Overpowered Magic

Anomaly Detection in the Original Trading Card Game

Abstract: In Magic the Gathering, cards have effects when played and an associated cost to play them. One measure of a good Magic the Gathering card is the cost to effect ratio, a ratio that is not easily quantified by people. We sought to apply an objective function to a subjective process, and analyze the difference between a predicted cost to play a card and the actual cost to play a card. By taking the difference between the predicted cost and actual cost, we hope to spot cards that are overpowered and underpowered for their actual in-game cost to play. We further hypothesized that there is a relationship between monetary cost to purchase the card on the second-hand market and this objective power rating.

Motivation: Magic the Gathering holds a special place in the hearts of many individuals who grew up playing the original trading card game created by [Wizards of the Coast] (WotC). However, as the game approaches its 30th anniversary of its release, players with greater years of experience playing the game have been able to determine which cards are more powerful and build decks to include them. This is partly because there are over 27,000 unique cards printed. While no player is able to store all of that information in their memory, many long-term players have a working model of what makes an overpowered or under-powered card. This gap of exposure can be a hindrance to newer players wishing to get into the game. The goal of our project is to provide new players with objective card analysis of what many long-term players most likely already innately know. We will serve this analysis up as a selection of cards to guide them through picking the best cards to put in their decks for a given card type (the role the card plays in the deck) and play legality (rules that dictate which cards are legal).

Data Sources: This project was completed using data from two main sources: Scryfall and TCGplayer.

Scryfall: Scryfall is a United States based company who host a REST-like API of card imagery and data. They focus on providing data tools that other content creators can use to build more things about Magic the Gathering. Their API is available at Scryfall API and the documentation for their API is available at Scryfall API Documentation. They also have a community of content creators on a Discord server who discuss game legalities, discuss updates to the API, and even create custom cards that they share amongst one another. As multiple applications send requests to the Scryfall API - it is recommended to download the bulk data of Magic the Gathering cards in a JSON file. The bulk data JSON file can also be requested through API methods, as it is updated every 12 hours. To minimize requests to Scryfall, we have elected to download the bulk JSON file. An example version with the first 100 rows of data is available at this link from the Github repository for Overpowered-Magic.

TCGplayer: TCGplayer is a group of individuals that build applications and websites to connect hobby gaming business with potential customers. Through their API, app developers are able to make new products for those in the hobby business community. Their data tracks Magic the Gathering cards, but other collectable cards such as Pokemon cards. The documentation for their API is available at [TCGplayer API Documentation]. In our project, we have made an example notebook which calls on the API for card price information, but to minimize the amount of API calls and not have the developer credentials that we have access to blocked from accessing their API - the bulk of our analysis comes from a JSON file. The file of card prices is provided by MTGJSON which is affiliated with TCGplayer as well as other collectable card websites. MTGJSON is an open-source project created and maintained by fans that catalogs all Magic the Gathering data. Data included in the MTGJSON repository, which can be found on their GitHub, includes recent pricing of cards from TCGplayer. The MTGJSON provides more information about their available data at MTGJSON All Prices Data Model and MTGJSON TCGplayers SKUs Data Model.

Data Manipulation: The data manipulation and cleaning methods for the Scryfall data and TCGplayer data are described separately below.

Scryfall: The bulk data JSON file of all Magic the Gathering cards was loaded into a python environment through the Pyspark library. This was the optimal library to use for data manipulation due to the size of the 'all_cards' dataset prior to cleaning.

```
In [3]: # Load in the Original All Cards Data Set JSON File and Turn it into a Pyspark Dataframe
        all_cards = spark.read.json('all-cards-20220403091133.json')
        all_cards.createOrReplaceTempView("all_cards")

        all_cards.count()

Out[3]: 356777
```

The data manipulation is summarized in six major steps:

- Data trimming for cards which can be used for our analysis and classification
- Extraction of card features from the inner schema of card information
- Play Type Determination
- Optimal Feature Count Selection
- Principal Component Analysis
- Boosted Gradient Regression/Classification

Data Trimming: The data was trimmed down from almost 360,000 cards to 14,059 cards. The first step was to trim the cards to only include cards with english print. Then land cards were removed from the dataset. Land cards provide a basic functionality of supplying 'mana' to players. There is no play cost to place a land card into play. Players place one land card into play at the beginning of their turn. There are five basic land types: Mountains, Islands, Forests, Swamps and Plains.

The initial cards released in 1993 included basic lands, but over time and looking for new play styles to introduce to the Magic community, some creature cards were given the ability to generate mana themselves. The names of these cards would typically be generic, however, some cards with this play style were also listed as a ‘land’ type. To ensure these cards remained in our dataset, any ‘potential lands’ were filtered from the data and then subsequently filtered to remove potential creatures within the ‘land’ cards. Once the list of ‘land’ cards did not include creature cards, they were removed from the overall dataframe of cards with an anti-join.

Magic also adapted with technology improvements. Playing the game does not require the purchase of physical paper cards and meeting with someone in person. Now, players can play virtually through online applications such as MTG Arena. Scryfall maintains an API for all players of Magic the Gathering, including those who play the game virtually - however, our analysis is focused on physical paper cards which have a monetary cost to purchase them through websites like TCGplayer. In the Scryfall JSON dataset, the card property ‘digital’ indicates whether a card was only offered digitally. These digital only cards were removed from our analysis.

Cards within the dataset can also have the same base attributes but have different ‘flavors’. A card’s ‘flavor’ is simply changes to artistic design or other attributes which are irrelevant to the in-game mechanics of a card. An example of this can be seen in the ‘Stranger Things’ themed set cards. Wizards of the Coast released a set of cards based on the popular Netflix series ‘Stranger Things’ in October 2021 through an IP collaboration. However, these exclusive ‘Secret Lair’ cards were offered for a limited time and did not have mechanically similar Magic the Gathering counterparts at the time of their release - which upset the player community. Their outcry led to WotC releasing MtG counterparts in March 2022 in the ‘New Capenna’ set. A similar example can be found in Godzilla Series Monster cards, however, the Godzilla cards had alternate artwork MtG cards at the time of their release.

The previously described card examples are presented on the following page to demonstrate the concept of card “flavors” - note that the abilities and in-game cost to play the cards are the same between the cards of each pair.

Our project is focused on identifying overpowered cards and multiple “prints” of the same card can increase the size of our dataset unnecessarily. Additionally, our project’s target user is a new player of the game who does not want to spend an excessive amount of money purchasing cards. IP related themed cards have a higher price than their MtG counterparts which may not fit the budget of a new player using our analysis. Scryfall keeps track of original cards through an ‘oracle_id’ - where the print style of one card will have the same ‘oracle_id’ as all of the other card prints of the same card. The card dataset was decreased to not include duplicate ‘oracle_id’s and therefore not include each alternate card print. But, by performing a groupby aggregation, the total number of prints of a card was added as an attribute to the card dataset - as the number of different prints of a card will influence its price.

Specific cards can also be ‘multi-faced’ either by being ‘split’ or ‘transform’. Cards that are ‘split’ include two different cards on the same card face, but only one portion of the ‘split’ card can be played. Cards that ‘transform’ will flip from one face to the other based on a specific requirement in the game. Although these cards prove to be interesting, they prove difficult to include due to both portions of the cards having the same ‘oracle_id’. Tracking the in-game costs of ‘transform’ cards is also difficult due to the requirements of a transformation not including a mana cost. Due to these differences for typical cards, as well as the relative number compared to the rest of the dataset (88 ‘split’ cards and 226 ‘transform’ cards), these ‘multi-faced’ cards were not included in our analysis, which contained 14,059 standard cards. To the side are example images of ‘split’ and ‘transform’ cards.

At this stage in the data cleaning process, a majority of the cards that were not of interest to our analysis had been removed from the dataset. But, we were particularly interested in the more popular legalities - also known as play styles. As the game developed, multiple play styles were developed by card creators, some of which were adapted by Wizards of the Coast - the official company that prints Magic the Gathering cards. However, some of these legalities are not ‘mainstream’ and therefore new players that purchase cards of these special legalities would most likely be unable to play with other players who have standard cards. An example pair of cards shown below are ‘Vanguard’ cards - cards which players picked at the beginning of the game that gave them a character identity along with specific hand size and life point adjustments at the beginning of the game. These styles of cards are not mainstream, and their legalities can include a total of 1,500 or less cards altogether. Therefore, we filtered the card dataset to only include cards which were legally able to play in at least one of the four most popular legalities: Modern, Pauper, Pioneer and Standard.

Feature Extraction: There were several features that needed to be extracted from each.

- Card Types Tokenization
- Keyword Tokenization
- Keyword Count
- Encoding of Categorical Features
- Card Color Count
- Casting Devotion
- Oracle Text Tokenization
- Oracle Devotion
- Oracle Choices
- Oracle Tapping
- Flavor Inclusion and Length

Each Magic the Gathering card has a card type, which is found on the card type line below the artwork. These card types dictate general rules of when or how a card can be played. While in game there are various supertypes and subtypes contained on the cardtype line, we treated all types as the same. We tokenizard all card types as individual features in our training dataset.

Keywords are a feature on a card that indicates an ability or attribute that when featured on a card interacts with the game in the same way. An easy way to think about this is keywords are shorthand for a larger set of rule text. An example would be the “vigilance” keyword on a creature card, meaning the creature is still able to block on an opponents turn even if that creatured was declared an attacker on the current turn. Cards can have no to many keywords on them. We tokenized each keyword on a card and provided a count for the number of keywords on a card in our training dataset.

Several categorical features already existed within the dataset, such as, converted mana cost (the generic cost to play a card), rarity (how rare the card is), power (if a creature card, what it’s “attack” strength is), toughness (if a creature card, what it’s defense is), colors (what colors are needed to play the card), and color identity (what color the card identifies with). These features were encoded into separate boolean features. The number of colors needed to cast the card, and the number of pip’s needed to cast the card (non-generic cost to play the card) were also captured.

The part that really defines a card is its oracle text, the large lower half of the card that contains particular rules and functions that are inherent to that card. We tokenized the text of the card, and used regular expressions to find particular features we wanted to capture. Some of these features were, if there was a devotion cost in the oracle text, if the oracle text contained options to pick from when the card was played, and if the card could be tapped (used in some way while card is on the board).

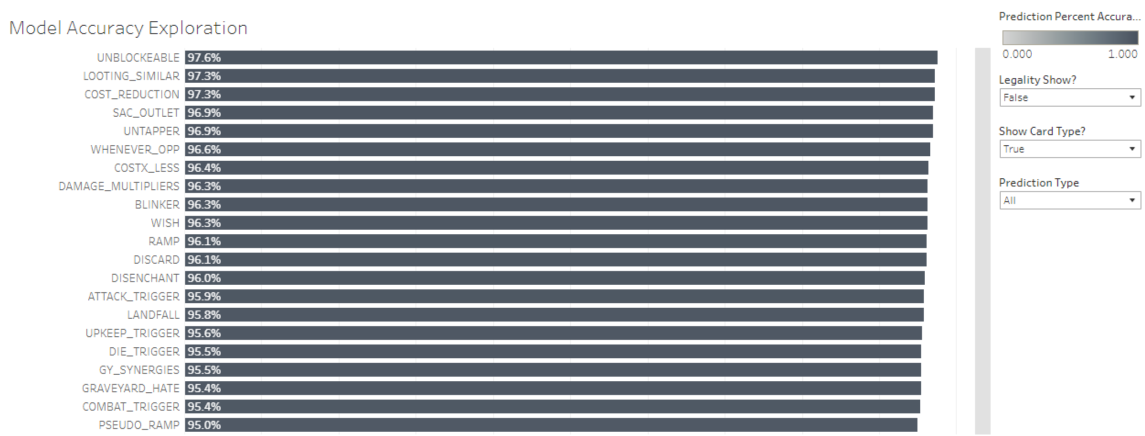
Lastly we looked at the flavor text of the card. This flavor text serves no function to the game, but rather builds upon the atmosphere and story of the game itself. The reason we created features to capture the inclusion, and length of flavor text is because it occupies space that could be used for oracle text. In this way, it might be perceived that cards with flavor text could be less impactful than cards without.

Play Type Determination: Within the game cards serve purposes or roles that are not strictly stated by the game rules or on the card. An example might be a ‘counterspell,’ with the defining feature of said card play type being that it negates an opponent's card play. A card can have multiple play types, and these play types are critical to one’s strategy in deck building and game play to win. We used code provided by Gabriel Pierobon, link here. Gabriel built out a couple 100 lines of code that used regular expressions to see if particular word patterns are contained within the oracle text of each card. This feature was key in allowing us to segment our models, and further give a user the ability to find a card they might be looking for to fit their deck.

Optimal Feature Count Selection: With a dataset that has over 600 features, we needed to find the number of features that were necessary to predict a feature for our given segments. To do this we used the recursive feature elimination function for each prediction (converted mana cost, color count, and devotion count), card play type, and each legality.

Principal Component Analysis: Using the dataset for the optimal count of features for each of the prediction segments, the main data frame was filtered to include the cards that fit that segment, and then had its principal components found based upon the total features. This was done for each prediction we wanted to make, each card play type, and each play legality. In total about 900+ card sets were created.

Boosted Gradient Regression/Classification: Boosted Gradient Regression/Classification models were trained to predict the converted mana cost, count of colors, and devotion for each card in each applicable legality and card play type. A boosted gradient model was used based upon our experience with how Magic the Gathering players typically talk about cards, in a short series of questions. What is the card's toughness? What keywords does it have? Is it instant speed? While we can not directly ask these questions as our dataset is now in its principal component form, we choose a model that behaves in a similar fashion. The results of our predictions can be seen here MTG Model Accuracy Exploration | Tableau Public.



TCGplayer: The TCGplayer data was acquired from MTGjson. One of the largest issues with finding the lowest price for a card was dealing with the set versions, card conditions, and potential multiple alternate artworks. We chose to find the last sale prices for each of these segments, and find the minimum for a given card id. This effort involved reading two JSON files from MTGjson, one for a list of sales of every card printed by condition, and another that allows us to link TCGplayer's ID to MTJSON's ID to the card's Oracle ID in scryfall. Grouping by the Oracle ID and taking the minimum price across all versions we then had a dataset that could be joined back to our main analysis that integrated the lowest card price by Oracle ID.

Analysis and Visualization:

Now that we have our three prediction components we created an equation that creates a positive number where the model over-predicts a component and a negative number for the opposite. The equation is as follows:

(Predicted Converted Mana Cost - Actual Converted Mana Cost) +
(Predicted Number of Colors - Actual Number of Colors) ^ 2 +
(Predicted Devotion - Actual Devotion) ^ 2

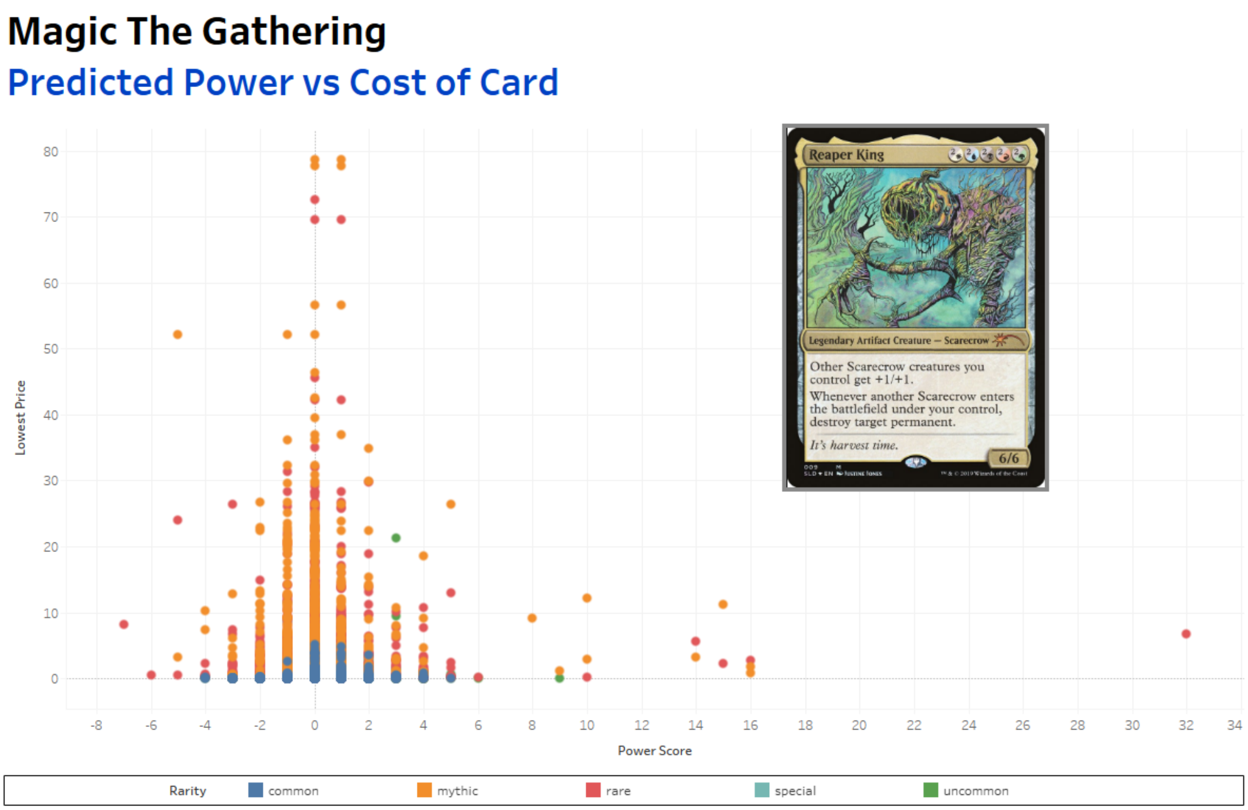
The relationship between each prediction is important. A card that costs one more to play generically is much less difficult to play than a card that is one more color to play, or has one more devotion to play. This has to do with the fact that as color and devotion increase specific resources in game are required to be present to play the card. Since we believe that cards that cost more to play should be stronger we are extremely interested in cards where the model over-predicts.

As an example, we have snapcaster mage, a card we assigned a power score of 2 because our model predicted this card should have four converted mana cost instead of two.

Comparatively look at coral eel. For the most part these cards have the same features. A one blue one generic creature spell with power two and toughness 1. The major difference between the two is what lies in the oracle text for snapcaster mage. Flash, the ability to play the card at any time, and an enter the battlefield effect that allows the player to play sorcery or instant cards they had already played in the game again. Snapcaster Mage is commonly seen as an overpowered card by players, our models and composite score were able to highlight that fact.

Visualization #1: Scatter Plot of Cards By Type and Rarity

In addition to allowing a user to visualize the most powerful cards and price for cards of a particular rarity type - allowing a new player to visualize the predicted card power across multiple card types in a legality is also beneficial for their decision of which cards to purchase. An interactive scatter plot was created with Tableau and posted on Tableau public which can be found at the following link (Price vs Power Dashboard) and allows the user to flip through card types by simply hovering over them to see a preview on the right hand side while observing where that particular card falls within the power scale.



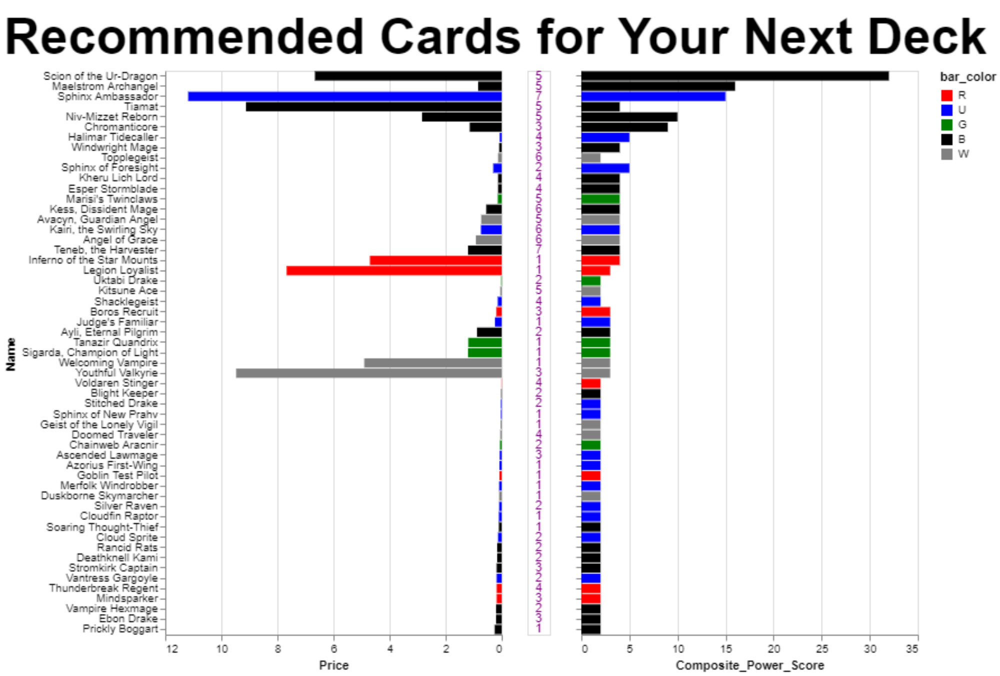
This scatter plot allows us to visualize the relationships between power and price and lets us identify which cards are more desirable for being 'overpowered' for the cost. We see from the scatter plot that the relationship between price and rarity is fairly predictable, that is, cards of higher rarity are typically worth more and priced higher. This scatter plot allows us to easily identify the cards which are overpowered by examining the data points with the highest x-values while simultaneously having the lower y-values. These cards represent the best value as they have the highest power score while being among the most affordable cards.

The vast majority of the cards that were analyzed were of either the ‘mythic’ or ‘rare’ designation so potential future work could involve the use of more cards at the other rarity levels to have a better understanding of how the rarity affects power.

From the scatter plot, we see a normal distribution with the majority of cards having a power value between -2 and 2 and was not something that was expected at the onset of the project. We believed that there would be a more even distribution of cards at the different power levels and is something that could be explored further in additional analyses where we could potentially increase the fidelity between levels -2 and 2 to allow for additional details and categorizations of cards.

<u>Visualization #2:</u> Interactive Bar Chart of Recommended Cards

We are able to take this information and encode it into a bar chart which plots both the composite power metric, the price of the card from TCGplayer, and even the converted mana cost for additional information. The chart is encoded with a tooltip to display the image of the card when the user is hovering over a bar on the graph to give the user a visual perspective of the card. The image below is a static version of the bar chart, but check out the interactive version of the bar chart in this [Google Colab Notebook](). We hoped to enable and allow other users to do their own exploration of the dataset we put together to help come to their own conclusions, or help them build their own deck. There are some obviouses misses such as the card manticore showing up as the most overpowered card in the pioneer legality when it sees next to no play, but other cards like cemetery gatekeeper in standard being an overpowered is an interesting find.



What could have worked better: While we believe our card power predictions did well in spotting powerful cards, when considering price, it does not capture the totality of the effect. Other casual legalities such as ‘commander’ are very popular and can drive up second hand costs even when the card is not seen as usable in a particular legality. Some under-powered cards have the ability to be cheated into play via effects of other cards. So while the out right cost to play the card when compared to the size of the effect is very high, the effect might be worth figuring out how to play via a different means than the out right cost on the card. Scarcity is another factor. There are a lot more common cards in the world than there are mythic cards. Year of print is another factor, Magic the Gathering has only gained in popularity over the years, more cards are printed now than they were earlier in the games lifespan. ‘Investing’ in collectable trading cards has been a trend in the past few years also. Lastly reprints are another consideration, the creators of Magic the Gathering have the ability to reprint particular cards which typically lowers the price, but failing to do so for particular cards forces up the price if the cards are particularly useful for certain strategies. In total, we think factoring for scarcity in some manner would have helped this analysis find a stronger signal in the relationship between cost and predicted power.

Statement of Work:

Taylor Druhot:

- Helped with feature extractions of Scryfall data source including oracle text tokenization, card play types, and other lesser extractions
- Created recursive feature selection functions to determine optimal feature counts for each segment of cards
- Created principle component analysis based upon the optimal features for each segment of cards
- Created and trained three gradient boosted model to predict various portions of our composite power metric
- Creation of prediction accuracy visuals
- Wrote or helped write all portions of the report
- Supplied API credentials for example notebook which calls TCGplayer API
- Designed and created chart examples (Seen Here, not used in final report)

Nicholas Ruloff:

- Initial cleaning of data from Scryfall and optimization of some feature extractions by incorporating Pyspark for faster analysis
- Created example notebook which call TCGplayer API for data - alternative dataset (MTGJSON) was used to avoid deactivation of API credentials
- Initial version and final draft version of data cleaning from MTGJSON as a representation for TCGplayer API
- Created First Draft of Visualization #1 - Interactive Scatter Plot of Cards by Type and Legality
- Created Visualization #2 - Interactive Bar Chart of Recommended Cards Visualization
- Transfer of notebook files into Google Colab Notebooks
- Transfer of smaller data sets to Github Repository
- Helped write all portions of report

Shyam Veerasankar:

- Helped with validation of feature extractions and determining which
- Wrote or helped write all portions of the final report
- Testing and validation of notebooks/visualizations
- Helped with validation of PCA and ensuring that results made sense to a non-magic player
- Created Viz 1 - Double Bar Chart - Not used in final report
- Created Viz 2 - Scatter Plot - Used in final report

Sources:

- All card images provided by Scryfall unless otherwise indicated
- Vanguard (card type) - MTG Wiki
- Stranger Things Is Not Magic: The Gathering…And That's Okay (youtube)
- "Every Stranger Things MTG reprint in New Capenna set boosters" https://dotesports.com/mtg/news/every-stranger-things-mtg-reprint-in-new-capenna-set-boosters (Dot Esports)
- Card Type Functions https://towardsdatascience.com/artificial-intelligence-in-magic-the-gathering-4367e88aee11

Workbooks:

Listed in order they need ran

- **Scryfall Data Cleaning (Final Version)**
  - Scryfall data cleaning and feature extraction
- **Feature Extraction**
  - Further feature extraction of the output from Scryfall Data Cleaning Final
- **FeatureLength_PCA_Model**
  - Recursive Feature Selection, Principal Component Analysis, Gradient Boosted Models Training and Predictions, Composite Power Score Creation
- **Card Price Data Cleaning**
  - Takes MTJSON TCGPlayer dataset and gets the latest lowest price for each card
- **Data Visualizations Notebook (Final)**
  - Workbook which generates interactive bar chart and scatter plot by using the Altair python library
- **TCGplayer Data Request and Cleaning**
  - Example workbook to show the API calls done in real-time
  - IMPORTANT NOTE: Not functional due to API credentials being replaced with fake credentials.