# An Overview of NLP Techniques

By: Silas Raye

# Anatomy of a Magic Card

**Card Name**

**Card Type**

**Artist Info**

**Card Cost**

**Set & Rarity**

**Rules Text**

**Flavor Text**

**Power & Toughness**



Thragtusk     4🌳

Creature — Beast

When Thragtusk enters the battlefield, you gain 5 life.

When Thragtusk leaves the battlefield, put a 3/3 green Beast creature token onto the battlefield.

*"Always carry two spears."*
—Mokgar, Kalonian hunter

— Nils Hamm

™ & © 1993-2012 Wizards of the Coast LLC 193/249

5/3

# Anatomy of a Magic Card



**Card Name** ———

**Card Type** +

**Artist Info** ———

Thragtusk  4 🌳

Creature — Beast  M13

When Thragtusk enters the battlefield, you gain 5 life.

When Thragtusk leaves the battlefield, put a 3/3 green Beast creature token onto the battlefield.

*"Always carry two spears."*
—Mokgar, Kalonian hunter

Nils Hamm
™ & © 1993–2012 Wizards of the Coast LLC 193/249

5/3

——— **Card Cost** ++

——— **Set & Rarity** ++

——— **Rules Text** +++

——— **Flavor Text**

——— **Power & Toughness** ++

# Anatomy of a Magic Card

Card Name

Card Cost
++

Card Type
+

Set & Rarity
++

Rules Text
+++

Artist Info



Lightning Bolt

Instant

Lightning Bolt deals 3 damage to any target.

085 R
SLD ★ EN ROBBIE TREVINO

TM & © 2020 Wizards of the Coast

# The Data - Magic: The Gathering Cards

| Name | Rules Text | Power | Toughness | Reserved | Booster | EDHRec Rank |
|---|---|---|---|---|---|---|
| "Ach! Hans, Run… | At the beginnin… | 0 | 0 | 0 | 1 | 0 |
| "Brims" Barone,… | When "Brims" Ba… | 5 | 4 | 0 | 1 | 24238 |
| "Rumors of My D… | {3}{B}, Exile a… | 0 | 0 | 0 | 1 | 0 |
| +2 Mace | Equipped creatu… | 0 | 0 | 0 | 1 | 13735 |
| A Display of My… | When you set th… | 0 | 0 | 0 | 0 | 0 |

# The Data - Magic: The Gathering Cards

| Loyalty | Price | Number of Printings | Card Cost | Number of Colors | Year |
|---------|-------|---------------------|-----------|------------------|------|
| 0 | 3.32 | 1 | 6 | 2 | 2004 |
| 0 | 0.02 | 4 | 5 | 2 | 2022 |
| 0 | 0.09 | 1 | 3 | 1 | 2017 |
| 0 | 0.02 | 1 | 2 | 1 | 2021 |
| 0 | 12.47 | 1 | 0 | 0 | 2010 |

# The Data - Magic: The Gathering Cards

| | Rarity | | | |
|---|---|---|---|---|
| Common | Mythic | Rare | Special | Uncommon |
| 0 | 0 | 1 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1 |
| 0 | 0 | 0 | 0 | 1 |
| 1 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 |

| | Legality | | |
|---|---|---|---|
| Standard | Pioneer | Modern | Commander |
| 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 |
| 0 | 1 | 1 | 1 |
| 0 | 0 | 0 | 0 |

# The Data - Magic: The Gathering Cards

| | | Category | | |
|---|---|---|---|---|
| **Instant** | **Land** | **Planeswalker** | **Sorcery** | **Token** |
| 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 |

# Word Counting

- This method…
  - Is simple
  - Counts the number of words in the rules text
  - Is computationally cheap
  - Is very lossy
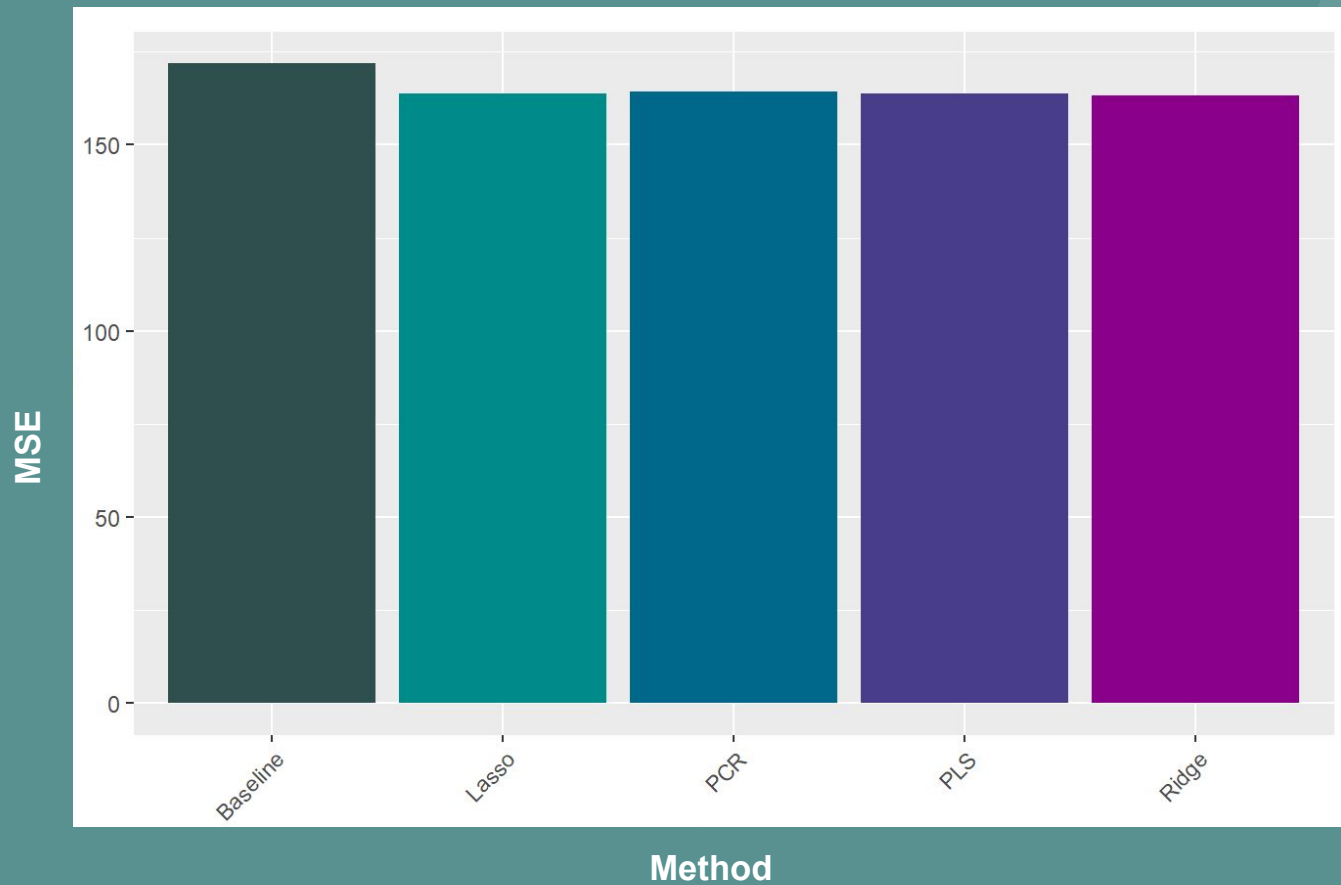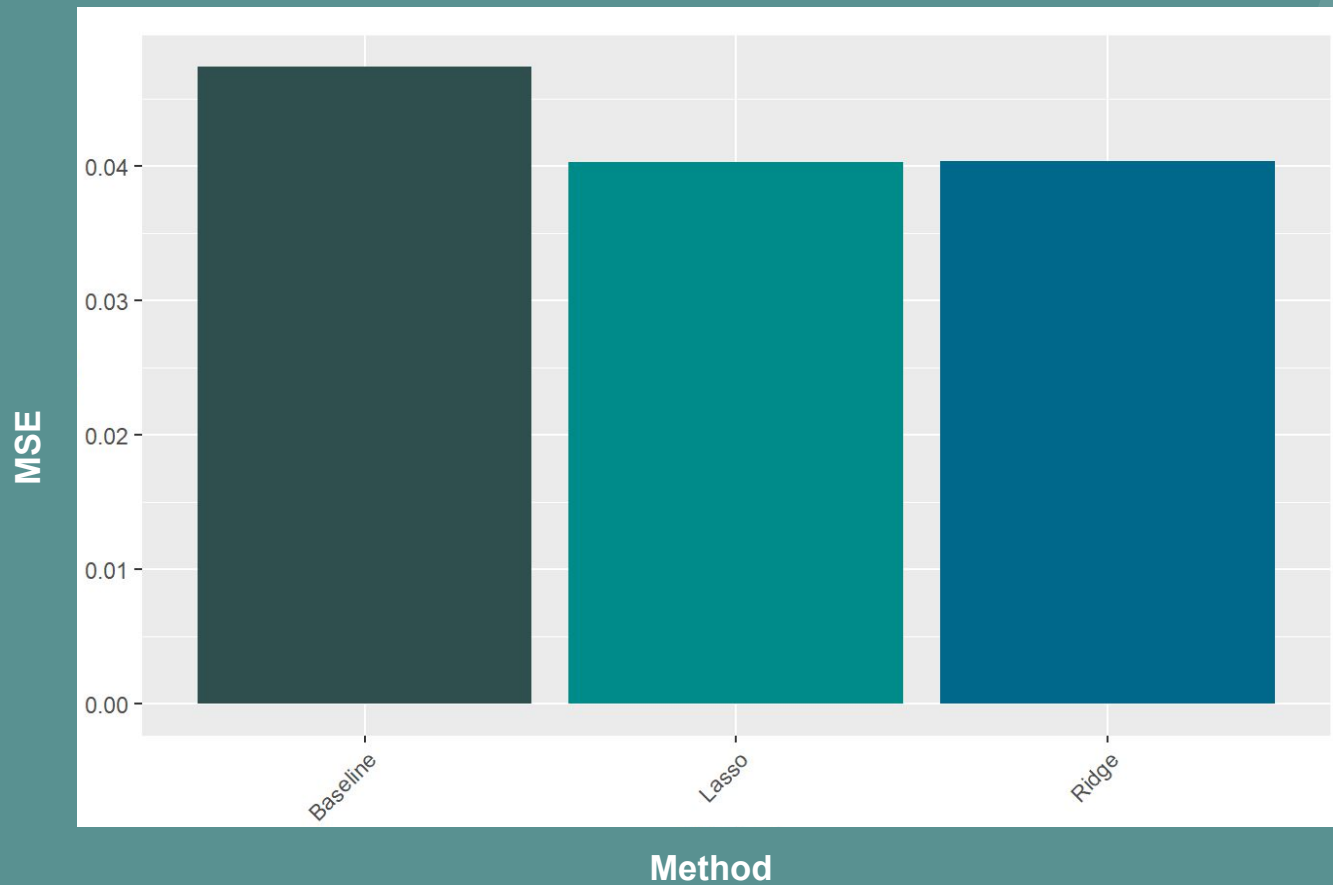
# **Phrase Counting**

- This method…
  - Is simple
  - Counts the number of occurrences of a phrase, like "draw a card", in the rules text
  - Is computationally cheap
  - Is very lossy
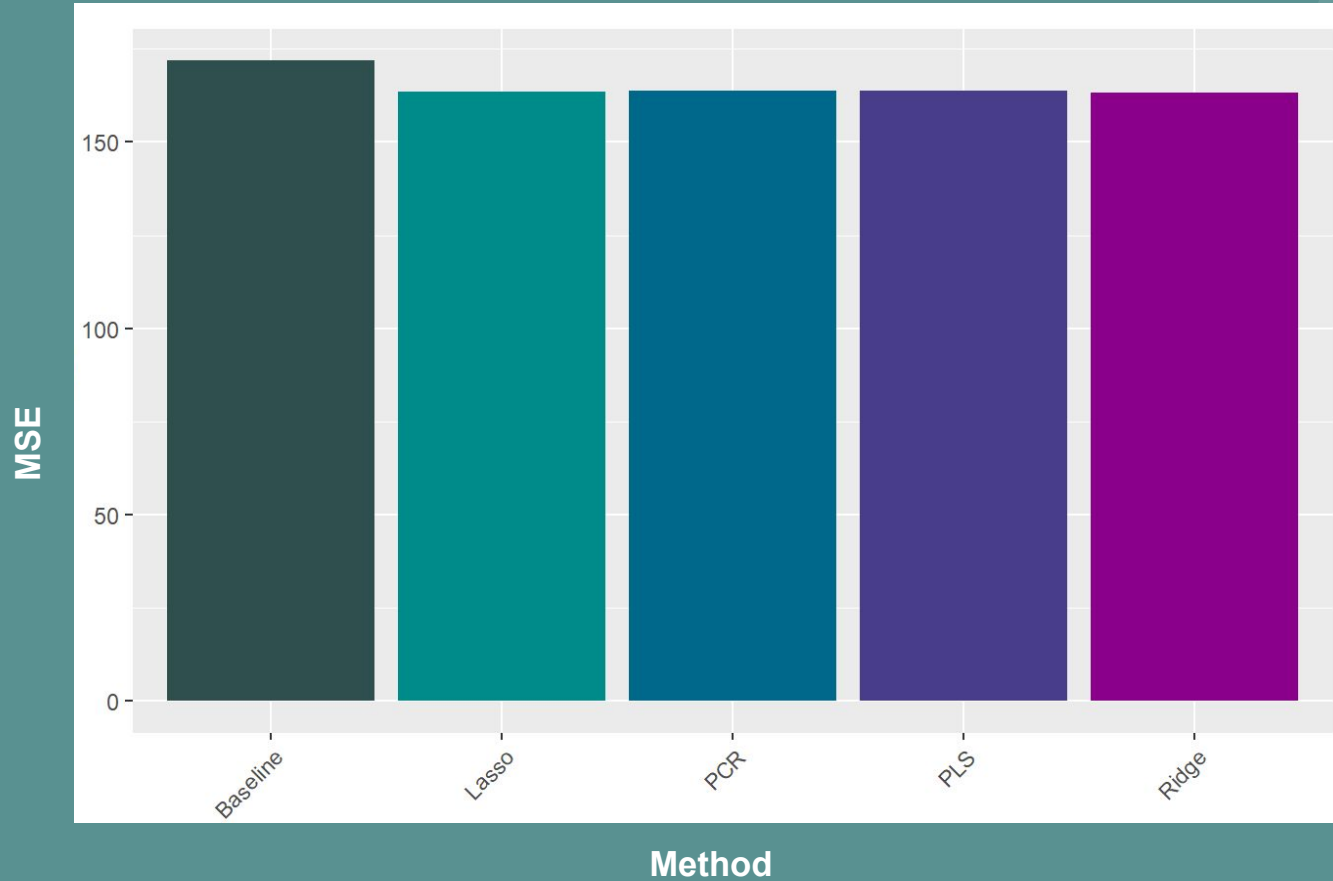
Categorical Phrase Counting
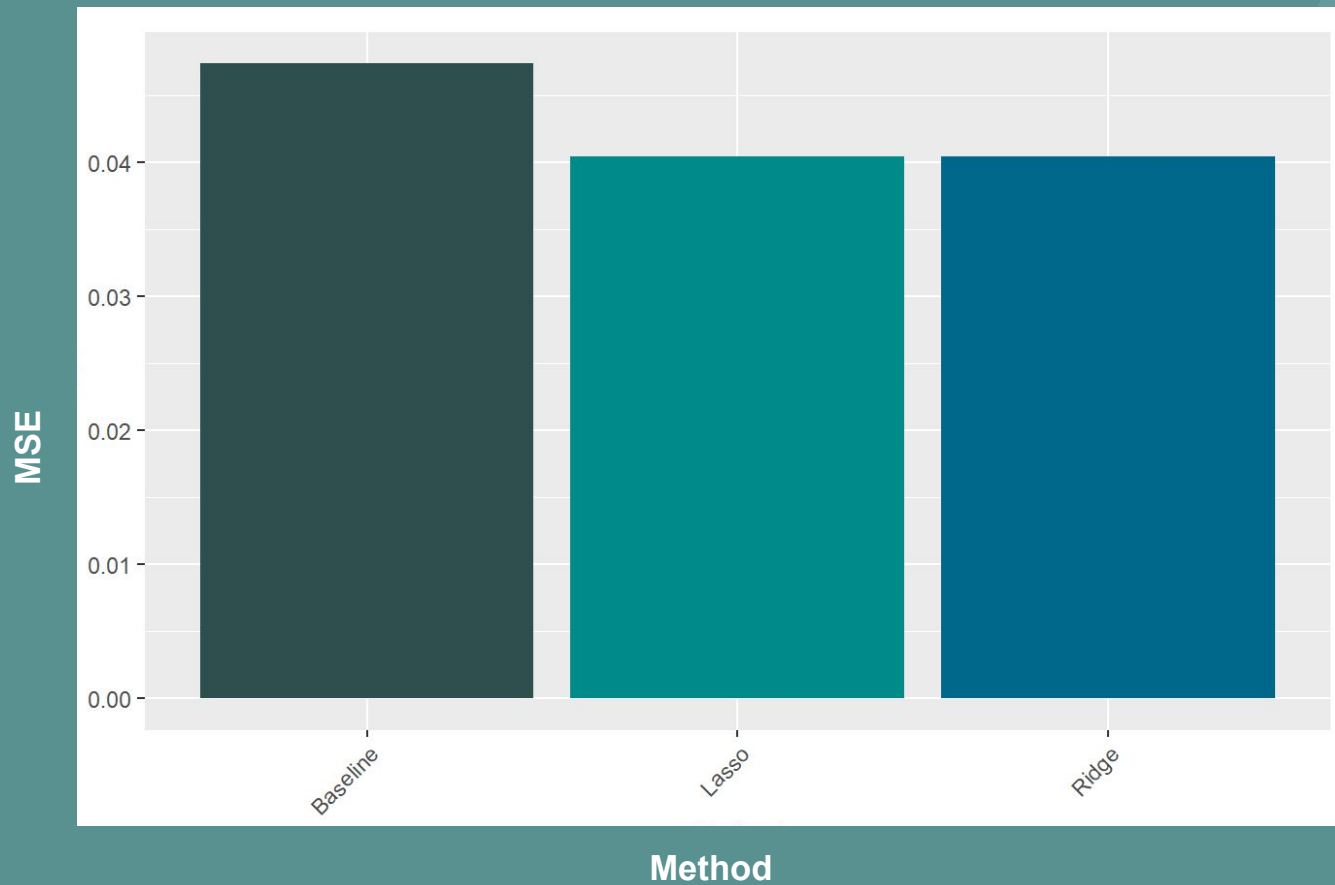
# TF-IDF

- This method is…
  - Complicated
  - Computationally expensive
  - Lossy
- Let's go over how it works…

# TF-IDF

- **Term Frequency:**

  TF(t, d) = (Number of times term t appears in document d) / (Total number of terms in document d)

- **Inverse Document Frequency:**

  IDF(t, D) = log(Total number of documents D / Number of documents containing term t)

- **Score:**

  TF-IDF(t, d, D) = TF(t, d) * IDF(t, D)

# TF-IDF Example



Card Name

Card Cost
++

Card Type
+

Set & Rarity
++

Rules Text
+++

Artist Info

# TF-IDF Example

| Card | Word | N | TF | IDF | TF-IDF |
|------|------|---|-----|-----|--------|
| Lightning Bolt | 3 | 1 | 0.125 | 2.203711087 | 0.2754638858 |
| Lightning Bolt | any | 1 | 0.125 | 2.252609716 | 0.2815762146 |
| Lightning Bolt | damage | 1 | 0.125 | 1.751213375 | 0.2189016719 |
| Lightning Bolt | deals | 1 | 0.125 | 2.080975284 | 0.2601219105 |
| Lightning Bolt | lightning | 1 | 0.125 | 6.391198981 | 0.7988998726 |
| Lightning Bolt | bolt | 1 | 0.125 | 3.452793202 | 0.4315991503 |
| Lightning Bolt | target | 1 | 0.125 | 0.9449543411 | 0.1181192926 |
| Lightning Bolt | to | 1 | 0.125 | 1.068746648 | 0.133593331 |
| | **Mean** | 1 | 0.125 | 2.203711 | 0.2754639 |
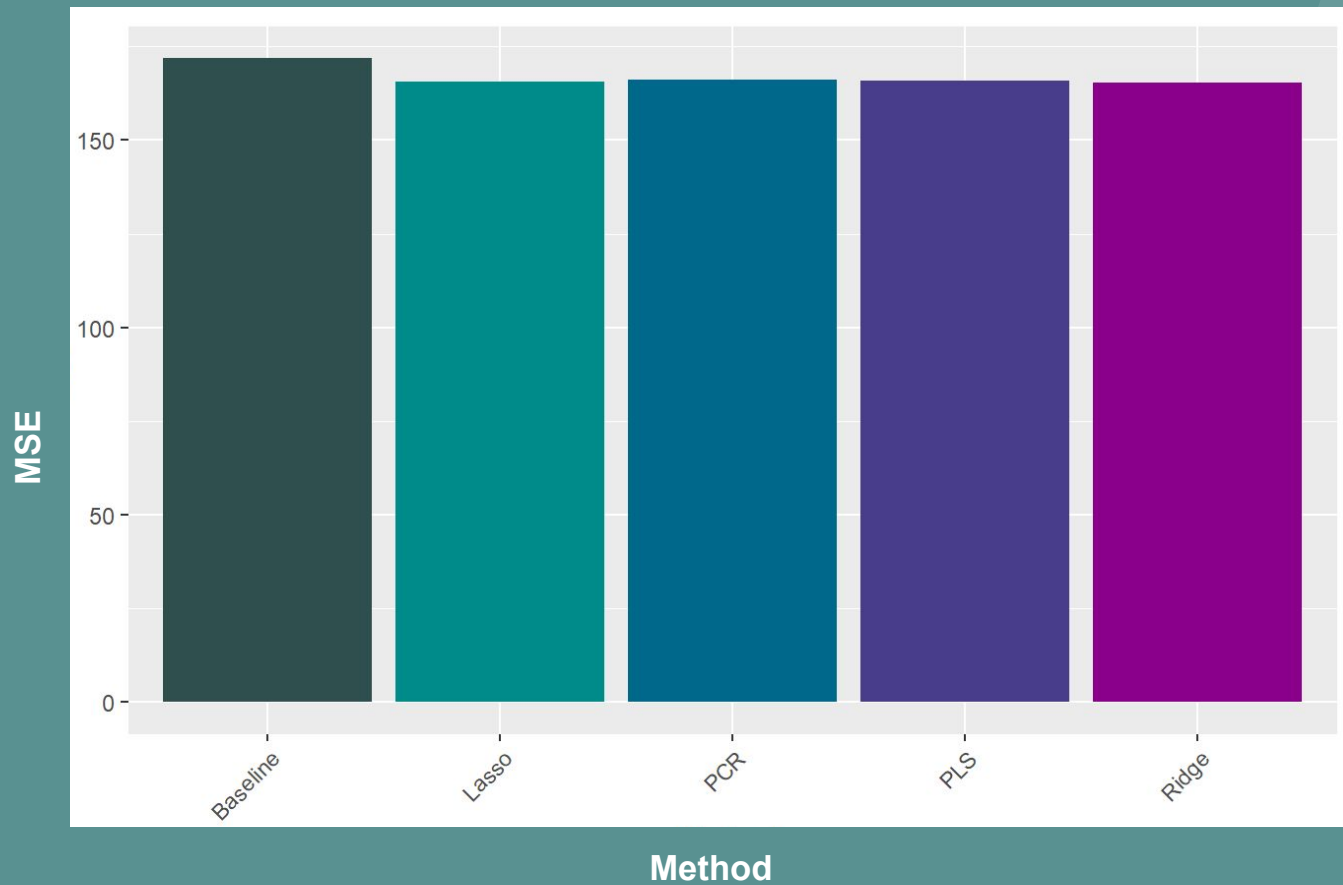
# Categorical TF-IDF Counting
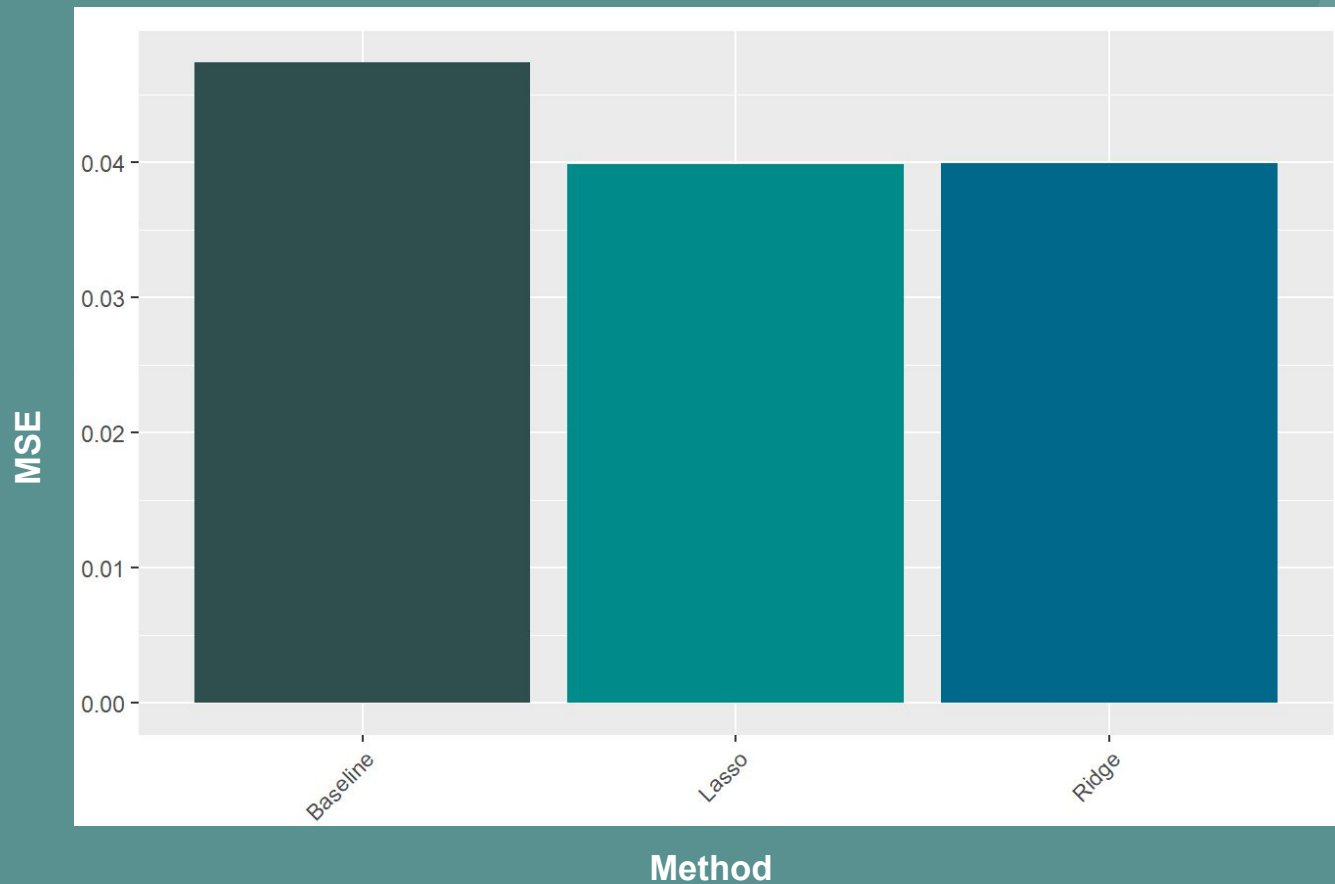
# Bag of Words

- This method is…
  - Complicated
  - Very computationally expensive
  - Somewhat less lossy
- Let's go over how it works…

# Bag of Words Example

| Card | Lightning | Bolt | Thragtusk | Battlefield | |
|------|-----------|------|-----------|-------------|---|
| Lightning Bolt | 1 | 1 | 0 | 0 | 11,500+ more columns |
| Thragtusk | 0 | 0 | 2 | 2 | |

21,100+ more rows

# Bag of Words
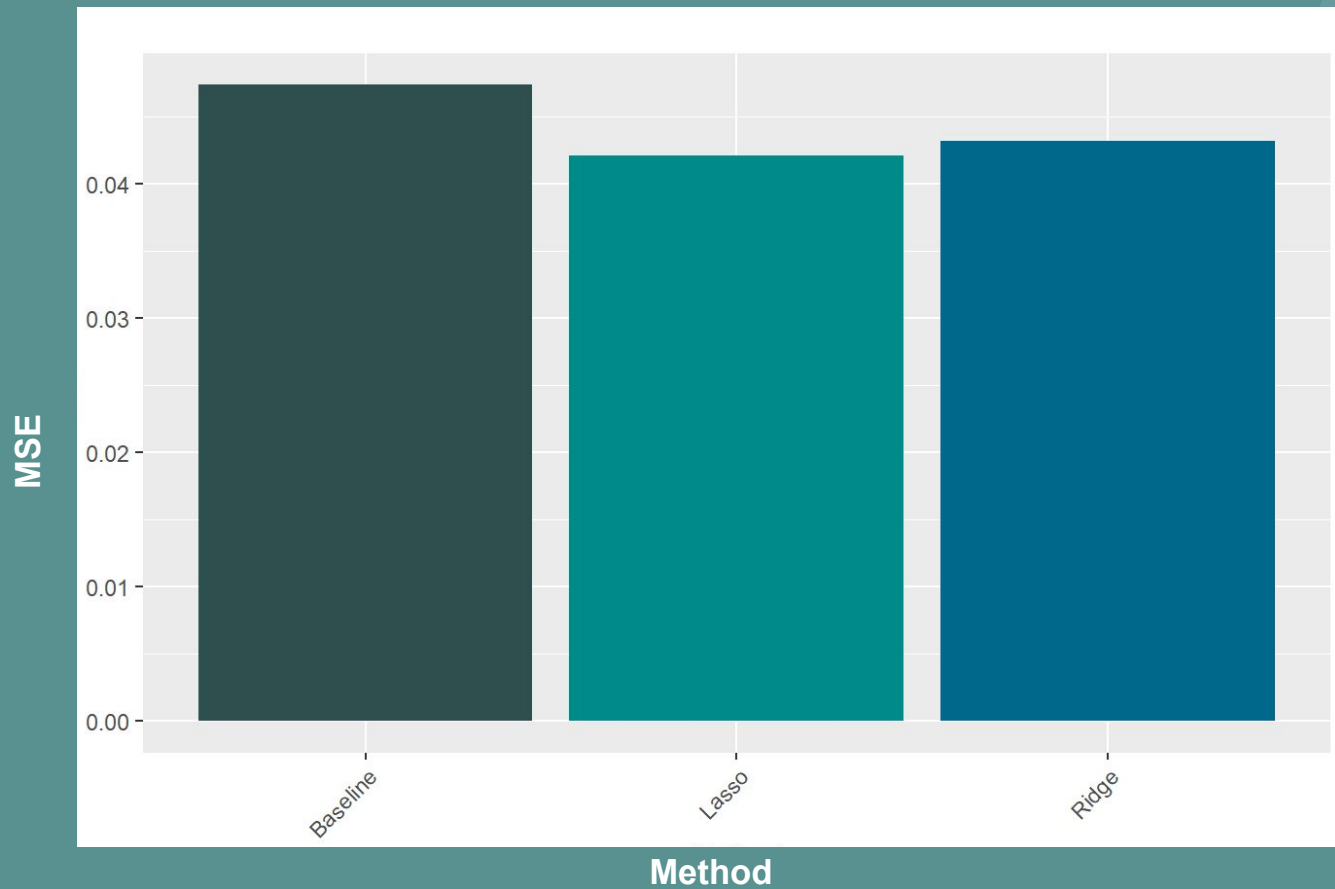
- As you can imagine, regression on over 11,500 words is inefficient
- We remove words with low variance (< 0.01)
  - This removes words that are less common while leaving words that are more common
- This cuts the number of words down to 222
- Like all the methods we've looked at so far, Bag of Words does not preserve the information stored in word order
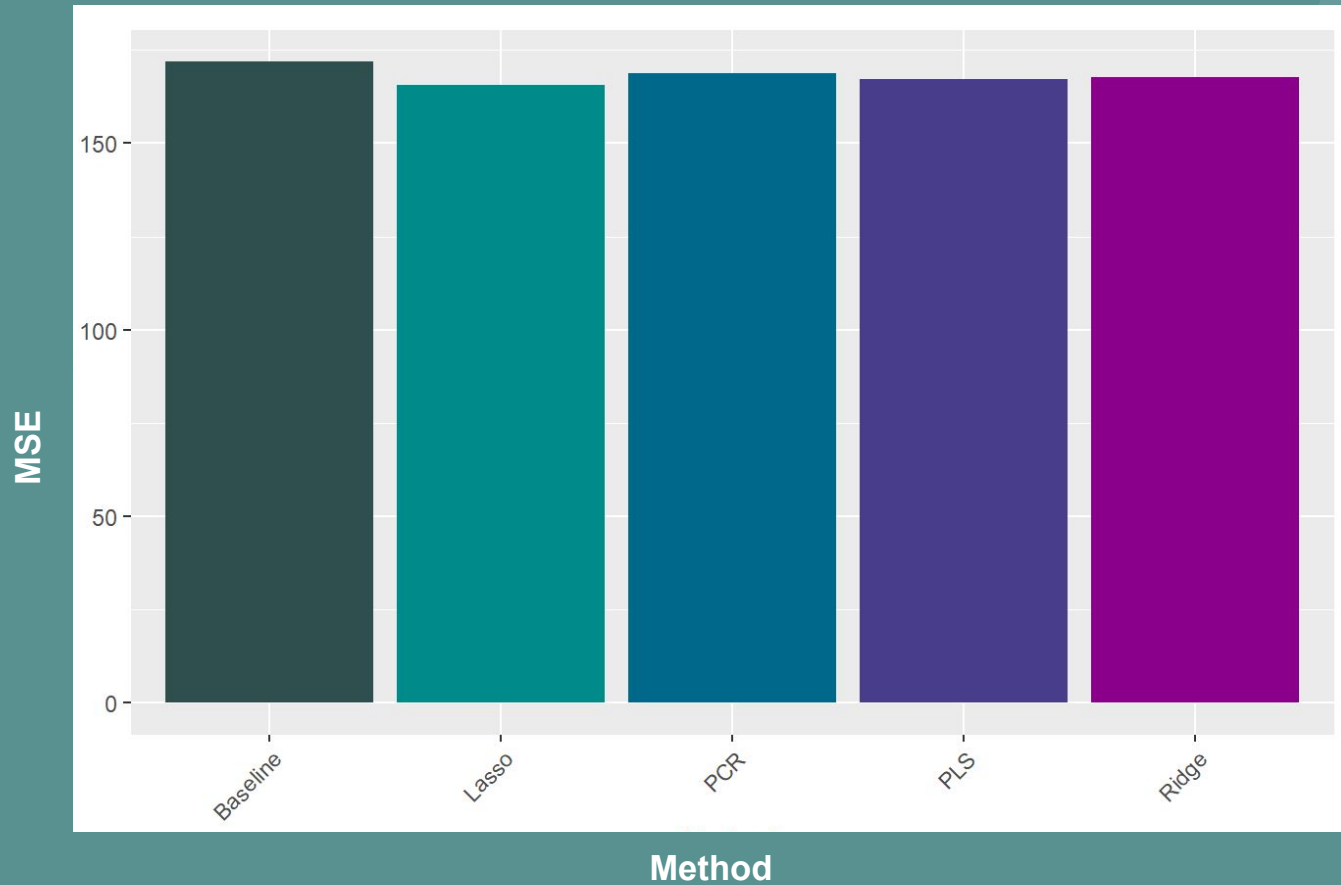
Categorical Bag of Words

# Fine Tuning ChatGPT

- The data was split into train and test
- Then the data was prepared in prompt/response format (jsonl)
- For example…

```
{"messages":[{"role":"system","content":"You are an expert AI system trained in predicting the price of trading cards." },
{"role":"user","content":"Guess the price of a trading card with the following statistics…
Name:\"Ach! Hans, Run!\"
Cost:6
Type:Enchantment
Text:At the beginning of your upkeep, you may say  \"Ach! Hans, run! It's the . . . \" and the
name of a creature card. If you do, search your library for a card with that name, put it
onto the battlefield, then shuffle. That creature gains haste. Exile it at the beginning of
the next end step.
Number of Printings:1" },
{"role":"assistant","content":"3.32"}]}
```

# Fine Tuning ChatGPT



| | | | |
|---|---|---|---|
| Training loss | 1.0269 | Trained tokens | 1,906,362 |
| Validation loss | 0.4668 | $ Cost | 15.25 |
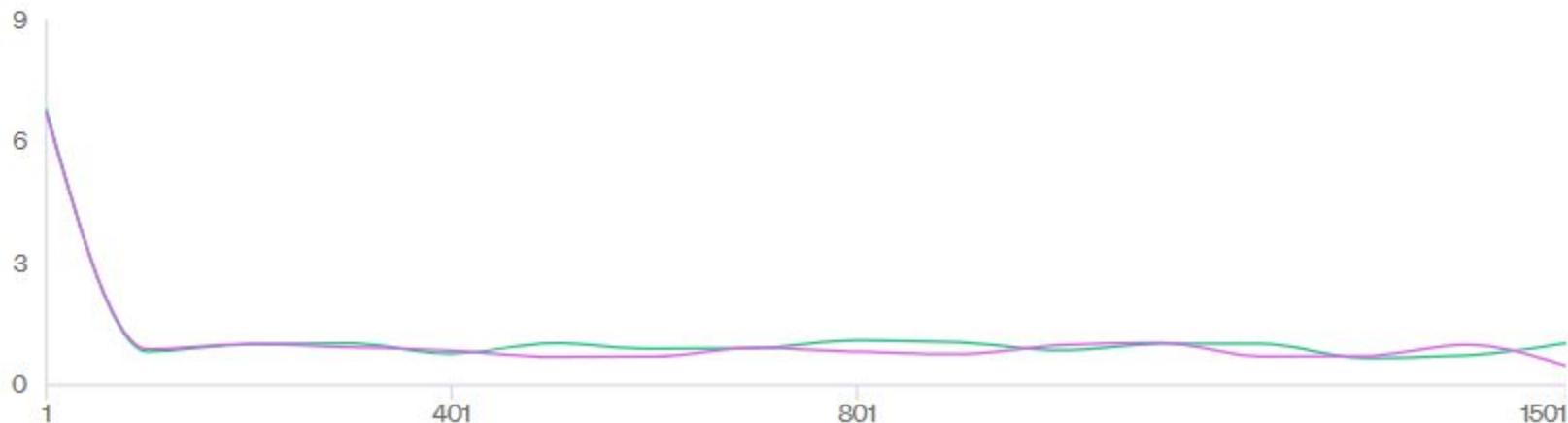
Step

**USER**

Guess the price of a trading card with the following statistics...

Name: Thragtusk

Cost: 5

Type: Creature - Beast

Text: When Thragtusk enters the battlefield, you gain 5 life.

When Thragtusk leaves the battlefield, create a 3/3 green Beast creature token.

Number of Printings: 9

Power: 5

Toughness: 3

**ASSISTANT**   0.15                                                    Real Price 0.20

**USER**

Guess the price of a trading card with the following statistics...

Name: Lightning Bolt

Cost: 1

Type: Instant

Text: Lightning Bolt deals 3 damage to any target.

Number of Printings: 37

**ASSISTANT**   0.55                                                    Real Price 0.51

**USER**

Guess the price of a trading card with the following statistics...

Name: Thragtusk

Cost: 5

Type: Creature - Beast

Text: When Thragtusk enters the battlefield, you gain 5 life.

When Thragtusk leaves the battlefield, create a 3/3 green Beast creature token.

Number of Printings: 9

Power: 5

Toughness: 3

**ASSISTANT**

0.26.5

- Sometimes doesn't work as shown above
- Also, I forgot to include rarity in the training data

# Questions…