

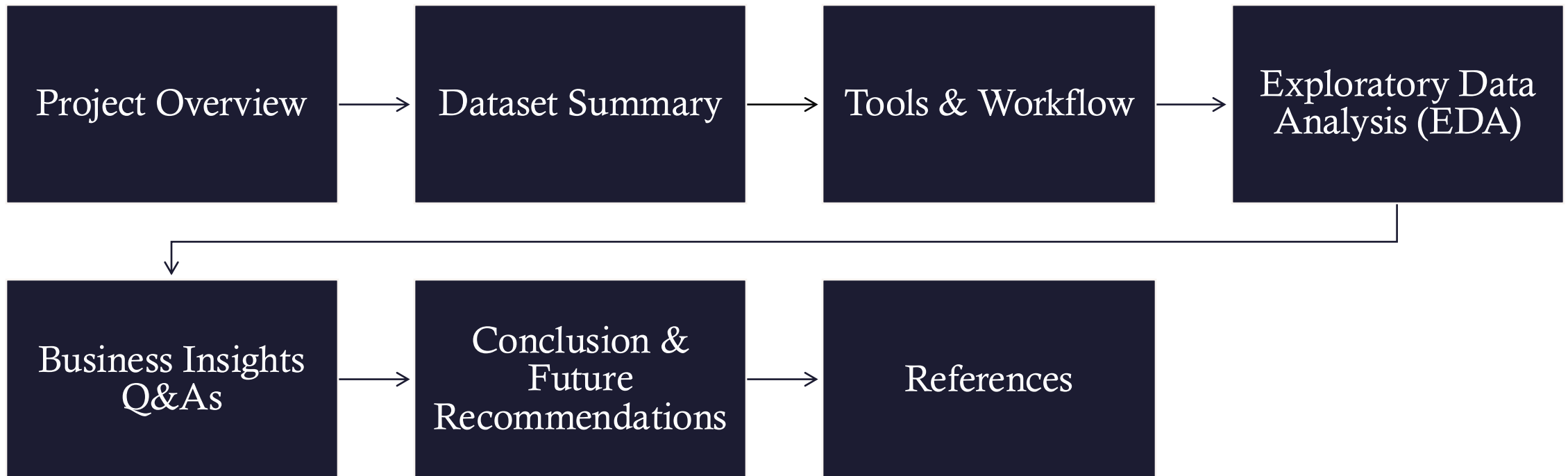


# DIAMOND QUALITY VS. PRICE TRADEOFF: FINDING THE BEST VALUE COMBINATIONS

By Chi (Alison) Dang

---

# AGENDA:



---

# PROJECT OVERVIEW



Client: General Customer  
who does not have much  
knowledge in diamonds



Business Problem: How can  
customers identify the best  
quality-to-price combinations  
in diamonds?



Goal: Analyze the tradeoff  
between price and quality  
dimensions.

---

---

# DATASET SUMMARY



Source: Diamonds from [Kaggle](#)



Dataset: ~54,000 diamonds with pricing and quality attributes



Key variables: carat, cut, color, clarity, depth, table, price, x, y, z



Adding variable: price\_per\_carat



Number of rows: 53,940 entries



Number of columns: 12

---

# TOOL & WORKFLOW

- **Tools:** Python (pandas, seaborn, matplotlib, numpy) in Google Colab
- **Workflow:** Data Cleaning & Transformation → Exploratory Data Analysis (EDA) → Heatmaps → Business Recommendation

Google Colab for Python 



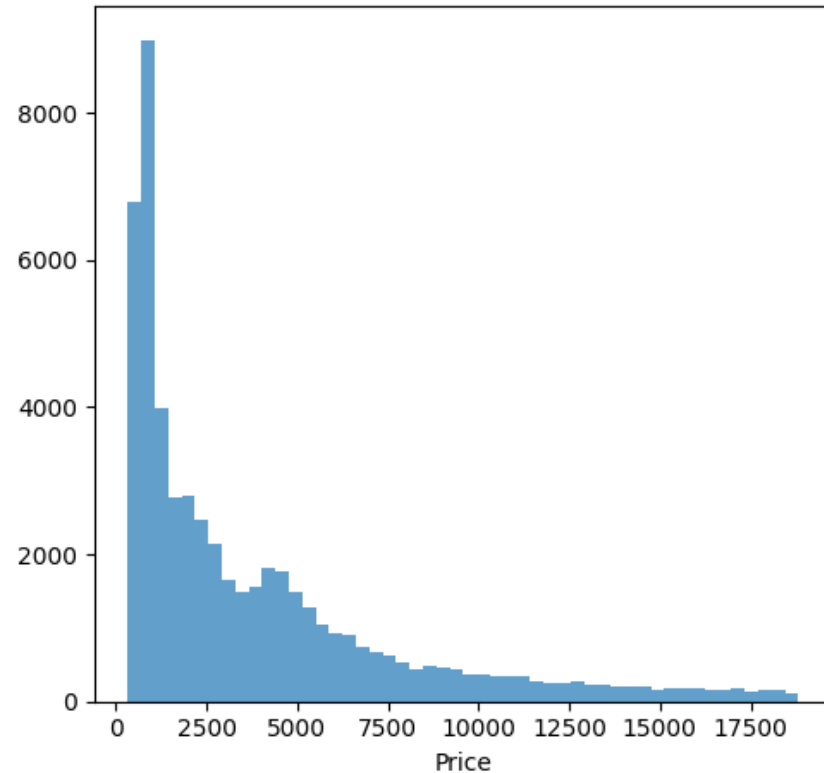
---

# EXPLORATORY DATA ANALYSIS (PT.1)

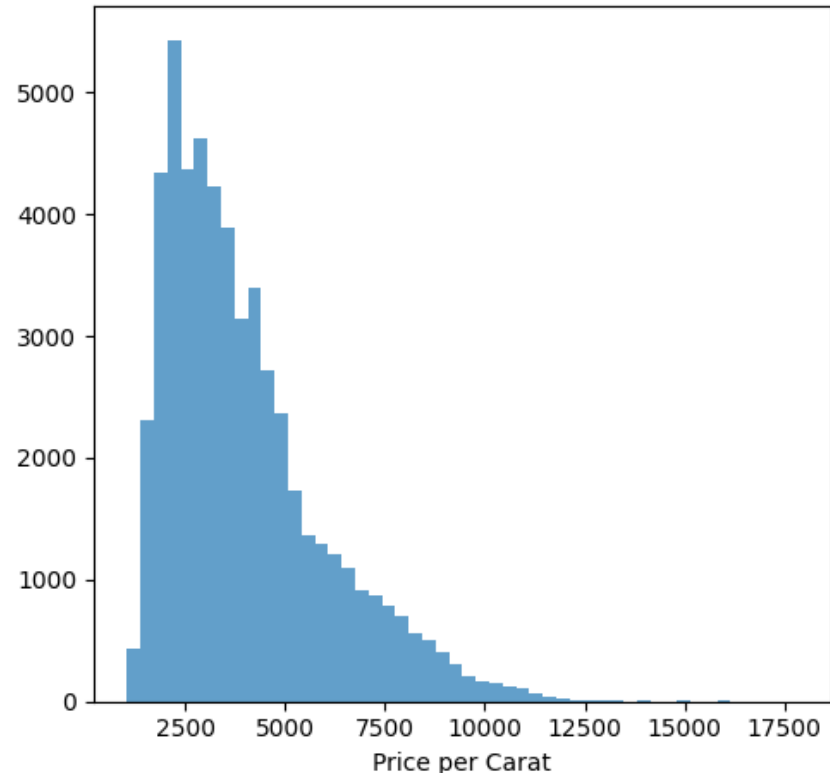
---

- All distributions below are heavily skewed.
- For the Price and Price per Carat have most diamonds priced under \$5,000 and price-per-carat values peaking between \$2,000 and \$3,000. And for the Carat distribution, we can see the peak between around 0.2 and 1.5 carat
- Fewer diamonds have extreme price per carat beyond \$7,000.
- => A majority of diamonds are lower-priced with a small carat, but a small number are extremely expensive.

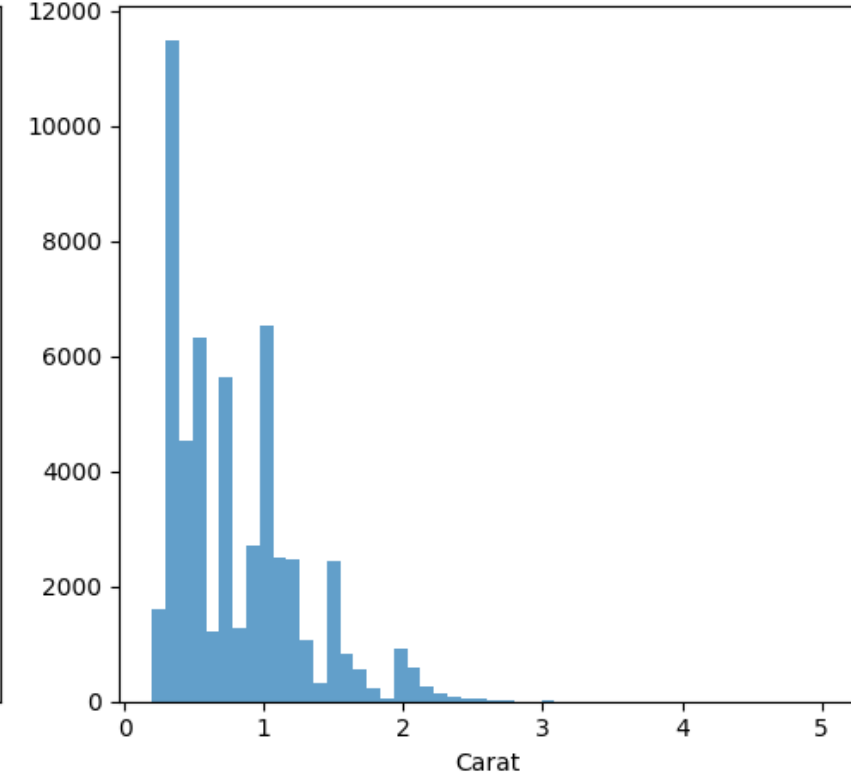
Price Distribution

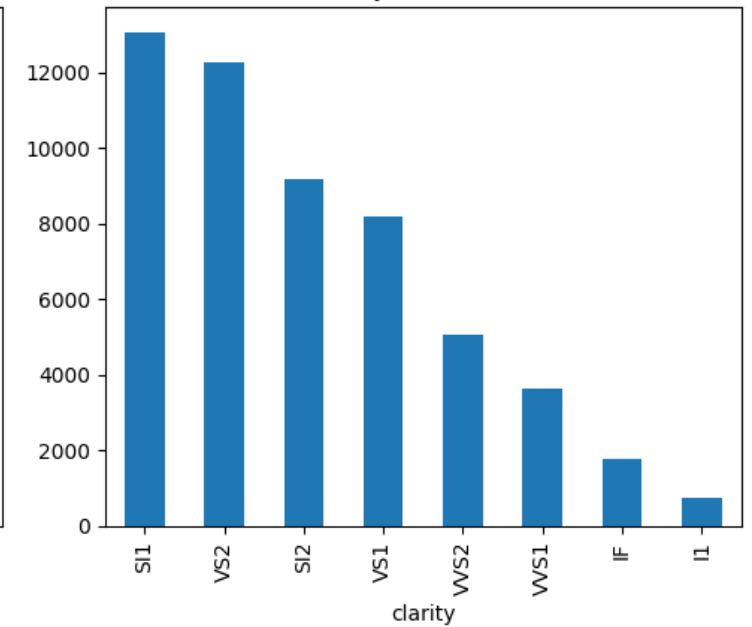
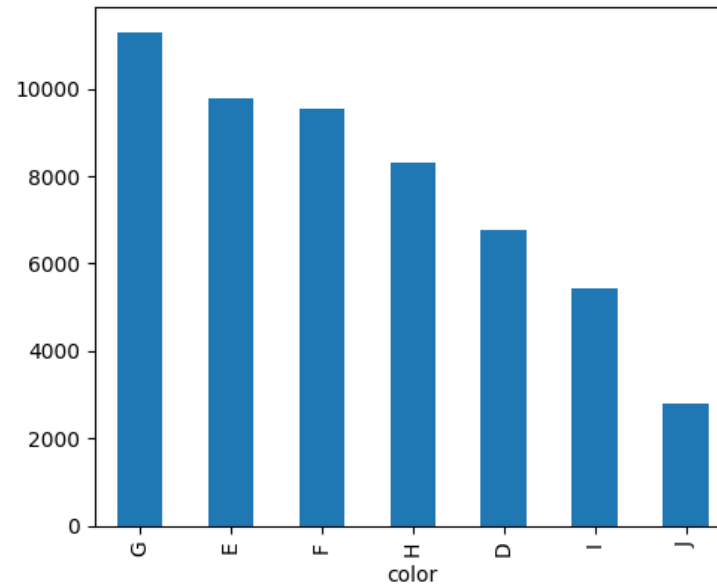
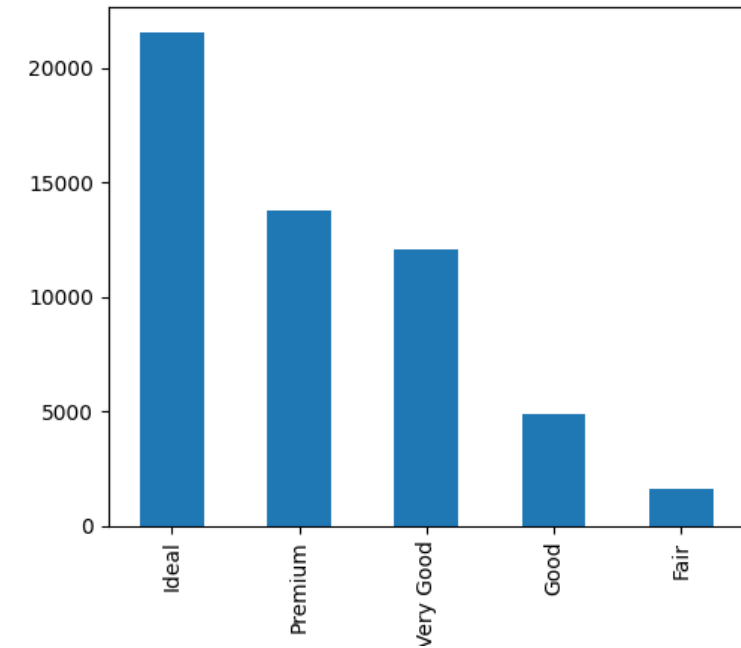


Price per Carat Distribution



Carat Distribution

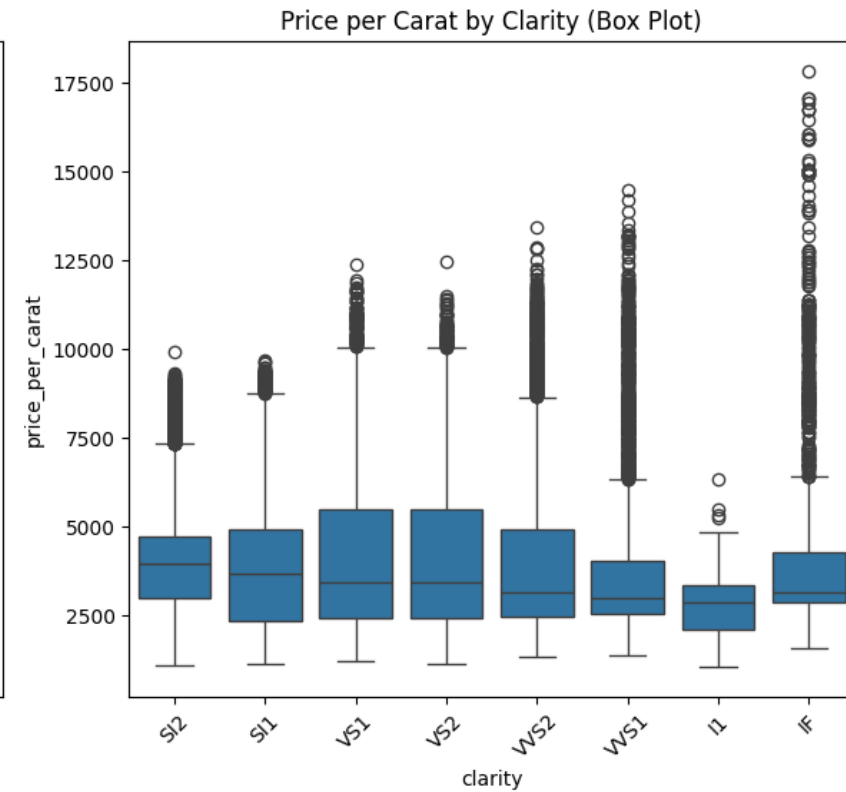
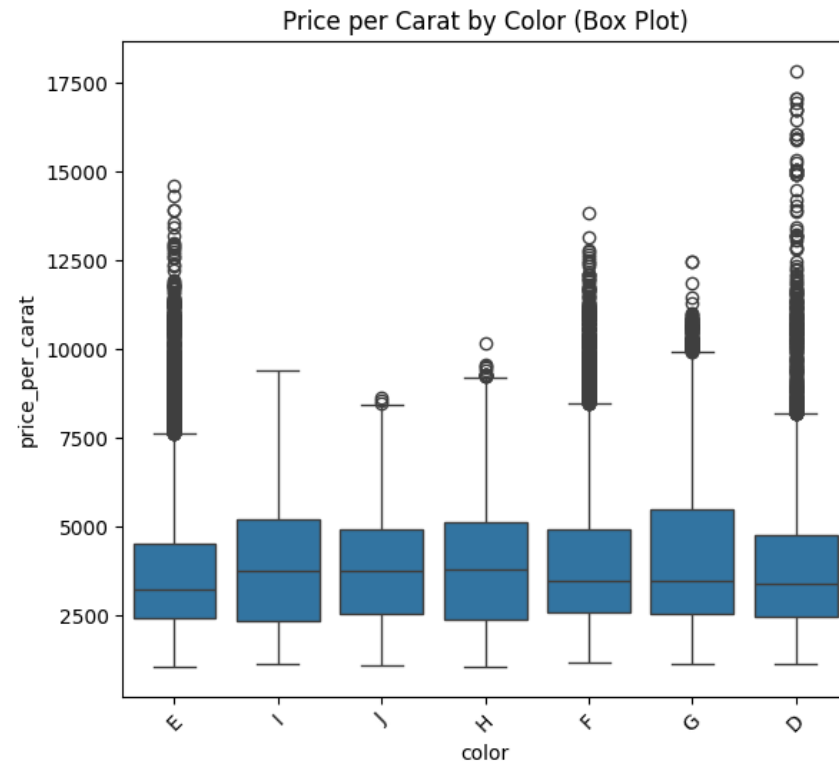
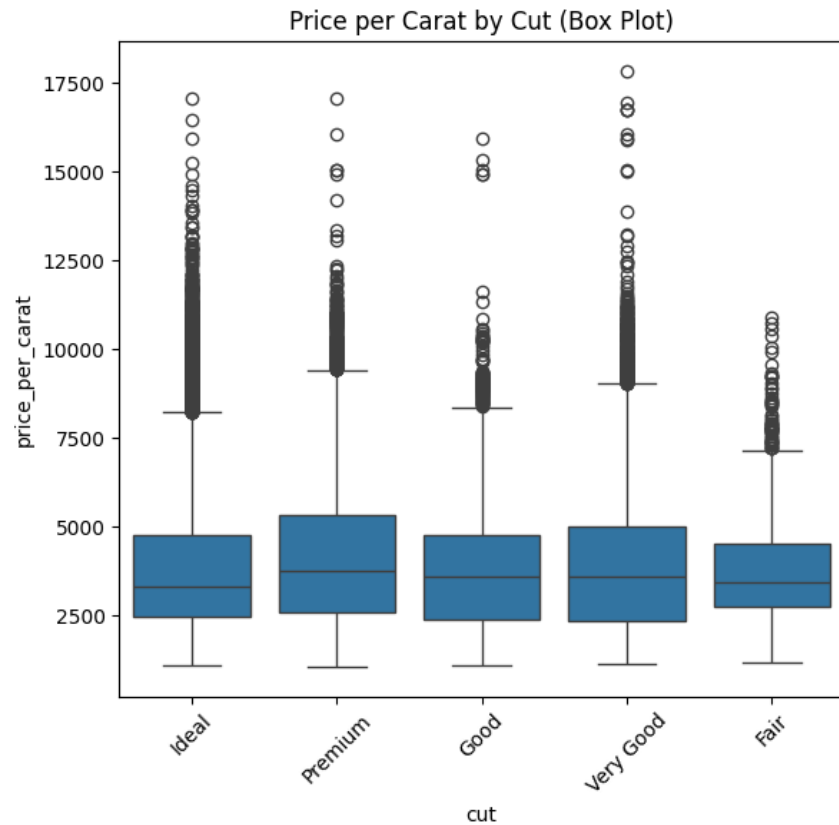




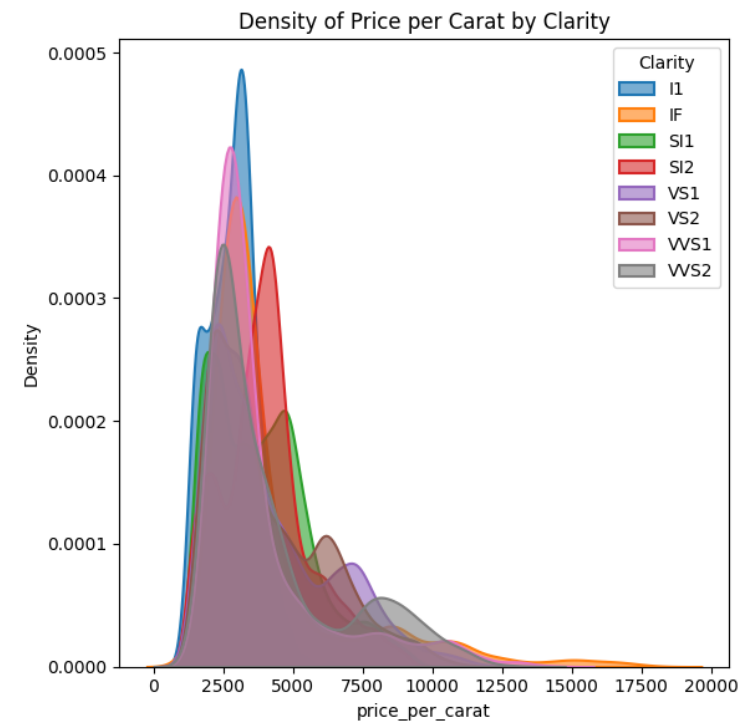
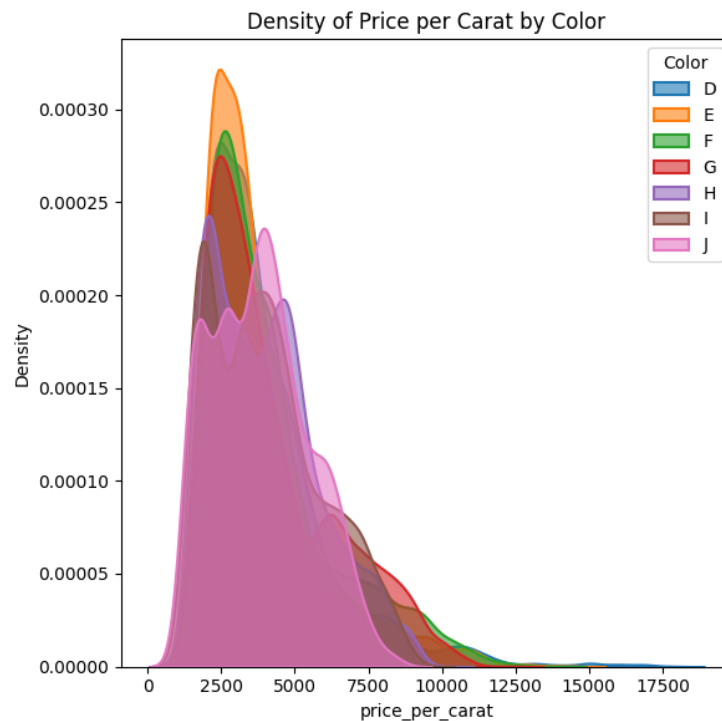
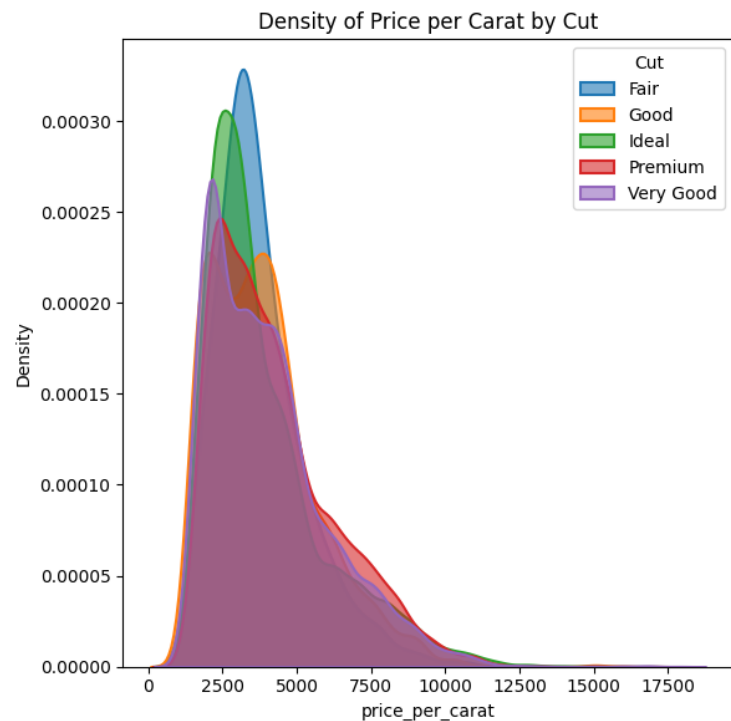
# EXPLORATORY DATA ANALYSIS (PT.2)

- “Ideal” cut is the most common, likely reflecting both demand and supply.
- Color G is most prevalent—suggesting it may be the sweet spot between visual quality and price. The relative scarcity of D may indicate high cost and low affordability, while J may be unpopular due to visible yellow tint.
- Mid-tier clarity grades (SI1, VS2) dominate, which may indicate a consumer preference for balancing price and perceived flawlessness. Higher clarity grades might be overkill for the naked eye but come at a premium.

# EXPLORATORY DATA ANALYSIS (PT.3)







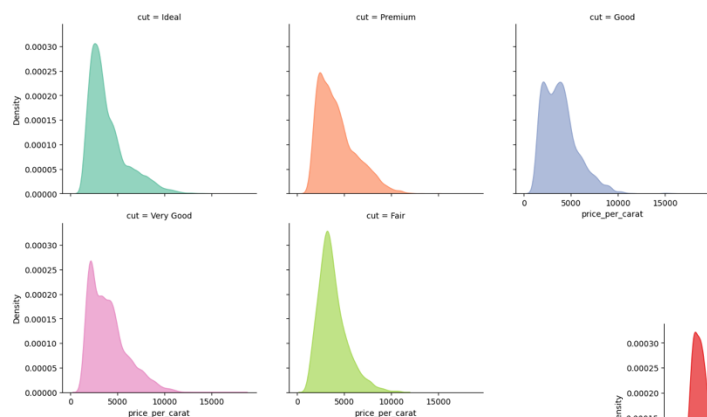
# EXPLORATORY DATA ANALYSIS (PT.4)

- "Ideal" and "Premium" cuts don't always justify the price gap over "Very Good"
- G and H color grades hit the sweet spot in color/price balance
- VS2 and SI1 clarity grades provide high quality at mid-tier pricing — without overspending for imperceptible benefits.

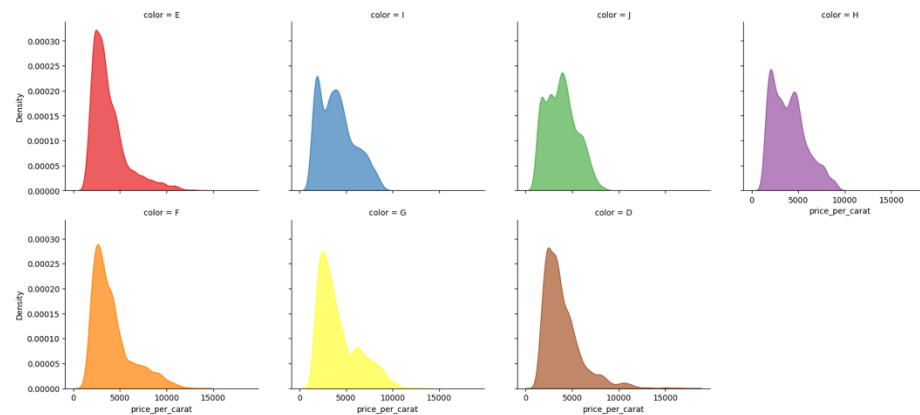
---

# EXPLORATORY DATA ANALYSIS (PT.5)

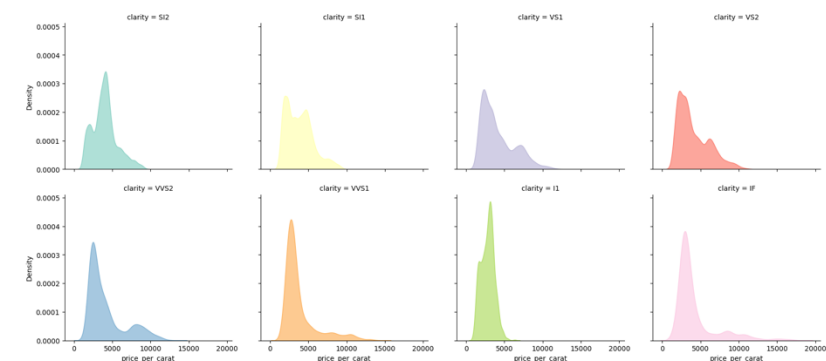
Price per Carat Distribution by Cut

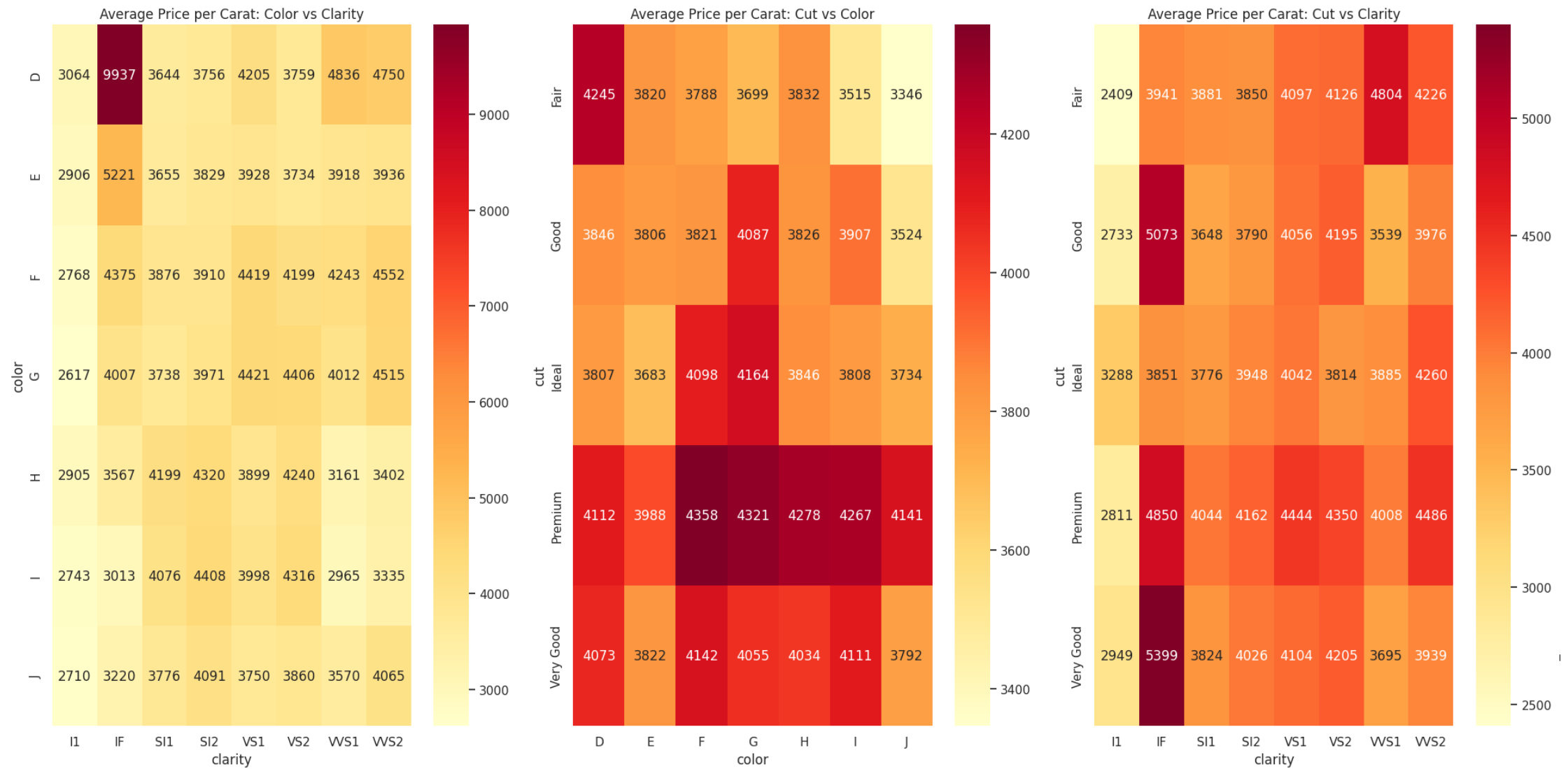


Price per Carat Distribution by Color



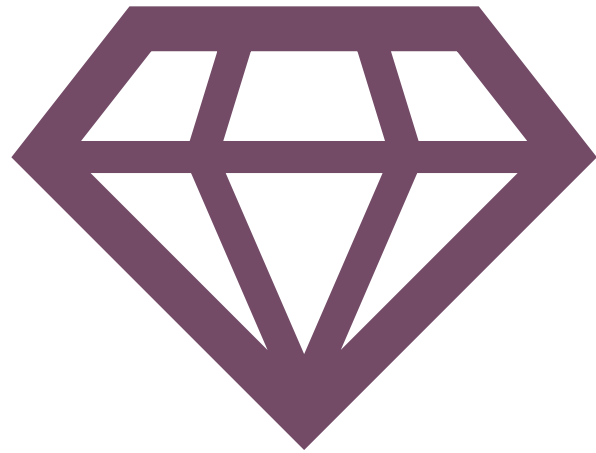
Price per Carat Distribution by Clarity





# EXPLORATORY DATA ANALYSIS (PT.6)

# BUSINESS INSIGHTS Q&A



- 1) Which combinations of clarity + color offer similar visual quality but significantly lower price?
  - If we pick the similar clarity of the diamond, we can choose **SI1** and **SI2** with the color from **D** to **G**, which is the best diamond color you can choose but with an average price below \$4000.
- 2) Is "ideal" cut worth the premium based on price/carats vs. "Very Good" or "Premium"?
  - The “**Ideal**” vs “**Very Good**” cut have the similar price range when it comes to color and clarity, so it's worth the premium. However, “**Premium**” is not worth the premium and the cut is between the “**Ideal**” and “**Very Good**,” but its price is more expensive than the other two.
- 3) Which overrated quality grade drive up cost without proportional visual improvement?
  - Grades like **IF**, **VVS1**, and **D color** tend to **inflate price** disproportionately to their actual visual improvement. Most buyers can achieve the **same aesthetic result** with **VS2–SI1 clarity** and **G–H color**, at **significantly lower cost**.

---

# CONCLUSION & FUTURE RECOMMENDATIONS



## Best Value Sweet Spot:

**Cut:** “Ideal”

**Color:** G or H

**Clarity:** SI1 or VS2



## Recommendations:

**Educate consumers** on visible vs. invisible differences

Build **interactive decision tools** (e.g., dashboards or recommendation engines)

Extend analysis to **brand, shape, or certification** if data available

---

---

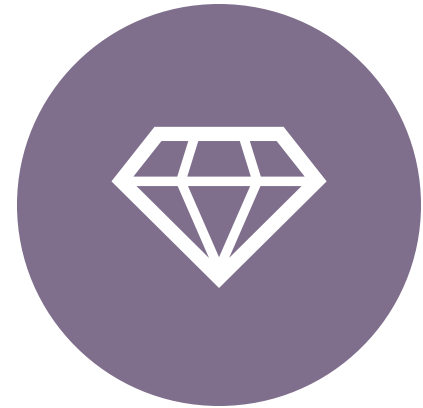
# REFERENCE:



DATASET: [KAGGLE](#)



TOOL: [GOOGLE](#)  
[COLLAB](#)



DIAMOND GUIDELINE:  
[TIFFANY & CO.](#), [DIA](#)

---