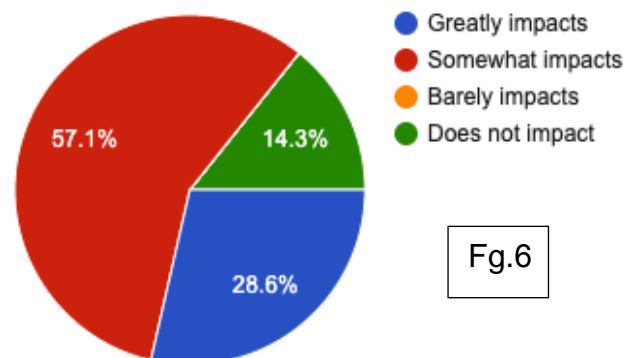
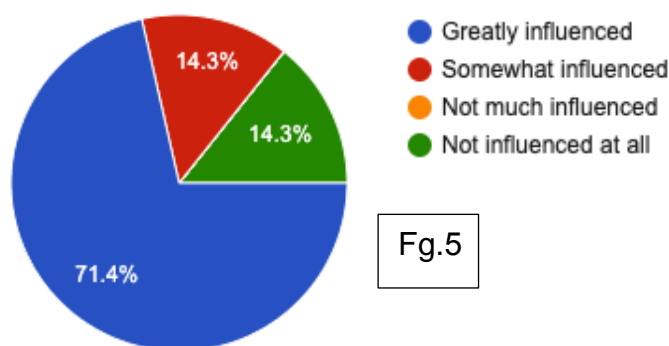
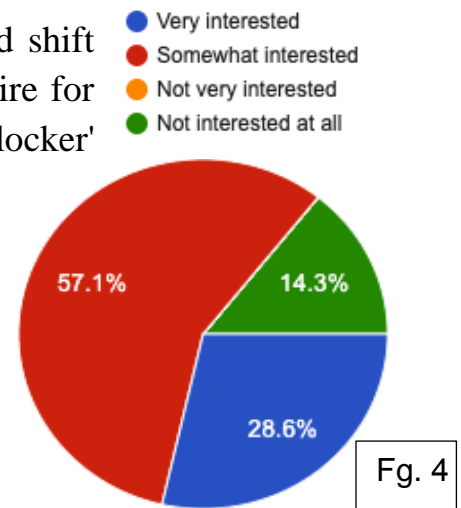


Fig.3

of consumer values. Moreover, the reputation of Indian diamonds—steeped in a tradition of esteemed craftsmanship—continues to resonate globally (Fig.5), leading to a prevailing perception of high quality among these gems.

As lifestyles and disposable incomes evolve, there is a marked shift toward lightweight and wearable designs, propelled by the desire for modern and versatile pieces that transition seamlessly from the 'locker' to the 'drawer'. The Indian diamond market, synonymous with opulent craftsmanship and cultural heritage, is adapting to this new era (Fig.6), as evidenced by an uptick in the online jewelry sector and a nuanced approach to retail experiences. Although the government's policies aimed at bolstering the sector have been met with mixed reactions, their impact on fostering industry growth is undeniable. In this evolving narrative, the cultural significance of diamonds remains unchallenged, particularly in matrimonial and festive contexts, while consumers' aspirations and the quest for personalized expressions through jewelry continue to sculpt the market's future contours.



Navigating through the intricate interplay of tradition and modernity, the Indian diamond industry stands at a juncture where growth and innovation intersect. As it embraces the challenge of adapting to the nuanced demands of a discerning, contemporary consumer base, the industry's resilience is poised to redefine the global landscape of luxury jewels. Its sustained vibrancy hinges on its ability to harmonize the enduring allure of diamonds with the ever-changing aspirations of its patrons, ensuring that these timeless treasures continue to sparkle across generations.

Survey Link - <https://forms.gle/e5gxdxgMnwcie3SY6>

Refer Chapter 10

- **Strategic Support and Indian Regulations:**

The Indian government has developed various policies and frameworks to support and regulate the diamond industry, an essential component of the country's export economy. Key measures include:

- 1. Regulatory Framework and Industry Support:**

- **Kimberley Process Certification:** To regulate and monitor the trade of rough diamonds, Indian traders must be members of the Gem & Jewellery Export Promotion Council (GJEPC), aligning with global efforts to curb the trade of conflict diamonds.
- **Special Economic Zones (SEZs):** Initiatives like the SEEPZ have been established to promote exports by providing conducive operational environments and fiscal benefits.

- 2. Economic Contributions and Initiatives:**

- The diamond industry significantly contributes to India's exports, with cut and polished diamonds forming a major portion of this export segment. The government has continually supported this sector through various economic reforms aimed at fostering growth and solidifying India's position as a leading exporter in the global market.

- 3. Infrastructure and Institutional Framework:**

- **Customs and Clearance Facilities:** Facilities like the Precious Cargo Customs Clearance Centre have been set up to enhance the efficiency of importing and exporting precious items including diamonds, streamlining processes and reducing operational delays.
- **Research and Development:** Institutions like the Gemmological Institute of India support the industry with essential research, development, and certification, ensuring that industry standards remain high.

- 4. Crisis Management and Recovery Support:**

- During global economic downturns and crises like the COVID-19 pandemic, the industry has seen targeted support from the government to help mitigate impacts and facilitate a swift recovery, aiding in the stabilization and subsequent growth of diamond exports.

- 5. Education and Skill Development:**

- The sector benefits from focused educational programs and skill development initiatives that aim to enhance craftsmanship and maintain high industry standards, which are crucial for sustaining growth and competitiveness in the global market.

CHAPTER 5

DATA ANALYSIS

- **DATA OVERVIEW:**

The dataset* comprises information on approximately 54,000 diamonds, encompassing attributes such as carat, cut, colour, clarity, depth, table, price, length (x), width (y), and depth (z). The target variable is the price of the diamonds, denoted in Indian Rupees (₹). (Conversion Rate - 1 United States Dollar = 83.30 Indian Rupee as on 5th April'24)

- **DATA MANIPULATION & PRE-PROCESSING:**

Before analysis, the dataset undergoes manipulation and pre-processing steps. This includes handling missing values, encoding categorical variables like cut, colour, and clarity, and scaling numerical features like carat, depth, table, and dimensions (x, y, z). Additionally, outliers are identified and removed to ensure data quality.

- **EXPLORATORY DATA ANALYSIS (EDA):**

Exploratory data analysis involves understanding the dataset's characteristics and relationships between variables. Graphical techniques such as violin plots, lmlots, and pair plots are employed to visualize distributions, correlations, and patterns within the data. EDA also includes statistical summaries and insights derived from the dataset.

- **MODEL BUILDING:**

Several regression models are employed to predict diamond prices based on the given attributes. These models include Linear Regression, Lasso Regression, Decision Tree Regression, K-Nearest Neighbours Regression. XGBoost Regression & Random Forest Regressor are used to build final models due to their robustness.

- **MODEL PERFORMANCE:**

The performance of each regression model is assessed using metrics such as mean squared error (MSE), mean absolute error (MAE), and R-squared (R²) score. The model with the best performance is selected as the final predictive model for diamond price estimation.

***Note** (Refer Dataset and Analysis Notebook) – <https://github.com/arhamk01/Indian-diamond-market-analysis-capstone>

CHAPTER 6 - Data Visualizations

Diamond Cut vs Price:

1. The graph illustrates the relationship between diamond price and cut grade, where cut grade signifies the quality of a diamond's proportions, symmetry, and polish, influencing its sparkle.

2. Generally, as cut grade improves from "Fair" to "Ideal," diamond price tends to increase, indicating that higher-quality cuts command higher prices.

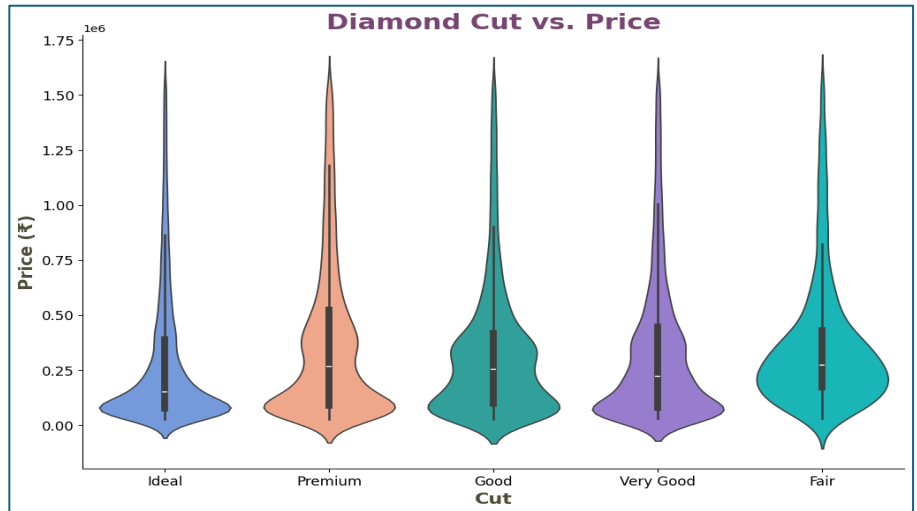


Figure 1

3. However, the price increase is not linear; it accelerates notably for diamonds with superior cuts ("Ideal" and "Premium") compared to those with lower-quality cuts ("Good," "Very Good," and "Fair").

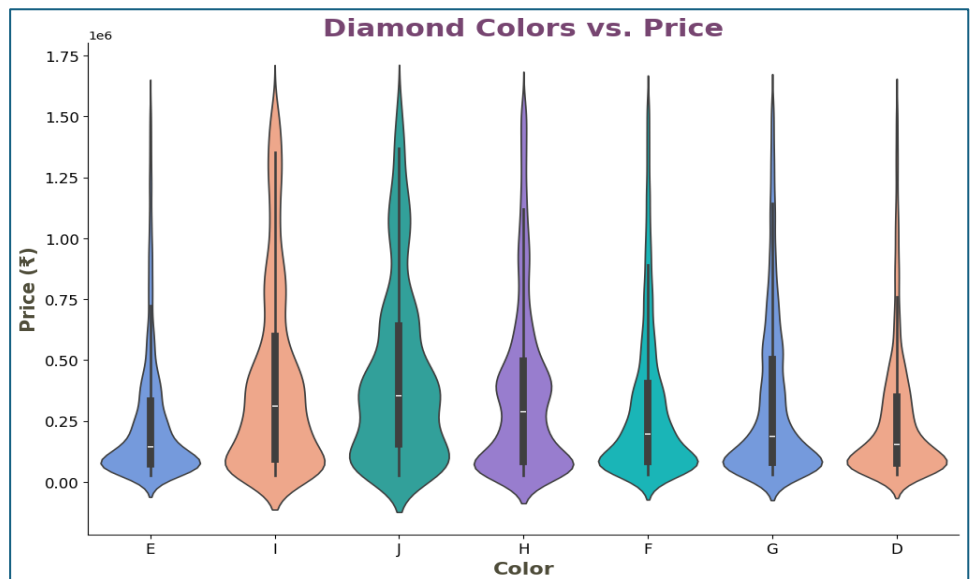
4. This suggests that the value difference between a well-cut diamond and an excellently cut one surpasses that between a lower quality cut and an average one, emphasizing the premium placed on superior cut quality in diamond pricing.

Figure 2

Diamond Colors vs Price:

1. The graph illustrates the correlation between diamond color grade and price, where the y-axis denotes diamond price per carat and the x-axis represents color grade from "E" (colorless) to "J" (faint yellow).

2. Generally, diamonds with higher color grades (more colorless) exhibit higher prices compared to diamonds with lower color grades (more yellow), indicating a downward trend in price as one moves from left to right on the x-axis.



Diamond Clarity vs. Price:

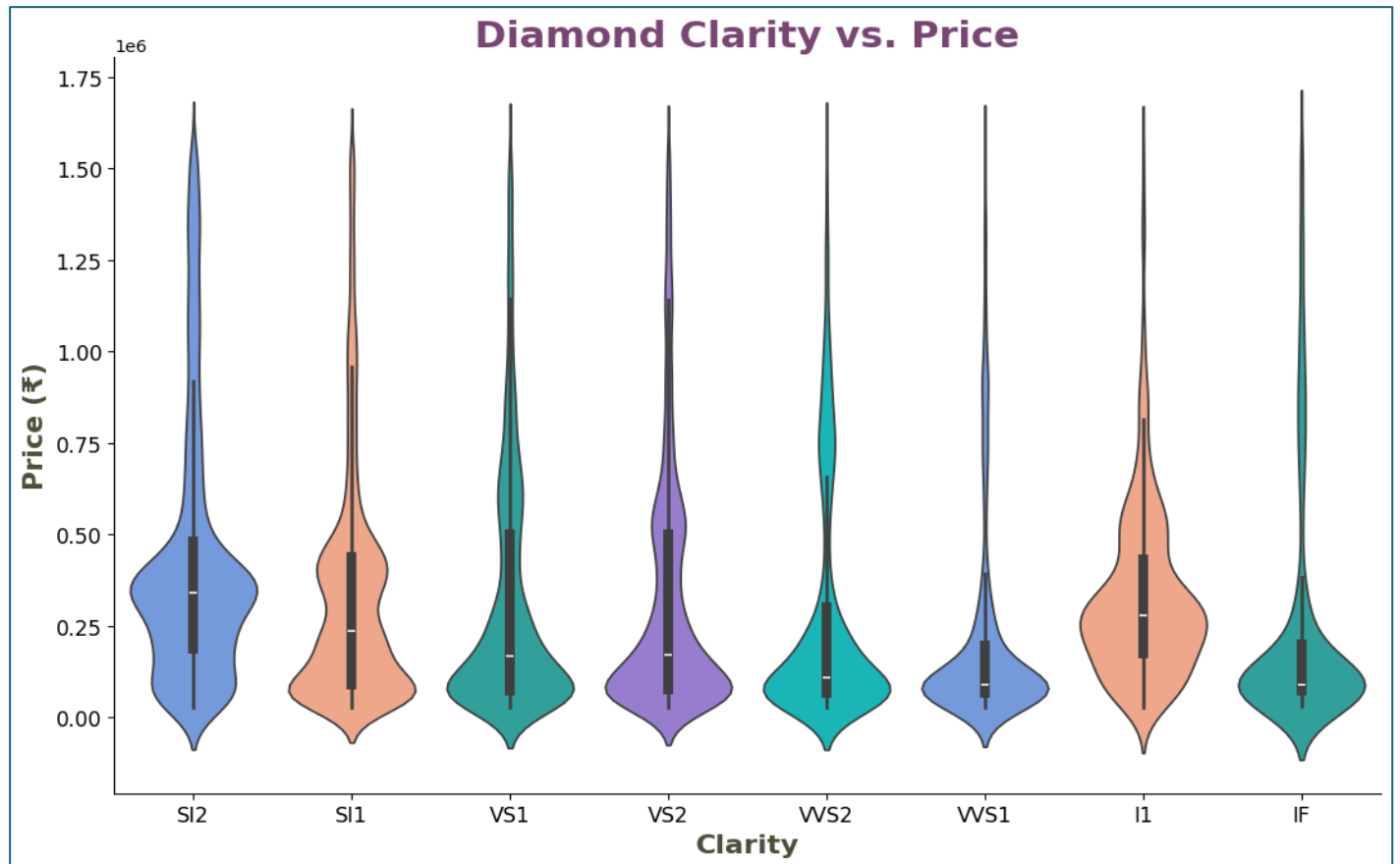


Figure 3

1. The graph illustrates the relationship between diamond clarity grades and prices, with higher clarity grades corresponding to higher prices, denoting fewer inclusions.
2. Prices tend to decrease as clarity grades decline from IF to I1, though differences are more noticeable at the upper end of the scale.
3. Notably, the price variance between adjacent clarity grades appears more significant at higher clarity levels and less pronounced at lower grades.
4. However, it's crucial to recognize that the actual diamond price is influenced by additional factors such as carat weight, cut, and color.
5. The graph suggests that diamonds with fewer visible inclusions command higher prices, indicating a preference for clarity among buyers.
6. While clarity plays a significant role in determining diamond prices, it is just one of several factors that contribute to the overall value and attractiveness of a diamond.

Pair-plot of Diamond Features by Cut :

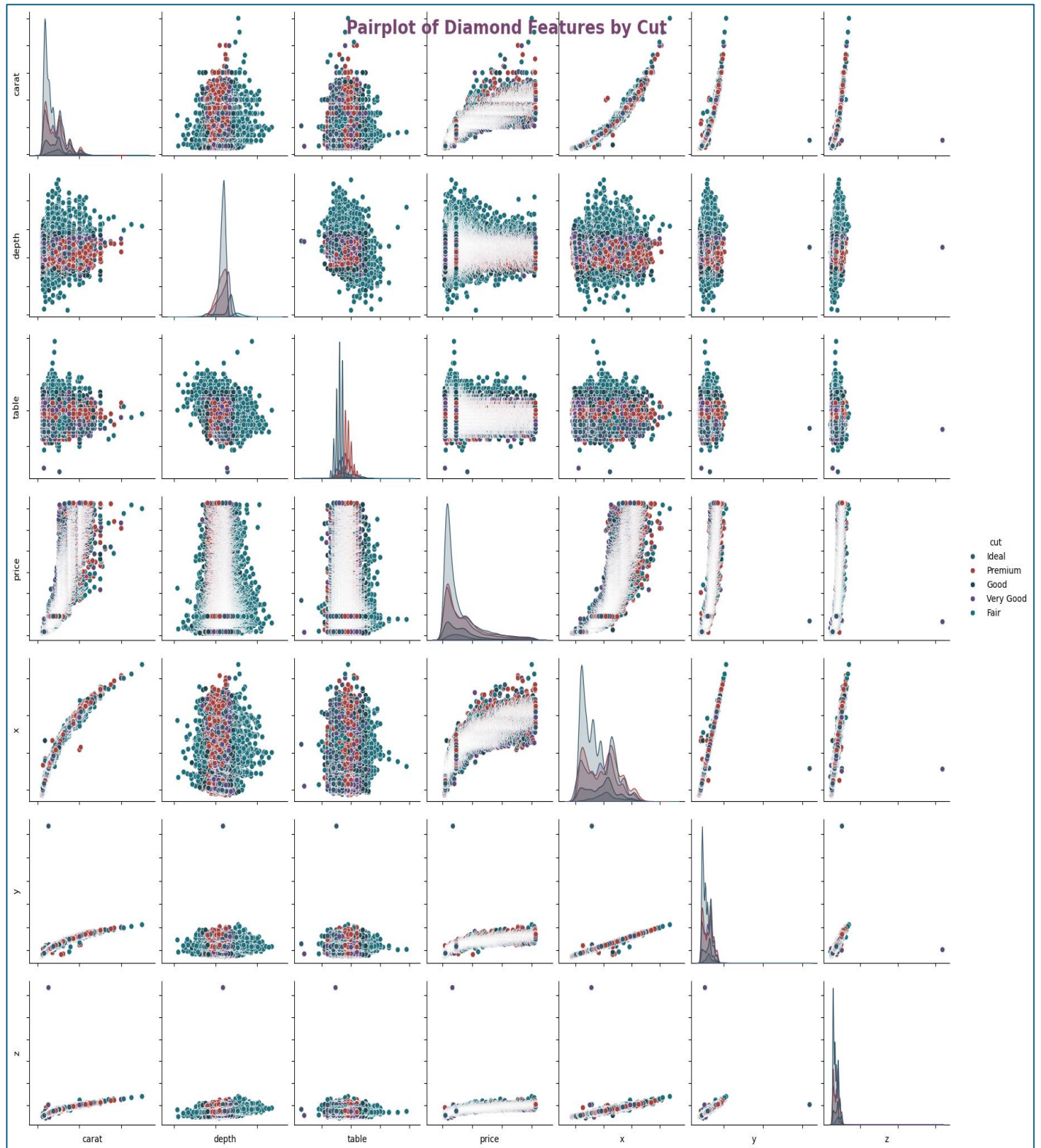


Figure 4

Observations from the Pair-Plot graph:

1. Carat vs. Price: The data illustrates a positive correlation between carat weight and price, indicating that larger diamonds generally command higher prices due to their increased rarity and desirability.
2. Carat vs. Depth: While no clear linear relationship is evident between carat weight and depth, a clustering of deeper diamonds with higher carat weights is observable on the right side of the graph. This suggests a potential trend towards deeper diamonds in larger size categories.
3. Carat vs. Table: Similar to carat vs. depth, there is no distinct linear correlation between carat weight and table size. However, the absence of a clear relationship indicates that variations in table size do not consistently correspond to changes in carat weight.
4. Price vs. Depth: Contrary to expectations, there appears to be no discernible correlation between price and depth. This observation suggests that diamonds with similar prices can exhibit considerable variation in depth, indicating that depth alone is not a decisive factor in determining diamond prices.
5. Price vs. Table: Similarly, no explicit correlation between price and table size is apparent. The lack of a consistent relationship implies that diamonds with comparable prices may possess differing table sizes, further emphasizing the multifaceted nature of diamond pricing.
6. Depth vs. Table: A tentative positive correlation between depth and table size is observed, hinting at a potential association between these two attributes. This suggests that diamonds with larger tables may also exhibit greater depths, albeit the correlation appears to be weak.
7. Cut Grade: The supplementary graphs depict the distribution of diamonds across various cut grades ("Ideal," "Premium," "Good," etc.) concerning other features such as carat, price, depth, and table. These visualizations provide insights into how different cut grades are distributed concerning other diamond characteristics.

It's essential to recognize that while these observations offer valuable insights into the relationships between individual diamond attributes, diamond pricing is influenced by a myriad of factors, including cut, clarity, and color. Therefore, a comprehensive understanding of these interrelated variables is crucial for accurately assessing diamond value and market dynamics.

Modelling the Data

1. Linear Regression: The linear regression model yielded a predicted price of approximately ₹12,44,880.09. Linear regression assumes a linear relationship between the independent variables (features) and the target variable (price). It's notable that this model provides a baseline performance for comparison with more complex models.
2. Lasso Regression: Lasso regression, a type of linear regression with regularization, produced a predicted price of around ₹12,14,116.68. Lasso regression penalizes the absolute size of the regression coefficients, promoting sparsity in the model and potentially improving its generalization performance.
3. Decision Tree: The decision tree model generated a predicted price of about ₹6,24,796.64. Decision trees partition the feature space into regions and make predictions based on the average target value within each region. This model can capture complex interactions between features but may suffer from overfitting.
4. Random Forest: Random forest, an ensemble learning method based on decision trees, resulted in a predicted price of approximately ₹4,59,895.50. Random forest builds multiple decision trees and averages their predictions, often leading to improved performance and robustness compared to individual decision trees.
5. K-Nearest Neighbours (KNN): The KNN model produced a predicted price of around ₹6,76,937.72. KNN predicts the target variable by averaging the values of its k nearest neighbours in the feature space. While simple and intuitive, KNN may struggle with high-dimensional data and requires careful selection of the hyperparameter k.
6. XGBoost Regressor: XGBoost, a gradient boosting algorithm, yielded a predicted price of about ₹4,57,400.07. XGBoost builds an ensemble of weak learners (decision trees) sequentially, with each tree learning from the errors of its predecessors. XGBoost is known for its high performance and is widely used in various machine learning competitions and applications.

```
LinearRegression: 124488.096250
Lasso: 121411.677374
DecisionTree: 62479.637783
RandomForest: 45989.502590
KNeighbors: 67693.716618
XGBRegressor: 45740.069656
```

CHAPTER 7 - Model Prediction

Output 2

I. XGB Regressor :

The XGB-Classifer model performed exceptionally well on the test data with a high R-squared value of approximately 0.9811 and an adjusted R-squared value of around 0.9810. These values indicate that the model can explain about 98.11% of the variance in the diamond prices, demonstrating its robust predictive performance.

```
R^2: 0.981064323419003
Adjusted R^2: 0.9810485026717667
```

II. Improving Accuracy With Random Forest Regressor :

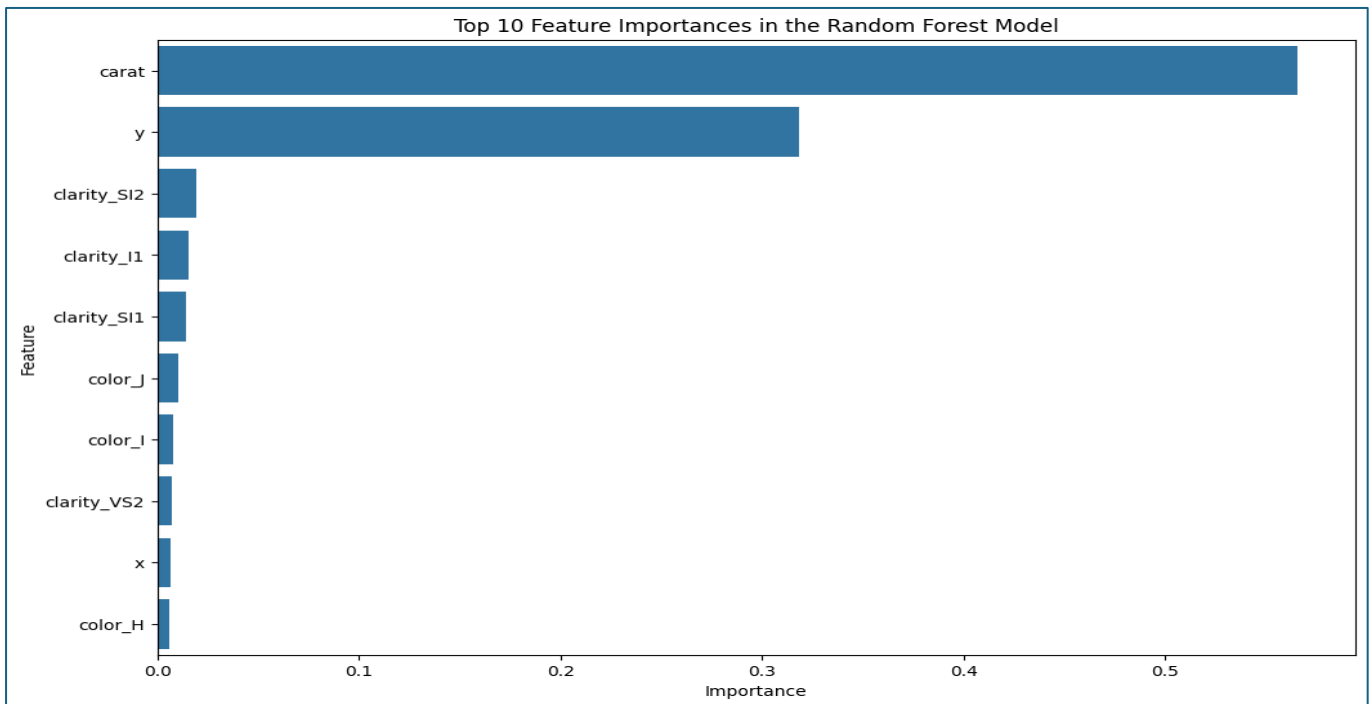
Output 3

The R^2 score of 0.983 indicates that the model explains about 98.3% of the variability in the diamond prices, which is an excellent fit.

```
• Mean Squared Error (MSE): 274,693.80
• R^2 score: 0.983
```

(The data has been split into training and testing sets, with 43,136 samples for training and 10,784 samples for testing. The categorical variables are converted into numerical format using one-hot encoding, resulting in a total of 26 features.)

- Feature Extraction :



Output 4

The feature importance analysis from the Random Forest model highlights the most influential factors in predicting the price of diamonds:

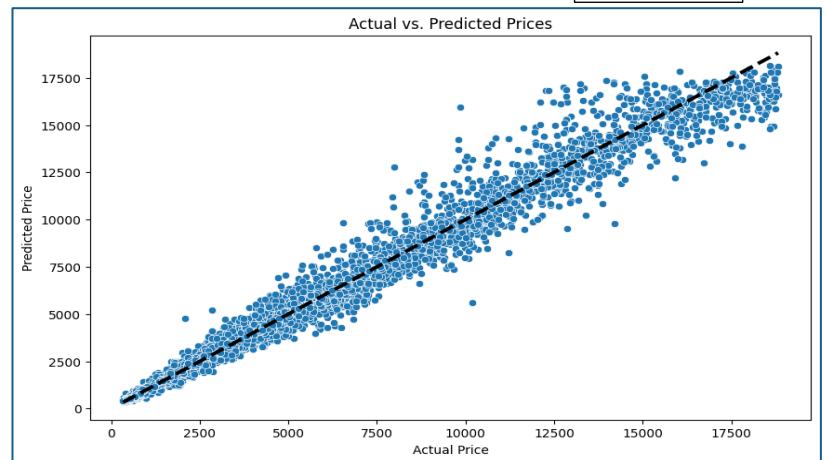
1. Carat (0.566): The weight of the diamond is the most significant predictor, accounting for approximately 56.6% of the predictive importance.
2. y dimension (0.318): The width of the diamond in mm is the second most significant, with about 31.8% importance.
3. Clarity SI2 (0.019) and Clarity I1 (0.016): Lower clarity grades also show significant importance in determining the price.
4. Clarity SI1 (0.014) and colour grades (especially Color J (0.010) and Color I (0.008)) also play important roles.

This suggests that the physical size (carat, dimensions) and clarity of the diamond are critical in setting its price, with carat being the most dominant feature.

Output 5

- Actual vs Predicted Prices :

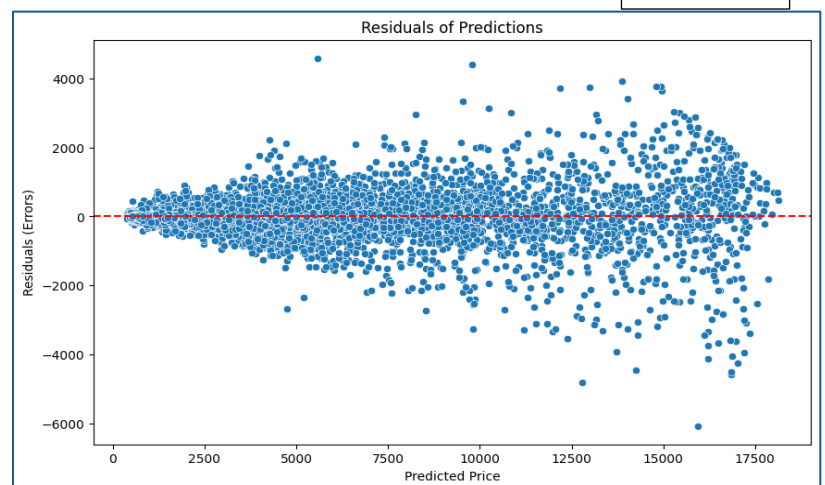
The plot indicates a strong predictive performance by the regression model, with most predicted prices aligning closely with the actual prices across the entire range of values. There is no evident pattern indicating systematic prediction errors, and the model seems to handle both lower and higher-priced diamonds well.



Output 6

- Residuals of Prediction :

The model's predictions generally align well with the actual prices, with residuals close to zero. There's no discernible pattern in the residuals, indicating consistent accuracy across predictions. Yet, outliers exist, particularly for higher predicted prices, where deviations from actual values are more noticeable.



CHAPTER 8 - RESULTS AND DISCUSSIONS

1. **Market Dominance:** The Indian diamond market, valued at USD 65.8 billion, asserts its dominance globally, with India leading as the largest manufacturer of cut and polished diamonds, exporting 93% of its production.
2. **Predictive Accuracy:** The RF Regressor model's exceptional performance, with an accuracy of 98.3% underscores its ability to accurately forecast diamond prices based on the given data.
3. **Multi-faceted Approach:** Utilizing both quantitative and qualitative methodologies, the research design provides a comprehensive understanding of the Indian diamond market, covering pricing dynamics, consumer behaviour, regulatory frameworks, and competitive dynamics.
4. **Data Collection:** Data collection from diverse sources, including industry reports, governmental databases, and market surveys, ensures the relevance and accuracy of insights into the nuances of the Indian diamond market.
5. **Interpretability:** Despite its complexity, the XGBClassifier & RF Regressor model offers insights into feature importance, aiding in understanding the significant factors influencing diamond prices.
6. **Ethical Standards:** Ethical considerations, such as data privacy and participant confidentiality, are maintained throughout the research process, ensuring integrity and reliability.
7. **Historical Evolution:** The historical evolution of the Indian diamond industry reflects its transformation from an informal sector to a well-organized community, contributing significantly to India's foreign exchange earnings and economic growth.
8. **Implications for Stakeholders:** The success of the RF Regressor model provides actionable insights for stakeholders, enabling informed decisions on pricing strategies, consumer preferences, and market competitiveness.
9. **Market Resilience:** With a projected CAGR of 3.4% and steady growth in production value, the Indian diamond market demonstrates resilience and potential for further advancement.
10. **Global Impact:** India's remarkable valuations, dominant manufacturing capabilities, and steady growth trajectory solidify its position as a formidable leader in the global diamond market, shaping the course of the industry and promising continued growth and excellence.

CHAPTER 9

RECOMMENDATIONS & FUTURE SCOPE OF STUDY

Recommendations:

1. **Continuous Market Monitoring:** It is advisable to conduct regular market surveillance to adapt pricing strategies in line with evolving consumer preferences, ensuring sustained competitiveness in the dynamic Indian diamond market.
2. **Technological Investment:** Allocating resources towards advanced technologies for diamond processing can optimize production processes, enhance product quality, and mitigate operational costs, thereby bolstering profitability.
3. **Product Portfolio Diversification:** Expanding the product portfolio to incorporate innovative designs and customizable options can cater to diverse consumer tastes, fostering market penetration and brand loyalty.
4. **Adoption of Sustainable Practices:** Embracing sustainable practices like ethical sourcing and eco-friendly manufacturing processes not only enhances brand image but also aligns with the rising demand for ethically produced diamonds.
5. **Strategic Collaboration:** Cultivating strategic alliances with industry peers and forging international partnerships can facilitate knowledge exchange, foster innovation, and unlock new market opportunities.

Future Scope of Study:

1. **Longitudinal Analysis:** Long-term studies tracking market trends, consumer behaviour, and industry dynamics over time can provide valuable insights into the evolving landscape of the Indian diamond market.
2. **Consumer Perception Studies:** In-depth qualitative research exploring consumer sentiments, preferences, and purchase drivers can uncover underlying motivations shaping diamond purchasing decisions.

3. **Technological Impact Assessment:** Evaluating the effects of emerging technologies, such as blockchain and artificial intelligence, on supply chain transparency, product authentication, and market dynamics within the diamond industry.
4. **Global Comparative Research:** Comparative analysis of the Indian diamond market with global counterparts can identify competitive advantages, challenges, and expansion opportunities in international markets.
5. **Policy Evaluation:** Analyzing the efficacy of government policies and regulations pertaining to the diamond industry, including trade policies and taxation, can inform policy reforms aimed at fostering industry growth and sustainability.

CHAPTER 10 - SURVEY QUESTIONNAIRE

1. How familiar are you with the concept of the 4Cs (Color, Cut, Clarity, and Carat) in diamonds? (Fig.1)
2. When considering a diamond purchase, which of the 4Cs do you prioritize the most?
3. Have you ever purchased a diamond based solely on its carat weight?
4. How important is the cut of a diamond in your purchasing decision?
5. On a scale of 1-10, how would you rate your understanding of diamond colour grades?
6. How often do you consider clarity as a factor when selecting a diamond?
7. Do you feel that you have enough information to make an informed decision about the quality of a diamond?
8. What sources do you trust for information when researching diamonds? (Fig.2)
9. Have you heard of lab-grown diamonds, and how do they influence your perception of natural diamonds? (Fig.3)
10. How likely are you to consider a lab-grown diamond for your next purchase?
11. Would you be interested in a diamond buyback scheme when making a purchase? (Fig.4)
12. Does the ethical sourcing of diamonds influence your purchasing decision?
13. How do you perceive the value of Indian diamonds compared to those from other countries? (Fig.5)
14. What is your preferred setting to buy diamonds? (e.g., online, traditional jewellery store, etc.)
15. Have recent global and local trends influenced your attitude towards diamond jewellery? (Fig.6)
16. How do you view the craftsmanship of Indian-made diamond jewellery?
17. What is your opinion on the current government policies related to the diamond industry?
18. Would you be willing to pay a premium for diamonds that are certified for their cut, colour, clarity, and carat?
19. How does the cultural and social significance of diamonds impact your purchase?
20. How has your interest in diamonds evolved with changes in disposable income and lifestyle?

Ch. 11 REFERENCES & BIBLOGRAPHY

- ⇒ <https://tradeflock.com/diamond-company-in-india/>
- ⇒ <https://www.gia.edu/doc/A-Review-of-the-Political-and-Economic-Forces-Shaping-Todays-Diamond-Industry.pdf>
- ⇒ https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2952702
- ⇒ https://www.researchgate.net/profile/Neeru-Jain/publication/378011604_A_STUDY_OF_JEWELLERY_DEMAND_IN_NON-URBAN_INDIAN_MARKETS/links/65c335c21bed776ae3359beb/A-STUDY-OF-JEWELLERY-DEMAND-IN-NON-URBAN-INDIAN-MARKETS.pdf
- ⇒ <https://m.economictimes.com/industry/cons-products/fashion/-/cosmetics/-/jewellery/indias-cut-and-polished-diamond-exports-likely-to-hit-five-year-low-in-fy24-says-careedge-ratings/articleshow/108290123.cms>
- ⇒ <https://finshots.in/archive/indias-diamond-problems/>
- ⇒ <https://m.economictimes.com/industry/cons-products/fashion/-/cosmetics/-/jewellery/indian-diamond-industry-to-resume-rough-diamond-imports-from-december-15/articleshow/105844731.cms>
- ⇒ <https://www.linkedin.com/pulse/shining-brilliance-exploring-dynamics-indias-diamond-market-wgnyf>
- ⇒ <https://blog.sarine.com/diamond-industry/diamond-industry-2019>
- ⇒ <https://www.undp.org/india/publications/losing-sparkle-impact-global-crisis-diamond-cutting-and-polishing-industry-india>
- ⇒ <https://www.vummidi.com/blog/diamond-jewelry-market-india-stats-facts/#:~:text=Valued%20at%20USD%2065.8%20billion,according%20to%20leading%20research%20firms.>
- ⇒ <https://www.linkedin.com/pulse/shining-brilliance-exploring-dynamics-indias-diamond-market-wgnyf>
- ⇒ <https://indbiz.gov.in/precious-opportunities-in-indias-diamond-industry/>
- ⇒ <https://www.polarismarketresearch.com/industry-analysis/india-diamond-cutting-and-polishing-market>
- ⇒ <https://www.statista.com/statistics/668071/india-diamond-production-value/>
- ⇒ <https://www.indiamart.com/>
- ⇒ https://giepc.org/news_detail.php?news=indian-govt-mandates-self-regulation-for-import-export-of-rough-diamonds
- ⇒ <https://pib.gov.in/PressReleasePage.aspx?PRID=1911554>
- ⇒ <https://giepc.org/solitaire/how-policy-reforms-paved-the-way-for-growth-of-the-indian-gem-jewellery-industry/>
- ⇒ <https://www.tpci.in/indiabusinesstrade/blogs/indian-diamond-industry-from-a-rough-patch-to-a-glittering-performance/>