

Board Game Recommendation System

...

Patrick Beben
Dataworks MD
April 13, 2023

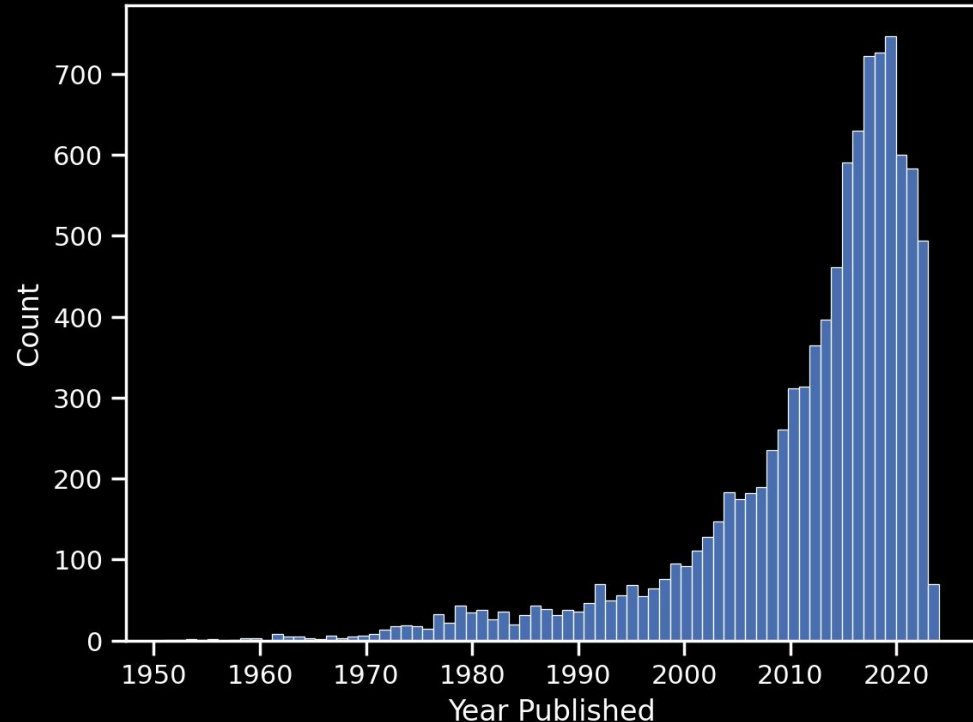
Overview

- Motivation and Background
- Getting and Prepping the Data
- Building our Model
 - TF-IDF on Long Description
 - Euclidean Distances on Categorical Data
- Comparing the Models
- Further consideration
- References and Resources



Board Games - An expanding hobby

- The boardgames hobby has seen tremendous growth in the past twenty years
- Pre-pandemic, over seven hundred new games were released every year.
- The number of choices are dizzying
- Discoverability is an issue



Gathering Data

- Boardgamegeek.com (BGG) was founded in 2000 and has been the online hub of board games on the internet
- Growing with the hobby, the database currently has over 350,000 games in it's database
- Users and BGG have a need to make more games they might be interested in discoverable
- Users would stay engaged longer and find the site more useful
- Take data from the top 10 thousand rated games in the database

Recommendations by Classifications



Recommendation by Classifications

- Each game in the BGG database has a number of classifications describing different aspects
- These correspond to categories, families and mechanics
- Recommendations will try to find games with similar sets of classifications

CLASSIFICATION

Type

[Abstract](#) 

Category

[Abstract Strategy](#)

Mechanisms

[Grid Movement](#)

[Pattern Movement](#)

[Square Grid](#)

[Static Capture](#)

Family

[Category:](#)
[Combinatorial](#)

[Traditional Games:](#)
[Chess](#)

[+ 27 more](#)

CLASSIFICATION

Type

[Family, Strategy](#) 

Category

[Economic](#)

[Negotiation](#)

Mechanisms

[Dice Rolling](#)

[Hexagon Grid](#)

[Income](#)

[Modular Board](#)

[Network and Route Building](#)

[+ 4 more](#)

Family

[Animals: Sheep](#)

[Components:](#)
[Hexagonal Tiles](#)

[Components: Wooden pieces & boards](#)

[Digital](#)

[Implementations:](#)
[Board Game Arena](#)

[Game: Catan](#)

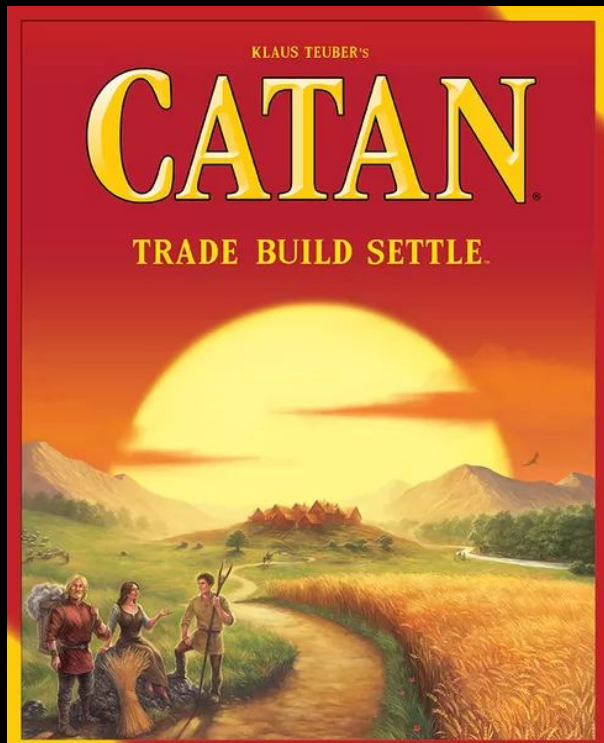
[Promotional: Promo Board Games](#)

Most Common Classifications

- The most common tag appear in just over 32% of listed games
- Our system will use ones occurring in more than 10 games
- System will turn each game into a vector with categories, mechanics, and families weighted 3, 2, and 1 respectively

	Frequency
Dice Rolling	3227
Card Game	2847
Hand Management	2584
Crowdfunding: Kickstarter	1995
Wargame	1912
Variable Player Powers	1864
Players: Two Player Only Games	1586
Fantasy	1571
Set Collection	1547
Open Drafting	1185
Hexagon Grid	1162
Area Majority / Influence	1130
Simulation	1111
Modular Board	1067

(Settlers of) Catan - Classification Vector

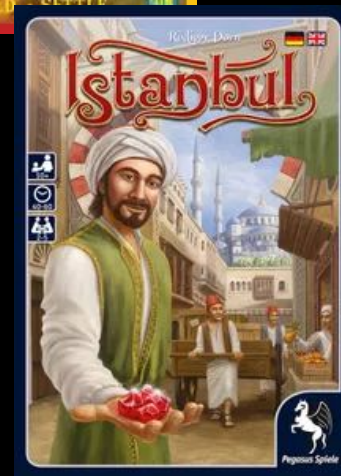
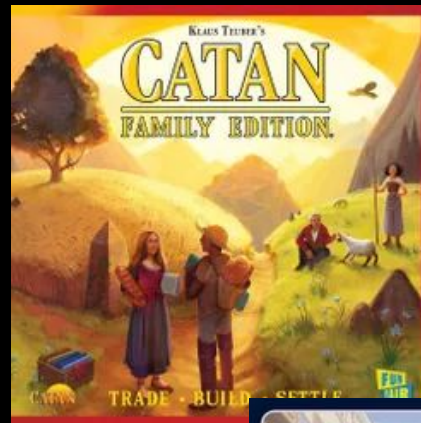


Economic	3
Negotiation	3
Dice Rolling	2
Hexagon Grid	2
Modular Board	2
Network and Route Building	2
Variable Set-up	2
Trading	2
Race	2
Income	2
Components: Wooden pieces & boards	1
Everything Else	0

Note: Some of categorizations have been filtered out since they are too rare

Catan Recommendations

1. Catan: Family Edition(2012) - 5.631:
 - Relive the rise and fall of Inkas, Catan style!
2. Catan Histories: Rise of the Inkas(2018) - 5.787:
 - Relive the rise and fall of Inkas, Catan style!
3. Offshore(2019) - 5.628:
 - Semi-cooperatively build oil reserves off the coast of Norway.
4. Vernissage(1993) - 5.577:
 - A game of speculation and dirty tricks in the art world.
5. Catan: Ancient Egypt(2014) - 5.788:
 - Gather your resources and build settlements in the fertile Nile Valley!
6. Istanbul(2014) - 7.401:
 - Send your assistants through the shops of Istanbul as you race to acquire rubies.
7. Franchise(2018) - 5.835:
 - Spread your franchise across USA to control regions & establish monopolies in cities.
8. De Kolonisten van de Lage Landen(2009) - 5.601:
9. Ponzi Scheme(2015) - 6.478:
 - Use funds from your investors to acquire industries, but can you pay them back?
10. Yak(2022) - 5.651:
 - Build a great and colorful stone tower to guide yaks in the Himalayas.



Recommendations by Description



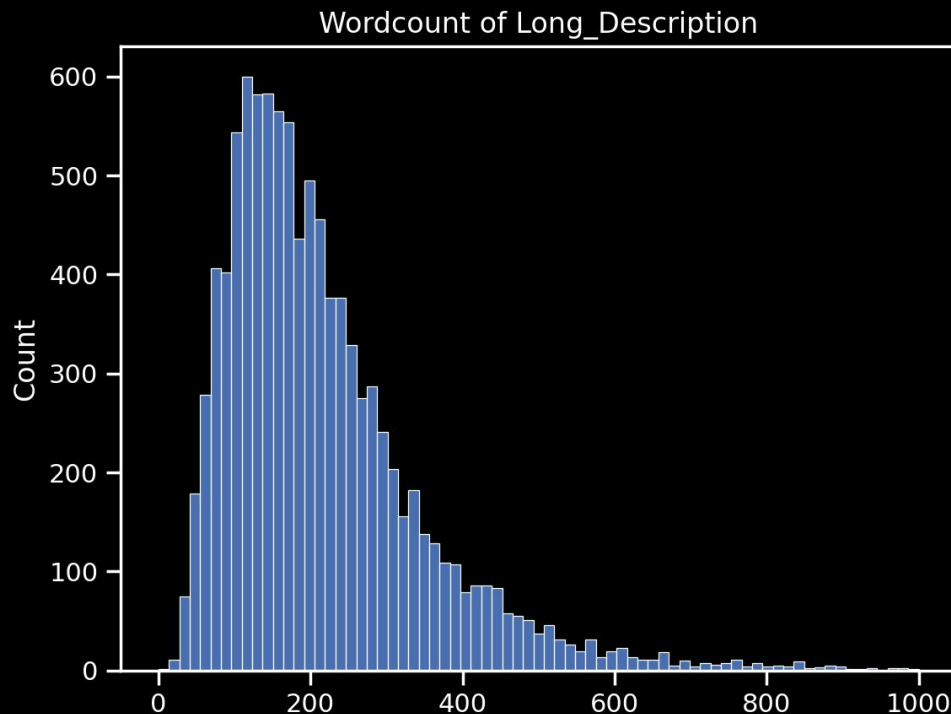
Recommendation Based on Descriptions

- Each game in the database has a lengthy description, sometimes multiple paragraphs
- Often written by publishers with some provided by users
- Build our recommendation system around TF-IDF (term frequency - inverse document frequency) vectorization of the descriptions
- We will measure the cosine similarity between vectors
- This system will work with the assumption that every description is written to highlight the important qualities of each game

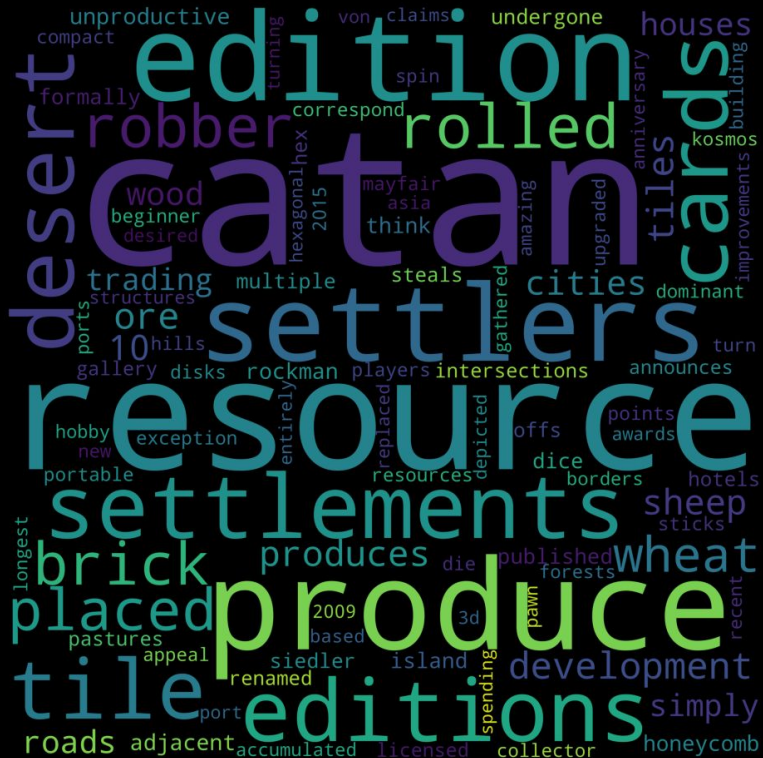


Detailed Descriptions

- The vast majority of games have longer descriptions of over two hundred words
- There are 49282 words in the vocabulary used in the descriptions



(Settlers of) Catan - TF-IDF

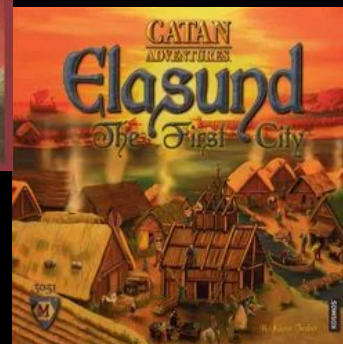
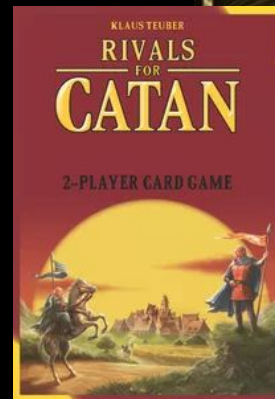


- TF-IDF measures a combination of how much a word occurs in a given description in addition to its rarity in other descriptions
- Here the important words seem to be: catan, resource, settlements
- Some words are thematic: Catan, settlers, desert
- Some are mechanic: resource, produce, cards, rolled

Catan Recommendations

1. The Settlers of Zarahemla(6778) - 5.795: nan
2. Simply Catan(24511) - 5.615: nan
3. Struggle for Catan(91534) - 5.85: Gather, trade and then spend resources to grow your own Catanian empire.
4. Settlers of Catan: Gallery Edition(38821) - 5.728: nan
5. Rivals for Catan(66056) - 6.595: Duel to build the best kingdom in different eras of Catan.
6. Catan: Family Edition(147240) - 5.637: nan
7. Elasund: The First City(19526) - 6.209: Build & fortify city together. Make sure that pirates will plunder others, not you.
8. Catan: Traveler – Compact Edition(172994) - 5.737: Collect and trade resources to build up the island of Catan in this portable version.
9. De Kolonisten van de Lage Landen(52825) - 5.602: nan
10. Catan Dice Game(27710) - 5.62: Roll to build settlements, roads, cities and hire knights.

Recommendations here are almost entirely dominated by Catan variants.



Results Comparisons

...

Chess

Categorizations

1. International Checkers(26952) - 5.614: Checkers on a 10 by 10 board.
2. SHŌBU(272380) - 6.356: Use your stones to push all your opponent's stones off any one of four boards to win!
3. Shogi(2065) - 6.331: Deploy captured pieces for rapid checkmates in Japan's traditional "general's game."
4. Chess960(12608) - 5.629: Disrupt book knowledge with a different force structure every game.
5. Camelot(5251) - 5.637: An abstract game of jumping and capture enlivened by medieval knights and men-at-arms

TF-IDF Similarity

1. No Stress Chess(19918) - 5.576: Play the world's greatest game, simplified for youngsters!
2. Martian Chess(19803) - 5.649: Capture your opponent's pieces before they cross the Canal and become yours.
3. Six Making(149910) - 5.715: Build chess figures to checkmate your opponent.
4. Bughouse Chess(14188) - 5.795: Team play chess variant where captures are given to allies.
5. Shuuro(38764) - 5.588: Discard the traditional chess set and build a custom army to conquer the board.

Scrabble

Categorizations

1. Upwords(1515) - 5.62: A 3-dimensional word game. Spell out, stack up and score high!
2. Super Scrabble(11182) - 5.679: nan
3. Scarabeo(12747) - 5.635: nan
4. A Little Wordy(335869) - 5.787: A sneaky tile-unscrambling game for aspiring geniuses and would-be word wizards.
5. SHH(164158) - 5.722: Cooperate in silence to spell words and rid your hands of as many letters as possible

TF-IDF Similarity

1. Scarabeo(12747) - 5.635: nan
2. Upwords(1515) - 5.62: A 3-dimensional word game. Spell out, stack up and score high!
3. Das perfekte Wort(326112) - 5.59: Find the perfect word for the perfect score.
4. Prolix(39635) - 5.626: nan
5. Kiitos(300509) - 5.562: Form words with letters that may or may not be in your opponents' hands.

Gloomhaven

Categorizations

1. Frosthaven(295770) - 5.716: Adventure in the frozen north and build up your outpost throughout an epic campaign.
2. Gloomhaven: Jaws of the Lion(291457) - 8.262: Vanquish monsters with strategic cardplay in a 25-scenario Gloomhaven campaign.
3. Bardsung(322524) - 5.802: Inspire the Songs. Live your Legend.
4. Middara: Unintentional Malum – Act 1(169427) - 6.972: Co-op dungeon crawler in unusual fantasy setting with massive campaign and storyline.
5. Mage Knight Board Game(96848) - 7.87: Build your hero's spells, abilities, and artifacts as you explore & conquer cities.

TF-IDF Similarity

1. Combat!(253652) - 5.771: A solitaire game of man to man combat in the 20th century.
2. Night of Man(169986) - 5.576: A card-driven, tactical game in post-alien-invasion-of-Earth.
3. Destinies(285192) - 7.073: Characters explore a medieval-fantasy land as they race to fulfill their destiny.
4. Avanti(91668) - 5.577: nan
5. Dungeons & Dragons: Temple of Elemental Evil Board Game(172220) - 6.367: Explore the monster-infested dungeons beneath the Sword Coast.

Conclusion

- Categorization consistently gives strong recommendations
- TF-IDF approach more consistently identifies similar theming sometimes to a fault of discovering something really new
- Content-based recommendation systems like this avoid the “cold start” problem of collaborative filtering
- Downside is the database needs a lot of maintaining for recommendations to be good.
- Ultimately, quality evaluation should come from subject matter experts and A/B testing

Further Development

- Incorporate BGG ratings into consideration
- Allow real-time querying of the API to get suggestions for games form outside of the top 10k
- Better interface/dashboard for exploring suggestions
- Refine categorical distance calculation to not penalize missing categories when comparing games

Anything already implemented by <https://recommend.games/> by Markus Shepherd

References and Resources

Machine Learning Basics with the K-Nearest Neighbors Algorithm

<https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761>

Beginners Guide to Content Based, includes vector based and TF-IDF

<https://www.analyticsvidhya.com/blog/2015/08/beginners-guide-learn-content-based-recommender-systems/>

BGG Games Dataset

<https://www.kaggle.com/datasets/threnjen/board-games-database-from-boardgamegeek>



Link to a Google doc with lots of more resources

Thank you!
Questions?

...

Patrick Beben
plbeben@gmail.com