



# Predicting Board Game Ratings

DSA 601 Spring 2021



# So... Why Board Games?

- Bring families and friends together - there are games for all ages and skill levels
- Sense of pride and accomplishment in gameplay and game collections
- Brought a lot of joy to homes during the COVID-19 pandemic



# Quick Stats

- Board gaming has become an incredibly popular hobby in the past few decades
- 75,867 new games since 2000
- Over 5,000 new games per year since 2014
- 7,158 new games just in 2019



# Main Data Source

Board Game Geek

<https://boardgamegeek.com/>



- Web scrape of 119,699 games
- 82 attributes
  - Number players, game type, complexity, play time, game cover art, etc.
- Games from 1500 BC to now

## QUICKBAR



Add shortcuts to your favorite games, designers, forums, or any other page within the Geek.

Add Current Page

## RECENTLY VIEWED








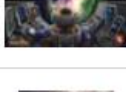
- Gloomhaven
- Deep Deep Dungeon
- Party Game
- Oceans
- Evolution: Climate
- Azul
- See All »

## THE HOTNESS

### GAMES

- Red Rising 2021
- Sleeping Gods 2021
- Res Arcana 2019
- Root: The Marauder Expansion 2021
- Outlive

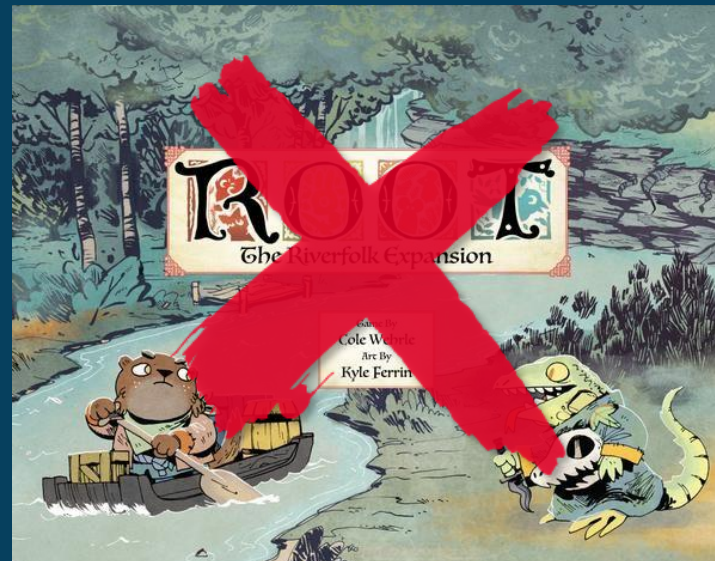
## Board Game Rank

	Title	Your Rating	Geek Rating	Avg Rating
1	 <b>Gloomhaven</b> (2017) Vanquish monsters with strategic cardplay. Fulfill your quest to leave your legacy!	N/A	8.553	8.79
2	 <b>Pandemic Legacy: Season 1</b> (2015) Mutating diseases are spreading around the world - can your team save humanity?	N/A	8.464	8.61
3	 <b>Brass: Birmingham</b> (2018) Build networks, grow industries, and navigate the world of the Industrial Revolution.	N/A	8.368	8.66
4	 <b>Terraforming Mars</b> (2016) Compete with rival CEOs to make Mars habitable and build your corporate empire.	N/A	8.286	8.43
5	 <b>Twilight Imperium: Fourth Edition</b> (2017) Build an intergalactic empire through trade, research, conquest and grand politics.	N/A	8.247	8.70
6	 <b>Through the Ages: A New Story of Civilization</b> (2015) Rewrite history as you build up your civilization in this epic card drafting game!	N/A	8.189	8.43
7	 <b>Gaia Project</b> (2017) Expand, research, upgrade, and settle the galaxy with one of 14 factions.	N/A	8.168	8.49
8	 <b>Gloomhaven: Jaws of the Lion</b> (2020) Vanquish monsters with strategic cardplay in a 25-scenario	N/A	8.166	8.89

# General Cleaning

- Use only games newer than 2010
- Remove any games marked “unpublished”
- Remove any “expansion to base game”

That leaves 39,600 games  
for modeling



# Adding Features: **Sentiment**



Natural language processing to determine positive, negative, or neutral sentiment from the Game Description

Vanquish monsters with strategic card play. Fulfill your quest to leave your legacy!  
[Gloomhaven]

Survive an alien-infested spaceship but beware of other players and their agendas.  
[Nemesis]

Solve puzzles to escape a haunted house, a submarine, and search for pirate gold.  
[Unlock! Mystery Adventures]

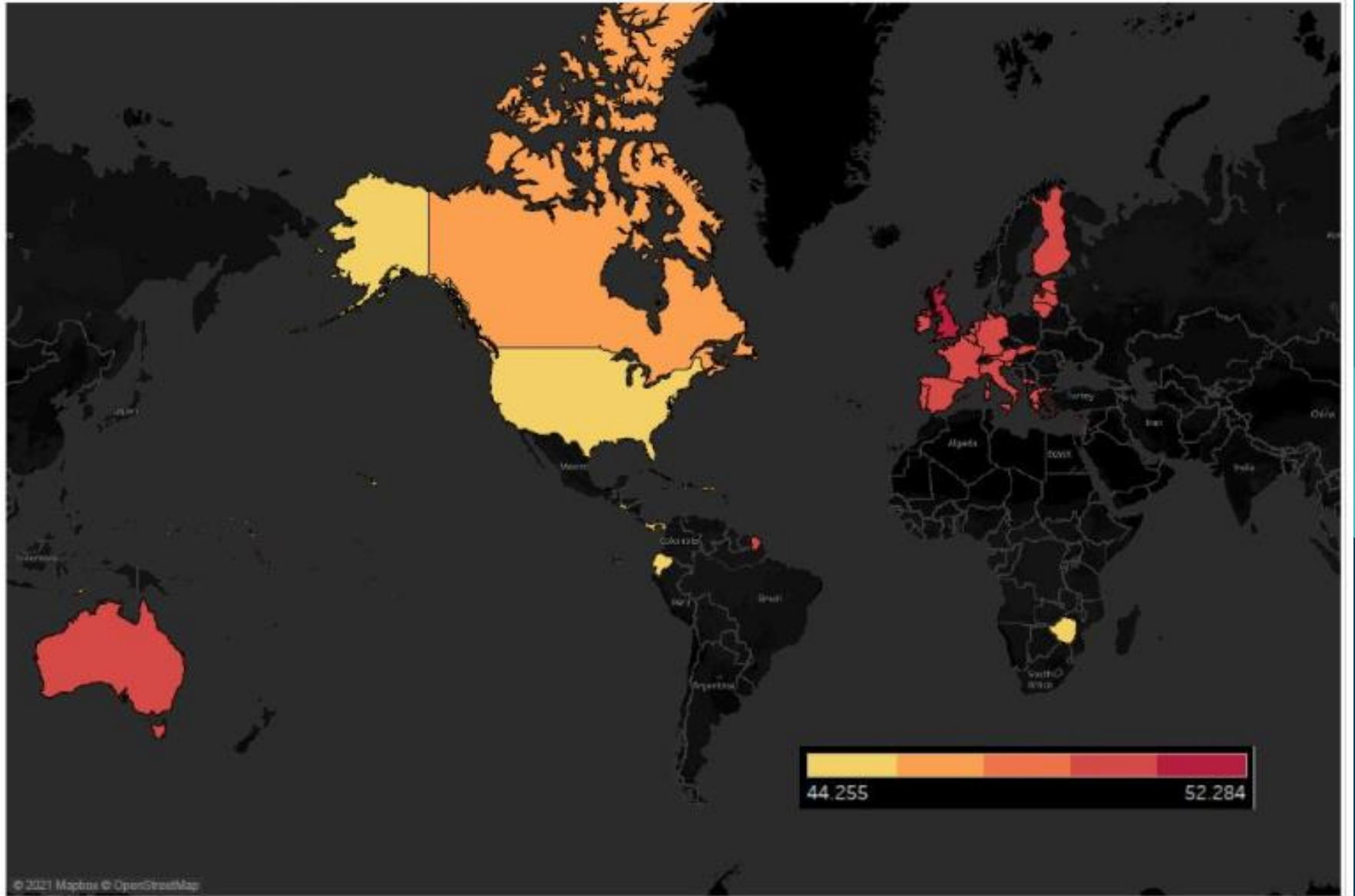
# Used Games:

What is the demand in different regions?

The UK seems to spend a lot of money on games!



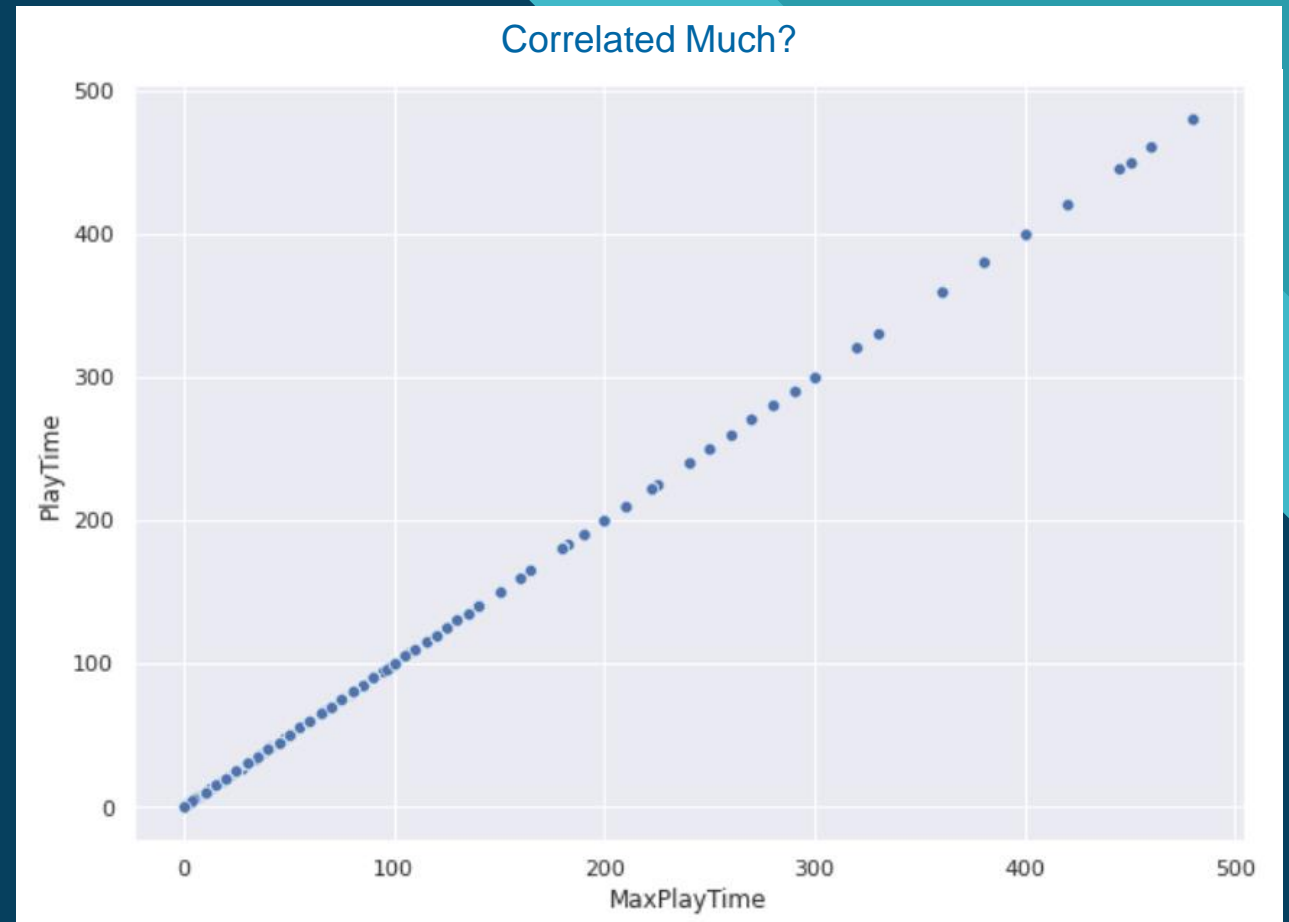
Ave Sale Price by Currency (Converted to USD)



# Correlations



- Used correlation matrix to find features that were highly correlated to each other
- If we left these in, they could add noise to our models and reduce accuracy
- **Play Time** and **Max Play Time** were 100% correlated, so dropped Max Play Time





# One-Hot Encoding

- Need to convert categorical features into a format that the models will understand
  - Sentiment (Positive, Negative, Neutral)
  - Year (2016-2020)
  - Categories (87)

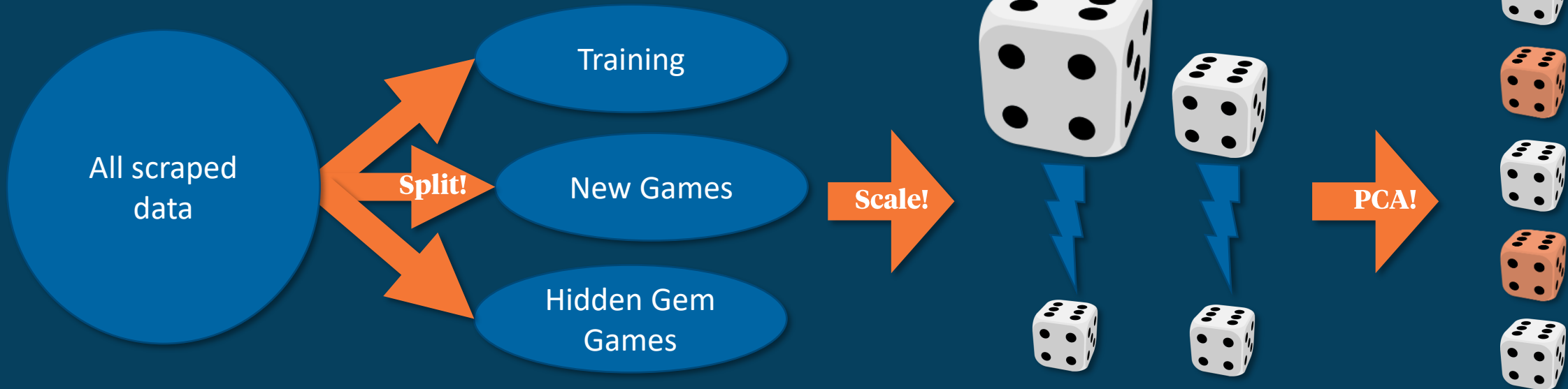
Name	Categories
Forbidden Waters	Adventure, Pirates
Mice & Mystics	Adventure, Animals



Name	Adventure	Pirates	Animals
Forbidden Waters	1	1	0
Mice & Mystics	1	0	1

# Split, Scale, PCA

- Training dataset with valid Geek Ratings separated from new data and gem data
- Data scaled so features are in the same “currency”
- PCA (Principal Component Analysis) to distill the most important features to help the models



# Many Different Models Used



## Predicting Geek Rating (Regression)

- Random Forest (x3)
- XGBoost
- Lasso (x2)
- Ridge (x2)
- Elastic net

## Predicting “Good or Bad Game”

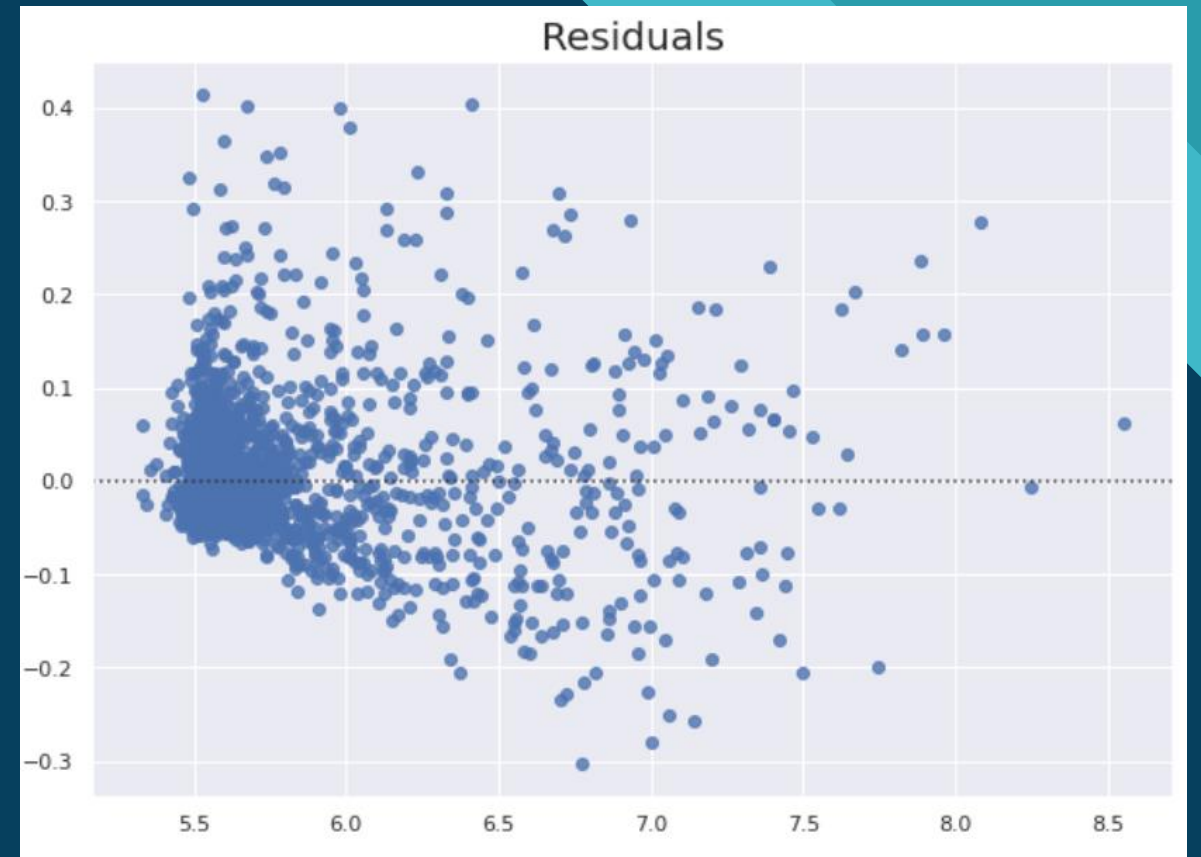
- Logistic Regression

## Classify into “Received Valid Geek Rating or Not”

- Decision Tree Classifier

# Random Forests 1 & 2

- Approx. 7% error in both training & validation
- With and without split, scale, PCA
- Geek Rating range 5.48 - 7.31



# Logistic Regression



- Geek Rating 6 or greater: “Good” game
- Less than 6: “Bad” game
- Approx. 5% error
- Used as a modifier to regression models - if a game received a “Good” label, this was like a gold star of recognition

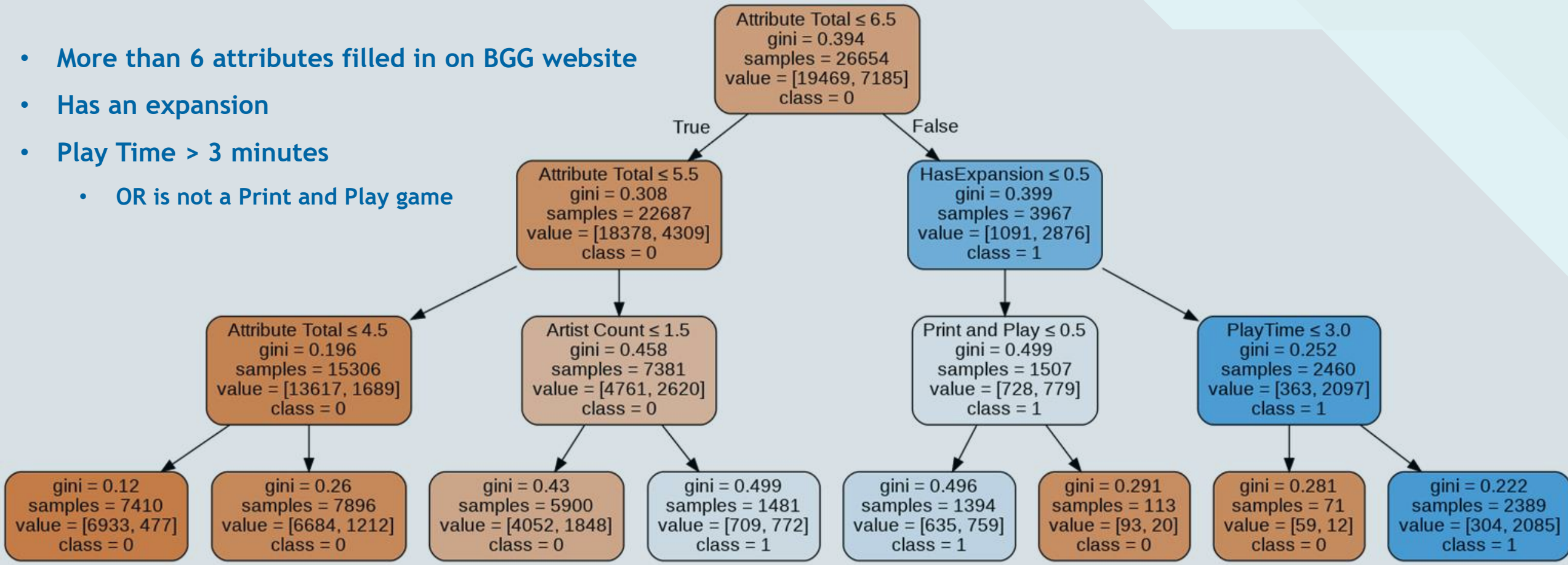
	Predicted Bad	Predicted Good	
Actual Bad	1,282	27	1,309
Actual Good	43	190	233
	1,325	217	

# Decision Tree Classifier: Received Valid Geek Rating or Not

- Model identified which attributes are most important for becoming a valid Geek Rated game
- 80% accuracy



- More than 6 attributes filled in on BGG website
- Has an expansion
- Play Time > 3 minutes
  - OR is not a Print and Play game




# Final Model

All models were informative, but we selected

## Random Forest Regression 1

- Least overfit
- Higher “riskier” Geek Rating predictions
- RMSE 0.11 (on average off by ~0.11 points)



# Power BI Dashboard Demo



# Welcome!

There are so many board games to choose from, yet so little time.

*What are the best games overall?*

*What games are the best for their price?*

*What are the highest rated 2 player games?*

With the power of machine learning and data science, this dashboard helps answer those questions and more.

There are 4 main pages:

**Current Games:** games with valid Geek Ratings (weighted average rating from 0-10 for games with 30 or more user ratings - this data was used to train the models)

**New Games:** new, upcoming, and unreleased games for 2021 and newer

**Hidden Gems:** games older than 2021 but never received a valid Geek Rating - the model was applied to these games to see if there are any games with potential

**Game Finder:** combines the above 3 datasets and provides filters of the various game attributes to help you find a good game to play!

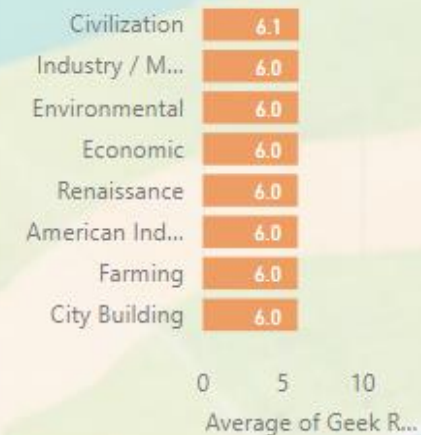
This dashboard was created by 4 students in the Data Science & Analytics master's program at Buffalo State College. It is part of the final project for Machine Learning Models in Python during the spring 2021 semester. Data was scraped from [boardgamegeek.com](https://boardgamegeek.com) for games published in 2010 and newer.

Happy gaming!

# Current BGG Geek Ratings

Name	Geek Rating	Year Published
Gloomhaven	8.55	2017
Pandemic Legacy: Season 1	8.46	2015
Brass: Birmingham	8.37	2018
Terraforming Mars	8.29	2016
Twilight Imperium: Fourth Edition	8.25	2017
Through the Ages: A New Story of Civilization	8.19	2015
Gaia Project	8.17	2017
Gloomhaven: Jaws of the Lion	8.17	2020
Star Wars: Rebellion	8.16	2016
Great Western Trail	8.12	2016
War of the Ring: Second Edition	8.10	2012
Spirit Island	8.10	2017
Scythe	8.08	2016
Terra Mystica	7.99	2012
7 Wonders Duel	7.98	2015
Concordia	7.96	2013
The Castles of Burgundy	7.95	2019
Wingspan	7.94	2019
A Feast for Odin	7.93	2016
Arkham Horror: The Card Game	7.93	2016

## Categories by Average Rating



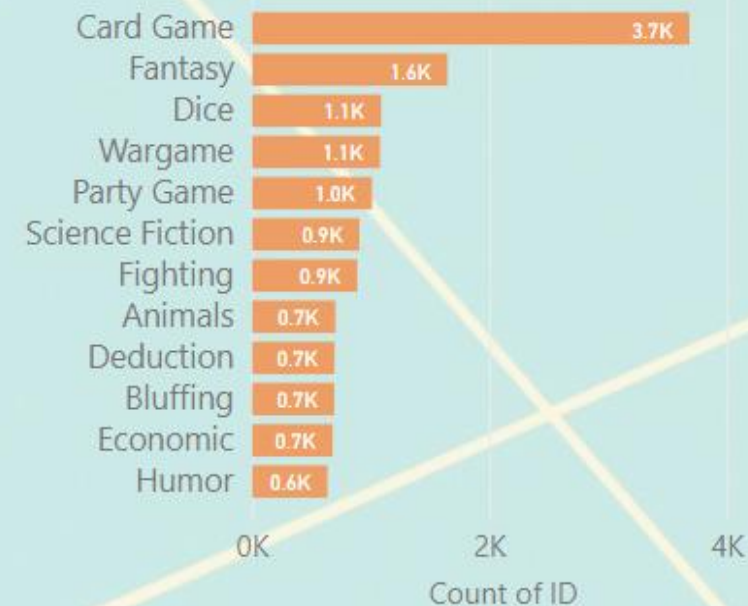
5.78

Average of Geek Rating

10K

Total Geek Rated Games

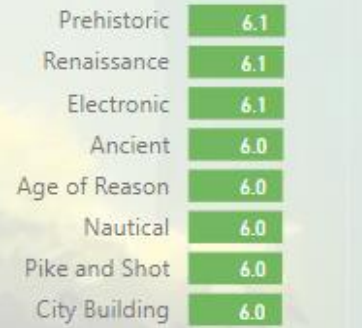
## Game Categories



# Which new games look promising?

Name	Predicted Rating	Year Published
Frosthaven	7.31	2021
Long Shot: The Dice Game	6.76	2021
Batman: The Animated Series Adventures â€” Shadow of the Bat	6.68	2021
Rallyman: DIRT	6.57	2021
Arkeis	6.57	2021
Kemet: Blood and Sand	6.52	2021
Dark Ages: Holy Roman Empire	6.47	2021
Endless Winter: Paleoamericans	6.45	2021
Castles of Mad King Ludwig: Collector's Edition	6.41	2021
Stadium: A Sport Odyssey	6.41	2021
The King's Armory: 7th Anniversary Edition	6.40	2021
Quest: Avalon Big Box Edition	6.39	2021
Merchants Cove	6.39	2021
Dominant Species: Marine	6.38	2021
Ankh: Gods of Egypt	6.38	2021
Alhambra: Big Box (Second Edition)	6.37	2021
Tinners' Trail	6.37	2021
Trinidad	6.34	2021
Castle Von Loghan	6.34	2021
Fluxx	6.22	2021

## Categories by Average Rating



0 5 10  
Average of Predicti...

5.81

Average Predicted Rating

1522

Total New Games

## Game Categories



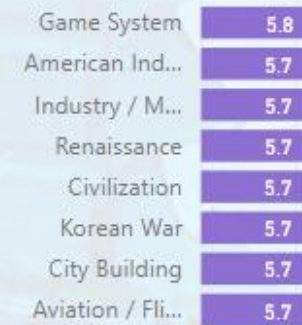
0 200 400

Count of ID

# Hidden Gems: Games that never received a valid Geek Rating

Name	Predicted Rating	Year Published
Mombasa (Limited Edition)	7.30	2015
Tiny Epic Kingdoms: 2nd Edition â€” Deluxe Edition	6.69	2015
SimplyClever.Cards Functions 2	6.67	2017
Age of Dogfights: WW1	6.66	2020
Dragon Ball Z: The Miniatures Game	6.66	2017
Cowboys II: Cowboys & Indians Edition	6.65	2020
GAMBONI	6.65	2017
SimplyClever.Cards Functions 1	6.64	2017
SimplyClever.Cards Derivatives 1	6.62	2017
All Things Zombie: Final Fade Out	6.58	2012
SimplyClever.Cards Division 3	6.58	2017
SimplyClever.Cards Derivatives 2	6.57	2017
Tiny Epic Kingdoms: Deluxe Edition	6.57	2014
SimplyClever.Cards Subtraction 2	6.56	2017
SimplyClever.Cards Division 2	6.55	2017
SimplyClever.Cards Multiplication 3	6.53	2017
Nemo's War: The Ultimate Edition	6.52	2020
SimplyClever.Cards Fractions 1	6.51	2017
SimplyClever.Cards Percentages 2	6.51	2017

## Categories by Average Rating



0 5 10  
Average of Predicti...

5.59

Average of Prediction RF1

71K

Total Hidden Games

## Game Categories



0K 5K 10K

Count of ID

**MinPlayers**     
**MaxPlayers**     
**Playtime (Hours)**     
**GameCategory** Children's Game   
**Label** All   
 Search by Name

Name	Current Geek Rating	Predicted Rating	Label	USD Average
Doublehead Kids	1.00	5.68	Gem	12.99
Get a Grip: The No Thumbs Challenge Game	1.00	5.50	Gem	10.00
Activity Franz Ferdinand	1.00	5.49	Gem	
Big Picture Apples to Apples	5.50		Current	
Blink of Death	1.00	5.55	Gem	
Charades: Kid Edition	1.00	5.49	Gem	
Chazz	1.00	5.52	Gem	
Doodle Delirium	1.00	5.51	Gem	
Dragon Bee	1.00	5.55	Gem	
Hand Made Wonders	5.49		Current	
HINT Junior	1.00	5.52	Gem	
JELL-O Jiggler Slap Game	1.00	5.65	Gem	
Kids Create Absurdity	1.00	5.48	Gem	
Moobie	1.00	5.51	Gem	
Morphology Junior	1.00	5.57	Gem	

**Total\_Converted\_USD\_Ave**