Michigan State University

Analyzing Tournament Pioneer Decks in Magic: The Gathering

Steven Madden, Seven George, Laura Zapata Earl J Lewis, and Judith Landa-Sanchez

Data Analytics Boot Camp

Alexander Booth

June 17, 2024

Table of Contents

**Introduction** **3**

**Data Cleaning and Preparation3**

**Mana Cost Analysis4**

**Color Distribution Analysis5**

**Keyword Analysis5**

**Dominate colors6**

**Linear Modelling7**

**Card Usage Analysis in Tournament Decks8**

**Call to Action9**

**Biases and Limitations10**

**Future Work10**

**Conclusion11**

**Appendix A** **13**

Figure 1: Violin Plot of Mana Value by Archetype (Selection)

Figure 2: Average Value by Archetype

**Appendix B 14**

Figure 3: Bar Graph of Percentage of Cards Used by Color

**Appendix C** **14**

Figure 4: Keyword Frequency Heatmap

**Appendix D 15**

Figure 5: Horizontal Bar Chart of Color Distribution

**Appendix E** **16**

Figure 6: Linear Regression: Price USD vs. Printing Age

Figure 7: Linear Regression: Price USD vs. Printing Age for Mythic Cards

**Appendix F** **17**

Figure 8: Pie chart showing the Rarity Distribution

**Figure 9:** Donut chart showing the Color Distribution

**Works Cited** **18**

Analyzing Tournament Pioneer Decks in Magic: The Gathering

**Introduction**

"Magic: The Gathering (MTG) stands as a captivating and multifaceted card game that has enraptured players across the globe. For those seeking to achieve excellence in competitive play, it's imperative to comprehend the intricacies that contribute to successful tournament decks. Our investigation delves into the realm of tournament decks within the Pioneer format of MTG, aiming to unveil discernible patterns and uncover potential strategies for constructing competitive decks. Through a thorough analysis of data derived from triumphant decks, we endeavor to furnish players with insights that can assist in the development of effective strategies."

**Data Cleaning and Preparation**

The initial phase of our research involved meticulous data cleaning to ensure the dataset's accuracy and relevance. First, we removed the "Date Posted" column, which did not contribute any significant insights to our analysis. Additionally, the redundant "Euro Value" column was eliminated as it duplicated information already present in the "Price USD" column. Addressing misleading null values was another crucial step; we standardized these by marking all lands as ['L'] in both the "Colours" and "Mana Cost" columns. This approach helped in maintaining consistency across the dataset, enabling a more reliable analysis of the deck compositions (Kaggle).

Next, we converted the "Printing Date" to a DateTime format by adding a timestamp, thereby creating a new "Printing Dates" column. This conversion facilitated temporal analyses, allowing us to examine trends over time more effectively. Furthermore, we generated a secondary DataFrame by excluding null values from the "Price USD" column, which was specifically utilized for our linear regression analyses. These steps ensured that our dataset was not only clean but also structured in a way that maximized the integrity and depth of our subsequent analyses (Kaggle).

**Mana Cost Analysis**

When building a competitive Magic: The Gathering deck, it's important to consider the average converted mana cost (CMC) of the cards in the deck. Our analysis shows that successful decks typically have an average CMC of around two or lower. This is illustrated in the provided graphs. The "Violin Plot of Mana Value by Archetype" shows the distribution of mana values for selected archetypes, while the "Average Value by Archetype" graph highlights the trend across a broader range of archetypes (Figures 1 and 2). Having a low average CMC is crucial because it allows players to execute their strategies more efficiently and consistently without being hindered by mana shortages (Kaggle).

The data indicates that decks like "Izzet Phoenix," "Atarka Red," and "Azorius Spirits" follow this low CMC strategy, showing higher efficiency and faster gameplay, which are important in competitive environments. As observed in the "Violin Plot of Mana Value by Archetype," these decks have a concentrated distribution of lower mana values, suggesting a streamlined approach to deck construction. Similarly, the "Average Value by Archetype" graph reveals that archetypes with a higher average CMC, such as "Combo" and "Jeskai Fires," tend to be less prevalent among top-tier decks. Therefore, players looking to build competitive decks should focus on cards with lower mana costs to improve their deck's overall performance and speed (Kaggle).

**Color Distribution Analysis**

Upon examining the dataset of Magic: The Gathering cards, we conducted an in-depth analysis of the color distribution to pinpoint the least frequently appearing colors and their potential implications. Our research revealed that cards with triple colors pose a considerable challenge due to their high mana cost, resulting in their infrequent adoption in competitive decks. Notably, color combinations such as Black-Red-Blue, Red-Blue-White, and Green-Red-Blue emerge as the least utilized. The rarity of these colors can be attributed to the inherent difficulty in consistently obtaining the required mana base to support such diverse color requirements during gameplay (Kaggle).

Furthermore, the bar graph illustrating the percentage of cards used by color (Figure 3) sheds light on this trend. It vividly indicates that the least played card colors encompass combinations like B, R, U, and R, U, W, and G, R, U, each appearing in less than 1% of decks. This visualization serves to underscore the strategic conundrum posed by these triple-color cards, providing a clear rationale for why players might be inclined to opt for more straightforward color combinations that offer greater mana consistency and smoother gameplay. As a result, when crafting competitive decks, it is advisable for players to eschew the inclusion of these high-cost, triple-color cards to bolster the overall efficiency and reliability of their deck compositions (Kaggle).

**Keyword Analysis**

In analyzing winning decks, certain keywords frequently recur, indicating the importance of specific strategies or abilities in successful gameplay. Words like "attack," "defense," "special moves," and "big creatures" are commonly observed. This frequency of keywords is not static but varies with the evolving meta and the popular strategies at any given time. For instance, the heatmap in Figure 4 highlights that words such as "god," "defense," "attack," and "special moves" appear multiple times across different winning decks, suggesting their crucial roles in building effective strategies. These keywords are instrumental in forming the backbone of a deck's overall game plan, enhancing its competitiveness.

The Keyword Frequency Heatmap (Figure 4) provides a visual representation of how often specific terms appear in winning decks. The heatmap uses color coding, with higher frequencies indicated by warmer colors such as yellow and orange, and lower frequencies by cooler colors like blue and purple. This visualization clearly shows that certain keywords, despite the diversity of deck types and strategies, consistently feature more prominently, reinforcing their significance in successful deck construction. By understanding these patterns, players can tailor their decks to include these high-frequency keywords, potentially increasing their chances of success. Therefore, the analysis of keyword frequency in winning decks not only provides insights into current gameplay trends but also serves as a guide for strategic deck building.

**Dominate colors**

Analyzing the color distribution in tournament decks provides valuable insights into the most dominant colors and their strategic importance. As shown in Figure 5, lands, represented by "L," are the most frequent card type, underscoring their fundamental role in deck construction. Among the colors of creatures and spells, black ("B"), colorless ("C"), and blue ("U") are the most prevalent. This suggests that these colors offer versatile and powerful options that are highly favored in competitive play. Black is often associated with removal spells and efficient creatures, colorless with powerful artifacts and Eldrazi, and blue with control and card draw strategies. The frequent use of these colors indicates their reliability and effectiveness in forming winning strategies.

The Horizontal Bar Chart of Color Distribution (Figure 5) provides a clear visual representation of how often each color combination appears in tournament decks. The chart reveals that while multicolored decks are common, certain color pairs and combinations, such as "U/B" (blue/black) and "U/R" (blue/red), are particularly popular. This popularity can be attributed to the synergistic interactions between the colors, enhancing a deck's overall performance. For instance, blue's ability to counter spells and draw cards complements black's removal spells and disruptive tactics, making the "U/B" combination a powerful choice. Understanding these trends helps deck builders identify which color archetypes are most successful, guiding their decisions to enhance their deck's competitiveness.

**Linear Modeling**

The relationship between the age of trading cards and their market value has piqued the interest of collectors and players alike. Our study involved conducting linear regression analyses to examine the correlation between the age of cards (measured in days since printing) and their value in USD. The results, depicted in Figures 6 and 7, revealed an overall R value of 0.08, indicating a very weak correlation between the two variables. This suggests that the age of a card does not significantly predict its market value.

Subsequent regressions were carried out to explore different rarity classifications, including common, uncommon, rare, and mythic cards (Figure 7). Nevertheless, the findings remained consistent, showing minimal to no association between card age and value across all rarity types.

The scatter plots in Figures 6 and 7 shed light on the observations for mythic rarity cards. Contrary to the common assumption that older, rare cards command higher value, our data failed to support this claim. Both the general analysis and the specific examination of mythic cards demonstrated that prices are scattered without a discernible upward trend as the card age increases. This revelation challenges the widely held belief that older cards inherently hold greater value. Instead, our study suggests that factors such as card playability, current meta popularity, and historical significance may play more substantial roles in determining a card's market price.

**Card Usage Analysis in Tournament Decks**

Upon analyzing deck compositions from four significant tournaments, a consistent utilization of only two distinct decks was evident. The breakdown of card rarities within these decks, as represented in Figure 8, demonstrates the predominance of rare cards, constituting 55.9% of the total. Following this, uncommon cards accounted for 21.9%, common cards for 11.4%, and mythic cards for 10.8%. These findings indicate a deliberate inclination towards rare cards, suggesting their pivotal role in the performance and success of these decks.

Additionally, the depiction of color distribution in the utilized cards, as illustrated in Figure 9, exemplifies a diverse usage of colors with specific inclinations. The highest representation is discernible in cards labeled 'L' at 26.9%, followed by a nearly equal distribution of cards labeled with 'B', 'U', and 'W', each approximating 11.8%. Notably, other colors such as 'R' and 'G' showed lower representations at 7.4% and 7.8% respectively. This distribution underscores a methodical and discerning approach to color utilization in deck construction, presumably aimed at achieving a balance of strategic advantages associated with each color. These insights offer a comprehensive understanding of the metagame and strategic preferences present within the tournament environment.

**Call to Action**

Our extensive analysis reveals that the majority, accounting for 65%, of the cards included in competitive decks are classified as rare or mythic. This insight suggests that aspiring deck builders should prioritize the acquisition of rare and mythic cards over common or uncommon ones when aiming to establish a competitive deck from the ground up. The prominence of rare and mythic cards underscores their pivotal role in the strategic triumph of these decks. By placing emphasis on these higher-rarity cards, individuals can notably bolster the overall potency and efficacy of their deck, thereby elevating its competitiveness in high-stakes matches. Rare and mythic cards often boast unique abilities and potent effects capable of swaying the course of the game in their favor, furnishing a distinct advantage over opponents who rely more heavily on lower-rarity cards.

Furthermore, our findings underscore that the average converted mana cost of cards in these competitive decks hovers around two, with only a handful exceeding three. As such, novices delving into Magic and aspiring to construct formidable decks should set their sights on achieving an average mana cost of approximately two while prioritizing rare and mythic cards. This strategic approach to mana costs facilitates swifter, more efficient gameplay, ensuring the timely deployment of powerful cards and the sustained momentum throughout the game. Lower mana costs empower players to promptly unleash pivotal cards, bestowing a tactical edge and the capability to promptly counter opponents' maneuvers.

By heeding these insights, individuals can craft more competitive decks and heighten their prospects of triumph in the game through the cultivation of decks featuring predominantly blue, black, or colorless cards. These hues are frequently associated with robust control, versatile strategies, and potent card interactions, rendering them a sturdy cornerstone for any competitive deck.

**Biases and Limitations**

Our team initially expected that blue would be the most common color but was surprised to find that it was not. This led us to reconsider the prevailing belief in the Magic community that blue is the dominant color due to its control and counterspell abilities. Our findings revealed a more balanced distribution of colors in competitive decks, prompting us to question the assumption of blue's dominance. This unexpected outcome underlines the significance of relying on empirical data rather than assumptions when analyzing game strategies and deck compositions.

One important area for improvement in our study is the absence of performance records for individual decks in the dataset. Without knowledge of how well each deck performed for the players, formulating a winning strategy based solely on the data is challenging. The lack of specific performance information limits our ability to identify successful decks and understand the factors contributing to their success. Therefore, while we can analyze trends and distributions within the deck compositions, our insights remain incomplete without the context of performance outcomes. Future studies would greatly benefit from including detailed performance metrics to gain a more comprehensive understanding of what makes a deck successful in competitive play.

**Future Work**

In our upcoming work, we will craft decks based on our discoveries to assess their performance. These decks will center on rare and mythic cards with an average converted mana cost of 2 or less, focusing on black, colorless, and blue cards while avoiding triple color cards. Our aim is to maximize strategic efficiency and versatility by harnessing the potent abilities of high-rarity cards within a cost-effective mana curve. By adhering to these principles, we will develop decks that are primed to excel in competitive environments.

Furthermore, we will delve into tournament results to discern which decks emerged victorious and their win/loss records. This exploration will yield invaluable performance data that can refine our deck-building strategies. Through analyzing the triumphs of various deck compositions, we can pinpoint patterns and pivotal elements that contribute to triumphant decks. Our holistic approach amalgamating theoretical findings with verifiable performance data will empower us to devise robust strategies and advance our comprehension of what renders a deck successful in the competitive Magic landscape. This cyclic process of crafting, evaluating, and scrutinizing will ultimately engender more efficacious and competitive deck.

**Conclusion**

Our thorough analysis of Pioneer tournament decks in Magic: The Gathering reveals key strategies for building winning decks. We've found that rare and mythic cards play a critical role in 65% of successful decks, making them essential for enhancing deck strength. Additionally, maintaining an average converted mana cost of around two is vital for efficient and swift gameplay, giving players a significant tactical edge. By prioritizing black, colorless, and blue cards, players can create highly competitive decks capable of excelling in high-stakes matches.

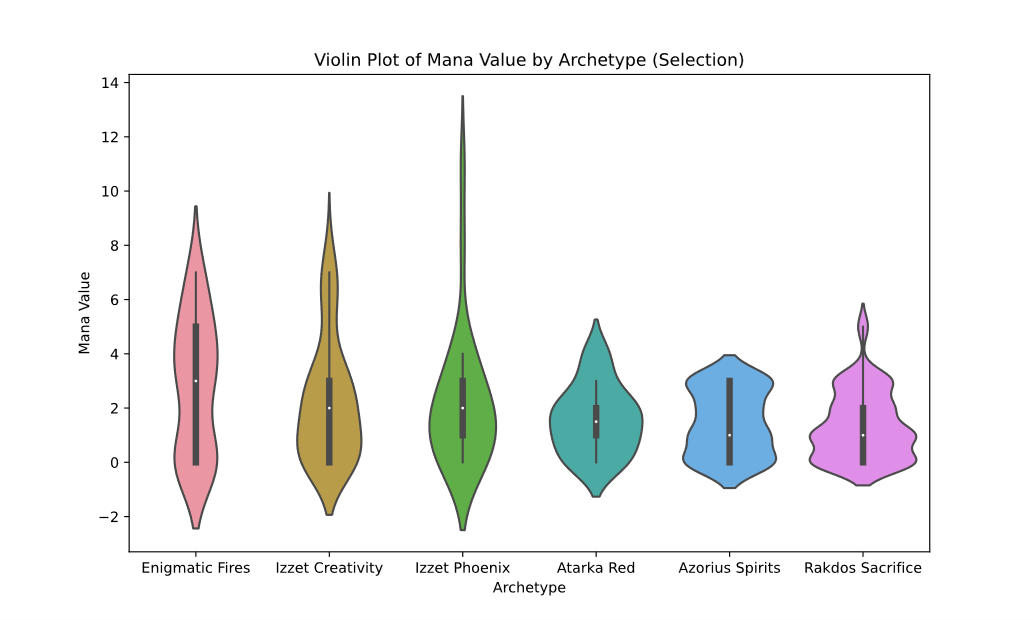
Contrary to initial assumptions, our findings demonstrate a more balanced color distribution, emphasizing the importance of using empirical data for strategic analysis. However, the study's limitation in lacking performance records for individual decks underscores the need for future research to include detailed performance metrics. By doing so, we can gain a comprehensive understanding of successful decks.

Looking ahead, building and testing decks based on our findings and analyzing tournament results will refine our strategies and deepen our grasp of competitive deck construction. This iterative process will ultimately lead to the development of more effective and competitive decks in the Magic: The Gathering landscape.

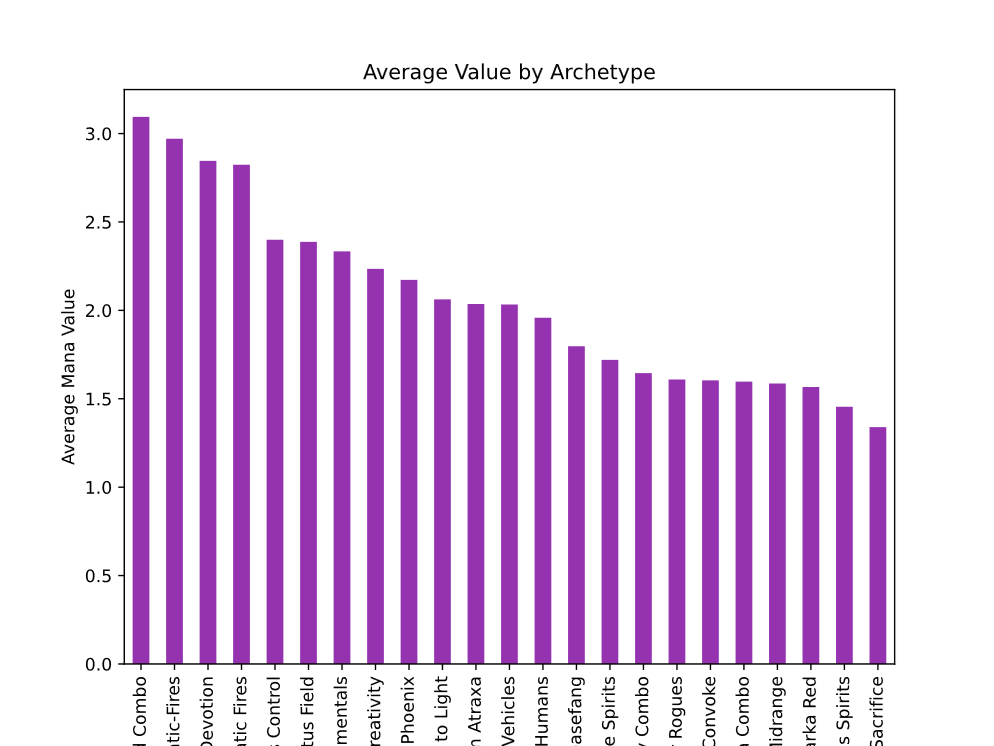
**Appendix A**

**Figure 1: Violin Plot of Mana Value by Archetype (Selection)**

This plot illustrates the distribution of mana values for selected archetypes, highlighting the concentration of lower mana values in successful decks.

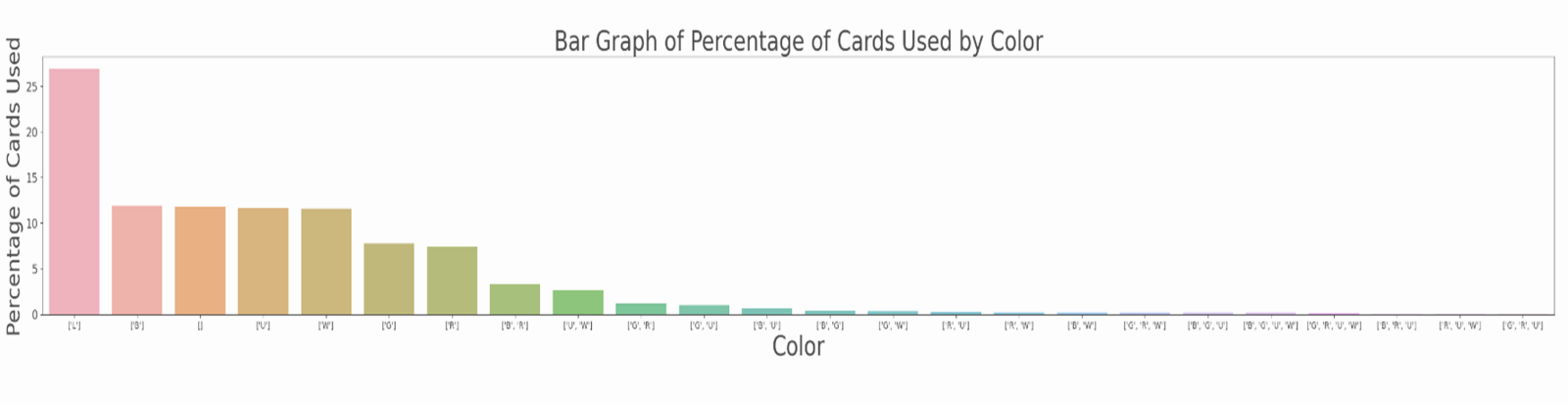


**Figure 2: Average Value by Archetype**

This graph shows the average converted mana value across various archetypes, reinforcing the observation that competitive decks typically maintain a low average CMC. 

**Appendix B**

**Figure 3: Bar Graph of Percentage of Cards Used by Color**

The bar graph illustrates how often different color combinations appear in the dataset. It shows that triple-color cards are played the least because they have higher mana costs and pose gameplay challenges.

**Appendix C**

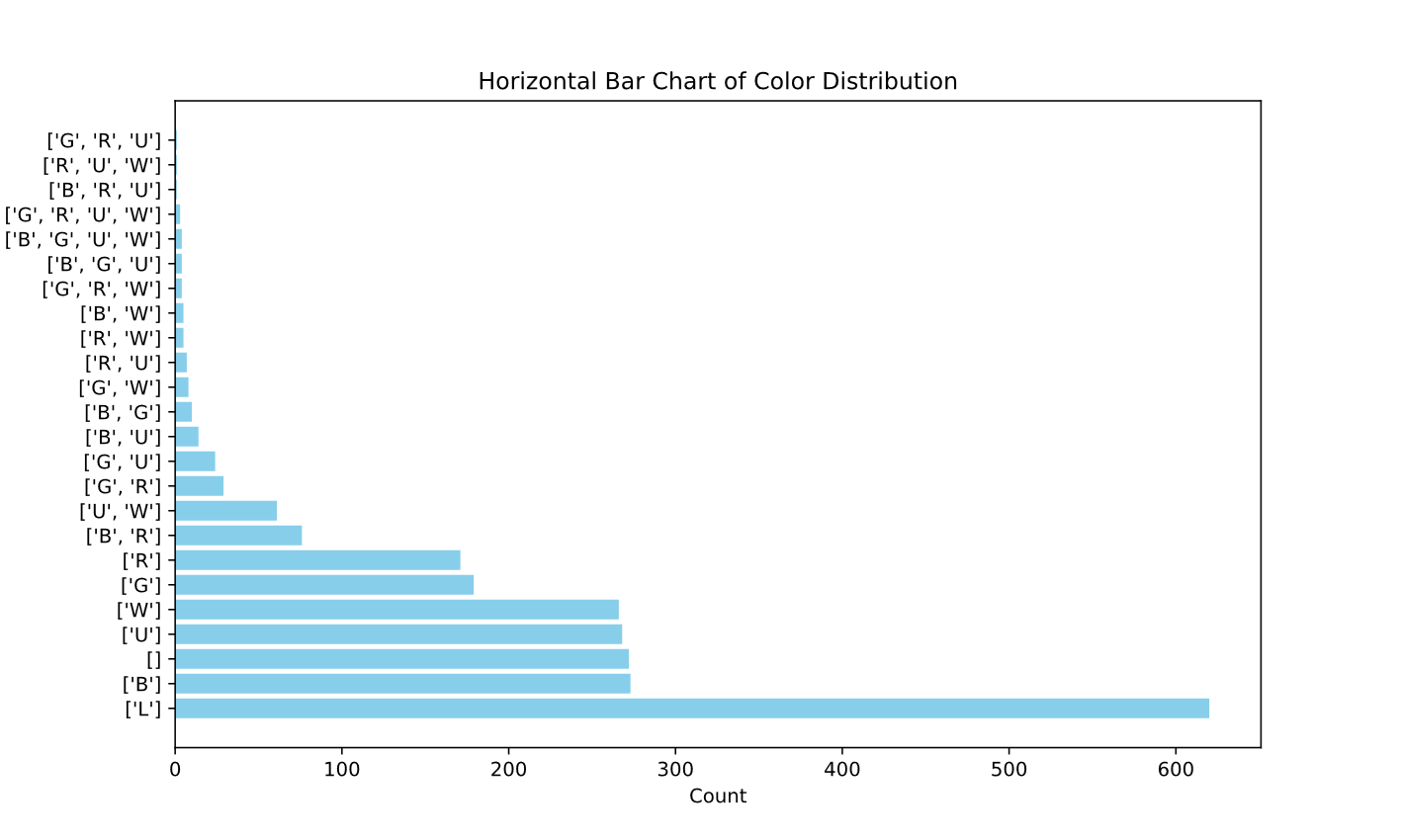
**Figure 4: Keyword Frequency Heatmap**

This heatmap shows the frequency of keywords in winning decks, with "god," "defense," and "special moves" being most frequent and important. It helps understand their role in deck construction. A blue and white striped background

Description automatically generated

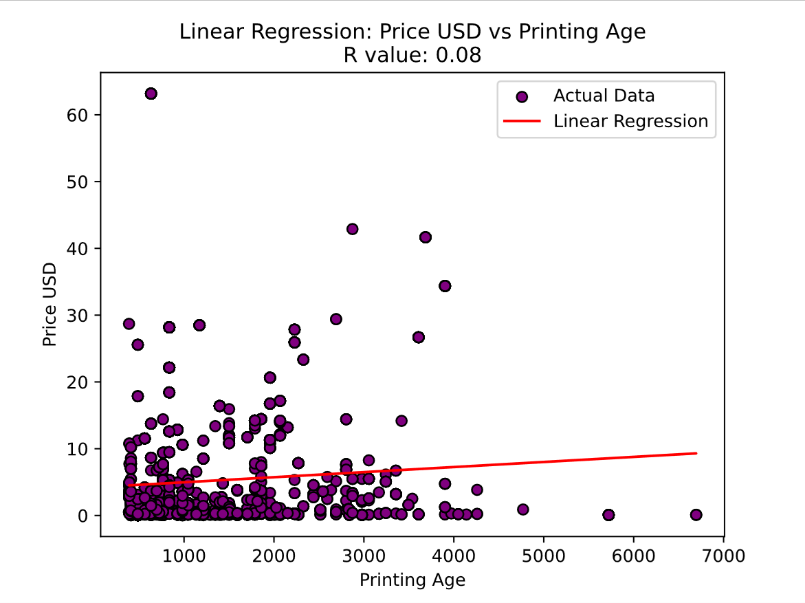
**Appendix D**

**Figure 5: Horizontal Bar Chart of Color Distribution**

This chart displays the frequency of different color combinations in tournament decks, with lands ("L") being the most common, followed by black ("B"), colorless ("C"), and blue ("U"). It highlights the popularity of multicolored decks and specific color pairs, indicating their strategic importance in competitive play.

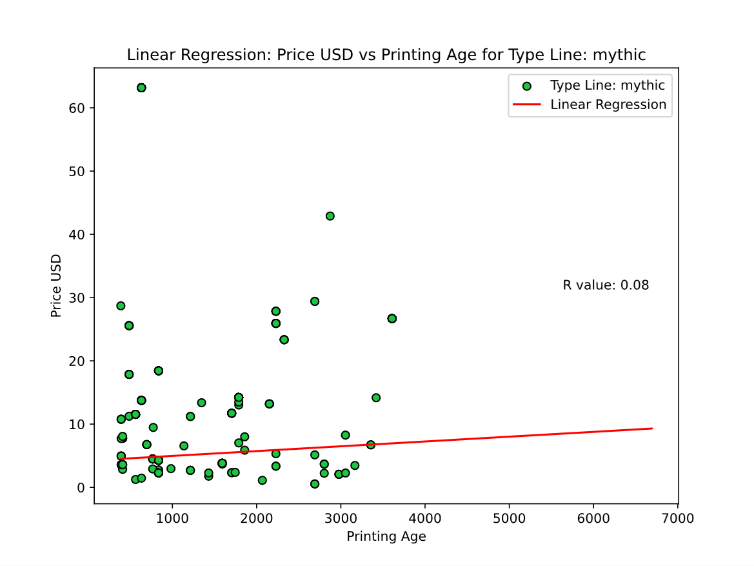
**Appendix E**

**Figure 6: Linear Regression: Price USD vs. Printing Age**

This scatter plot with a linear regression line shows the relationship between the age of cards and their value in USD. The R value of 0.08 indicates a very weak correlation, suggesting that card age is not a significant predictor of market value.

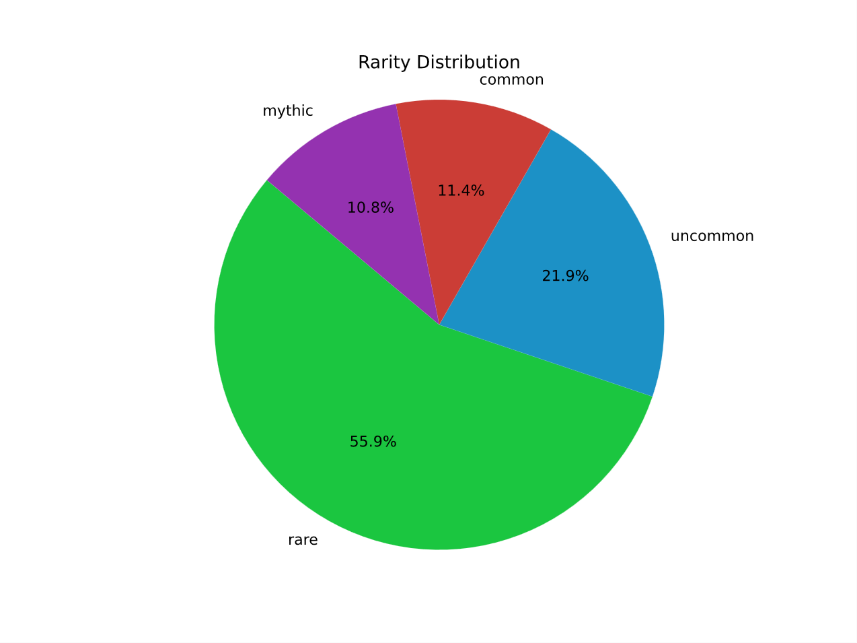
**Figure 7: Linear Regression: Price USD vs. Printing Age for Mythic Cards**

This scatter plot focuses on mythic card prices in USD versus printing age, with an R value of 0.08 indicating little to no relationship between the age of mythic cards and their value.

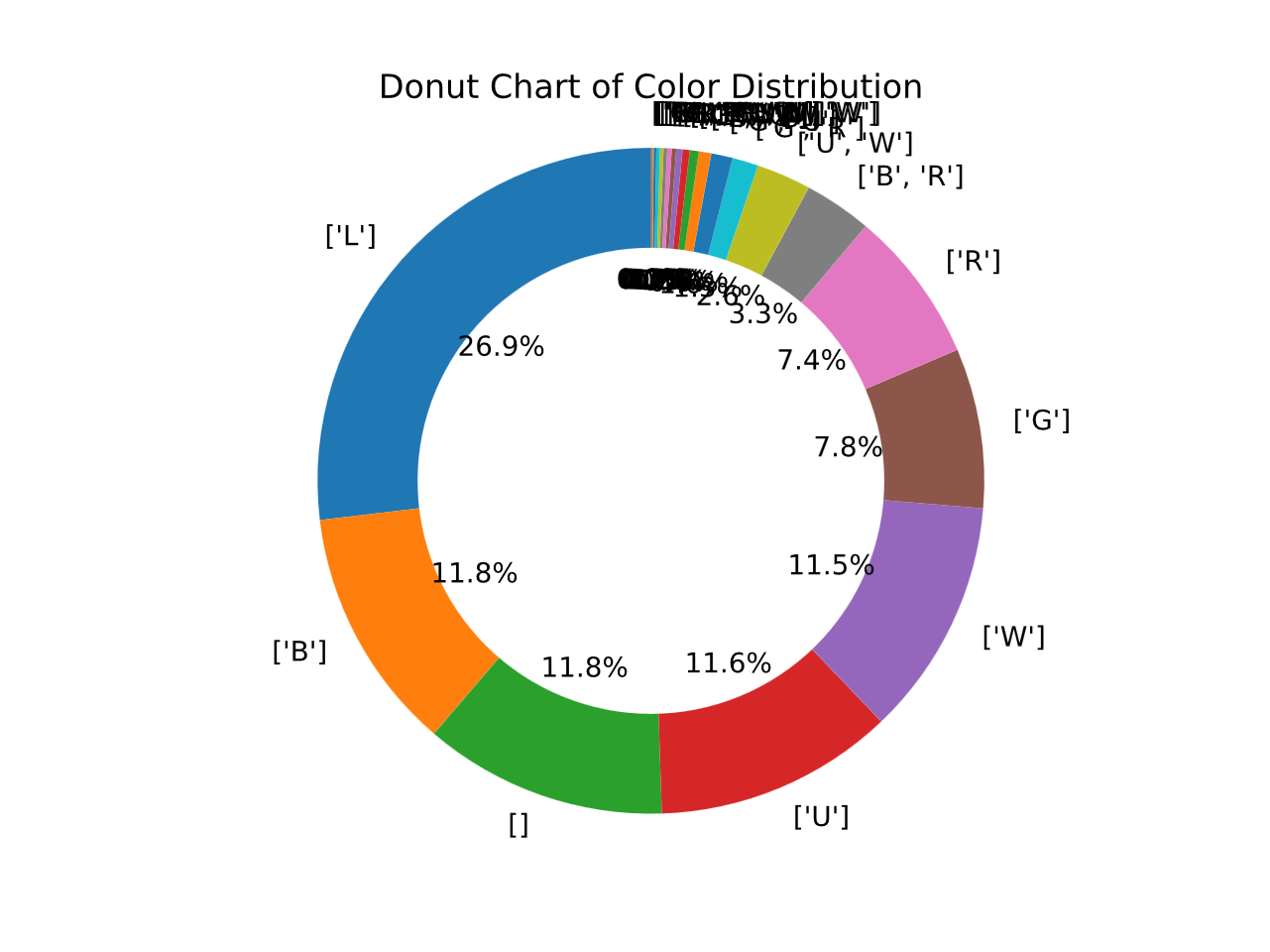


**Appendix F**

**Figure 8:** Pie chart showing the Rarity Distribution of cards used in the tournament decks.



**Figure 9:** Donut chart showing the Color Distribution of cards used in the tournament decks.



**Works Cited**

Booth, Professor. “Python Time to Age - Time Stamp.” *Stack Overflow*, 1 Sept. 1954, stackoverflow.com/questions/508727/python-time-to-age.

George. “Magic: The Gathering - Winning Pioneer Decks.” *Kaggle*, 24 June 2023, www.kaggle.com/datasets/scarfsman/magic-the-gathering-winning-pioneer-decks.

Yi, Mike. “A Complete Guide to Heatmaps.” *Atlassian*, www.atlassian.com/data/charts/heatmap-complete-guide. Accessed 3 June 2024.