COSC 526: Analysis of Factors Affecting Card Price

Jihun Kim
University of Tennessee, Knoxville

William Parham
University of Tennessee, Knoxville

***Abstract*—With constantly fluctuating prices within the Trading Card Game (TCG) markets, it can be difficult to invest constantly in products. Due to our shared interest in card games, we developed a way to solve this problem within the Pokémon Trading Card Game. This project first determines whether rarity is the sole contributor to a card's price; then it expands upon this idea to determine if and what other factors, such as the quality of a card (or the "grade"), influence a card's price, allowing you to invest smarter and more confidently.**

Keywords: K-Means, TCG, Grade, Rarity

I. Introduction

Trading card games are known for having explosively volatile markets, and Pokémon is no exception. Every new release causes new product prices to soar through the roof while re-igniting prices for existing sets. By catching the eye of new, novice collectors and luring seasoned veterans back for more, the market is constantly changing.

This makes Pokémon a particularly difficult market to invest in since predicting price trends can be difficult. Due to our shared interest in trading card games, we aim to solve this by applying the data analytics tools we learned in Data Engineering to help analyze the price trends of Pokémon sets.

Our project breaks this down into a three-step process of Data Acquisition, Data Cleaning, and Data Analysis.

II. Methodology

A. Combine Datasets

Since our data came from multiple sources, we first needed to merge them into a single cohesive data set for analysis. To facilitate our work in Python and PySpark, we decided to combine the price trends dataset [2] with the specific card information dataset [1] and export the result as a CSV file.

B. Scrape Web Decklists

In order to effectively determine whether or not meta decklists affected card prices, we needed to understand what meta deck construction looked like. To do this, we scraped a website dedicated to tracking meta Pokémon decklists [3]. We took this data and combined it with the card information dataset to create an additional dataset that we could compare against the original.

C. Clean Dataset

To effectively use our data, we needed to clean it. This process involved coalescing all of our data into one master set. This allowed us to use different analytical tools and methods, from trend plotting to K-means clustering. In order to clean our data, we had to convert all dates into mm-dd-yy format, convert prices into a monthly average for comparison, and convert rarity from a textual format to a numerical format. This allowed us to have a format baseline against which to analyze our data.

D. Analyze the Data Set

1) Heatmaps: Heatmaps provided a quick and effective way to visulaize overall trends in the data, such as average set prices, grade card prices, and rarity overtime. However, they were less effective at capturing outliers, which became a major factor in understanding our data.

2) Scatterplots: After recognizing the importance of outliers in our dataset, we introduced visualizations that were better suited to highlighting them. Scatterplots were particularly useful, allowing us to track extreme price values over time and identify the rarities most commonly associated with these outliers (primarily Ultra Rare and Rainbow Rare cards).

3) Line Graphs: After creating the scatterplot, we experimented with a line graph to capture the broader trends over time while still accounting for outliers. This graph effectively showed a monthly comparison of different rarities. Encouraged by the clarity provided by the graph, we prioritized outliers in my analysis.

4) Boxplots: By overlaying a boxplot on our scatterplot, we can see how major outliers affect the average and weight of our line graphs. Viewing these in conjunction shows what cards have a major impact on rarity and how these cards change in price on a month-to-month basis.

Since we knew that outliers were so important, we wanted to cluster our dataset and see which cards fell into affordable, moderate, and expensive price ranges.

5) K-Means: Clustering provided the perfect medium for this. Using the K-means algorithm with a K-value of 3 chosen in an elbow fashion, we categorized cards into cheap, medium, and expensive cards based on the price data.

However, the initial clustering results were unexpected. Most cards were categorized as medium or expensive, which contradicted our prior knowledge that most cards were cheap. This prompted us to examine the data more closely.

Upon re-examining our data, we realized that this graph is correct, however, our scales are disproportionate. The expensive outliers were dragging the graph so far upward that we overly compressed the most affordable cards. To better visualize this, we performed the same k-means treatment but plotted our data on a log scale as shown on figure 1, which visualized our results much more cleanly, and proved our assumption correct that the vast majority of cards fall into the affordable range.

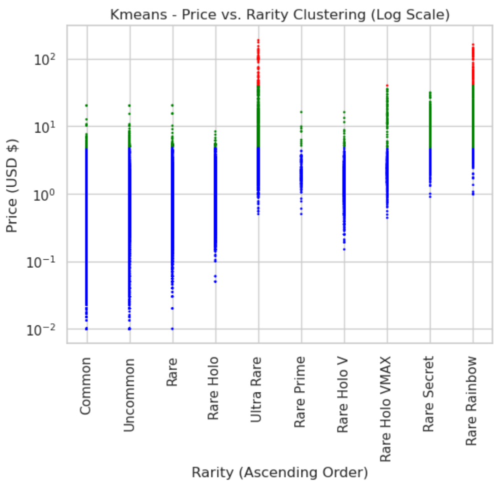

Fig. 1. K-Means Clustering with Log Scale

III. Key Results

1) Rarity: First, we answered the question of whether or not a higher rarity impacts the price of a card. Our results found this to be generally true, except a few outliers as discussed previously.

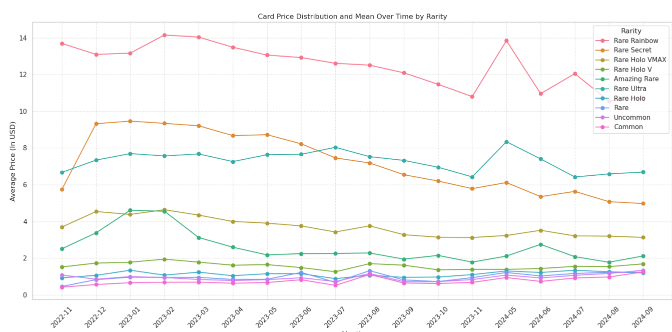

Fig. 2. Rarity Price Trends

As figure 2 illustrates, cards that are generically rarer simply get opened less and garner a higher price because there are fewer in circulation. However, since we already determined that there are outliers we wanted to investigate what caused these outliers and how this, in conjunction with rarity, influenced a card's price.

2) Grade: Our first intuition is that since Pokémon is largely a collector's TCG, the card's grade would greatly influence the card's price. Our data proved this intuition correct. However, this was generically true across all rarities, which made it difficult to explain why some common rarity cards could be worth as much as a rainbow rarity card. If both had equal grades, they generally followed the rarity-to-price trend.

3) Meta Decks: Our second hypothesis was that meta decklists with high competitive viability would drive up the prices of the included cards. To determine this, we compared a dataset of meta cards to our original dataset of all cards.

However, the most expensive card in the meta decklist dataset was a Cinderace VMAX, which barely broke the threshold for moderately priced cards. After the Cinderace, all other cards quickly fell into the affordable range of the clustering graph. This led to the conclusion that meta viability does not greatly impact a card's price. This gave us a hint that the card's text may not matter, which led to our next hypothesis.

4) Illustrator: Our final hypothesis is that the illustrator impacts the price of a card. The reasoning behind this is that if the competitive context of a card didn't matter, that meant we needed to look in the direction of collectability. To do this, we plotted card prices by illustrator over a monthly period.

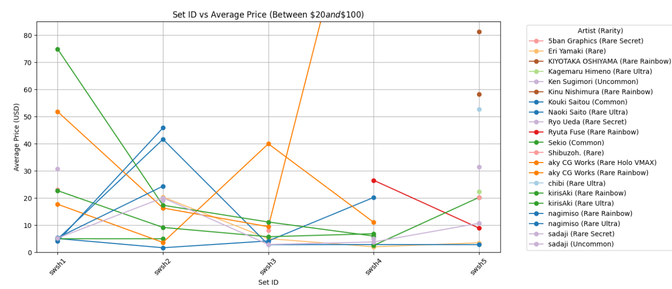

Fig. 3. Illustrator Price Trends

This allowed us to conclude that the illustrator seems to have some impact on a card's price due to their unique or interesting art style it is still difficult to clearly understand the degree of impact they have on a card.

IV. Conclusion

Our analysis revealed several key factors that influence Pokémon card prices:

- **Rarity**: Even though there are some notable exceptions, cards with higher rarity generally have higher prices.
- **Grade**: For identical cards, those with higher grades are traded at higher prices.
- **Meta Deck Inclusion**: Being included in competitive meta decks has limited impact on a card's market price. Most competitively viable cards remain relatively low in value.
- **illustrators**: Certain illustrators tend to be associated with higher card prices, though not universally.

V. Future Work

In future research, we plan to develop predictive models that estimate card prices based on historical market trends. These models could help forecast price fluctuations and identify potential investment opportunities.

References

[1] adback03. pokemon-tcg-data. https://github.com/PokemonTCG/pokemon-tcg-data, 2025.

[2] Aviortheking. Tcgdex pricing history - pokemon tcg cards database. https://github.com/tcgdex/price-history, 2025.

[3] Robin Schulz. Decks – Limitless — limitlesstcg.com. https://limitlesstcg.com/decks, 2017. [Accessed 08-05-2025].