Topic Model of Current Magic: The Gathering Metagame (Commander Format)

Maggie Schweihs

Spring 2019

Contents

Background

Magic: the Gathering is a strategy card game owned by Wizards of the Coast with an estimated 20 million players worldwide as of 2015 (Duffy 2015). Over 15,000 unique cards (relikter 2016) have been created for the game. Players can build their own decks using cards that are legal for the format they are playing. For one particular format, called Commander or EDH, almost all 15,000 cards are eligible for inclusion in a player's deck.

A commander deck consists of 100 cards, pursuant to the following rules ("Magic: The Gathering Comprehensive Rules" 2019):

- Each deck must contain exactly 100 cards, including its commander.
- Other than basic lands, each card in a Commander deck must have a different English name.
- A card can be included in a Commander deck only if every color in its color identity is also found in the color identity of the deck's commander.

The ecosystem of Commander decks is vast and wild. Players rely on creativity, available resources and tournament deck lists to craft a seemingly infinite variety of decks. Naturally, certain types of decks become popular based on "what everyone else is playing" ("What Is the Metagame?" 2007) and what everyone is playing to beat everyone else. This is called the metagame. Many websites report deck statistics and metagame analysis. One such website is MTG Goldfish. They derive their metagame analysis from current MTG tournament games. MTGgoldfish.com publishes a list of commander metagame decks, along with the cards in the decks and the prices of the cards.

A MTG card looks like the following image. Each card has a name, cost, type, and power/toughness. Most cards, with the exception of basic lands, have text. The card text that makes up each Commander deck in the metagame is the interest of this analysis.

(*Image of a Magic: The Gathering Card with Descriptive Text*, n.d.)

Motivation

Each card has a rich story. I suspect that by modelling the topics of the text, we can see the story of the current Commander-format metagame. I also suspect that the topics will naturally be grouped by color combination.

Related Work

Some work has been done analyzing MTG card text. One study (Zilio, Prates, and Lamb 2018) outlines the methodology used to train neural nets to predict a card type based on imagery. They also trained neural networks to generate card text to match an image.

In addition, several researchers looked into the ability to use artificial intelligence to play MTG. They showed that the games' outcomes were non-computable: "Magic: The Gathering does not fit assumptions commonly made by computer scientists while modeling games. We conjecture that optimal play in Magic is far harder than this result alone implies, and leave the true complexity of Magic and the reconciliation of Magic with existing theories of games for future research," (Churchill, Biderman, and Herrick 2019). The framework of rules that leads to this conclusion is largely buried in the text on each individual card.

Dataset Dscription

The dataset is a compilation of Commander deck data scraped from MTGgoldfish.com using a python script (./scripts/mtg_scraper2.py) and an MTG software development kit (sdk).

Variables

The data is saved as a JSON file and contains the following information:

- Deck ID
- Deck name
- Number of decks of type reported to MTG Goldfish
- Percent of Metagame represented by a deck
- Deck price (paper deck)
- Deck price (online deck)
- Cards:
- Name
- Mana cost
- Colors
- Text

The card colors are related to the flavor of the text. Angels and knights are white; Dragons, volcanos and goblins are red. In this analysis, the variables of interest are the deck name, the colors, and the text.

Methodology

The following steps were taken to perform the topic modelling analysis:

Find the Data

The website and data requirements were analyzed.

- MTGGoldfish.com is a leading source for MTG tournament deck lists. This is among the best sources for analyzing the current metagame of any MTG format.

Web Scraping with Python

I build a webscraper in Python 3.5 using beautiful soup. The deck data was saved as a JSON file.

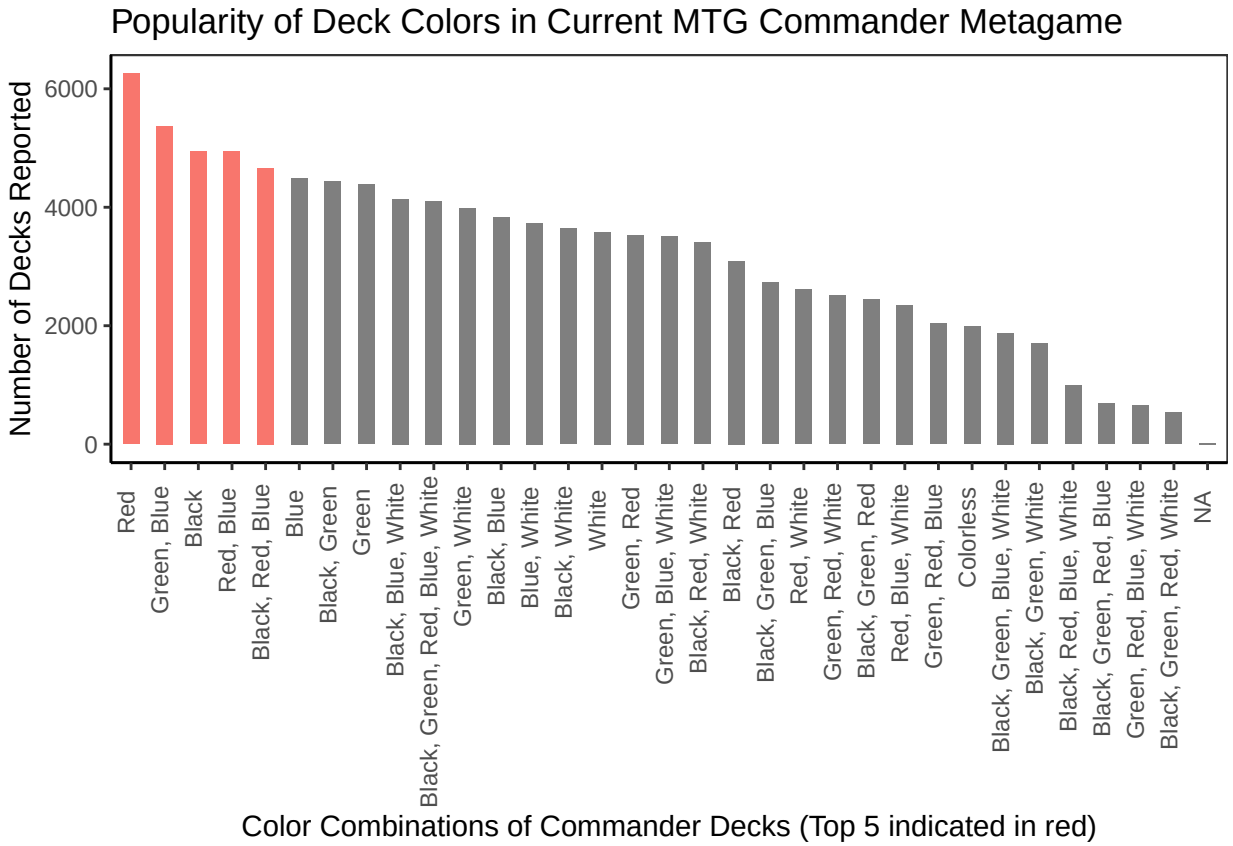
Process Data

The JSON file was loaded into R and processed into a data frame

Initial Data Exploration

I grouped decks according to color-type [1].

[1] In the Commander format, there is one card that is the "Commander" of the deck. This card has special play rules and properties. The colors that appear on the Commander dictate which colored cards are allowed to be included in the deck. For example, in a deck with a Commander of red/blue type, a player cannot have white, green, or black cards; only red and blue cards are allowed.



Visually determine which are the 5 most popular color combinations. Will the topic model detect these color combinations?

- Black,Red,Blue
- Red,Blue
- Black
- Green,Blue
- Red

Prepare Data

Prepare data for topic modeling using the function `makeFlexTexChunks` from Chapter 13 of Jockers' text.

Simple Topic Modeling

Perform initial simple topic modeling with 33 topics using the stoplist2.csv file as the stoplist. Several MTG specific words were added such as cards, creature, creatures, spell, library, battlefield, etc. [2]

words	term.freq	doc.freq	
1	flying	776	208
2	vigilance	2680	1091
3	deathtouch	1825	960
4	lifelink	2047	1025

[2]Output from this point forward is copied and pasted from the Virtual Lab R Studio due to problems installing mallet on my Mac. The code is in the RMD file, set to not run.

words	term.freq	doc.freq	
5	beginning	11194	3158
6	step	5542	2491

Train the Model

Set the parameters and train the model

```
topic.model$setAlphaOptimization(40,80)
topic.model$train(400)
```

Explore the Model

The top ten words in group 1:

words	weights	
create	create	0.13426572
token	token	0.11429650
green	green	0.06896267
tokens	tokens	0.05525709
control	control	0.04719498
sacrifice	sacrifice	0.03634215
saproling	saproling	0.03113279
beginning	beginning	0.02375286
copy	copy	0.02300867
dies	dies	0.01879157

The top ten words in group 2:

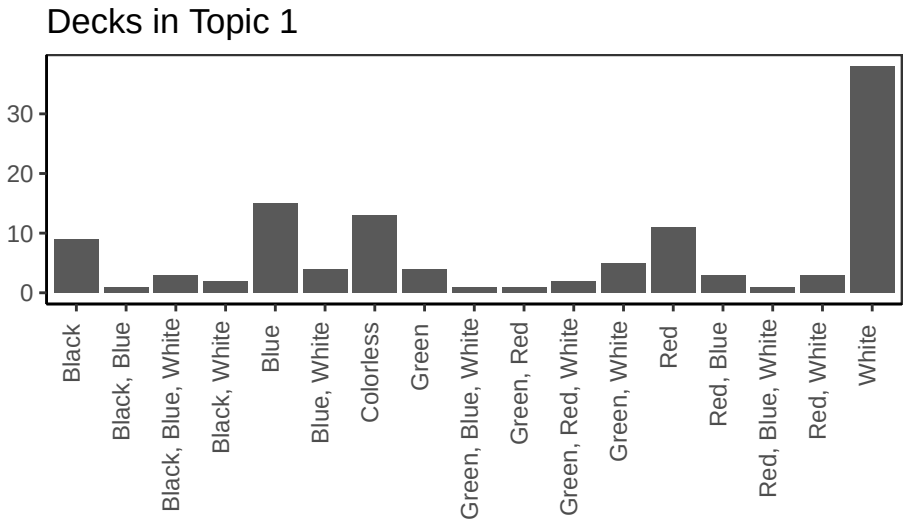
words	weights	
enchanted	enchanted	0.08977330
enchant	enchant	0.07081737
control	control	0.05614461
damage	damage	0.02960630
permanent	permanent	0.02295001
aura	aura	0.02251591
enchantment	enchantment	0.02049008
return	return	0.02043220
owner's	owner's	0.01724876
draw	draw	0.01525188

Visualizations and Findings

To determine if the topic models separated the text by color, we will look at the topic probability. The colors of the decks decks with a greater than 50% chance of belonging to each topic are represented in the bar chart followed by the topic cloud corresponding to the topic.

Topic 1

Figure 1:



This topic consists of mostly mono-colored decks, and predominantly white decks.

Topic 2

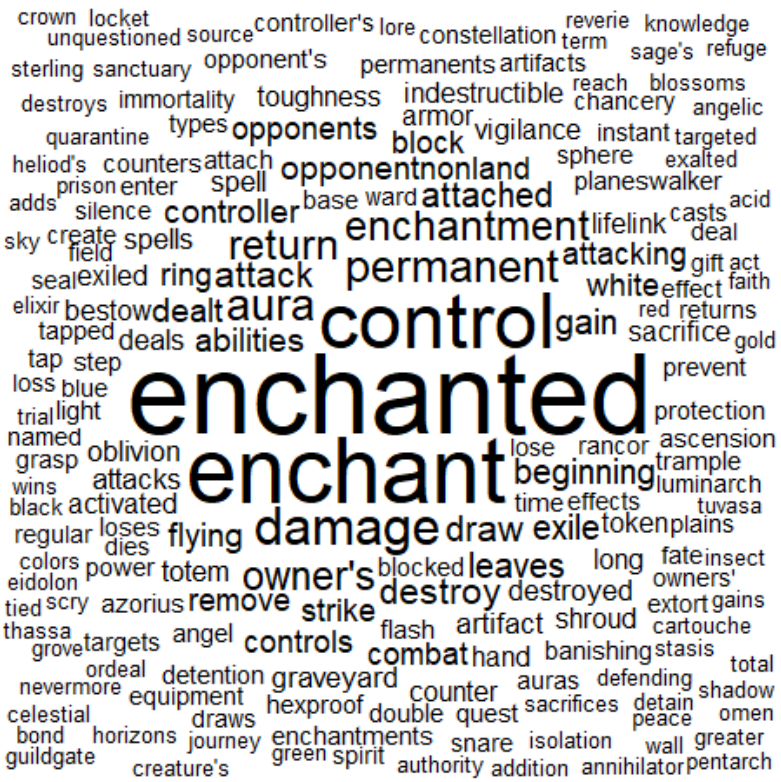
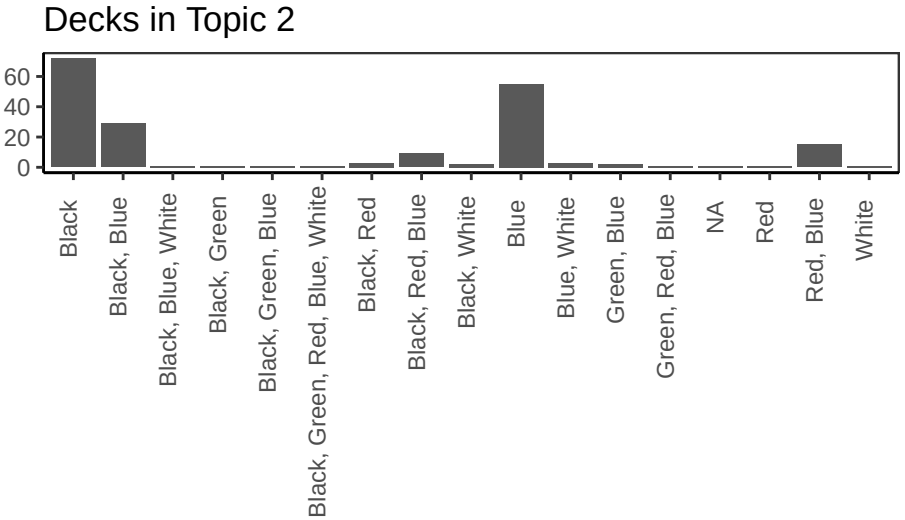


Figure 2:



Topic 2 is composed of a lot of blue, a color associated with 'control', which can be seen in the wordcloud.

Topic 3

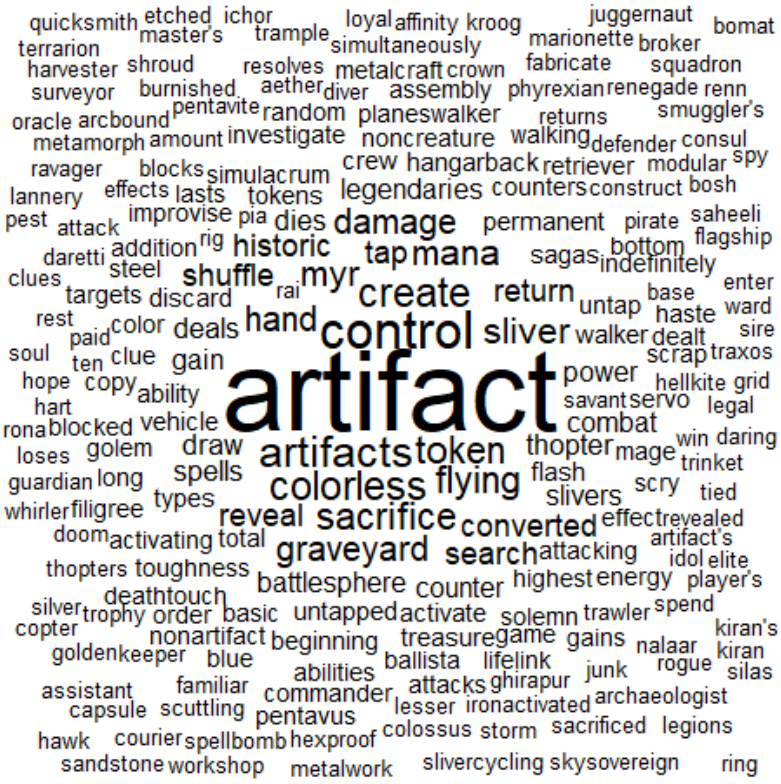
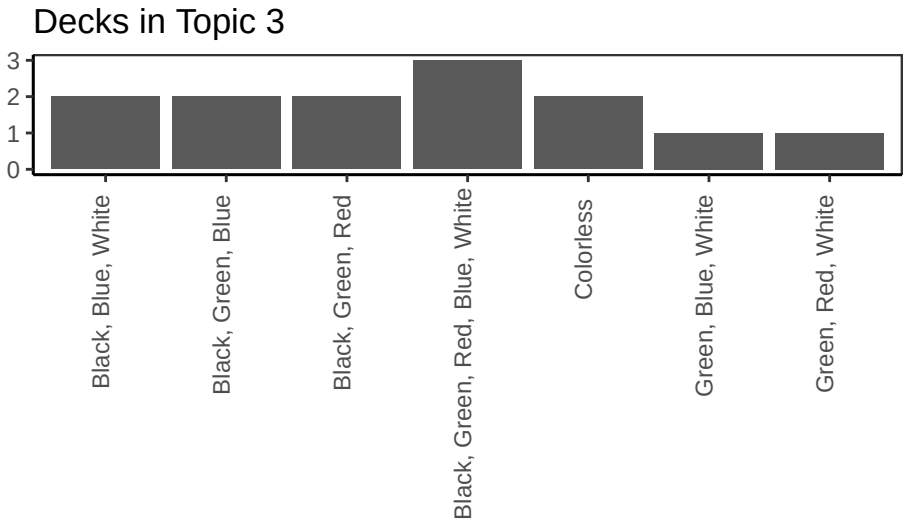


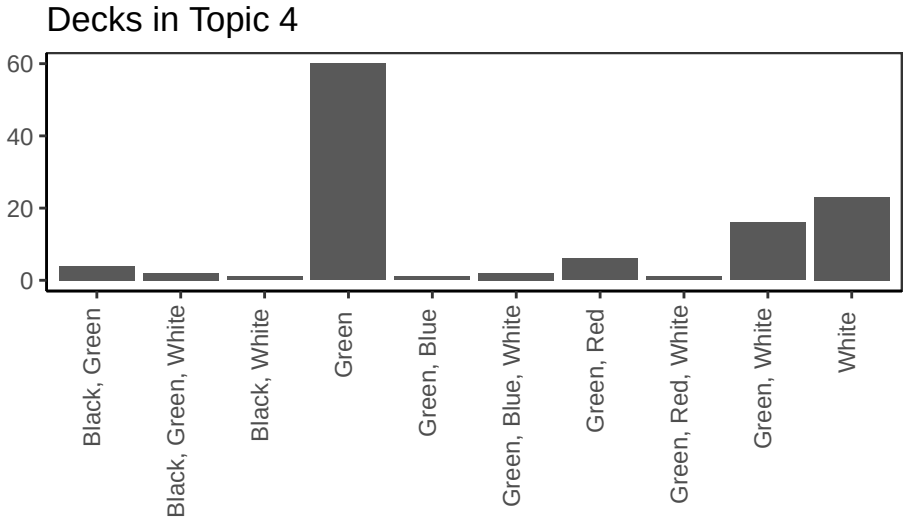
Figure 3:



Topic 3 consists of mostly 3-colored decks and colorless decks. Colorless is a special type that is often associated with "Artifact" type cards, so it is no surprise to see this represented in the wordcloud. It could also be true that decks with more than two colors utilize alot of artifacts.

Topic 4

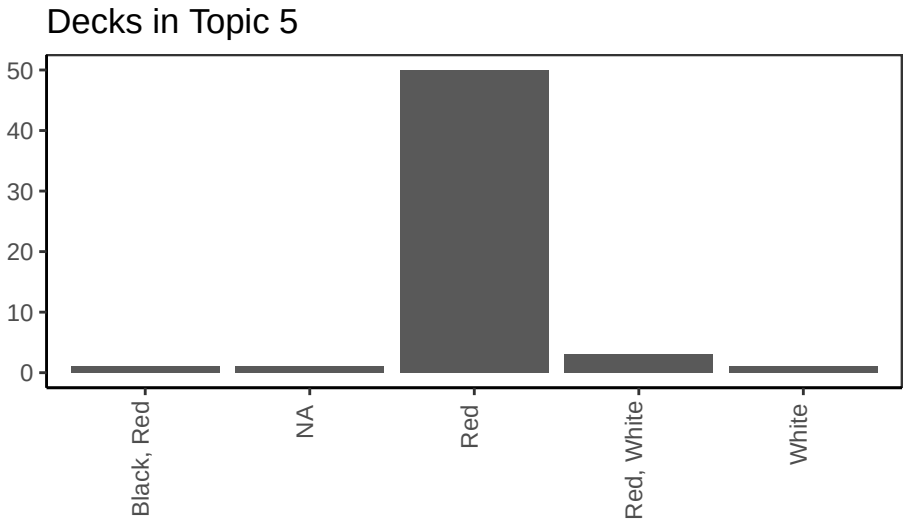
Figure 4:



Suprisingly, in the topic 4 wordcloud, we see the word "graveyard" which is often associated with black. However, one popular game-mechanic is "ressurection" in which things are brought back from the dead (ie., from the graveyard). This may actually be a green, life-giving, mechanic.

Topic 5

Figure 5:



Topic 5's wordcloud tells the story of the aggressive red deck: control, block, callous, attack, power, champion, etc.

Conclusion

After modeling deck data for the current Commander-format metagame, patterns and stories emerged and could be seen in the wordclouds. There is room for further exploration into this dataset. The data collected from MTGgoldfish could be used to analyze the network of cards, connected to each other via decks.

References

Churchill, Alex, Stella Biderman, and Austin Herrick. 2019. "Magic: The Gathering Is Turing Complete."

Duffy, Owen. 2015. "How Magic: The Gathering Became a Pop-Culture Hit – and Where It Goes Next." *The Guardian*, July. https://www.theguardian.com/technology/2015/jul/10/magic-the-gathering-pop-culture-hit-where-next.

Image of a Magic: The Gathering Card with Descriptive Text. n.d. *Magic*. https://magic.wizards.com/en/magic-gameplay.

"Magic: The Gathering Comprehensive Rules." 2019. *Magic: The Gathering*. https://media.wizards.com/2019/downloads/Magi
20190125.txt.

relikter. 2016. "How Many Cards Exist in Magic, Not Including Reprints?"

"What Is the Metagame?" 2007. *MAGIC*. https://magic.wizards.com/en/articles/archive/magic-academy/what-metagame-2007-01-06.

Zilio, Felipe, Marcelo Prates, and Luis Lamb. 2018. "Neural Networks Models for Analyzing Magic: The Gathering Cards." *Lecture Notes in Computer Science*. Springer International Publishing, 227–39. doi:10.1007/978-3-030-04179-3_20.